Encyclopedia of Sustainability Science and Technology Series
Editor-in-Chief: Robert A. Meyers

SPRINGER REFERENCE

James W. LaMoreaux *Editor*

Environmental Geology

A Volume in the Encyclopedia of Sustainability Science and Technology, Second Edition

Springer

Encyclopedia of Sustainability Science and Technology Series

Editor-in-Chief
Robert A. Meyers

The Encyclopedia of Sustainability Science and Technology series (ESST) addresses the grand challenge for science and engineering today. It provides unprecedented, peer-reviewed coverage in more than 600 separate articles comprising 20 topical volumes, incorporating many updates from the first edition as well as new articles. ESST establishes a foundation for the many sustainability and policy evaluations being performed in institutions worldwide.
An indispensable resource for scientists and engineers in developing new technologies and for applying existing technologies to sustainability, the Encyclopedia of Sustainability Science and Technology series is presented at the university and professional level needed for scientists, engineers, and their students to support real progress in sustainability science and technology.
Although the emphasis is on science and technology rather than policy, the Encyclopedia of Sustainability Science and Technology series is also a comprehensive and authoritative resource for policy makers who want to understand the scope of research and development and how these bottom-up innovations map on to the sustainability challenge.

More information about this series at https://link.springer.com/bookseries/15436

James W. LaMoreaux
Editor

Environmental Geology

A Volume in the Encyclopedia of Sustainability Science and Technology, Second Edition

With 231 Figures and 32 Tables

Springer

Editor
James W. LaMoreaux
PELA GeoEnvironmental
Tuscaloosa, AL, USA

ISBN 978-1-4939-8786-3 ISBN 978-1-4939-8787-0 (eBook)
ISBN 978-1-4939-8788-7 (print and electronic bundle)
https://doi.org/10.1007/978-1-4939-8787-0

This Springer imprint is published by the registered company Springer Science+Business Media, LLC, part of Springer Nature.
The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Series Preface

Our nearly 1000-member team recognizes that all elements of sustainability science and technology continue to advance as does our understanding of the needs for energy, water, clean air, food, mobility, and health, and the relation of every single aspect of this vast and interconnected body of knowledge to climate change. Our Encyclopedia content is at a level for university students, professors, engineers, and other practicing professionals. It is gratifying for our team to note that our online First Edition has been heavily utilized as evidenced by over 500,000 downloads which of course is in addition to scientists' utilization of the Encyclopedia and individual "spin-off" volumes in print.

Now we are pleased to have a Living Reference on-line which assures the sustainability community that we are providing the latest peer-reviewed content covering the science and technology of the sustainability of the earth. We are also publishing the content as a Series of individual topical books for ease use by those with an interest in particular subjects, and with expert oversight in each field to ensure that the second edition presents the state-of-the-science today. Our team covers the physical, chemical and biological processes that underlie the earth system including pollution and remediation and climate change, and we comprehensively cover every energy and environment technology as well as all types of food production, water, transportation and the sustainable built environment.

Our team of 15 board members includes two Nobel Prize winners (Kroto and Fischlin), two former Directors of the National Science Foundation (NSF) (Colwell and Killeen), the former President of the Royal Society (Lord May), and the Chief Scientist of the Rocky Mountain Institute (Amory Lovins). And our more than 40 eminent section editors and now book editors, assure quality of our selected authors and their review presentations.

The extent of our coverage clearly sets our project apart from other publications which now exist, both in extent and depth. In fact, current compendia of the science and technology of several of these topics do not presently exist and yet the content is crucial to any evaluation and planning for the sustainability of the earth. It is important to note that the emphasis of our project is on science and technology and not on policy and positions. Rather, policy makers will use our presentations to evaluate sustainability options.

Vital scientific issues include: human and animal ecological support systems, energy supply and effects, the planet's climate system, systems of agriculture, industry, forestry, and fisheries and the ocean, fresh water and human communities, waste disposal, transportation and the built environment in general and the various systems on which they depend, and the balance of all of these with sustainability. In this context, sustainability is a characteristic of a process or state that can be maintained at a certain level indefinitely even as global population increases toward nine billion by 2050. The population growth, and the hope for increase in wealth, implies something like a 50% increase in food demand by as early as 2030. At the same time, the proportion of the population that lives in an urban environment will go up from about 47% to 60%. Global economic activity is expected to grow 500%, and global energy and materials use is expected to increase by 300% over this period. That means there are going to be some real problems for energy, agriculture, and water, and it is increasingly clear that conflicting demands among biofuels, food crops, and environmental protection will be difficult to reconcile. The "green revolution" was heavily dependent on fertilizers which are manufactured using increasingly expensive and diminishing reserves of fossil fuels. In addition, about 70% of available freshwater is used for agriculture. Clearly, many natural resources will either become depleted or scarce relative to population.

Larkspur, CA, USA
March 2019

Robert A. Meyers, Ph.D.
Editor-in-Chief

Volume Preface

The discipline of environmental geology provides academicians, government entities, scientists, researchers, land use planners, and the general public with a comprehensive approach to planning for natural resource utilization to ensure protection of persons and property. This *Environmental Geology* volume of the encyclopedia examines unique approaches to tackling widely known problems and highlights the variety of issues and techniques which the ever-changing discipline of environmental geology must incorporate in day-to-day cases.

This volume is divided into categories to address these concerns. Two of the "hottest topics" are in "Land and the Environment." These relate to karst, or limestone, terranes that are subject to sinkholes affecting infrastructure and water supplies and lands that are affected by subsidence from overdrafting. Unique applications of environmental geology are being developed to address these problems and involve the public in doing so.

"Water and the Environment" covers a broad range of topics, including surface water geochemistry, impacts on groundwater from radioactive wastes, and groundwater modeling. Two other aspects include intrusion of saline waters in coastal areas impacting water quality, and mine water inrush leading to loss of life. Water in loess also has serious potential impacts in urban areas in terms of damage to infrastructure and potential harm to human life. Another area deals with mineral and thermal water. The bottled water industry is burgeoning and creating conflicts in relation to who owns or needs the water and the utilization of plastic bottles to package the water.

How do all of the above relate to "Environmental Remediation and Sustainability"? Much is being published to meet the demand of this thirst for knowledge. This includes the Second Edition of Springer's Encyclopedia of Sustainability Science and Technology series of which this volume is a part, and the relatively new journal *Sustainable Water Resources Management* edited by me and published by Springer. Topics addressed in this category vary from desertification and its encroachment on arable lands to impacts of offshore drilling on marine life. Geochemical modeling is being applied across many of these areas to help develop cost-effective alternatives. Carbon sequestration is receiving much interest and research funds to explore ways to remove CO_2 from the atmosphere. As population shifts from lands that are increasingly becoming more arid to cities, more resources are needed to sustain them. With this migration of stakeholders, more planning is required for sustainable resource management and minimization of infrastructure failures by preventing them before they happen, or, if necessary, remediating them afterward.

"Construction and the Environment" incorporates the need for and building of water supplies, wastewater treatment facilities, offices, dwellings, and grounds to handle this upwell of population. In this regard, the balance between the need to provide people with these and still preserve the environment must be addressed. This movement of people also means many more people are susceptible to natural and manmade disasters. Guidelines have been and are continuing to be developed to minimize damages and injury or death from same. In addition to development of scientific and technical guidelines, the workforce and the public must be better educated about risks and how to be better prepared to deal with them whether by mining, landslides, sinkholes, flooding and sediment transport, dam failures, or others.

"Earthquakes and Volcanoes" are natural disasters which have significant, if not catastrophic, impacts on civilization and the environment. Technology is still in its infancy in warning the public about these dangers, although new and innovative ways of doing so are in development. Recently, potential risks from manmade earthquakes or induced seismicity are being addressed.

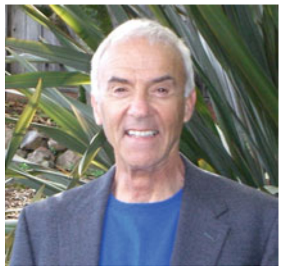
The thread throughout all of the above topics is the need for increased communication among scientists, engineers, policy-makers, regulators, and other stakeholders. Improvements are being made in technology to address these needs, and environmental geology provides a vital framework for doing so.

Tuscaloosa, AL, USA
March 2019

James W. LaMoreaux
Volume Editor

Contents

About the Editor-in-Chief



Dr. Robert A. Meyers
President: RAMTECH Limited
Manager, Chemical Process Technology, TRW Inc.
Postdoctoral Fellow: California Institute of Technology
Ph.D. Chemistry, University of California at Los Angeles
B.A., Chemistry, California State University, San Diego

Biography

Dr. Meyers has worked with more than 20 Nobel laureates during his career and is the originator and serves as Editor in Chief of both the Springer Nature *Encyclopedia of Sustainability Science and Technology* and the related and supportive Springer Nature *Encyclopedia of Complexity and Systems Science*.

Education

Postdoctoral Fellow: California Institute of Technology
Ph.D. in Organic Chemistry, University of California at Los Angeles
B.A., Chemistry with minor in Mathematics, California State University, San Diego

Dr. Meyers holds more than 20 patents and is the author or Editor in Chief of 12 technical books including the *Handbook of Chemicals Production Processes*, *Handbook of Synfuels Technology*, and *Handbook of Petroleum Refining Processes* now in 4th Edition, and the *Handbook of Petrochemicals*

Production Processes, now in its second edition (McGraw-Hill), and the *Handbook of Energy Technology and Economics*, published by John Wiley & Sons; *Coal Structure*, published by Academic Press; and *Coal Desulfurization* as well as the *Coal Handbook* published by Marcel Dekker. He served as chairman of the Advisory Board for *A Guide to Nuclear Power Technology*, published by John Wiley & Sons, which won the Association of American Publishers Award as the best book in technology and engineering.

About the Volume Editor

Dr. James W. LaMoreaux is Chairman of P.E. LaMoreaux & Associates, Inc. (PELA), an international consulting firm providing services in hydrogeology, geology, environmental sciences, and engineering. Dr. LaMoreaux serves as Editor in Chief of the international journals *Environmental Earth Sciences*, *Carbonates and Evaporites*, and *Sustainable Water Resources Management* published by Springer Nature of Heidelberg, Germany. He also serves on the advisory boards for SpringerBriefs and Theses and the Water Program and is Editor of the following Springer Book Series: Environmental Earth Sciences; Cave and Karst Systems of the World; Advances in Karst Science; and Professional Practice in the Earth Sciences. Previously he served as Editor of the newsletters of the American Institute of Hydrogeologists, the Environmental Institute for Waste Management Studies, and the Alabama Water Environment Association.

Dr. LaMoreaux serves as Chairman of the International Association of Hydrogeologists (IAH) Commission on Mineral and Thermal Water, a member of the IAH Karst Commission, and President of the IAH US National Chapter. He formerly served on the Board of Directors of the American Ground Water Trust. He is the author and/or editor of numerous technical publications, including the following books: *Environmental Hydrogeology*; *Legislative History of the Superfund Amendments and Reauthorization Act*; *Survival and Sustainability: Environmental Concerns in the 21st Century*; *Advances in Research in Karst Media*; *Hydrogeological and Environmental Investigations in Karst Systems*; *Management of Water Resources in Protected*

Areas; *H2Karst Research in Limestone Hydrogeology*; and *Thermal and Mineral Waters: Origin, Properties and Applications*.

Dr. LaMoreaux also served as editor of the *Environmental Geology* volume of Springer's Encyclopedia of Sustainability Science and Technology series. He is currently serving in the same position for the online version and the second edition to be published in 2019.

Dr. LaMoreaux currently serves as a member of the Ground Water Committee of the Water Environment Federation (WEF). He has served on the Ground Water, Hazardous Waste, Industrial Waste, Public Education, and Program Committees of the Water Environment Federation (WEF) and represented Alabama as State Director to WEF. He served as a member of the Science Advisory Committee for the US Environmental Protection Agency (USEPA)-funded Urban Waste Management and Resources Center at the University of New Orleans. He received his M.S. and Ph.D. from Syracuse University in 1970 and 1976, respectively.

Contributors

Stefan Anderberg Lund University Centre for Sustainability Studies, Lund University, Lund, Sweden

Ömer Aydan Department of Civil Engineering, University of the Ryukyus, Nishihara, Okinawa, Japan

Caitlin Barnes Educational Research, Oklahoma State University, Stillwater, OK, USA

Barry Beck Karst, P.E. LaMoreoux & Associates, Inc, Oak Ridge, TN, USA

José Joel Carrillo-Rivera Instituto de Geografía, UNAM CU, Coyoacán, Mexico

Pedro José Depetris Academia Nacional de Ciencias, Córdoba, Argentina

Shuning Dong Xian Branch, China Coal Technology and Engineering Group Corp., Shannxi, China

Jens Gutzmer TU Bergakademie Freiberg, Institute of Mineralogy and Helmholtz Institute Freiberg for Resource Technology, Freiberg, Germany

Todd Halihan Boone Pickens School of Geology, Oklahoma State University, Stillwater, OK, USA

Greg Hartman Hartman Associates Inc, Waterway Engineering and Sediment Remediation, Redmond, OR, USA

Martin Hovland Centre for Geobiology, University of Bergen, Bergen, Norway

Statoil ASA, Stavanger, Norway

Asadullah Kazi Isra University, Hyderabad, Pakistan

James W. LaMoreaux P.E. LaMoreaux & Associates, Inc, Tuscaloosa, AL, USA

Judith Lancaster Division of Earth and Ecosystem Sciences, Desert Research Institute, Reno, NV, USA

Bo Li Key Laboratory of Karst Environment and Geohazard Prevention, Guizhou University, Guiyang, China

Peiyue Li School of Environmental Science and Engineering, Chang'an University, Xi'an, Shaanxi, China

Key Laboratory of Subsurface Hydrology and Ecological Effects in Arid Region of the Ministry of Education, Chang'an University, Xi'an, Shaanxi, China

Chongxuan Liu Southern University of Sciences and Technology, Shenzheng, China

Pacific Northwest National Laboratory, Richland, WA, USA

Rui Ma School of Environmental Studies, China University of Geosciences, Wuhan, China

Jörg Matschullat TU Bergakademie Freiberg, Institute of Mineralogy and Helmholtz Institute Freiberg for Resource Technology, Freiberg, Germany

Bashir Memon P.E. LaMoreaux & Associates, Tuscaloosa, AL, USA

Petar T. Milanović Belgrade, Serbia

Harry Moore Tennessee Department of Transportation, Blaine, TN, USA

David Mouat Division of Earth and Ecosystem Sciences, Desert Research Institute, Reno, NV, USA

Curtis M. Oldenburg Energy Geosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

M. Adrián Ortega-Guerrero Centro de Geociencias, UNAM Juriquilla, Querétaro, Mexico

Samira Ouysse Cadi Ayyad University, Marrakech, Morocco

Adam Porowski Stable Isotope Laboratory, Institute of Geological Sciences Polish Academy of Sciences (INGPAN), Warszawa, Poland

Hui Qian School of Environmental Science and Engineering, Chang'an University, Xi'an, Shaanxi, China

Key Laboratory of Subsurface Hydrology and Ecological Effects in Arid Region of the Ministry of Education, Chang'an University, Xi'an, Shaanxi, China

Scott Thomas Division of Earth and Ecosystem Sciences, Desert Research Institute, Reno, NV, USA

Nick Varley Colima Exchange and Research in Volcanology, Faculty of Science, Universidad de Colima, Colima, Mexico

Craig Vogt Craig Vogt Inc, Environmental Consultants, Hacks Neck, VA, USA

Qiang Wu National Engineering Research Center of Coal Mine Water Hazard Controlling, China, China University of Mining and Technology, Beijing, China

Chunmiao Zheng Department of Geological Sciences, University of Alabama, Tuscaloosa, AL, USA

Southern University of Sciences and Technology, Shenzheng, China

Wanfang Zhou ZeoEnvironmental, LLC, Knoxville, TN, USA

Chen Zhu Department of Geological Sciences, Indiana University, Bloomington, IN, USA

Part I

Introduction

Environmental Geology: Introduction

James W. LaMoreaux
P.E. LaMoreaux & Associates, Inc, Tuscaloosa, AL, USA

The world's population continues to increase and as it does greater demand is placed on the finite resources that the earth provides. Geographic areas once judged inadequate or inappropriate for construction, water and wastewater management, or transportation routes are being used for development.

Environmental geology is the discipline which brings together these issues in a holistic approach to land use planning and utilization. This section of the encyclopedia examines some of the problems, techniques, and solutions for managing resources so the balance between development and the environment can be preserved.

Natural disasters have varying impacts depending upon the proximity of the disaster to population centers and the magnitude and extent of the event or hazard. In all cases, the primary purpose of studying these phenomena is to minimize loss of human life and secondarily to minimize loss of infrastructure and impacts on present and future developments.

Each type of hazard impacts the environment in different ways. Some pose more danger to human life than others, and some pose more hazards to infrastructure and ongoing development. This section on environmental geology will address selected types of hazardous impacts to provide a better understanding of how they occur, what can be done to minimize risk and how to plan better to address these concerns in the future.

Impacts of natural disasters including earthquakes and tsunamis and occurrence of sinkholes and landslides can have more devastating effects due to increases in world population. Two entries address earthquake issues: ► "Earthquake Faulting: Ground Motions and Deformations" and ► "Construction Planning: Environmental Impact of Foundation Studies and Earthquake Issues." Earthquakes have occurred throughout history and left their traces as manifestations of geological processes on the geological record.

Many lessons have been learned from analyzing the damages caused by earthquakes. Guidelines have been developed for design of earthquake resistant structures under different foundational and geologic conditions. Although education of the public has saved lives, it continues to be a critical need, as is research to develop better prediction and warning mechanisms.

Sinkholes and landslides are both naturally occurring and man-induced. Both hazards can cause significant damages to various structures. Protective measures can minimize the risks involved by preparing sinkhole maps which delineate the low, medium, and high risk areas subject to subsidence. If decisions are made to develop in sinkhole and landslide prone terrains, then site-specific investigations of the geology and hydrogeology can determine where to best seat structures into bedrock to minimize damages. Engineering design can address the added stresses such hazards pose and incorporate additional measures so that buildings and infrastructure can better withstand the risk.

Loess is fine-grained sediment deposited by the action of the wind on the surface of the earth. This type of soil occurs primarily in arid and semi-arid regions where rainfall is intermittent or infrequent. Due to the fine-particle structure of this soil, it is susceptible to erosion, landslide, and other potential hazards. Characteristically, loess may destabilize under certain conditions. The study of this unique terrane is of increasing importance in areas of rapid growth. ► "Water in Loess" offers insight into problems and solutions, which may occur in this specialized environment.

► "Karst Terrane and Transportation Issues" discusses different types of sinkholes, some of the causes and how to protect against occurrence.

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_932

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media, LLC, part of Springer Nature 2018, https://doi.org/10.1007/978-1-4939-2493-6_932-3

Methods for remediating damage are often unique applications of engineering geology. The study of sinkhole causation can aid in preventing future losses in sinkhole-prone terranes. ▶ "Land Subsidence in Urban Environment" reviews an emerging cause of subsidence in the burgeoning population centers of the world. Better management of water resources is required to prevent further subsidence from occurring in urban areas. Optimal utilization of water supplies must be implemented to correct existing situations and to minimize subsidence in the future.

Karst environments vary widely. Karst areas may be extremely complex and can require complex solutions. When problems with infrastructure occur in karst, detailed studies are required to establish risk-based methods for remediation. ▶ "Remediation in Karst" provides a discussion of engineering geology in karst areas and methods to minimize failures.

Overuse of water supplies in concert with drought conditions and land degradation is creating desertification conditions rapidly. Desertification affects arid, semiarid, and dry subhumid regions. Occurring on every continent, desertification impacts more than two billion people. Often these conditions occur across country borders and within areas of the world where tensions are already strained for economic, political and religious reasons. ▶ "Desertification and Impact on Sustainability of Human Systems" discusses the activities of the United Nations Convention to Combat Desertification (UNCCD) and emphasizes the need for sustainable development at the local level. Stakeholders in these inherently fragile ecosystems must be brought together to arrive at solutions that will allow continued production of ecosystem goods and services.

Not only do arid conditions require sustainable development, but urban areas do also. In ▶ "Natural Resource Flows and Sustainability in Urban Areas," cities are increasingly recognized as the most important global sustainability challenge of this century. Here too, stakeholders are critical in regard to resource access and distribution. Studies of urban metabolism are required to explore interactions among resource flows, urban transformation processes, waste streams, and quality of life. Results of these studies must be correlated to address the challenge of combining urban planning with sustainable resource management.

An integral part of continued growth and development of urban areas is detailed in ▶ "Groundwater Salinity Due to Urban Growth." Understanding of groundwater is critical in regions where increasing water demand and shortages in water supply affect available and proposed infrastructure. Knowledge of water accessibility in terms of quantity and quality can help reduce the risk of saline intrusion in coastal areas. Socioeconomic and environmental development programs require adequate political decision-making and planning at local and regional levels. Proper space planning implies that proposed land use is in agreement with what land "attitude" and its environment can provide without putting in danger sustainability.

▶ "Geochemical Modeling in Environmental and Geological Studies" is a powerful tool for investigations of water quality, as well as many other fields. Modeling is indispensable for research and investigations of environmental sustainability, science, and technology. It provides quantitative evaluation of complex processes and can predict the extent and consequences of geochemical reactions in the order of tens of thousands of years.

▶ "Mineral and Thermal Waters" have been used for drinking, bathing, medical, or religious purposes throughout time. Today, spa and bottled water industries continue these themes with modern applications. Depending on the locale, there are usually specific guidelines and requirements for bottled waters. Thermal waters utilized for geothermal energy and electrical power generation are specialized branches of science with exciting applications.

The study of surface water geochemistry provides a baseline guide for assessing future impacts to springs, streams, lakes, and reservoirs from both natural and manmade factors. Specific geologic formations have specific mineral signatures which under the influence of weathering and runoff make their individual contributions to the water chemistry. Atmospheric precipitation may also influence water chemistry, and atmospheric temperature variations can influence chemical bonding. ▶ "Fresh Water Geochemistry: Overview" discusses the mechanisms and processes of chemical change.

An excellent example of the utility of geochemical modeling in environmental sustainability, science, and engineering is illustrated in geological carbon sequestration. Much research is currently being performed on injection of carbon dioxide (CO_2) into deep geological formations as a climate mitigation tool. ▶ "Geologic Carbon Sequestration: Sustainability and Environmental Risk" provides an overview of the research and ideas about the future. National and local regulations will require risk assessments of wellbore integrity, well injectivity, and long-term performance in the order of thousands of years. The geology of each site differs and performance assessments are necessary at all stages of CO_2 storage operations: site assessment and selection, design, installation, operations and monitoring, and closure and post closure.

▶ "Induced Seismicity" reviews the phenomenon associated with injection wells since the 1960s. Fluid induced earthquakes can be potentially generated if an existing fault is overexposed to an increase in pressure from fluids, typically injected wastewaters from petroleum production. Ways to understand and better manage this effect are examined.

Drilling into deep geological formations is also taking place offshore in an activity known as Oceanic Hydrocarbon Investigation (OHI). OHI is currently increasing its production volume in an energy hungry society, and the quest for finding and exploiting underground hydrocarbon resources is being continuously improved. This is particularly true in relation to underwater mapping, inspection, and monitoring of the environment. ▶ "Marine Life Associated with Offshore Drilling, Pipelines, and Platforms" states that OHI has inflicted oil-spills on marine and coastal environments. At large, however, OHI appears to be environmentally friendly. This notion has been documented by extensive seafloor mapping and annual visual inspections of platforms, pipelines, and other infrastructure.

Environmental concerns in marine environments are also being addressed as relates to dredging of navigable waterways. Dredges of various designs have been used for many years to create and maintain navigable waterways to move people, goods, and materials. ▶ "Dredging Practices and Environmental Considerations" discusses the history of dredging from the times of the pyramids when people used long-handled dipper shovels, to when productivity gains from using animals increased digging power. In the 1800s, development of electric and steam power units enabled construction of mechanical, backhoe, and pipeline dredges. Hydraulic technology in the 1960s, utilizing hydraulic winches and hydraulic rotary cutter drives, facilitated removal of finer grade sediment. As technology improved, so did concern for the environment. Today there is much debate and research on the impacts of dredging. This concern has heightened with increased awareness of contaminated sediments.

Navigable waterways integrally relate to dam construction and movement of people, goods, and services. Construction of dams increases awareness of environmental impacts. Since time immemorial, people have looked for ways to tame surface waters to prevent floods and use water for different purposes, including transportation, irrigation, water supply, and electricity production. ▶ "Dam Engineering and Its Environmental Aspects" discusses the history of dams and their primary role to store or to divert waters. Many dam projects have been controversial or potentially disastrous. Failure of dams and their environmental impacts are the main reasons for controversy and anxiety. This is particularly true in karst environments which tend to be more sensitive to the influence of dams and reservoirs. With increasing demand on water resources and electric power, how to maintain the balance between the necessity for development and preservation of the environment is a significant concern in many areas.

Energy continues to weave a thread through this section of the encyclopedia as the chapter on mining deals with mining of coal. Also addressed are impacts from mining of rare earth minerals, gold and copper among others. ▶ "Mining and Its Environmental Impacts" details the demands of the growing world population for increasing amounts of raw materials. To find acceptance and support in society, however, any future mining activity will demand state of the art environmental management and must incorporate sustainable development. A history of mining is provided, as well as an overview of the environmental effects of mining, differentiated by its relevant phases: exploration, exploitation and processing, decommissioning and rehabilitation.

Underground mining in geologic areas with a relatively high groundwater table or groundwater under pressure is an activity with high hazard potential. Detailed site-specific studies and unique engineering solutions and techniques are required to minimize risks to the environment and insure human safety. ► "Mine Water Inrush" provides an overview of these complex problems and discusses methods of classification and study with a goal of prediction, prevention, or remediation.

Nuclear energy has its own inherent concerns which are increased in times of natural disasters. ► "Groundwater Impacts of Radioactive Wastes and Associated Environmental Modeling Assessment" discusses the long-term risks to human health and environment from radioactive waste contamination in soil and groundwater. Major sources of radioactive wastes and their impact on groundwater contamination are reviewed along with the major biogeochemical processes that control the transport and fate of radionuclide contaminants. The evolution of mathematical models designed to simulate and assess the transport and transformation of radionuclides is also described. Understanding the fate and transport of radionuclide contaminants can improve the ability to forecast contaminant destination and select cost effective remediation technologies.

Perhaps the greatest energy source known to man is from volcanoes. ► "Volcanoes of Mexico" includes an introduction to volcanoes: the reasons they form, where they are located, and the hazards presented by different types of eruption. Mexico is one of the world's most volcanic regions and a summary of the famous Mexican volcanoes is presented. Volcanic eruptions can be divided into two broad categories: explosive or effusive. Various factors combine to determine how the magma emerges, the key ingredient of which is water. With the exception of ballistics, which do not reach very far from the volcano, volcanic hazards can present a major risk to populations living close by. Volcanologists investigate deposits of previous eruptions to better understand what might happen in the future. Hazard maps can be created which indicate where eruptive products might accumulate.

► "Infrared Thermographic Imaging in Geoengineering and Geoscience" is concerned with the prediction of the time, location, and magnitude of earthquakes through implementation of an imaging technique utilizing thermographic cameras to detect radiation in the long-infrared range of the electromagnetic spectrum and then to produce images of that radiation, called thermograms. The development and interpretations of this technique are especially useful in construction siting and risk assessment studies.

Conclusion

Water, energy, minerals, and the environment all play key roles in development and future growth of the world and its population. No single aspect can be addressed without considering another. This section of the encyclopedia on environmental geology places these different factors into perspective so that scientists, politicians, economists, planners, and stakeholders can work together to develop solutions that provide an acceptable balance to development and the environment. Information on projected advances in each discipline or topic is included at the conclusion of each chapter as a glimpse toward the future.

Part II

Land and the Environment

Karst Terrane and Transportation Issues

Harry Moore[1] and Barry Beck[2]
[1]Tennessee Department of Transportation, Blaine, TN, USA
[2]Karst, P.E. LaMoreaux & Associates, Inc, Oak Ridge, TN, USA

Article Outline

Glossary

Carbonate rocks Bedrock that is composed of calcium carbonate, such as sedimentary rocks like limestone and dolostone.

Cave A natural cavity or series of galleries that is produced by solution of limestone and is large enough to be entered by man.

Centerline A term that refers to the center of a road or highway and runs the length of the road or highway.

Collapse A term used in conjunction with the formation of sinkholes; the subsidence of land or bedrock due to an underlying cavity.

Ditch line A term used in roadway design and construction that refers to a drainage ditch that is parallel to the roadway and located on either side of the road and sometimes in the median of the road.

Gabion basket A basket configuration constructed from double-twist wire meshing; typically the basket is backfilled with rock or gravel to form a semirigid retaining wall.

Geomembrane Typically a synthetic man-made material that is usually impervious to water penetration, e.g., a swimming pool liner; used in roadway construction to reinforce embankments and line ditches.

Groundwater That part of the subsurface water which is in the zone of saturation; also called phreatic water.

Grout A term that refers to thin fluid mortar; commonly used in construction and often used to fill in cavities in limestone to produce a stable foundation for a road, bridge footing, or dam.

Karst A landscape underlain by carbonate rocks, such as limestone and dolostone, that exhibit such features as sinkholes, sinking streams, caves, and subsurface drainage, e.g., central Kentucky Mammoth Cave region.

Lined ditch A term used in highway design and construction that refers to a drainage ditch that has an engineered "bottom" such as concrete, asphalt, or geomembrane.

Ponor A karst feature where a surface stream disappears into the subsurface, usually via a sinkhole or cave entrance.

Recharge A term that usually refers to the replenishing of a substance, usually used in association with the groundwater table, i.e., recharging the groundwater.

Regolith The layer or mantle of loose weathered rock and soil debris that forms the surface of the land and rests upon hard bedrock material; most commonly associated with the soil horizons overlying bedrock.

Reinforced concrete Concrete that is strengthened by using steel bars.

Roadway alignment The general configuration and direction of a roadway.

Harry Moore has retired
Barry Beck: deceased

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_206

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_206-4

Rock pad A term used in construction that refers to a layer of quarried and/or crushed rock that is used to form a hard stable layer or foundation and also provides for drainage of water seeps.

Sinkhole A circular or funnel-shaped depression in the landscape of a limestone region usually formed by the solution of limestone and the subsurface erosion of the overlying regolith; can occur rapidly and catastrophically causing great damage to roads and structures.

Solution Typically the change from a solid, such as limestone into the liquid state by its combination with a liquid; usually associated with karst.

Definition of the Subject

As roads, highways, bridges, and railroad corridors are built to better transport our society, the terrane that they are located in may dictate design and construction procedures necessary to produce a stable and safe transportation route. One such terrane is karst, a landscape that is typically underlain by carbonate rocks such as limestone and weathers chemically and physically to produce features such as sinkholes and caves.

Designing, constructing, and maintaining these transportation systems that are located in karst terrane require proactive techniques to better stabilize the road or bridge. Too often a reactive approach to dealing with karst issues is pursued by transportation agencies. Being more proactive toward karst issues can lessen maintenance costs and also protect the environment.

Introduction

Dissolution of bedrock in areas underlain by carbonate rocks (usually limestone or dolomite) results in a terrane characterized by sinkholes, sinking streams, underground (cave) streams, and springs. (The entire landscape formed in soluble rock areas is known as a karst terrane. The term *terrane* is used rather than *terrain* to include subsurface features as well as surface features.) This terrane is usually referred to as karst. Groundwater in these settings is more susceptible to contamination because surface water may pass directly into the subsurface with little or no filtration by soil. Because karst groundwater typically flows through relatively large fractures and conduits within the bedrock, it may transport contaminants rapidly from points of recharge (such as sinkholes) to distant cave streams, water wells, springs, and surface streams.

A karst landscape can be composed of gently rolling hills and valleys textured with sinkholes, cave entrances, sinking streams, and outcroppings of weathered limestone (Fig. 1) as is typically found in Kentucky and Tennessee and the Virginia/West Virginia/Pennsylvania sections of karst. Dramatic topography that exhibits tall "haystack"-type karst terrane as found in the Li River section of southwest China (Fig. 2) can be impressive and a daunting landscape for highway construction.

Existing roads in karst areas usually experience the sudden collapse of a sinkhole or the flooding of a sinkhole basin crossed by a road. In addition, issues of groundwater pollution often go unchecked or not addressed by highway agencies as well as developers and city planners.

To the geologist and geotechnical engineer, caves are more than just curious features because they reveal information about the groundwater movement and solution processes that act on carbonate rocks. The patterns in which the solution cavities and resulting caves develop generally indicate the attitude of the bedrock, fracture density, and groundwater movement. Cave passage patterns, sinkholes, and springs are important features that may contribute to an understanding of the local groundwater flow. The recognition of areas of active karst subsidence and collapse is of considerable importance to those engaged in the design and implementation of highways, especially the construction of infrastructure.

Some of the problems that have resulted from human modification of karst into subdivisions, highways, and commercial properties include the subsidence of building foundations and the collapse of lawns due to leaking swimming pools and inappropriately located septic tanks (Fig. 3). The collapse of highway surfaces (Fig. 4), ditch lines,

Karst Terrane and Transportation Issues, Fig. 1 Features such as sinkholes and cave entrances such as this Roane County, Tennessee cave entrance, often characterize karst terrane

Karst Terrane and Transportation Issues, Fig. 2 A karst region in China is characterized by remarkable landscape relief as shown here along the Li River near Guilin, China

and bridge foundations and numerous instances of flooding are also karst-related problems triggered by human construction activity. In most instances, it is the impact of human construction activity and alteration to the surface drainage that induces the collapse of a highway or house foundation or results in the flooding of a commercial strip mall.

Solution cavities developed in karst terrane characteristically have numerous open subsurface flow paths. Accordingly, any contaminating fluids are able to be rapidly dispersed along difficult-to-predict routes. This type of groundwater contamination from highway runoff is becoming a greater concern as transportation routes are expanded into karst areas.

This is particularly so in karst areas where toxic or hazardous spills along highways can directly flow from the highway into sinkholes and cave systems that recharge groundwater aquifers. Everyday highway runoff including tar, gasoline, diesel fuel, asbestos, and metals from the roadway can also contaminate groundwater supplies where the runoff flows into sinkholes affecting the groundwater quality.

Most of the remedial activity concerning karst problems has been reactive in nature. In most

Karst Terrane and Transportation Issues, Fig. 3 Sinkholes can occur most anywhere in karst areas, as evidenced by this sinkhole near Morristown, Tennessee

Karst Terrane and Transportation Issues, Fig. 4 Karst terrane can impact our highways as shown in this photo of a sinkhole along I-40 near Knoxville, Tennessee

instances it is highway maintenance personnel that tend to remediate these sinkhole collapse and runoff problems. The reactive approach to dealing with karst problems involves responding in an emergency situation.

Remedial action is usually adopted in response to karst problems. Remedial concepts used for correcting karst problems include bridging, drainage, and relocation measures. Bridging a collapse with a rock fill or a concrete structure may be one alternative. Trying to solve flooding problems may require the use of existing sinkholes for drainage outlets (possibly resulting in groundwater contamination) or even constructing special ditches to bypass sinkhole areas to access a nearby stream. Sometimes moving a section of road or relocating a house is the course of action required.

If proper consideration is given to karst problems in advance of a construction project, then proactive measures can be taken that involve the understanding of karst processes. By simply avoiding a karst area when planning a highway or developing a commercial zone or residential subdivision, one can save future agony as well as dollars for the home owner, developer, or government agency. If avoidance is not possible, then there must be certain integrated measures taken that address not only potential subsidence but also flooding and groundwater contamination.

In general, transportation issues in karst terrane deal mostly with sinkhole collapse problems, sinkhole flooding, groundwater pollution, and environmental issues dealing with caves and cave inhabitants. How these issues are dealt with usually falls into reactive measures and proactive measures.

The Character of Karst

As would be expected, the character of karst can be varied and unusual and is often dramatic. Caves and sinkhole collapse are the most common karst characters and are often the features encountered when designing and constructing highways. Sinkhole plains, disappearing streams, resurging streams, subsurface drainage, and delicate ecosystems are also characteristic of karst and present tenuous issues when locating or constructing transportation infrastructure.

The geologic character of the bedrock also plays an important role in the development of karst features. In addition to lithology, structural characteristics such as fracture density, folding, and faulting also contribute to the development of karst. Long linear belts of karst tend to develop on folded and faulted strata, whereas flat-lying strata tend to develop wider, dendritic-type expanses of karst. In addition, highly fractured strata provide more avenues for solution activity of groundwater and therefore develop into a karst landscape quicker than less fractured strata. There are numerous studies that describe the relationship between geologic structure and karst development such as [1–7].

Cave development in karst areas, like those in the East Tennessee region, central Kentucky, and other karst areas around the USA, follows the more soluble zones or lines of weakness within the host bedrock. These weak areas are generally parallel to zones of fractures (joints and faults) or bedding planes of the rock. Solution enlargement of these weak areas of fractured rock generally coalesces into interconnecting passages that form intricate caverns, some of which are profusely decorated with cave formations (speleothems or dripstone) [8, 9].

Surface expression of the subsurface conditions in karst areas is usually manifested as depressions and sinkholes (Fig. 5). The sinkholes vary in size from several feet to several hundred feet across and up to tens of feet in depth. Some of the sinkholes may have active swallets or openings into the cave environments, while others may be simply silted in and covered with vegetation. Additionally, numerous rock outcrops are present, as well as sinking streams, cave entrances, and springs. Mostly all of these conditions are present along highways in the karst areas of the USA.

Sinkhole development along highways usually results from the collapse of the residual clay soil into cavities developed in the subsurface soil due to erosion of the residual clay. As these "soil cavities" enlarge and approach the surface, the remaining soil bridge over the cavity loses strength and collapses forming a typical sinkhole collapse. The eroded soil is flushed down into the solution cavities in the bedrock. This type of induced sinkhole failure has been previously described by Donaldson [10], Jennings, et al. [11], Moore [12, 13], and Newton [14, 15].

The majority of sinkhole collapse incidents experienced along most highways in karst areas are usually the soil collapse type (Fig. 6a, b, c, d, and e) as defined by [13] for sections of East Tennessee karst. Extremely rare are occurrences of bedrock collapse into large open caves; however, they do occur.

Karst Features and Highway Construction

There are a number of features that characterize the karst landscape that can, and often do, present problems for highway engineering projects. Features such as subsidence depressions (no visible cave opening) and sinkholes should be considered in any planning study. Likewise, known cave systems, springs, "dry valleys," and areas of numerous limestone outcrops should be closely studied. Ponors, sinkholes that receive surface streams, and flooded or water-filled sinkholes and depressions should be avoided if at all possible. Large karst valleys known as uvalas and

Karst Terrane and Transportation Issues, Fig. 5 Naturally occurring sinkholes often expose limestone bedrock as shown in this photo from Middle Tennessee

poljes may be overlooked due to their enormous size relative to other karst features.

Building roadways over existing sinkholes and other karst features is obviously not the preferred engineering choice. However, when building roads across karst areas, these features are often unavoidable. The construction practice of dumping soil into a depression or sinkhole without proper investigation and design procedures will usually result in subsidence of the fill.

Collapse

One of the most recognizable features of karst is sinkholes. The result of human activity on the karst landscape often includes subsidence or collapse of a building, roadway, house, or even a yard (Fig. 7). When subsidence develops beneath structures or highways, it may have a catastrophic impact resulting in high repair costs.

The construction and maintenance of highways and appurtenant structures in karst terrane often will trigger sinkhole collapse, subsidence, and flooding. The collapse of highways in karst areas has been well documented in the literature [12, 16–19].

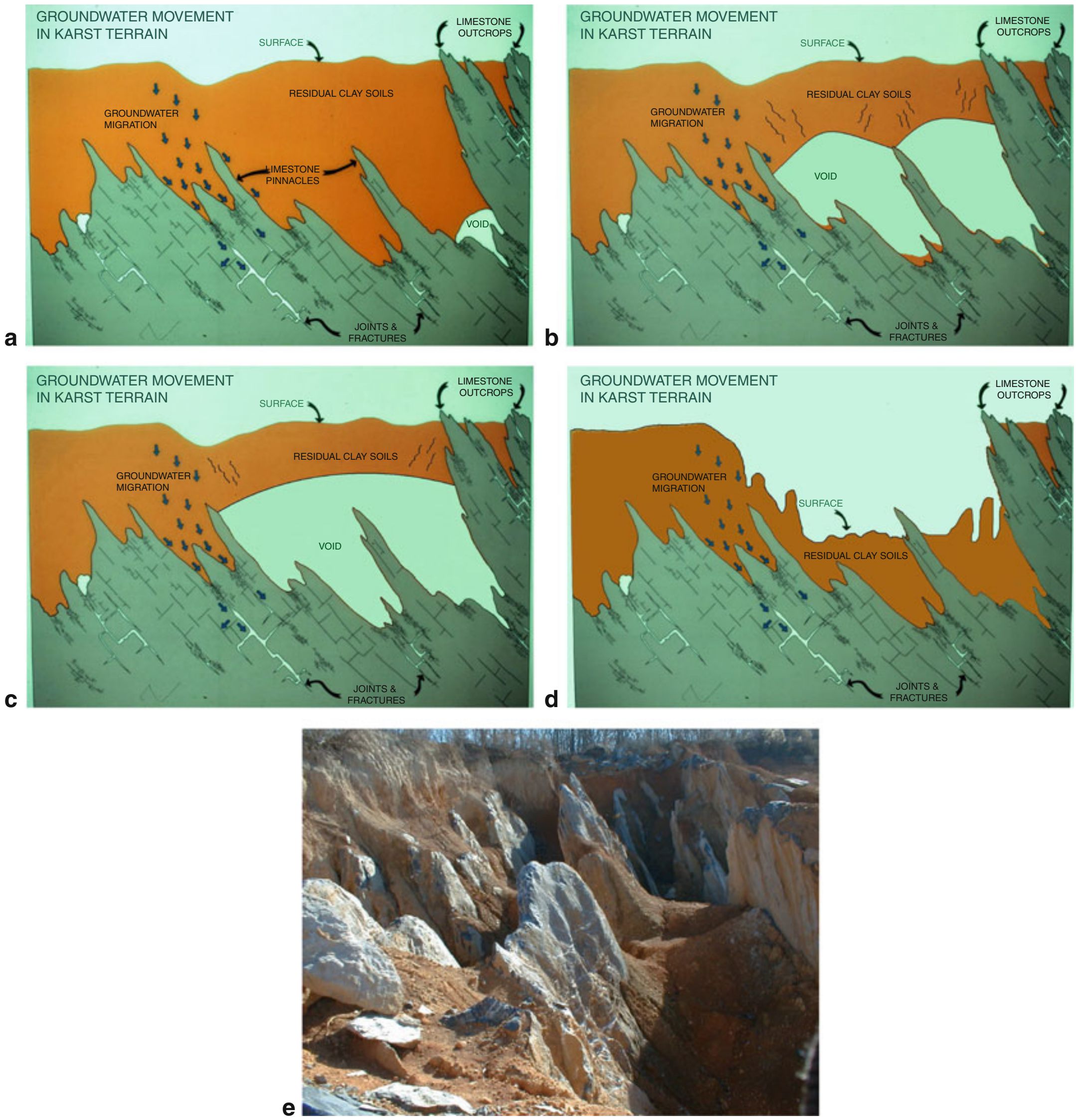
The relationship between construction activity and the occurrence of sinkholes has also been documented [14, 20, 21]. More specifically, Moore [12, 13] relates the occurrence of sinkhole collapses to highway construction, particularly the grading and ditching operations where surface and subsurface water flow paths have been altered. Studies conducted by Moore [13] cited above show that over 75% of the sinkhole collapses that have been documented along the roadways by the Tennessee Department of Transportation (TDOT) in East Tennessee were found to have occurred in the drainage ditch of the roadway.

Sinkhole collapse and flooding problems along Tennessee highways have been described by Royster [22] and Moore [12, 23]. In addition, Sowers [24], Newton [19], Foose and Humphreville [25], and Amari and Moore [26] have detailed possible geotechnical solutions to subsidence, flooding, and groundwater contamination in the Valley and Ridge Province from Alabama to Pennsylvania.

In 2003 Moore updated the 1987 study by analyzing 163 cases of sinkhole collapse incidents in East Tennessee between 1969 and 2002. Of the 163 sinkhole incidents studied, 86.5% of the sinkhole occurrences were located in highway ditch lines [18]. The 2003 study also supported the findings of the 1987 study by showing that of the ditch line collapse incidents analyzed, 93% also involved unlined ditches (the same percentage disclosed in the 1987 study).

Sinkhole activity often occurs in episodes as a response to increased groundwater recharge or drawdown. Sometimes these sinkhole occurrences can be quite dramatic, impacting not only roadways but also houses and buildings.

The 2003 sinkhole study by Moore documents numerous sinkhole occurrences over a 32-year



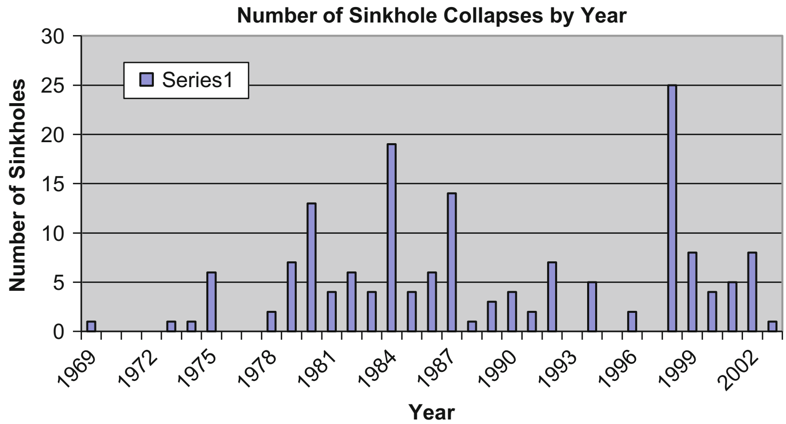
Karst Terrane and Transportation Issues, Fig. 6 This sequence of drawings shows a hypothetical development of a cover collapse-type sinkhole, common in much of the karst areas of the eastern USA: (**a**) shows void developing in the residual soil covering tilted limestone; (**b**) shows the void becoming larger and spanning between two rock pinnacles; (**c**) shows a larger soil bridge over the void; (**d**) shows the result of the soil bridge collapsing into the void below, causing a cover collapse sinkhole; and (**e**) shows a real example of the collapse as depicted in the above drawings

period along the state highways in eastern Tennessee.

There have been four major episodes of sinkhole collapse based on TDOT office records during the years from 1969 through 2002: 1980, 1984, 1987, and 1998. Historically, from 1977 to 1987, a total of 74% of highway-related sinkhole collapses occurred in roadway ditch lines [13]. The 1987 study revealed that of 74% of the ditch line collapses, 93% occurred in unlined (sod or clay) ditches. Overall, from 1969 through 2002, 86.5% of the sinkhole collapse incidents were in ditch line locations, again with 93% of those ditch line collapses occurring in unlined ditches (Table 1).

Karst Terrane and Transportation Issues, Fig. 7 Some sinkholes are often filled with trash and other debris which leads to the contamination of the surrounding groundwater. (**a**) This sinkhole collapse occurred in the summer of 2002 along I-640 in Knoxville, Tennessee. (**b**) This collapse was located in the median of I-75 in Anderson County, approximately 20 miles north of Knoxville. Note the marked area in the grass outlining the sinkhole boundary which was measured to be approximately 30 ft in diameter. (**c**) A number of ditch-line type sinkhole collapse problems occurred along highways in East Tennessee in the winter of 2002–2003. This is along U.S. 321 in Blount County. (**d**) Some of the recent sinkhole collapse problems were just open pit shafts, as deep as 10 to 12 ft. (**e**) TDOT maintenance forces were used to correct many of the collapse problems, such as this one on an I-140 off ramp near Alcoa, Tennessee. (**f**) This large collapse occurred in the ditch line of an off ramp along I-140 in Blount County, near the Knoxville airport in March of 2002

Karst Terrane and Transportation Issues, Table 1 This table shows the annual number of sinkhole collapse incidents from 1969 through 2002 (163 sinkhole collapse cases studied)



All major highways including Interstate routes experienced sinkhole collapses within the past 5 years. One ditch line collapse on I-75 in Anderson County was measured to be over 30 ft in diameter. Most of the sinkhole collapses were on the order of 10–12 ft in diameter and up to 8–10 ft in depth. No particular geologic formation was found to be more dominant than another in areas where the sinkholes had formed. However, it was observed that those areas where excavation had approached the pinnacle-type soil/rock interface (within 8–10 ft) experienced more collapses than other areas where there was no excavation or very deep soils that existed below the road grade.

Precipitation Data

Annual precipitation data for the middle areas of East Tennessee was obtained from the Tennessee Valley Authority and the National Weather Service in Morristown, Tennessee, to see if there was a correlation between precipitation events and increased sinkhole collapse incidents along the highways. It was hypothesized that during high rates of precipitation, there would be a corresponding increase in sinkhole collapses.

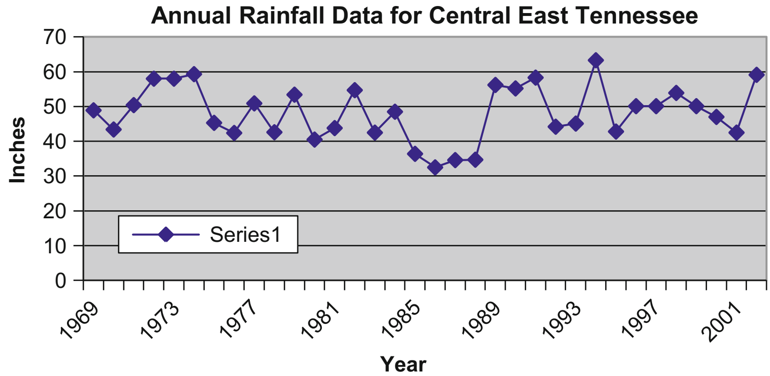
With the general average annual precipitation rate for the middle areas of East Tennessee being approximately 48 in., there have been six periods of above-average precipitation during the years from 1969 to 2002. These include the following: 1972–1974, 1979and 1982 (as one period), 1989–1991, 1994, 1996–1999, and 2002. The average annual precipitation of these high years was 59.9 in. There was also one major period of below-average precipitation (1985–1988) where the 4-year average was 34.53 inches (Table 2).

As was first surmised, there appears to be some link to precipitation. However, it is not what was first believed. There has been a general perception that the sinkholes tend to occur during very wet periods. And generally the data tend to support this. However, our data also shows a sinkhole occurrence spike during low periods of precipitation.

Obviously, the occurrence of sinkholes during high levels of precipitation is usually predicted due to the increase in surface and subsurface erosion. Not clearly understood is the occurrence of sinkholes during drought conditions. A correlation of the drought conditions may be made to mining or quarry operations where those operations dewater or lower the local groundwater table causing induced sinkhole collapses to occur. There are a number of case histories where quarrying adjacent to highways in karst areas resulted in induced sinkhole incidents [14, 27, 28].

It is believed that the lowering of the water table results in a loss of support to the roofs of cavities in bedrock that were filled with water. This also applies to residual clay soils that overlie openings in the cavernous bedrock that were filled with water before the decline of the water table [29]. As drought conditions lower the water table,

Karst Terrane and Transportation Issues, Table 2 This chart illustrates the annual rainfall data (in inches) for the central region of East Tennessee from 1969 to 2002



the same mechanism causes the migration of the residual soils from near the surface to lower areas within the cavernous bedrock causing subsidence and collapse to occur. This may explain the increase in sinkhole collapse incidents during dry periods.

This study revealed that from 1969 to 2002 where 163 cases of highway sinkhole collapse were recorded, approximately 86.5% of the sinkhole occurrences were located in highway ditch lines. This study also supported a previous study by the author (1987) that of all the ditch line collapse incidents, 93% involved unlined ditches.

It is readily evident that channeling surface water into *unlined* ditches or even retention basins or "holding ponds" will increase the incidence of sinkhole collapse. Recognizing the impact of development (road building, subdivisions, shopping malls, etc.) on karst environments in advance of construction can result in minimizing or even avoidance of the sinkhole problem altogether. Providing some element of impervious lining (geomembrane, pavements) for all drainage ditches in karst areas will greatly limit the incidence of induced sinkhole formation (Table 3).

Warning Signs of Collapse

Catastrophic sinkhole collapses can engulf entire houses and sections of roadways. Predicting sinkhole development can be extremely challenging.

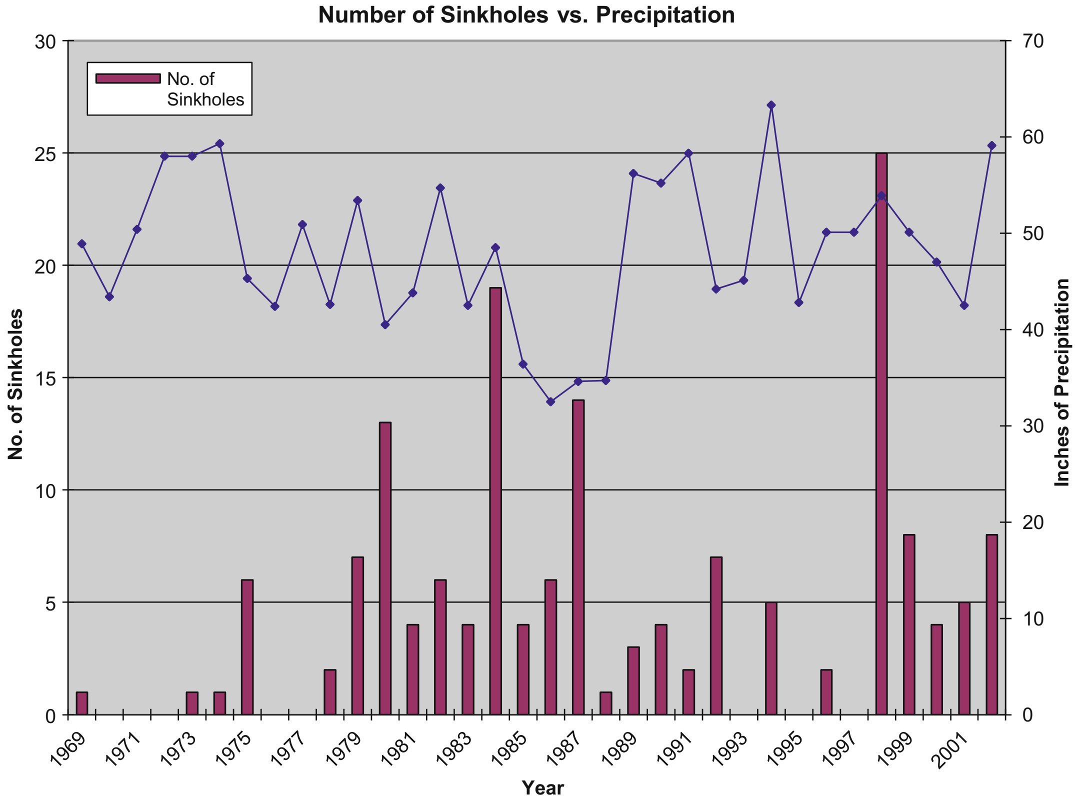
However, evidence of subsidence is sometimes evident prior to catastrophic sinkhole formation. One of the initial signs of collapse sinkhole development is the initially minor subsidence of an area. The effects of minor or initial subsidence are often evident. Minor pooling of surface water may occur in normally dry areas. Shallow puddles of water on a highway surface or along a ditch line are also initial signals of possible collapse.

Signs of settlement around houses and other structures may indicate subsidence. Cracks in house foundations, bridge abutments, or piers are signs of subsidence. The tilting of trees or other vegetation also indicates movement of soils. This type of evidence is often associated with landslides but can develop in association with catastrophic sinkhole collapse.

Under most conditions these early warning signs may develop over a considerable amount of time (months or even several years). In construction areas, induced sinkholes may occur suddenly without any indication of an impending collapse.

The formation of a sinkhole in a natural setting or on a construction site may occur in a sequence of settlement, significant subsidence, and collapse. Settlement of the ground may be noticed at first where minor movement of the soil, on the order of a few centimeters, may occur. Subsidence usually follows with drastic and very noticeable downward warping of the ground surface but with no large-scale cracks or scarps, sometimes many meters in diameter, but with no actual surface collapse of soil or bedrock. Finally, a sinkhole results from the collapse of the overlying soil (or bedrock) into the void below. This sequence may occur over a long period of time (months to years) or may occur quickly in a matter of hours.

Karst Terrane and Transportation Issues, Table 3 This table illustrates the relationship between annual precipitation rates and annual sinkhole occurrences in East Tennessee



Sinkhole Formation in Roadway Ditches

The greatest number of karst problems that develop along highways in East Tennessee involves subsidence and collapse of the drainage ditch (Fig. 8). These karst problems are referred to as being "induced" by some type of construction activity as opposed to being naturally developed [14].

In most highway design situations, roadway ditches with gradients of 3% or less are designed to be left barren or sodded. Only when the ditch line gradient exceeds 3%, according to most standard roadway design procedures, are the ditches paved or treated with an impervious material (i.e., concrete, polyethylene, geofabrics).

Within karst areas, the problem with untreated ditches often results when runoff accumulates in the ditch line and percolates into the underlying soil and rock. As the surface water migrates downward, clay particles are taken from the soil and carried into existing solution channels in the top of the bedrock and into the underlying bedrock conduits and caves. A cavity develops in the soil mass (above the bedrock) and enlarges from the bedrock toward the surface over a variable period of time (Fig. 9). Where the soil can no longer span the subsurface void, then the soil mass collapses into the underlying cavity, and the collapse progressively works upward to form a surface depression or a collapse sinkhole [30].

Construction projects are notorious for disregarding the surface runoff and usually result in serious siltation problems. In addition, the pooling of surface runoff in recently graded or excavated karst areas can result in sinkhole collapse. A significant aspect to consider is the new drainage path along the grade of the roadway (a drainage ditch) that concentrates the runoff and increases infiltration into the subsurface. An example of this process occurred on a highway

Karst Terrane and Transportation Issues, Fig. 8 Sinkholes often form along highway ditch lines as surface water is concentrated there leading to the subsurface erosion of the soil which forms these sinkholes

Karst Terrane and Transportation Issues, Fig. 9 Soil voids usually form over the limestone bedrock before the soil collapses forming a cover collapse sinkhole

construction project in Montgomery County, Tennessee, where surface runoff from a rain event was allowed to pond around newly constructed bridge pier footings on I-24. The concentration of water on the freshly excavated ground (which happened to be in an active karst area) triggered a collapse (Fig. 10) of residual clay soils around the pile-supported footings [22].

Drainage

Other types of karst features and hazards are associated with drainage issues. Flooding and groundwater pollution problems are two major concerns associated with drainage in karst.

Flooding of highways constructed across sinkholes is a relatively common occurrence in some karst areas. A number of sinkhole flooding problems of this type have been studied and documented (Tennessee: [12, 23, 31, 32]; and Royster [22]; Kentucky: [33]).

In most instances of highway flooding, the subject road was constructed across (or into) the sinkhole following the natural topography (Fig. 11). Prior to construction, land use may have been agricultural with little runoff and the sinkhole did not flood. Many such sites are now urbanized and the sinkhole can no longer adequately drain the increased runoff.

Karst Terrane and Transportation Issues, Fig. 10 When roadway bridges are constructed in karst terrane, sinkholes can develop beneath the supporting structures as shown here along I-24 in Montgomery County, Tennessee

Karst Terrane and Transportation Issues, Fig. 11 Roads built across and in sinkholes may flood if the subsurface drainage becomes blocked as shown in this photo of Prosser Road in Knoxville, Tennessee

Groundwater pollution due to the infiltration of highway surface runoff into a sinkhole feature has become a major environmental concern associated with highway routes in karst. While a number of researchers have studied the effects of highway runoff on the surface environment, few have dealt with the effects on the subsurface environment. Quinlan and Ewers [34] and Quinlan and Ray [35] discuss the effects of surface runoff on the Mammoth Cave System and their efforts to dye-trace the karst boundary. In their discussions, it was pointed out the necessity to map and dye-trace all sinkholes connected to roadways that traverse through the Mammoth Cave National Park.

Stephenson and Beck [36] made a review of literature regarding the quality of highway runoff and its potential impacts on karst groundwater. They found that little attention was being given by researchers to the issue of groundwater contamination by highway-related surface runoff and that the highway runoff could contain high concentrations of hydrocarbons and other vehicle-related contaminants.

In 1997, Stephenson and Beck describe the results of a Federal Highway Administration Pooled Fund Study (that involved 15 state DOTs including Tennessee) where a prototype filtration system (Fig. 12) was designed to treat highway runoff in an interchange of I-40 and

Karst Terrane and Transportation Issues, Fig. 12 Surface runoff from roadways can be filtered to remove pollutants before the water reenters the groundwater system through sinkholes

I-640 in the eastern part of Knoxville, Tennessee. This study described the attention that is required by infrastructure designers to prevent or lessen the pollution of the groundwater system in karst areas.

Reactive Approaches and Remedial Measures

Historically, the remedial approach to karst-related highway subsidence and flooding problems has typically been reactive in nature. Road crews commonly repair the sinkhole collapse after it has happened.

Remedial measures used in correcting karst-related subsidence problems reactively (after they occur) may be divided into three areas: bridging, drainage, and relocation. The techniques used in bridging include graded rock and shot rock pads and fills, graded rock backfill, concrete structures, and grouting [12, 37, 38]. Remedial measures used in drainage include alteration of the existing drainage by the use of paved or geomembrane-lined ditches, special overflow ditches, plastic overlays, pumps, horizontal drains, and the maintenance of obstructed sinkhole entrances. Surface drainage filter systems have been researched and are beginning to be employed to treat surface runoff from highways in karst areas. Although not widely used, relocation of the highway facility around a karst area may also be a viable corrective measure.

Types of Bridging Remedial Approaches

The treatment of karst-related subsidence and collapse features involving engineered structures can include methods as varied and site specific as the sinkholes themselves. Many remedial treatment techniques involve a method of "bridging" to provide adequate stability and bearing capacity.

There are a number of different approaches to bridging that can be applied. These include, but are not limited to, the following: conventional bridge spans, "muck trestles" (bridge built on grade), drilled shafts, riprap backfill, rock pads, grouting, concrete slabs, geofabrics, and geogrids. In some instances, a combination of these procedures may be needed [9].

Where field conditions warrant, a concrete structure may be used to bridge or span the collapse sinkhole area. This remedial measure may entail the construction of (1) a free spanning highway concrete or steel bridge, (2) a concrete bridge with the deck constructed on the ground surface (commonly referred to as a "muck trestle"), (3) drilled cast-in-place concrete piers to support a concrete slab or bridge section, or (4) a concrete slab placed over the subsurface opening [9].

Although a very costly approach, a free spanning bridge constructed over the sinkhole area is very effective in eliminating future problems. This approach should be considered when crossing very large and deep sinkholes or in environmentally sensitive areas such as Mammoth Cave National Park.

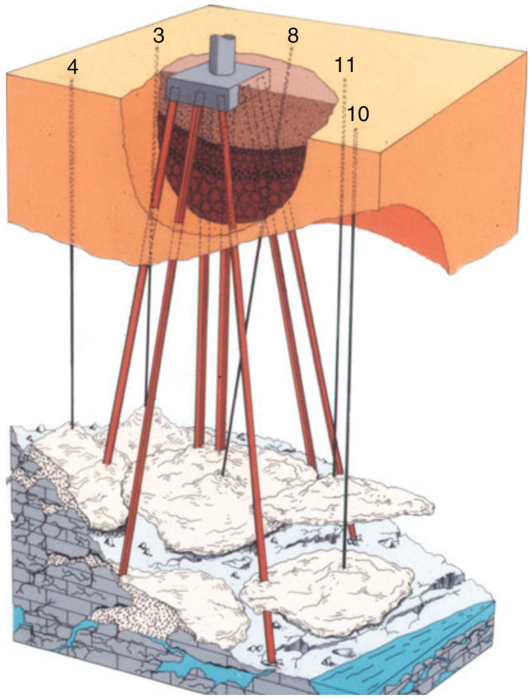
The "muck trestle" bridge has been considered for a severe karst-related flooding and collapse problem in Hamblen County along State Route 34 near Morristown. This concept was applied to a railroad bridge in the same vicinity of Hamblen County where a sinkhole collapse was endangering a rail line [39].

An often recommended measure for structural foundations in karst strata is the use of drilled cast-in-place piers (caissons). Where soluble limestone strata occur beneath a river channel, numerous cavities may exist which are either clay filled or open. In some instances, a series of solution cavities may extend downward to depths of over 30.4 m (100 ft). Karst strata that extend as much as 8–10 m (25–30 ft) below the surface can make excavation prohibitive and drilled cast-in-place piers more viable.

Another type of bridging concept commonly used in roadway applications is to use concrete grout to seal up cavernous strata. Grouting is often used at the soil/rock interface to close up openings that the soil may filter into – this is referred to as "cap grouting." Royster [22, 40] used this concept to repair a bridge footing along I-24 in Montgomery County, Tennessee, where a collapse occurred beneath a pier footing pile cap (Fig. 13). Another instance of using grouting to repair subsidence and sinkhole collapse problems is detailed in Mellett and Maccarillo [41] where the authors warn that grouting can divert underground water flow into other areas, and new sinkholes may occur nearby as a result. As shown in the above example, caution should be used when attempting to use grouting as a method of stabilization of collapse problems.

The Tennessee Department of Transportation applies another effective type of bridging concept which uses rock riprap to backfill the collapse. The interlocking action of the riprap backfill, often referred to as "chunk rock" backfill, and the placement of the backfill on the top of the in-place bedrock effectively bridge the solution cavity (Fig. 14). This concept also provides a free draining structure for groundwater and surface drainage.

These inverse filters using "chunk rock" backfill are designed to prevent overlying materials from filtering down into the interstitial spaces of the clean rock backfill and solution cavity resulting in subsidence or even another collapse.



Karst Terrane and Transportation Issues, Fig. 13 "Cap grouting," where concrete grout is pumped down to the soil-bedrock interface to seal the rock surface and prevent the migration of soil down into cavities in the bedrock, was used on a bridge footing in Montgomery County, Tennessee, along I-24

A bridging technique, commonly used in Tennessee, consists of a limestone graded rock pad placed to span an area of intense solution activity (Fig. 15). This technique was used on a section of Tennessee State Route 1 (U.S. 11-W) in Hawkins County where a new four-lane section of State Route 1 was constructed over an area of active sinkholes. This rock pad provided stability for the highway by allowing groundwater to flow freely from the sinkholes to a special channel and box culvert.

The rock pad constructed over this sinkhole area has a minimum thickness of 1.5 m (5 ft) and meets the following rock specifications [9]:

Rock pad material consists of sound, nondegradable limestone with a maximum size of 0.9 m (3 ft) and is free of shale and/or clay; at least 50% (by volume) of the rock is uniformly distributed between 0.3 m and 0.9 m (1 and 3 ft) in diameter, and no greater than 10% (by volume) is less than 5 cm (2 in.) in diameter; the rock material is roughly equi-dimensional; thin slabby material is not accepted.

Karst Terrane and Transportation Issues, Fig. 14 Clean graded rock is used to backfill a sinkhole collapse along Interstate 640 in Knoxville, Tennessee, by the Tennessee DOT

A geotextile filter fabric should be used to separate the rock pad from the overlying fill material (usually soil). The filter fabric specifications should be determined based on the grain size of the overlying fill material.

Drainage-Related Remedial Approaches

The development of proper drainage systems for highways constructed across karst terrane is commonly overlooked, often necessitating reactive approaches to karst drainage problems which involve ditch line collapse and flooding.

A successful technique of drainage control involves the use of lined drainage ditches along roadways in karst areas. This includes using materials such as asphalt, concrete, or geomembrane-lined drainage ditches. Studies [13, 18] have shown that the majority of collapse-type karst problems occur in unlined ditches. Unlined ditches are usually sodded at best and have gradients of less than 3% and more often 1% or less.

Unpaved ditches also provide for increased seepage which may result in subsurface erosion and finally collapse. Bare soil ditches and even sodden ditches are prone to this type of collapse failure (Fig. 16).

Usually, water collecting in an unlined ditch will seep into the underlying soil eventually reaching a conduit developed at the soil-bedrock interface below the ditch. Erosion continues until

Karst Terrane and Transportation Issues, Fig. 15 A rock pad is constructed over naturally occurring sinkholes on a road project in Tennessee. The rock pad provides not only stability for the road but allows surface drainage to flow into or out of the sinkhole, which had bedrock exposed and not soil

Karst Terrane and Transportation Issues, Fig. 16 Unlined drainage ditches in karst areas usually bear the brunt of sinkhole formation as shown in this photo of a sinkhole along State Route 72 in Tennessee

Karst Terrane and Transportation Issues, Fig. 17 Geomembrane is used for drainage ditches along some highways in karst areas of Tennessee. Most are covered with riprap or concrete paving

the soil over the cavity collapses, the result of the erosion being a collapse sinkhole induced by seepage from the drainage ditch. The solution is to line the ditch with an impervious material at the outset, during construction.

Materials that have historically been used to line ditches include sodding, asphalt, riprap, and concrete. Concrete-paved ditches have historically been used in these situations and perform moderately well.

In recent years, new emphasis has been put on the use of synthetic geomembranes (Fig. 17). Their use as ditch liners in karst areas is strongly recommended due to the following factors: flexibility, imperious character, strength against puncture, and ease of installation.

The geomembrane liners can be installed with a soil or sod cover for aesthetics and UV protection. Riprap can also be placed atop the geomembrane (careful placement is required). In emergency situations, polyethylene sheeting can be used until a more permanent repair can be affected. Finally, if a concrete ditch is desired, the concrete should be applied over a geomembrane liner. Geomembranes that tend to perform well in these applications are 60 mil thickness HDPE and PVC membranes.

Relocation

In some instances, relocation of the highway is the best alternative to solving the karst-related problem. This may simply involve shifting the roadway slightly to miss a specific sinkhole or cave entrance. In other cases, the relocation may be several hundred feet from the original alignment and cost thousands of dollars for additional right-of-way.

Proactive Measures

Being proactive in regard to dealing with karst in highway design and construction is to identify the problem and formulate a design that will remediate that problem before construction begins. If a roadway alignment has to be located in sinkhole areas, then design roadway embankments and cut sections so that the impact of the karst on the roadway will be lessened or eliminated. Avoidance measures and some combination of drainage and bridging methods are usually the best direction to take in a proactive approach to designing and constructing highways in karst areas.

Avoidance Measures

The avoidance of karst areas by transportation facilities is recognized as a highly desired approach in highway planning; however, the extensive suburban development necessitating new highway construction commonly limits highway corridors to less desirable land. Identification of the most severe karst areas (with respect to drainage, sinkhole collapse, and other environmental impacts) during location and design phases of highway planning can lessen the impact of karst features on the project and result in a better designed and more stable roadway facility.

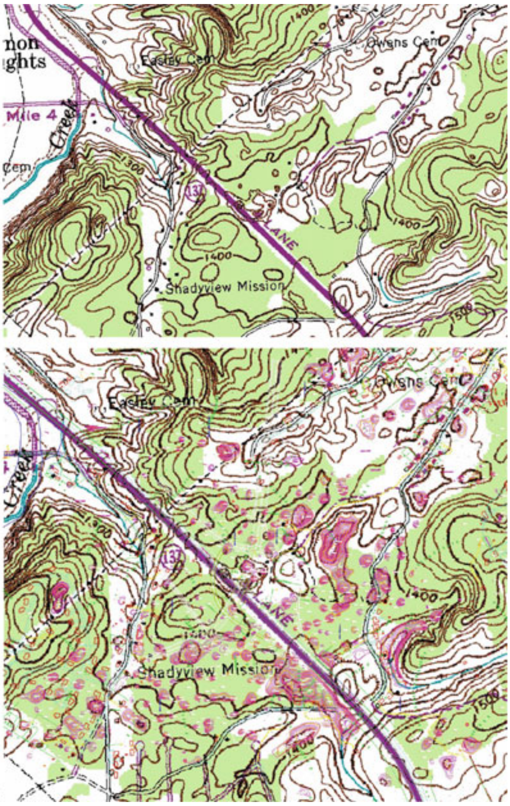
Avoidance is only effective where karst is active and recognizable. The construction of new roads with associated increases in runoff along the new landscape can trigger subsidence and flooding problems where previous land uses were in relative equilibrium with the karst.

Locating new roads in areas prone to karst development requires a great dependence upon the geologic information that is available. The geotechnical information gathered during the investigative phase of a proactive approach to highway design is most valuable in determining route location selection relative to karst. Maps and other information that delineate such karst areas (derived from a proactive investigation) can be used to avoid problem areas. The use of geophysical methods such as resistivity, seismic, and radar can be used to aid in the characterization of subsurface conditions in karst areas [42].

Caution should be exercised when using standard 7.5 min USGS topographic maps to identify sinkhole-prone areas. The standard 20 ft (6 m) contour interval used with the 7.5 min quadrangle rarely reveals the true number of sinkholes present. Hubbard [43] researched this issue in the karst regions of the Valley and Ridge of Virginia where he consistently identified more sinkholes on the ground than were depicted on selected 7.5 min quadrangle maps. In one instance, Hubbard describes a 7.5 min quadrangle with a 20 ft contour interval that revealed 55 sinkhole features, while ground field mapping identified 533 sinkhole features.

In another case that relates this concept to East Tennessee, Moore [44] studied a proposed location of an interchange along I-181 in Sullivan County, Tennessee, situated in the Valley and Ridge Province and underlain by carbonates. Within an approximately 100 acre site, the standard 7.5 min maps with 20 ft contour intervals revealed some 30 or so sinkhole features. After aerial mapping using a 1 m contour interval, a total of 177 sinkholes and 7 cave entrances were identified at the proposed site (Fig. 18a, b), which was later rejected by the Tennessee Department of Transportation for the interchange location.

Roadway designs may involve the use of broad horizontal curves to miss a sinkhole area or cave complex resulting in a more scenic facility. An example of this process is the Pellissippi Parkway Extension (I-140) in Knox County, Tennessee. The extension of a limited-access, four-lane highway facility, “the Pellissippi Parkway” (I-140), involved the possible location of corridors in active karst areas of Knox and Blount counties in East Tennessee. Geotechnical involvement in the location and design phases of this highway development process resulted in the recognition of several unstable areas of karst and led to the adoption of remedial and preventive design measures which were incorporated into the construction plans [23].



a b

Karst Terrane and Transportation Issues, Fig. 18 Using topographic maps is a good source for finding sinkhole terrane. However, the scale of map and the contour interval can be deceiving as shown in the map above. Map (**a**) shows an area on a 20 ft contour scale with some 30 sinkholes developed; map (**b**) shows the same area on a 3 ft contour scale showing over 100 sinkholes

Ultimately, the route selected was the most favorable geologic setting of the four proposed routes. Significant consideration by administrative officials was given to the geologic evaluation of the project, a measure of the progress in recognizing and addressing karst processes and teamwork within the geologic and engineering professions.

Mapping karst areas for planning use by government highway organizations (state DOTs) is very proactive and effective in reducing the impact of karst on proposed transportation routes. One such example is the South Knoxville Boulevard Extension (SR 71) in Knox County, Tennessee. The Tennessee Department of Transportation (TDOT) is planning a new roadway alignment (SR 71) that crosses portions of South Knox County in East Tennessee. The proposed corridor is located in a section of the Valley and Ridge Province of East Tennessee where several ridges and valleys will be crossed, as well as creeks, roads, subdivisions, and rural lands. The proposed corridor connects the current terminus of SR 71 at Moody Ave. in South Knoxville to John Sevier Highway (SR 168) (Fig. 19).

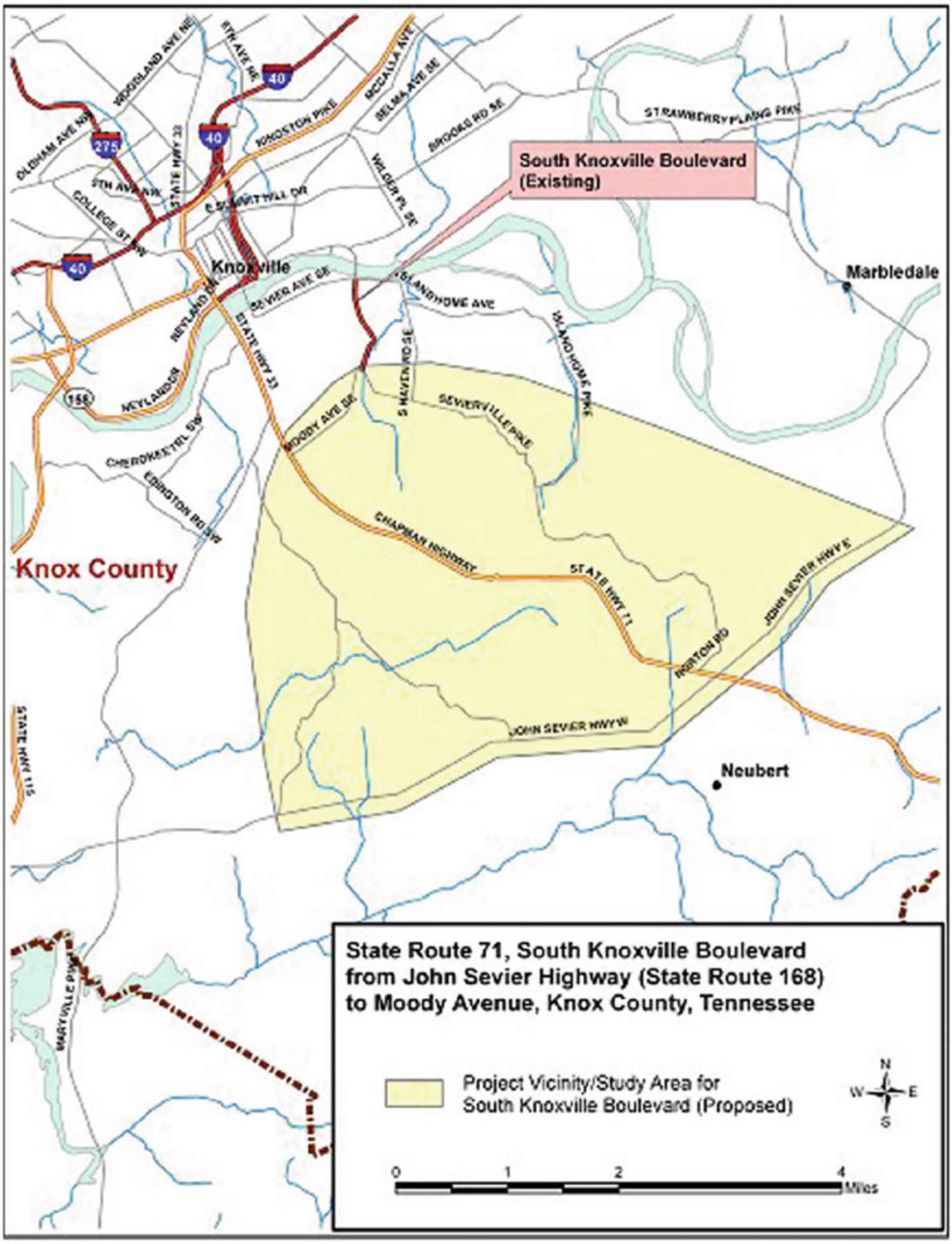
In an attempt to properly evaluate the potential for geologic hazards along the project, an effort was made to locate all sinkholes and caves within the project area, karst being the primary geologic hazard identified. To this effort, a karst map and a map of one cave were completed (Figs. 20 and 21).

The proposed SR 71 corridor is situated in the rolling to hilly topography of the Valley and Ridge Province of East Tennessee and includes a portion of South Knox County. Geologically, the corridor is situated in terrain underlain by folded and faulted middle Ordovician sedimentary strata composed of several major rock types including limestone, shale, sandstone, and siltstone.

The weathering of the rock strata has produced a knobby terrain that follows the outcrop pattern of the specific formations, especially the Holston and Chapman Ridge formations. The weathering has also produced classic karst topography characterized by sinkholes, disappearing streams, and caves, including Meades Quarry Cave (Fig. 22).

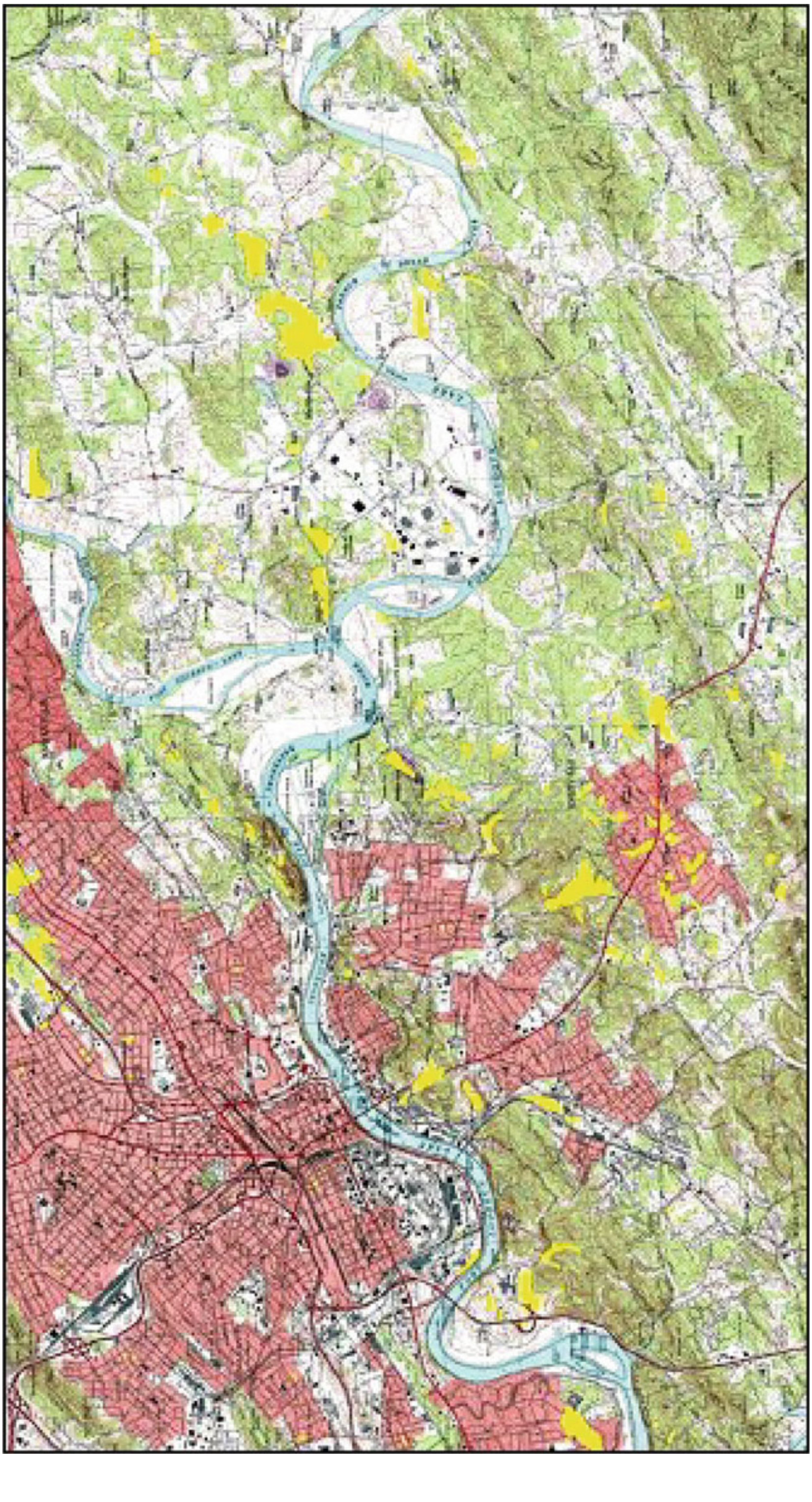
A geohazard evaluation was conducted by the Tennessee DOT Geotechnical Engineering Section for the proposed SR 71 route. This study was an effort by the Tennessee DOT to identify the geologic hazards along the corridor and to avoid or reduce the hazard impact on the roadway facility, the environment, and the public. The main geohazard found along the proposed SR 71 study area was karst.

The study revealed that the corridor is characterized by features such as sinkholes, caves and cave entrances, sinking streams, and outcroppings of weathered carbonate rock (limestone and dolostone). The recognition of areas of active karst subsidence and collapse is of considerable importance to the design and implementation of the proposed highway infrastructure, SR 71.

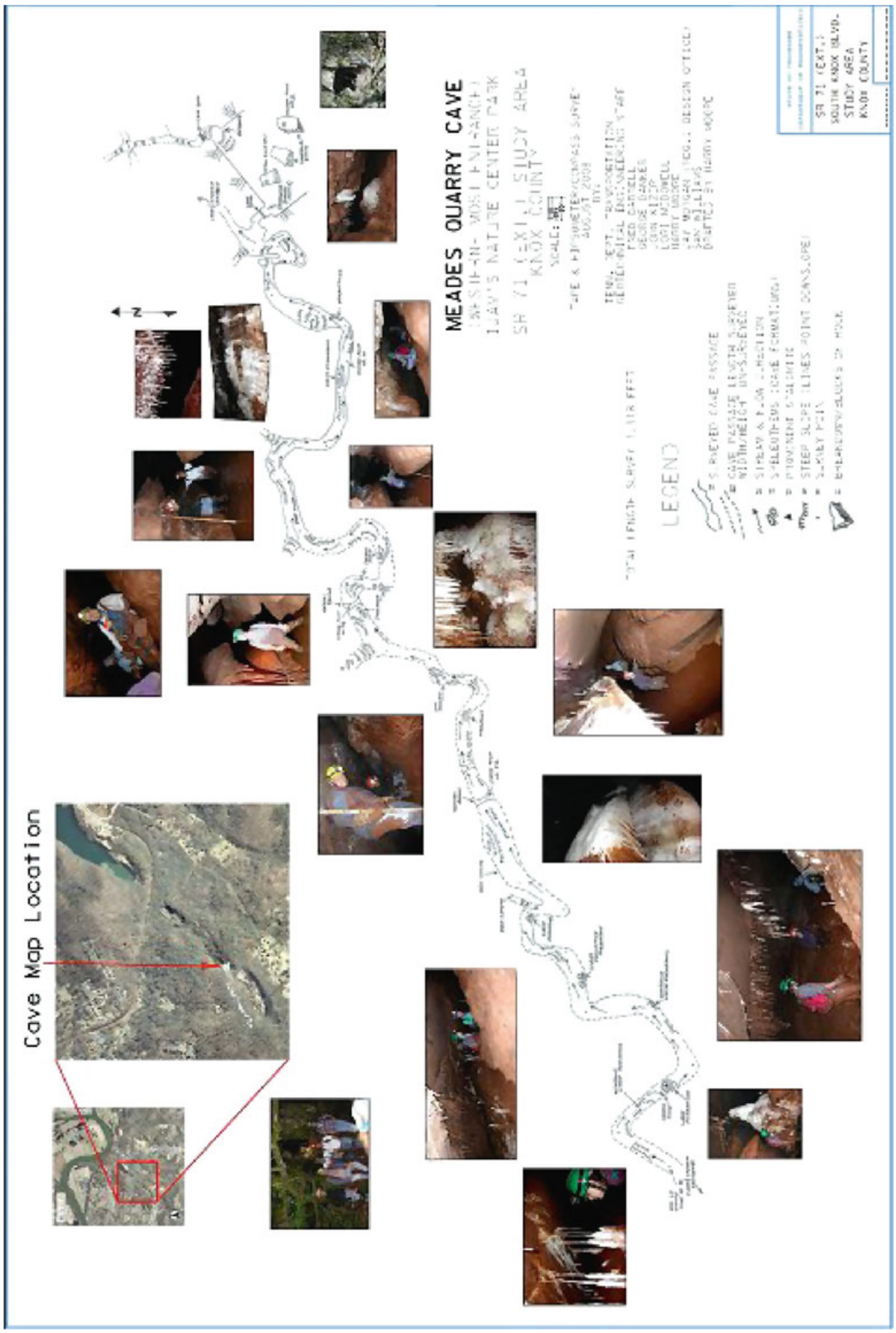


Karst Terrane and Transportation Issues, Fig. 19 This map shows the study area for the extension of the SR 71 – South Knoxville Blvd. in Knoxville, Tennessee. The study area contains a large karst sinkhole area shown in Fig. 23

TDOT initiated a karst geohazard inventory and assessment to assist in evaluating and selecting a satisfactory alignment within the corridor study area. This study involved the location of caves and sinkholes within and adjacent to the proposed corridor. In addition, the caves were visited and preliminarily evaluated as to their geotechnical and environmental importance. The caves that were found to be of significance to the proposed stability of the roadway or have environmental and archeological significance were surveyed to determine their lateral and vertical



Karst Terrane and Transportation Issues, Fig. 20 This map shows sinkholes (*yellow*) that were found in the study area for the proposed SR 71 – South Knoxville Blvd. Extension in Knoxville, Tennessee



Karst Terrane and Transportation Issues, Fig. 21 Cave maps, such as this detailed cave map of Meades Quarry Cave in South Knoxville, Tennessee, can be used by highway designers to avoid impacting fragile cave environments when planning and designing future roads

Karst Terrane and Transportation Issues, Fig. 22 This photo shows the man-made entrance to Meades Quarry Cave which was opened by marble quarrying operations in South Knoxville in the early 1900s

extent. Mapping of the caves was performed by both TDOT Geotechnical Engineering staff and local speleologists.

Caves and sinkholes offer special challenges relative to both the physical and environmental issues of highway development. The karst geohazard inventory study, which included mapping karst features, disclosed the presence of both sinkholes and caves, some of which may have detrimental structural, geologic, and environmental issues for alignment and grade design considerations.

Efforts to avoid karst areas with roads and other developments are difficult at best, especially since the "good" land is already mostly developed, leaving only the geologically undesirable land for current development. Geologically undesirable land includes such areas as karst terrane steep terrain, very rocky terrain, landslide-prone areas, soft ground, and areas prone to acid mine drainage, to name a few.

In planning new highway corridors, certain constraints must be known and when identified are usually displayed by mapping. The Tennessee Department of Transportation, Geotechnical Engineering Section, is involved in studying proposed roadway corridors in East Tennessee in an effort to identify geohazard areas.

The mapping of karst areas generally involves the following procedure as outlined by Moore and McDowell [46]:

- In general, 7.5 min topographic maps (that use a 6 m (20 ft) contour interval) are used to locate karst features such as sinkholes (closed depressions on the contour map), caves, springs, and sinking streams. Once these karst features are identified and located on the topographic maps, then a field reconnaissance is performed to field check the features to make sure that they are there.
- Afterward, the sinkholes and other karst features are enhanced on the topographic maps, and subjective boundaries are drawn to encircle these areas. Typically, these encircled areas are identified as "areas of high concentrations of sinkholes" and/or "areas of numerous cave openings." Actual cave entrances are not plotted on the final geohazard map due to access issues and private owner protection.
- The geohazard areas are then expressed as outlined patterns on topographic maps to better illustrate the geohazard relative to the surrounding landscape. In addition, the proposed corridor route is overlain on the geohazard map. This map is then used by the roadway planners to better locate the final roadway centerline.

In addition to the surface mapping of sinkholes, it is becoming more important to map the caves where they may exist in close proximity to the proposed roadway. By knowing spatially where the cave passages are located, a more accurate design of proposed roadway cut slopes can be made. This would prevent the unnecessary opening of a cave system to the surface, and benefit the cave biota, such as bat colonies and salamanders to name a few.

The use of experienced cavers (usually cavers who are members of the National Speleological Society, NSS) in combination with engineering survey crews provided the best results for locating the cave passages spatially with respect to the proposed roadway. It is anticipated that mapping of cave passages will become increasingly mandatory as society continues encroaching onto and into the karst environment.

After a review of available geologic data and field investigations, the SR 71 Extension corridor in South Knoxville, Tennessee, was found to be located within several strike belts of karst. The belts of karst trend in a northeast to southwest direction, reflecting the strike of the underlying bedrock. In addition, the rock strata are folded into a large structural anticline resulting in the bedrock forming a “U”-shaped pattern in the South Knoxville community (Fig. 23).

Numerous sinkholes and caves were found to be located in these karst areas. A map of the most intensive sinkhole areas was prepared in order to better assess the corridor terrain. A field reconnaissance of the study area was made in an effort to locate as many of the sinkholes as possible. Many were overgrown and difficult to locate. A few of the caves that were identified were found to be located outside of the numerous sinkhole zones (outlined on the attached karst sinkhole map) but within typical karst terrane.

The possible impacts on the karst environment from constructing the proposed road alignment (or any other structure, building, subdivision, etc.) include sinkhole collapse, sinkhole flooding, groundwater contamination, and environmental effects on the cave and subsurface-dwelling wildlife, such as bats and salamanders.

Mapping Subsurface Karst: Meades Quarry Cave Mapping Initiative

Due to the uncertainty of the existence of Meades Quarry Cave within the proposed study area, it was decided to map the westernmost portion of the cave system. The purpose of mapping the cave was to determine if the general trend of the cave and cave stream is toward the sinkhole area around Old Sevierville Pike and Red Bud Drive, which is within the study area of the proposed parkway extension.

The presence of the Berry Cave salamander (*Gyrinophilus gulolineatus*, a subterranean amphibian listed as a potential threatened species – see Fig. 24) in the Meades Quarry Cave stream has made the cave an important issue of the overall study of the proposed SR 71 extension. The Berry Cave salamander derives its nutrition from debris and organics that are flushed into the sinkholes by rain events which in turn recharge the cave stream in which the salamander lives. Adversely affecting the sinkholes would directly affect the salamanders.

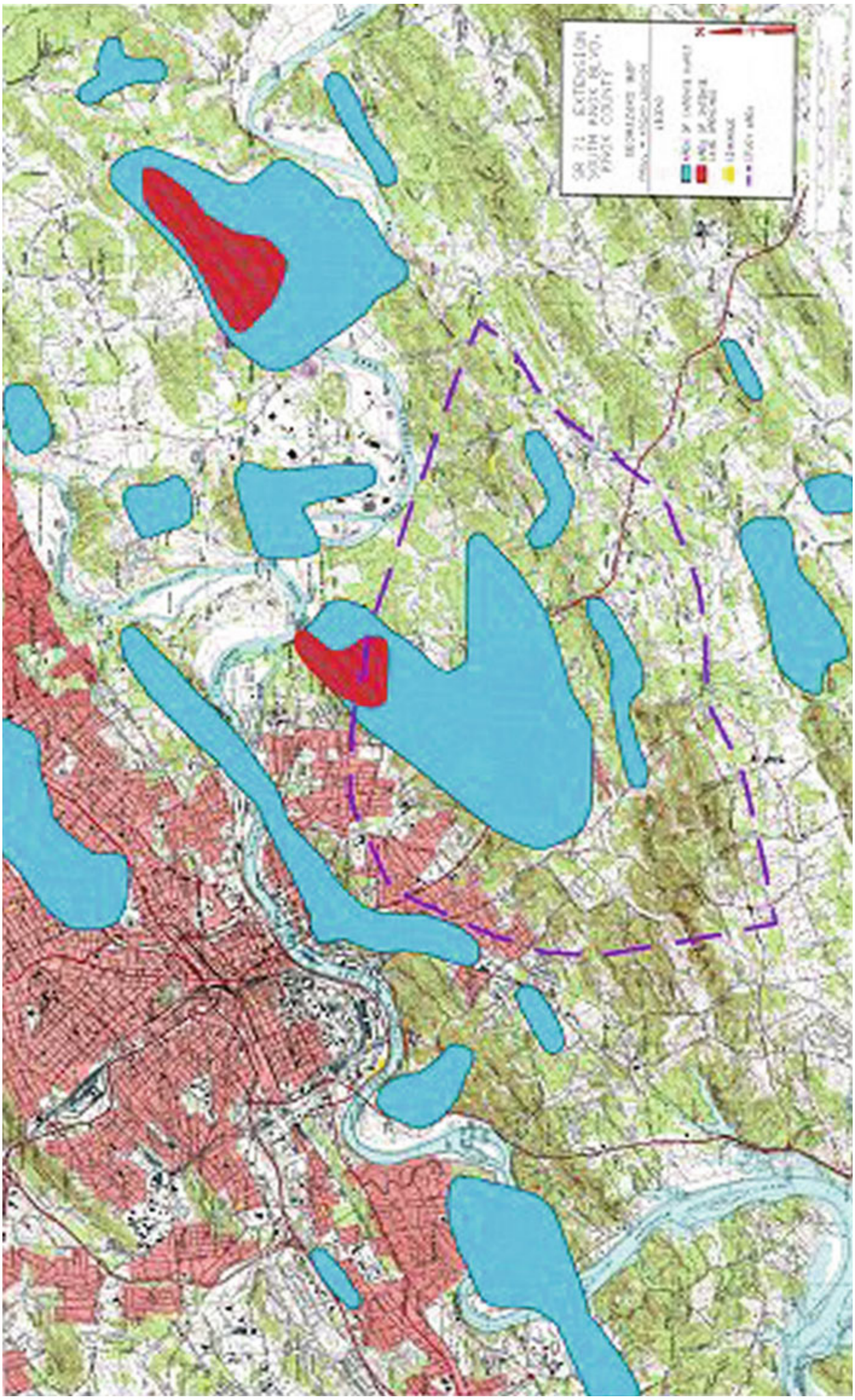
As a result of the concern over the salamander species found in Meades Quarry Cave, it was decided by Tennessee DOT that a mapping effort of the cave would be completed. In August of 2008, the Geotechnical Engineering Section initiated an effort to map the westernmost portion of the Meades Quarry Cave system (Figs. 25, 26, and 27). This section of the cave system has its entrance in the floor of an old quarry pit which exposed the cave during quarrying operations in the 1930s and 1940s.

As a consequence of our investigation and mapping initiative, it was interpreted that the stream passage of Meades Quarry Cave does indeed lie within the study area of the proposed SR 71 highway extension. In addition, it was decided that a dye trace study was needed to further delineate the groundwater recharge area for Meades Quarry Cave system.

Dye Trace Study

A proven and reliable method of identifying subsurface drainage areas in karst is to use dye tracing. Many studies have disclosed the valuable usefulness of dye tracing (Quinlan and Ray [34, 35, 48, 49]). To that end, the South Knoxville SR 71 highway project study has initiated a dye trace study of the proposed highway route in a sinkhole-prone section of the corridor.

As noted in the above discussion, the Meades Quarry Cave system of solution channels and stream(s) is located within the study area for the proposed extension of SR 71. Also, it was disclosed by local University of Tennessee



Karst Terrane and Transportation Issues, Fig. 23 This map shows the karst and sinkhole area that was mapped by the Tennessee DOT Geotechnical Engineering Section while studying the proposed extension for SR 71 – South Knoxville Blvd. in Knoxville, Tennessee. The dashed area shows the actual study area (base geologic information from Hardeman [45])

Karst Terrane and Transportation Issues, Fig. 24 This photo of the Berry Cave salamander was made in Meades Quarry Cave in South Knoxville, Tennessee. Cave environments are fragile places and harbor many rare and endangered species of wildlife (Photo by Matt Niemiller [47] University of Tennessee; from: http://www.herpetology.us/niemiller/)

Karst Terrane and Transportation Issues, Fig. 25 Mapping caves is an important process in studying the location of future transportation routes like highways and railroads; shown is Meades Quarry Cave mapping by Tennessee DOT geologists

Karst Terrane and Transportation Issues, Fig. 26 Meades Quarry Cave contains a perennial stream that supports the environment for the Berry Cave salamander

Karst Terrane and Transportation Issues, Fig. 27 Cave speleothems decorate a passage in Meades Quarry Cave in Knoxville, Tennessee. Shown are white soda straws connecting with white stalagmites

biology researchers that the Meades Quarry Cave contains a population of Berry Cave salamander which is a rare species of animal (Moore and McDowell [46]).

The US Fish and Wildlife Service and the Tennessee Wildlife Resources Agency has noted the presence of the Berry Cave salamander in the Meades Quarry Cave system and requested that TDOT do an adequate dye trace study of the sinkholes in the area of the proposed project to better understand the geohydrology of the Meades Quarry Cave system. The dye trace study is to better differentiate which sinkholes actually recharge the cave stream that the Berry Cave salamander lives in, so that TDOT can avoid any impact on the cave system with the construction of the proposed SR 71 extension. A result of the dye trace study will be an effort to minimize the impact of highway runoff on the local karst system. This will be accomplished with the use of runoff filtration systems.

In summary, the study of the karst of the SR 71 project found that the proposed corridor has numerous karst features that are found within the area topography. Mapping surface and subsurface karst features can greatly aid the evaluation of proposed roadway corridors. A result of this study was the development of a surface karst map which showed areas of numerous sinkholes, which are interpreted as areas of potential future sinkhole development.

The construction of road projects, as well as industrial sites, shopping malls, and subdivisions will be problematic in karst areas and should be avoided when possible. However, if avoidance is not possible, then minimization of the impact of construction on the sinkhole and cave environments should be the objective. Such design measures as minimal cut and fill construction and minimal alterations in the surface drainage of an area (both surface and subsurface) are recommended.

Another result of the karst geohazard study was the development of a karst planning map and the decision that mitigation of the impacts on the karst areas will need to be included with the roadway design and construction plans once the road project is approved. Proactive measures such as surface drainage filtration systems for sinkholes, impermeable lining for ditch lines (to prevent new sinkholes from forming), the use of graded rock embankments (for stability and groundwater impacts), and the use of structural bridges over sensitive sinkholes are recommended to be considered and/or employed during the roadway design phase.

The groundwater contamination issue is an important topic for the SR 71 project due to the potential impact on the rare Berry Cave salamander which lives in the Meades Quarry Cave system. Once the groundwater dye trace results are obtained by the Environmental Division (as discussed in the above narrative), then a more appropriate evaluation of the karst groundwater drainage can be made.

Surface water runoff filtration systems, as proposed for the SR 71 project in South Knoxville, have recently been constructed on a Tennessee DOT roadway project in Hamblen County, Tennessee, where the runoff empties into a sinkhole (Fig. 28). The filtration system design was based on results of a FHWA Pooled Fund Study that involved filtering highway runoff in karst areas. Tennessee DOT was a partner in the FHWA Pooled Fund Study [36, 50].

It is important to understand that there will always be an impact on the karst environment (sinkholes, caves, wildlife, and groundwater) when construction occurs in karst areas. In some places, this impact may be significant where caves and sinkholes are exhumed and/or filled in with embankment material and where surface drainage is directed into sinkholes that empty into the caves and groundwater systems. These impacts can be lessened by appropriate and judicious mitigation during the design and construction phase of the project.

Proactive Drainage and Bridging Concepts

The proactive approach to karst-related drainage problems requires attention to several design-related items. These include lined ditches, rock pads, overflow channels, sinkhole opening improvement/protection, curbs for embankment sections, and drainage wells. The following discussion summarizes the treatment of karst-related drainage problems [44]:

- *Lined ditches* – The single most important item that can be implemented to prevent future sinkhole collapse occurrence is the use of lined drainage ditches. Types of liners that tend to function the best include 60 mil PVC and/or HDPE geomembrane and concrete and asphalt materials.

- *Rock pads* – Rock pads beneath embankments using clean riprap limestone may be used for bridging depressions and sinkholes (Fig. 29).
- *Overflow channels* – This concept involves the construction of a lined channel or pipe from a negative drainage area (sinkhole) to a positive draining system.
- *Sinkhole opening improvement/protection* – This concept involves improving the runoff flowing into subsurface cavities by removing debris and trees from around the throat of a

Karst Terrane and Transportation Issues, Fig. 28 This shows the sinkhole filtration system constructed on a Tennessee DOT road project in Morristown, Tennessee. The filter is in the long rectangular structure near the top center of the photo. The sinkhole receiving the filtered runoff is in the top left of the photo

Karst Terrane and Transportation Issues, Fig. 29 Shown is a roadway embankment that is underlain by an engineered rock pad across a sinkhole in Knoxville Tennessee sinkhole and protection of the cavity opening from siltation and debris using such methods as siltation barriers, debris catchment fences, riprap, gabion siltation barriers, and concrete structures (Fig. 30).

- *Drainage wells* – An additional concept that has been implemented in some areas is the use of injection drainage wells [12, 22, 51–53]. These types of storm water drainage wells are required to be permitted by the state and are known as Class V injection wells.

When implementing any of the above drainage concepts, it is imperative that a maintenance program be established and monitored in order to reduce future problems.

Another proactive method of approaching karst drainage problems is the use of tracer studies (dye, pollen, spores, etc.) in identifying the subsurface drainage regime [48, 49, 54]. As discussed earlier (and in more detail) under proactive remedial measures – investigative measures – the use of tracers placed in the subsurface (via springs, wells, sinkholes, etc.) enables karst hydrologists to define the subsurface drainage basins and recharge areas.

The information gained through tracer studies is extremely useful in both design development and litigation, despite its time-consuming nature.

The use of sinkholes for drainage is often considered in roadway design. In a proactive methodology for environmentally sensitive roadway design in karst, the best approach would be to *avoid placing drainage into a sinkhole*, particularly using the sinkhole as a wastewater disposal feature. The use of sinkholes as drainage features is basically groundwater contamination by design. Eventually, hazardous waste will be spilled and the karst aquifer will become contaminated.

Concerns about groundwater contamination must be addressed by either a filtration system, a retention/sediment basin, or both. A study by Stephenson et al. [50] showed that using a pilot-researched peat-moss filtration system on highway runoff on the I-40/I-640 interchange in Knoxville, Tennessee (Fig. 31), effectively reduced the highway runoff contaminants by 90–99% [55]. This type of treatment of highway runoff is highly recommended for high-volume highways

Karst Terrane and Transportation Issues, Fig. 30 When roadways are built in sinkhole areas, sinkhole may need to be used for draining the runoff. Shown here is one such sinkhole that was cleaned of debris and protected against erosion and siltation with riprap and straw bales

Karst Terrane and Transportation Issues, Fig. 31 This shows the completed and functioning filtration system for a sinkhole that receives runoff from the I-640/I-40 interchange in East Knoxville, Tennessee. This is the prototype for such filtration systems that was the result of a Federal Highway Administration Pooled Fund Study on filtering highway runoff in karst areas

where the roadway runoff recharges karst aquifers via sinkhole features.

As discussed earlier under reactive remedial measures, there are a number of sinkhole improvement techniques that may be employed to ease the impact on the groundwater. These include siltation fences, trash racks, perforated and filter cloth wrapped standpipes, gabion revetment, erosion control geotextiles, riprap, and concrete structures such as box culverts and inlet structures.

Regardless of the type of method used to treat the sinkhole that is used for drainage, it is imperative that a maintenance program be established for the drainage facility. Without maintenance of the sinkhole improvement facility, silt and trash will eventually plug the sinkhole resulting in flooding.

Future Directions

The collapse of roadway surfaces, drainage ditches, and bridge foundations and numerous instances of flooding are karst-related issues and problems triggered by human construction activity on the environment. The greatest number of karst problems that develop along highways involves subsidence and collapse. This may involve the road or bridge footing itself or a drainage ditch. In some instances, the construction can trigger sinkhole formation outside of the construction limits of a project.

Remedial measures used in correcting karst-related subsidence problems may be divided into three areas: bridging, drainage, and relocation. Avoidance measures and some combination of drainage and bridging methods are usually the best direction to take in a proactive approach to designing and constructing highways in karst areas.

Innovative and cost-effective remedial concepts for solving karst-related geotechnical problems require modifications and refinement of the standard design to insure proper results to site-specific conditions. Proactive involvement by the geologic and engineering profession will be necessary to insure the success of karst-related remedial design concepts proposed for highways.

Bibliography

1. Black TJ (1984) Tectonics and geology in karst development of Northern Michigan. In: Barry B (ed) Proceedings of the 1st multidisciplinary conference on sinkholes, 15–17 Oct 1984, Orlando, pp 87–91
2. Hubbard DA (1984) Sinkhole distribution in the central and northern valley and ridge province, Virginia. In: Barry B (ed) Proceedings of the 1st multidisciplinary conference on sinkholes, 15–17 Oct, Orlando, pp 75–78
3. Myers PB, Perlow M (1984) Development, occurrence, and triggering mechanisms of sinkholes in the

carbonate rocks of the Lehigh Valley, Eastern Pennsylvania. In: Barry B (ed) Proceedings of the 1st multidisciplinary conference on sinkholes, 15–17 Oct 1984, Orlando, pp 111–115
4. Mylroie JE (1987) Influence of impermeable beds on the collapse of bedrock voids in the vadose zone. In: Karst hydrogeology: engineering and environmental applications. Proceedings of the 2nd multidisciplinary conference on sinkholes, Orlando, pp 95–99
5. Price DJ (1984) Karst progression. In: Barry B (ed) Proceedings of the 1st multidisciplinary conference on Sinkholes, 15–17 Oct 1984, Orlando, pp 17–22
6. Veni G (1987) Fracture permeability: implications on cave and sinkhole development and their environmental assessments. In: Karst hydrogeology: engineering and environmental applications. Proceedings of the 2nd multidisciplinary conference on sinkholes, Orlando, pp 101–105
7. White WB, White EL (1987) Ordered and stochastic arrangements within regional sinkhole populations. In: Karst hydrogeology: engineering and environmental applications. Proceedings of the 2nd multidisciplinary conference on sinkholes, Orlando, pp 85–90
8. Moore HL (1994) A geologic trip across Tennessee by interstate 40. University of Tennessee Press, Knoxville, p 339
9. Moore HL (2006) A proactive approach to planning and designing highways in East Tennessee karst. Environ Eng Geosci XII(2):147–160
10. Donaldson CW (1963) Sinkholes and subsidence caused by subsurface erosion. In: Proceedings, 3rd regional conference for Africa on soil mechanics and foundation engineering, pp 123–125
11. Jennings JE, Brink ABA, Louw A, Gowan GD (1965) Sinkholes and subsidence in the transvaal dolomites of South Africa. In: Proceedings of the 6th international conference of soil mechanics, pp 51–54
12. Moore HL (1981) Karst problems along Tennessee highways: an overview. In: Proceedings of the 31st annual highway geology symposium, Austin, pp 1–28
13. Moore HL (1987) Sinkhole development along 'untreated' highway ditchlines in East Tennessee. In: Beck B, Wilson W (eds) Karst hydrogeology: engineering and environmental applications. Proceedings of the 2nd multidisciplinary conference on sinkholes, Orlando, pp 115–119
14. Newton JG (1981) Induced sinkholes: an engineering problem. J Irrig Drain Div ASCE 107(IR3):175–185, Proceedings Paper 16343
15. Newton JG (1984) Natural and induced sinkhole development – Eastern United States: international association of hydrological sciences proceedings, third international symposium on land subsidence, Venice
16. Mathis H, Wright E, Wilson R (1985) Subsidence of a highway embankment on karst terrain. In: Proceedings of the 36th annual highway geology symposium, Clarksville, pp 14–27
17. Moore HL (1976) Drainage problems in carbonate terrain of East Tennessee. In: Proceedings of 27th annual highway geology symposium, Orlando, pp 112–131
18. Moore HL (2003) Recent sinkhole occurrences along highways in East Tennessee, a historical perspective. In: Proceedings of the 54th annual highway geology symposium, Burlington, 24–26 Sept 2003, pp 46–55
19. Newton JG (1976) Early detection and correction of sinkhole problems in Alabama, with a preliminary evaluation of remote sensing applications. Federal Highway Administration. HPR Report No. 76, Project 930-070, p 83
20. Ketelle RH, Newton JG (1987) Inventory of karst subsidence in the valley and ridge province of east Tennessee. In: Proceedings of 2nd multidisciplinary conference on sinkholes, Orlando, pp 25–29
21. Newton JG, Tanner JM (1987) Case histories of induced sinkholes in the Eastern United States. In: Karst hydrogeology: engineering and environmental applications. Proceedings of 2nd multidisciplinary conference on sinkholes, Orlando, A.A. Balkema, pp 15–23
22. Royster DL (1984a) Use of sinkholes for drainage. In: Construction and difficult geology: karstic limestone, permafrost, wetlands, and peat deposits; Trans. research board record 978, Washington, DC, pp 18–25
23. Moore HL (1984) Geotechnical considerations in the location, design, and construction of highways in karst terrain – 'The Pellissippi parkway extension', Knox-Blount Counties, Tennessee. In: Sinkholes: their geology, engineering, and environmental impact. Proceedings of the 1st multidisciplinary conference on sinkholes, Orlando, pp 385–389
24. Sowers G (1976) Foundations bearing in weathered rock. Proceedings speciality conference on rockeng. For foundations and slopes 2, University of Colorado, 15–18 Aug 1976
25. Foose RM, Humphreville JA (1979) Engineering geological approaches to foundations in the karst terrain of the Hershey Valley. Bull Assoc Eng Geol XVI(3): 355–381
26. Amari D, Moore HL (1985) Sinkholes and gabions: a solution to the solution problem. In: Proceedings of the 36th annual highway geology symposium, Clarksville, pp 47–68
27. Benson RC, Kaufmann RD, Yuhr LB, Martin D (1998) Assessment, prediction and remediation of karst conditions on I-70, Frederick. In: Proceedings of the 49th annual highway geology symposium, Prescott, 10–14 Sept, pp 307–312
28. Newton JG (1973) Sinkhole problem along proposed route of Interstate 459 near Greenwood. USGS circular # 83, 63 p
29. Newton JG (1971) Sinkhole problems in and near Roberts Industrial Subdivision, Birmingham. USGS circular #68, 42 p
30. Drumm EC, Yang MZ (2005) Preliminary screening of residual soil stability in karst terrain. Environ Eng Geosci XI(1):29–42

31. Kemmerly PR (1981) The need for recognition and implementation of a sinkhole-floodplain hazard designation in urban karst terrains. Environ Geol 3:281–292
32. Mills HH, Starnes DD, Burden KD (1982) Predicting sinkhole flooding in Cookeville. Tenn Tech J 17:1–20, (Tennessee Technological University), Cookville
33. Crawford NC (1981) Karst flooding in urban areas, Bowling Green, Kentucky. In: Proceedings of the 8th international congress of speleology, Western Kentucky University, Bowling Green
34. Quinlan JF, Ewers RO (1981) Hydrogeology of the mammoth cave region, Kentucky. Geological Society of America, Cincinnati field trip no. 8, pp 457–482
35. Quinlan JF, Ray JA (1990) Groundwater remediation may be achievable in some karst aquifers that are contaminated, but it ranges from unlikely to impossible in most: II. Implications for the mammoth cave area of long-term tracer tests and universal, nationwide failure in goal attainment by scientists, consultants, and regulators: In Karst hydrology. Proceedings of mammoth cave national park's first annual science conference: U.S. National Park Service, Mammoth Cave National Park, Kentucky, 17–18 Dec 1990, pp 113–114
36. Stephenson JB, Beck B (1995) Management of the discharge quality of highway runoff in karst areas to control impacts to groundwater – a review of relevant literature. In: Beck BF (ed) Karst geohazards – engineering and environmental problems in karst terrane. Proceeding of the fifth multidisciplinary conference on sinkholes and the engineering and environmental impacts of Karst, Gatlinburg, 2–5 Apr 1995, pp 297–321
37. Moore HL (1988) Treatment of karst along Tennessee highways. In: Sitar N (ed) Geotechnical aspects of karst terrains: exploration, foundation design and performance, and remedial measures. ASCE geotechnical special publication no. 14, pp 133–148
38. Sowers GF (1984) Correction and protection in limestone terrain. In: Beck B (ed) Sinkholes: their geology, engineering and environmental impact. Proceedings of the 1st multidisciplinary conference on sinkholes, Orlando, pp 373–378
39. Vance SJ (2003) Perceived risk versus cost in Karst remediation: a case history. In: Beck B (ed) Sinkholes and the engineering and environmental impacts of karst. Proceedings of the 9th multidisciplinary conference on sinkholes. ASCE geotechnical special publication no. 122, pp 652–658
40. Royster DL (1984b) Unpublished drawing of cap grouting on I-24, Montgomery Co., Tennessee. In: Geotechnical engineering section files, Tennessee Deptartment of Transportation, Nashville
41. Mellett JS, Maccarillo BJ (1995) A model for sinkhole formation on interstate and limited access highways, with suggestions on remediation. In: Barry B (ed) Karst geohazards, engineering and environmental problems in Karst Terrane. Proceedings of the 5th multidisciplinary conference on sinkholes, Gatlinburg, 2–5 Apr, pp 335–339
42. Benson RC, Yuhr L, Kaufmann R (2003) Some considerations for selection and successful application of geophysical methods. In: 3rd international conference on applied geophysics, geophysics 2003, Orlando, 8–12 Dec 2003
43. Hubbard DA (2003) Use of regional sinkhole mapping for sinkhole susceptibility maps. ASCE geotechnical special publication no. 122: sinkholes and the engineering and environmental impacts of Karst. In: Barry B (ed) Proceedings of the 9th multidisciplinary conference on sinkholes, Huntsville, 6–10 Sept, pp 61–71
44. Moore HL (2004) Preliminary geotechnical investigation of proposed interchange, I-181 with ext of SR 126, Sullivan, Co., unpublished report files, geotechnical engineering section, TDOT Region One, Knoxville Office
45. Hardeman WH (1966) Geologic map of Tennessee. Tennessee Division of Geology, Department of Environment and Conservation, Nashville
46. Moore HL, McDowell L (2008) The development and use of Karst maps in the location of highways in East Tennessee. In: Beck BF (ed) Sinkholes and the engineering and the environmental impacts of karst. Proceedings of the 11th multidisciplinary conference on sinkholes, 22–26 Sept 2008, Tallahassee. ASCE geotechnical special publication no. 183, pp 680–693
47. Niemiller M (2010) Photo of berry cave salamander. Accessed on 2008 from http://www.herpetology.us/niemiller/
48. Quinlan JF, Ray JA (1981) Groundwater basins in the mammoth cave region, Kentucky showing springs, major caves, flow routes, and potentiometric surface. Friends of Karst occasional publication, p 2
49. Quinlan JF, Ray JA (1981) Groundwater basins in the mammoth cave region, Kentucky: plate 1 in central Kentucky cave survey bulletin no. 1, Center for Cave and Karst Studies, Western Kentucky University, 1:138,000 map
50. Stephenson JB, Zhou WF, Beck B, Green T (1997) Management of highway storm water runoff in karst areas – baseline monitoring and design of a treatment system for a sinkhole at the I-40/I-640 interchange in eastern Knoxville, Tennessee. In: Proceedings of the 48th annual highway geology symposium, Knoxville, 8–10 May 1997, pp 24–34
51. Crawford NC, Groves CG (1984) Storm water drainage wells in the karst areas of Kentucky and Tennessee. U.S. Environmental Protection Agency and Center for Cave and Karst Studies, Bowling Green, Kentucky, p 52
52. Crawford NC, Groves CG (1995) Sinkhole collapse and groundwater contamination problems resulting from storm water drainage wells on karst terrain. In: Beck B (ed) Karst geohazards. Proceedings of the 5th multidisciplinary conference on sinkholes, Gatlinburg, pp 257–264

53. Reeder PP, Crawford NC (1989) Potential groundwater contamination of an urban karst aquifer – Bowling Green, Kentucky. In: Beck BF (ed) Engineering and environmental impacts of sinkholes and karst. Proceedings of the 3rd multidisciplinary conference on sinkholes and the engineering and environmental impacts of karst, Orlando, 2–4 Oct 1989, pp 197–206
54. Crawford NC (2003) Karst hydrogeologic investigation for proposed Kentucky trimodal transpark. In: Barry B (ed) ASCE special publication no. 122, sinkholes and the engineering and environmental impacts of Karst, Huntsville, 6–10 Sept, pp 385–403
55. Beck BF, Stephenson JB, Wanfang Z, Smoot JL, Turpin AM (1996) Design and evaluation of a cost effective method to improve the water quality of highway runoff prior to discharge into sinkholes. In: Proceedings of the 1996 Florida environmental expo, Tampa, 1–3 Oct 1996, pp 155–164

Land Subsidence in Urban Environment

M. Adrián Ortega-Guerrero[1] and
José Joel Carrillo-Rivera[2]
[1]Centro de Geociencias, UNAM Juriquilla, Querétaro, Mexico
[2]Instituto de Geografía, UNAM CU, Coyoacán, Mexico

Article Outline

Glossary

Aquifer unit Aquifer unit is a geological formation, part of a formation, or a number of formations that provide water substantially and in an adequate quality for the expected usage.

Aquitard Aquitard is a geological formation that although insufficiently in producing water as an aquifer unit does, the volume of water that it allows to be released from storage may provide an adverse environmental impact as subsidence.

Compressibility Relates the change in volume, or strain, induced in a soil under an applied stress.

Effective stress Represents the stress transmitted to the full-saturated soil skeleton when a force per unit area (total normal stress) is transmitted in a normal direction across the measuring plane.

Hydraulic conductivity Hydraulic conductivity is the rate of water that an unit volume of aquifer material may allow through under a unit hydraulic gradient such value is a function of its degree of saturation attaining it maximum at 100% saturation, and is also function of water density.

Hydraulic head Hydraulic head is the sum of two components: the elevation of the point of measurement and the pressure head (assuming nil flow velocity).

Specific storage Specific storage of a saturated aquifer is the volume of water that a unit volume of aquifer releases from storage under a unit decline in hydraulic head.

Void ratio Void ratio is the ratio of the volume of voids to the volume of solids.

Definition of the Subject and Its Importance

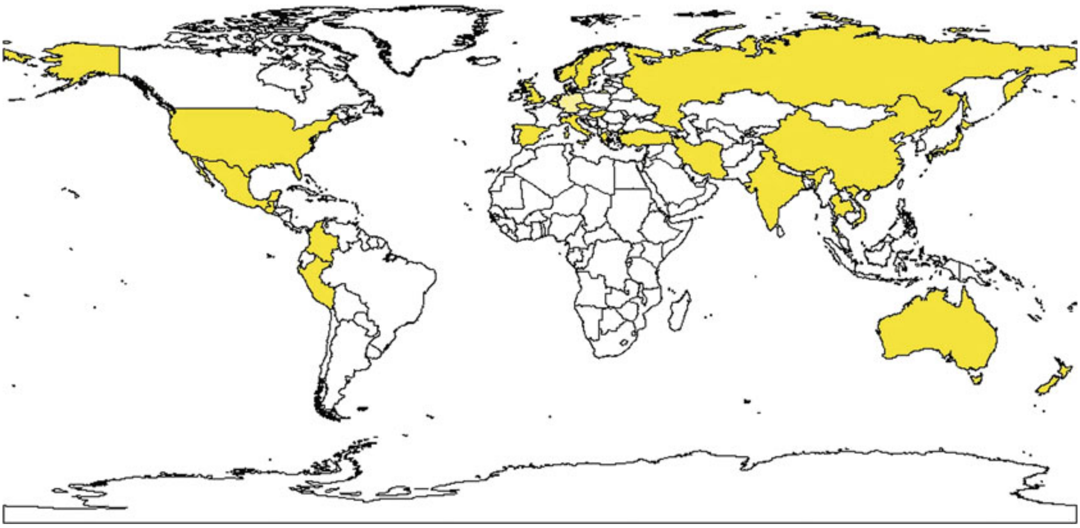
Progressive or sudden decrease in ground surface elevation, defined as land subsidence, has risen due to different issues related to the urban environment. These issues involve primarily the geologic evolution of a region that determines the type of rock formations, presence of natural raw material, morphology, soil, groundwater presence, and ecological complexity where a city might develop. The evolution of the urban environment is closely linked to the extraction of one or more natural raw materials, which under particular situations may expose urban populations to risk from land subsidence and negatively affect integrity of housing, urban infrastructure, public and private property, and also involve costs of damage in the order of several millions of dollars.

Subsidence is a global problem (Fig. 1). More than 200 cases have been reported in the world (USGS International Survey of Land Subsidence Database) and many more may be reported in the near future as the causes and evolution are better understood.

Due to the importance and implications of land subsidence, the United Nations Educational, Scientific and Cultural Organization (UNESCO) included this topic to be studied under the

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_440

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC, 2012
https://doi.org/10.1007/978-1-4419-0851-3_440



Land Subsidence in Urban Environment, Fig. 1 World map showing countries with problems of land subsidence

International Hydrological Programme (IHP). Each country has the responsibility of collecting, maintaining, and interpreting the basic related data (fluid extraction, mining, geotechnical properties, etc.) and associated land subsidence rates, usually with the participation of professional organizations in engineering represented by the areas of soil mechanics, groundwater hydraulics, geological and geophysical sciences, as well as consulting companies, and research groups at universities.

Introduction

The scope of land subsidence is wide. Although several of the processes involved are included in this entry, emphasis is given to the slow response of the soil due to the extraction of subsurface fluids, especially groundwater related to the water supply for large urban environments.

Causes of Land Subsidence

Land subsidence is caused by different processes on the crust of the Earth resulting from natural or anthropogenic activities such as: (1) direct withdrawal of fluids (water, gas, geothermal, oil, brine); (2) mining (ore deposits, construction earth materials) and tunneling; (3) application of water to unconsolidated moisture deficient soils; (4) dewatering of soils with high organic content usually located in groundwater discharge zones; (5) loading by engineering structures (buildings, bridges); (6) collapse of subsurface cavities in soluble rocks or soft volcanic deposits; (7) geologic loading; and (8) tectonic deformation.

Land subsidence may behave as a gradual soil settling or as a sudden sinking or collapse. Gradual settling of the land is usually due to withdrawal of fluids, application of water to unconsolidated moisture deficient soils, dewatering of organic soils, geologic loading, and tectonic deformation. Whereas, sudden sinking, is usually associated to collapse of the roof of a cavity formed due to underground mining, to the dissolution of soluble rocks (such as limestone) or to water flow under low consolidated volcanic materials, e.g., the well-covered Guatemalan event [1]. Gradual subsidence may affect areas from tens to hundreds of square kilometers, whereas, sudden sinking involves smaller areas in the order of hundreds of square meters. The predictability and the impacts on humans, their property, and urban infrastructure also differ from one case to another.

Different rates of land subsidence or vertical deformation causes development of earth fissures with different urban consequences, and may create risk for humans and damage to housing, buildings, and urban infrastructure, and may also represent more rapid pathways for contaminants to reach groundwater.

Table 1 shows 27 countries where different problems of land subsidence have been reported. From them, groundwater extraction and collapse due to underground mining and karst development represent the highest risk to human population, their properties, and urban infrastructure.

Land subsidence response is particularly critical in the urban environment where population and urban infrastructure are concentrated (housing, buildings, roads, surface and underground transportation system, distribution lines for water, oil, gas, among others).

From the mean causes for land subsidence presented above, the extraction of natural earth fluids (water, oil, gas) that fill pores of underlying sediments and fractures of rocks is paramount. These fluids perform different important services

Land Subsidence in Urban Environment, Table 1 Countries with land subsidence response

Main land subsidence origin	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Australia		✓			✓	✓		
Belgium		✓						
China	W,G,Gt		✓					
Colombia		✓						
Czech Republic								✓
Germany	B					✓		
Greece	W							
Guatemala						✓		
Hungary	W	✓	✓					
India		✓						
Iran	W							
Israel	W					✓		
Italy	W,G,Gt,B	✓						✓
Japan	W,G	✓		✓	✓			✓
Mexico	W			✓	✓			
Netherlands	W,G	✓		✓	✓			✓
New Zealand	B,Gt							
Norway	W							
Peru								✓
Russia	W	✓						
Spain	W					✓		
Sweden	W			✓				
Thailand	W			✓	✓			
Turkey		✓						
United Kingdom	W,B	✓				✓		
United States	W,B,G,O	✓	✓	✓	✓	✓		✓
Vietnam	W	✓						

Adapted from USGS International Survey on Land Subsidence, land subsidence proceedings 1991, 1995 *W* Water, *G* Gas, *O* Oil, B Brine, Gt Geothermal
(1) Direct withdrawal of fluids (water, gas, geothermal, oil, brine)
(2) Mining
(3) Application of water to unconsolidated moisture deficient soils
(4) Dewatering of soils with high organic content
(5) Loading by engineering structures
(6) Collapse of subsurface cavities in soluble rocks or volcanic rocks
(7) Geologic loading
(8) Tectonic deformation

to humans, such as domestic, industrial, and irrigation water; oil and gas for energy generation and combustible formulations; and others. Groundwater extraction represents the more common and critical process of gradual land subsidence (Table 2). Countries like Japan, Mexico, and the United States have reported large-scale land subsidence affecting populated urban environments, such as Mexico City; Tokyo and Osaka; Santa Clara Valley and San Joaquin Valley in California; Houston-Galveston, Texas; Las Vegas, Nevada; and South-Central Arizona. For example, at the beginning of the twenty-first century, observed subsidence response in Mexico City has been recorded to be in the order of 0.01–0.50 m/year. There is local variation in subsidence rates, but at some sites soil consolidation is causing cumulative lowering of the land subsidence of several meters.

A classic example in the technical and scientific literature is the land subsidence response in downtown Mexico City and surrounded areas. The city is built on fine lacustrine sediments, overlying a regional aquifer composed of both granular and fractured volcanic units, in a natural closed basin. More than 50 m^3/s of groundwater are extracted from this aquifer unit causing cumulative land subsidence of more than 10 m by 2010 after more than 150 years of groundwater extraction. In the Chalco Plain, the site of a past lake on the south eastern outskirts of Mexico City, land subsidence reached 15 m in 2010 after 35 years of extraction. Large-scale fractures have also developed associated with land subsidence, increasing the risk to inhabitants. The Chalco Plain suggests the result of one of the most critical land and water use practices carried out to encourage unplanned urban growth; the extracted groundwater is not used locally, but transferred to add to the water supply of the larger Mexico City area. Overpumping is causing severe damage to the local hydraulic infrastructure. In one case, anurban-industrial sewage canal broke due to subsidence which affected the health and property of thousands of people, and had repair and recovery costs of several million dollars.

Land Subsidence in Urban Environment, Table 2 Recorded total land subsidence due to groundwater extraction

Land	Locality subsidence (m)	Area affected (km^2)
Central Valley CA, USA	9.0	13,500
Houston TX, USA	2.75	12,170
Eloy ARZ, USA	3.6	8,700
Tokyo, Japan	4.6	2,400
Po Valley, Italy	3.0	780
City of London, Great Britain	0.35	450
Venice, Italy	0.14	400
Mexico City, Mexico	10.0	225
Chalco Plain, near Mexico City	15.0	60

Related Issues in an Urban Environment

From the international experience, different issues have been identified in an urban environment associated to land subsidence. Better understanding of these issues would permit reduction of risk and associated costs to urban population as well as to private and government property.

Geology

Geology is the science that studies the interior of the Earth and its origin, composition, and evolution. The geological evolution of a particular region determines the type of rock formations and soil characteristics where cities are built. The presence and distribution of fluids (groundwater, oil, gas, geothermal) also depend on the geological evolution. Geological understanding of the urban environment is therefore essential to determine the risk for either gradual soil settling or as a sudden sinking or collapse.

Different disciplines of geology study particular topics associated to geologic evolution and natural raw materials. Among these, hydrogeology, and in particular the flow system theory, proposes a wide system view for the analysis of groundwater in the environment which could assist in the definition of processes involved in land subsidence.

Properties of the Soils

Soils are formed by physical, chemical, or biological destruction (weathering) of rock units assisted by erosion processes and other factors. Physical weathering takes place in situ by the action of wind, water, or glacier that disintegrate a rock unit (alternate freezing and thawing in cracks in the rock), where the resulting grains or particles retain the same mineralogical composition as that of the parent rock. Chemical weathering, on the other hand, produces changes in the mineral composition of the parent rock, where the presence of water, carbon dioxide, and oxygen play a major role in the chemical reactions. Biological activity plays an important role in the disintegration of a rock unit and in the formation of soils from that of microorganisms to large individuals through, for example, allopathic reactions or through the roots of trees. Clay minerals (<2 μm) resulting from chemical weathering constitute basic structural units of silicon and oxygen (silica tetrahedron), which may combine with alumina and hydroxyl (alumina octahedron) to form combined sheets. Depending on the packing organization of the silica sheets, alumina sheets, and presence of water molecules or ions such as potassium, magnesium, or iron, different clay minerals are formed with specific physical and chemical properties. Some of these clay minerals have important volumetric changes depending on the water content.

Particle sizes in soils usually vary between less than 0.001 mm to more than 100 mm. Most types of soil consist of a graded mixture of particles between specific limits, this mixture is known as a porous media. The presence of clay minerals usually exerts a major influence on the properties of a soil, an influence out of all proportion to their percentage per weight in the soil [2]. Determination of particle size analysis and physical properties (plasticity, liquid and plastic limits, natural water content, porosity, void ration, compressibility, etc.) of the fine-grained particles (between 0.06 mm and less than 0.001 mm) constitutes an important issue to understand land subsidence.

Among the many different physical properties of a soil, compressibility (α) relates the change in volume, or strain, induced in a soil under an applied stress and may be represented as:

$$\alpha = \frac{-dVt/Vt}{d\sigma e} \tag{1}$$

where

−dVt is the reduction in the total volume of the soil mass

Vt is the total volume of the soil mass

dσe is the change in effective stress

Table 3 shows the range of value of compressibility for different porous materials and rocks large compactions occur in clayey soils due to the larger compressibility (a).

Compressibility controls the magnitude of land subsidence under loading by engineering structures or geologic loading.

Fluid Withdrawal and Properties Variation in Time

Extraction of fluids (water, gas, oil, geothermal) causes a depressurization of the granular materials or fractured rocks that is manifested as compaction. The amount and related importance of parameters of the geologic media and fluid (liquid or gas) vary from one case to another.

To illustrate the hydraulic and geomechanical parameters and to understand the groundwater flow system, the following example for Mexico City is presented. In this case, land subsidence is associated with groundwater extraction, where lacustrine clayey deposits overly a sandy aquifer unit. The transient hydraulic response and consolidation of the lacustrine clayey deposits can be demonstrated by using a coupled, one-dimensional, groundwater flow nonlinear deformation, fine element model [4]. Other aspects of the nonlinearity of

Land Subsidence in Urban Environment, Table 3 Compressibility factors of different geological materials and water

Material	Compressibility (α) (m^2/N or Pa^{-1})
Clay	10^{-6}–10^{-8}
Sand	10^{-7}–10^{-9}
Gravel	10^{-8}–10^{-10}
Jointed rock	10^{-8}–10^{-10}
Sound rock	10^{-9}–10^{-11}
Water	4.4×10^{-10}

Modified from Freeze and Cherry [3]

parameters for the Mexico City aquitard were presented by Rivera et al. [5]. The 1-D groundwater flow equation that solves for the hydraulic head (h') distribution within an aquitard in terms of the stress-dependent parameters of hydraulic conductivity (K') and specific storage (Ss') is:

$$\frac{\partial}{\partial z}\left(K'(e)\frac{\partial h'}{\partial z}\right) = Ss'(e, \sigma_e)\frac{\partial h'}{\partial t} \qquad (2)$$

This model incorporates empirical expressions that relate K' (L/T) and Ss' to the soil mechanics parameters of void ratio (e) and effective stress (σ_e) (M/LT^2) as:

$$de(\sigma_e) = C_c \log\left(\frac{\sigma_{eo} + d\sigma_e}{d\sigma_{eo}}\right) \qquad (3)$$

Equation 3 represents the change in void ratio (e) as a function of a change in effective stress ($d\sigma_e$); C_c is the compression index obtained from the slope of the linear portion of the e-log σ_e plot, and σ_{eo} is the effective stress under current hydraulic head conditions prior to the subsequent pore pressure change:

$$Ss'(e, \sigma_e) = \frac{\rho g C_c \log\left(\frac{\sigma_{eo} + d\sigma_e}{\sigma_{eo}}\right)}{d\sigma_e(1.0 + e_o)} \qquad (4)$$

Equation 4 provides an expression for Ss' in terms of void ratio and total stress:

$$dK'(e) = Ko'(e)\left(10^{de/m} - 1\right) \qquad (5)$$

Equation 5 relates the change in hydraulic conductivity of the aquitard (K') to variations in void ratio (e); where K'_o is the initial hydraulic conductivity and (m) represents the slope of the e-log K' plot [6].

An adequate evaluation of the land subsidence process requires a good understanding of the regional and local groundwater flow system function [7], and a precise quantification of evolution of groundwater flow direction and the hydraulic and geomechanical properties of the porous media in both aquifer units and aquitards in time and space. When the groundwater flow system function is not well understood or ignored, negative consequences have occurred due to arbitrary groundwater extraction. For example, negative changes in groundwater quality could be associated with either the overlying aquitard pore water [8] or the uprising thermal groundwater from deep regional flow systems [9, 10].

Underground Construction (Tunneling) and Solid Extraction

Tunneling is a common practice in the urban environment to conduct fresh or residual water, gasoline, gas, oil, or underground transportation systems. Solid raw material extraction such as from ore deposits, coal, or geologic materials for construction often use tunneling to get access and to extract materials. Surface settlements during tunnel construction depend on a number of factors that include geological conditions, presence of groundwater and geotechnical conditions, design of the tunnel (geometry and depth) and excavation-construction methods. Shallow tunnels will generally have a greater effect on surface structures than deep ones.

Underground construction requires detailed knowledge of the geology, hydrogeologic and geotechnical properties of rocks and soils for a long-term design that reduces the possibilities of sudden settlements. Existing underground constructions can also be analyzed and instrumented to quantify any rate of land subsidence associated to tunneling. Predictive models can be developed to identify zones of potential collapse.

Environmental Effects

Land subsidence is manifested as changes in ground surface elevation that progressively develop topographic depressions where meteoric water may accumulate, affecting native vegetation and fauna, among other. A good example of this environmental effect is a shallow lake that is forming in a topographic depression caused by the land subsidence in the middle of the Chalco plain, in the Basin of Mexico near Mexico City [11]. Land subsidence is causing loss of agricultural land and encroachment of a shallow lake toward the rapidly expanding urban sprawl areas. To control the accumulation of runoff and sewage, drainage canals are being deepened and increased in length. Pumping stations for canal water have been put into operation to control the growth of the surface water body during the rainy season. Land subsidence is also causing damage to sewage canals and their periodic rupture produce consequent flooding in nearby urban areas.

Large-scale fractures in the Chalco Plain have also developed as a consequence of groundwater extraction; these fractures are important due to their hydraulic implications in the groundwater flow and solute transport controls in the lacustrine aquitard to the underlying aquifer [12].

Urban Planning and City Growth

Several issues related to the unplanned growth of Mexico City have accompanied its development since historical times. The Aztecs were the initial dwellers that constructed the city of Tenochtitlan harmoniously with the natural conditions of the prevailing lake system. However, the construction and expansion of the city after the Spanish conquest in 1,521 have included a continuous struggle with water management issues instead of reaching a coexistence with nature. As population increased, flooding due to rainfall made surface water pooling undesirable, so the lakes were desiccated by the turn of the twentieth century through an artificial opening of the basin. Groundwater emerged in mid 1800s as a solution to the water requirements of the city. Population and industrial activities in the City grew rampantly by the second half of the twentieth century. The total population reached 14,987,000 inhabitants [13], so groundwater has been the major component. By 1905, more than 70% of the water supply to the city was from groundwater. Dramatic subsidence incidents reported by the mid 1950s reached 0.44 per year, mainly in the downtown area. By the 1980s, groundwater extraction in Mexico City proper was stopped and shifted to the surrounding areas, such as Chalco Plain, to provide water to Mexico City, as well as for the increasing population of its urban sprawl and increased prominent economic activities.

Monitoring of Land Subsidence

Global Positioning System (GPS) in combination with Total Topographic Stations are traditionally used to develop detailed topographic elevation surveys of the ground surface. Differences in ground surface elevation obtained for different periods of time would permit measurement of the rates and distribution of land subsidence. Yearly topographic surveys are usually recommended. Radar Interferometry is an indirect satellite technique that has also been used to track changes in ground surface elevation.

Economics

Costs and potential costs of damage to private and government property (urban infrastructure) caused by land subsidence are becoming important issues. Insurance against health implications, both physical or psychological, or property damage or loss in some subsidence-prone areas of the world involve millions of dollars to the urban infrastructure and private property.

Economic studies now consider the benefits of fluid, earth materials, or ore deposits extraction against the costs to repair damages caused by land subsidence in the urban centers and surrounding areas.

Future Directions

Prediction of land subsidence and its control represent one of the main scientific challenges in the urban environment. Prediction depends first on understanding the subsidence events occurring on the crust of the Earth resulting from natural processes such as collapse of subsurface cavities in soluble rocks or soft volcanic deposits, geologic loading, and tectonic deformation or from anthropogenic activities, for example, direct withdrawal of fluids, mining and tunneling, application of water to unconsolidated moisture deficient soils, dewatering of soils with high organic matter, and loading by engineering structures. Accurate predictions also depend on the quality and density of longterm instrumentations to measure deformation of the soils and rocks and particularly laboratory and field instrumentation to determine hydraulic and geomechanical parameters involved into the physical processes of land subsidence.

Understanding and quantification of the evolution of the groundwater flow systems represent an important area for research. Long before human intervention, the local environmental system took thousands of years to develop. Location of urban centers usually began in groundwater discharge areas where water was easily available. However, as groundwater and other fluids are extracted from underground, groundwater quality as well as its temperature and flow direction changes with time as extraction progresses. An understanding of these effects and their impact, not only of the obtained physical characteristics of the water, but the imposed constrains on the development of subsidence could became important issues to incorporate into subsidence studies as to understand related processes and possible options for their control. Due to the scale and time involved in the reaction of these systems, long-term evaluation and monitoring programs are needed.

Conclusion

The particular phenomena of land subsidence is related to geologic evolution, properties and distribution of sediments, rocks fluids and ore deposits, in addition to the urban setting, evolution, and planning. Ground-water extraction represents one of the most important factors causing land subsidence in the urban environment. Therefore, it is important to determine the functioning of the groundwater flow systems at regional and local scales. It is also important to consider that groundwater extraction has different effects on the groundwater flow regime if it is tapped in a recharge, a transit, or a discharge area. The type of flow hierarchy, local, intermediate, or regional [7], to be induced into the particular extraction site could trigger different subsidence responses.

Bibliography

Primary Literature

1. National Geographic (2010) http://news.nationalgeographic.com/news/2010/06/100601-sinkhole-in-guatemala-2010-world-science/
2. Craig RF (1987) Soil mechanics. Van Nostrand Reinhold, New York
3. Freeze RA, Cherry JA (1979) Groundwater. Prentice-Hall, Englewood Cliffs, p 604
4. Rudolph DL, Frind EO (1991) Hydraulic response of highly compressible aquitards during consolidation. Water Resour Res 27(1):17–30
5. Rivera A, Ledoux E, de Marsily G (1991) Non-linear modelling of groundwater flow and total subsidence in the Mexico City aquifer-aquitard system. In: Land subsidence proceedings of fourth international symposium of land subsidence, 200. International Assocation of Hydrological Sciences, Gentbrugge, pp 45–58
6. Lambe TW, Whitman RV (1969) Soil mechanics. Wiley, New York
7. Tóth J (1999) Groundwater as a geological agent: an overview of the causes, processes, and manifestations. Hydrogeol J 7:1–14
8. Ortega GA, Cherry JA, Rudolph DL (1993) Large-scale aquitard consolidation near Mexico City. Ground Water 31(5):708–718
9. Edmunds WM, Carrillo-Rivera JJ, Cardona A (2002) Geochemical evolution of groundwater beneath Mexico city. J Hydrol 258:1–24
10. Huizar-Alvarez R, Carrillo-Rivera JJ, Angeles-Serrano G, Hergt T, Cardona A (2004) Chemical response to groundwater extraction southeast of México City. Hydrogeol J 12:436–450
11. Ortiz-Zamora DC, Ortega-Guerrero MA (2010) Evolution of long-term land subsidence near Mexico City: review, field investigations, and predictive simulations. Water Resour Res 46:W01513. https://doi.org/10.1029/2008WR007398
12. Ortega GA, Rudolph DL, Cherry JA (1999) Analysis of long term land subsidence near Mexico City: field investigations and predictive modeling. Water Resour Res 35(11):3327–3341
13. AIC (1995) El Agua y la Ciudad de México. Academia de la Investigación Científica, Academia Nacional de Ingeniería, Academia Nacional de Medicina, National Academy of Sciences (through the National Research Council), p 364
14. Bouwer H (1978) Groundwater hydrology, series in water resources and environmental engineering. Mc Graw-Hill, Sydney, p 480
15. Carrillo-Rivera JJ (1998) Monitoring of exploited aquifers resulting in subsidence, example: Mexico City. Studies and reports in hydrology No 57. In: Van Lanen HAJ (ed) Monitoring for groundwater management in (semi-)arid regions. UNESCO, Paris, pp 151–165

Books and Reviews

Carrillo N (1947) Influence of artesian wells in the sinking of Mexico City. In: Carrillo VN (ed) Comision Impulsora y Coordinadora de la Investigacion Cientifica, Anuario 47. Secretaria de Hacienda y Credito Publico, Mexico City, pp 7–14, 1969

Dassargues A, Schroeder Ch, Li XL (1993) Applying the Lagamine model to compute land subsidence in Shanghai. Bull Eng Geol (IAEG) 47(1):13–26

Figueroa GE (1987) Structural stability problems of wells and aquifers. en: Workshop on leaky aquifer mechanics, conference proceedings. Universidad Nacional Auto'noma de México, Instituto de Geofisica, México, pp 53–61

Figueroa GE (1989) Mecanismos de producción de grietas inducidos por la explotación del agua subterránea. Academia Mexicana de Ingeniería, Alternativas Tecnológicas 29, México, pp 33–48

Juárez-Badillo E (1975) Constitutive relationships for soils. In: Symposium on recent developments in the analysis of soil behaviour and their application to geotechnical structures. University of New South Wales, Kensington, pp 231–257

Juárez-Badillo E, Figueroa GE (1984) Stresses and displacements in an aquifer due to seepage forces (one-dimensional case). J Hidrol 73:259–288

Gambolati G, Freeze RA (1973) Mathematical simulation of the subsidence of venice, 1, theory. Water Resour Res 9(3): 721–733

Helm DC (1976) One-dimensional simulation of aquifer system compaction near Pixley, California, 2, stress-dependent parameters. Water Resour Res 12:375–391

Herrera I, Figueroa GE (1969) Integrodifferential equations for systems of leaky aquifers. Water Resour Res 5(4):900–904

Herrera I, Rodarte L (1973) Integrodifferential equations for systems of leaky aquifers and implications, the nature of approximate theories. Water Resour Res 9(4):994–1005

Herrera IR, Yates R, Henart JP (1982) Estudio de hundimiento y balance de acuíferos subterráneos en la Ciudad de México. In: Proyecto elaborado para el Departamento del Distrito Federal por el Instituto en Investigaciones Aplicadas. Universidad Nacional Autonoma De Mexico, México

Hiriart F, Marsal RJ (1969) The subsidence of Mexico City. In: Volumen Nabor Carrillo, Comision impulsora y coordinadora de la investigacion Cientifica, Anuario 47. Secretaria de Hacienda y Credito Publico, Mexico City, pp 109–147

Lambe TW, Whitman RV (1969) Soil mechanics. Wiley, New York

Neuman SP, Witherspoon PA (1969) Applicability of current theories of flow in leaky aquifers. Water Resour Res 5(4):817–829

Neuman SP, Preller C, Narasimhan TN (1982) Adaptive explicit-implicit quasi three-dimensional finete element model of flow and subsidence in multiaquifer systems. Water Resour Res 18(5):1551–1561

Stamatakos JA, Connor CB, Martín RH (1997) Quaternary basin evolution and basaltic volcanism of Crater Flat, Nevada, From detailed ground magnetic surveys of the little cones. J Geol 105:319–330

Telford WM, Geldart LP, Sheriff RE (1990) Applied geophysics, 2nd edn. Cambridge University Press, Cambridge/New York

Part III

Water and the Environment

Fresh Water Geochemistry: Overview

Pedro José Depetris
Academia Nacional de Ciencias, Córdoba, Argentina

Article Outline

Glossary

Aquifer An underground layer of water-bearing permeable rock, rock fractures, or porous unconsolidated materials (gravel, sand, or silt) from which groundwater can be obtained by means of a water well.

Colloid A mixture in which one substance of microscopically dispersed insoluble particles is suspended throughout another substance. The dispersed-phase particles have a diameter of approximately between 1 and 1000 nanometers (1 nm = 10^{-9} m).

Congruent dissolution A mineral or salt is completely dissolved in water, adding elements to the solvent in the same proportions that existed in the original solid.

Conservative elements Cl^-, SO_4^{2-}, NO_3^-, Ca^{2+}, Mg^{2+}, Na^+, and K^+ are considered *conservative* in the sense that their concentrations are unaltered by changes in pH, temperature, or pressure, assuming that no precipitation or dissolution of solid phases or biological transformations occur within the ranges normally found near the surface of the Earth.

Denudation Involves the processes that cause the wearing away of the Earth's surface by moving water, by ice, by wind, and by waves, leading to a reduction in elevation and in relief of landforms and landscapes.

Exogenous cycle The sum of processes triggered by forces on or above the Earth's surface.

Hydrolysis A chemical reaction wherein a water molecule and a reactant exchange functional groups resulting in two end products, one containing the hydrogen ion and the other the hydroxyl anion.

Incongruent dissolution A mineral that does not dissolve entirely and leaves a solid residue which usually differs chemically from the original solid.

Riparian zone Interface between land and a river or stream; plant habitats and communities along the river margins and banks, characterized by hydrophilic plants.

Solubility product It is the equilibrium constant for a simple solubility reaction. More precisely, it is the product of the concentrations of the ions, with each concentration raised to a power equal to the coefficient of that ion in the balanced equation for the solubility equilibrium.

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_969

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_969-1

Definition and Significance of Surface Fresh Water Geochemistry

Geochemistry is commonly divided into a number of specialties. A customary division is between *high-temperature* and *low-temperature geochemistry*. The former is concerned with the processes that occur at and above the temperature of hydrothermal systems (i.e., 350–400 °C) and, therefore, is directly related to the geological fields of metamorphic and igneous petrology. The latter is usually understood as the subdiscipline that addresses to those physicochemical processes which take place at temperatures up to ~100°C. It focuses on processes occurring in the atmosphere, in the oceans, and in the *critical zone* (CZ), which is defined as the "portion of the Earth's surface that includes the atmosphere, the biosphere, the pedosphere, and the lithosphere interfaces" [1]. In a more descriptive fashion, the CZ is the "the Earth's permeable surface, from the top of the trees to the bottom of the groundwater zone, most of which involve reactions with, in, or mediated by water" [2] (Fig. 1). It is clear then that low-temperature geochemistry involves such topics as chemical weathering, soil chemistry, mineral dissolution and precipitation, sedimentary chemistry, and diagenesis, along with other, more restricted fields. *Surface fresh water geochemistry* concerns the use of the geochemical background to study the characteristics and the processes that govern the chemistry of rivers, streams, and lakes (i.e., fresh water). It is, therefore, a subset of low-temperature geochemistry and involves the study of the *natural* mechanisms that control fresh water (i.e., sufficiently dilute water to be potable, e.g., $<1000\ \mathrm{mg\ l^{-1}}$) chemistry. The *human-made* alterations that impact on the quality of fresh water are discussed within the field of *environmental geochemistry.*

Introduction

The roots of geochemistry as an independent scientific field are linked to early chemists. It was a motivation for such pioneers probing into the essence of natural substances, including minerals, rocks, and water. Chemists like Robert William Boyle (1627–1691), John Dalton (1766–1844), and Jöns Jacob von Berzelius (1779–1848), all considered forefathers of modern chemistry, were intellectually interested in the philosophical study of nature and the physical universe, which were prevailing knowledge foci before the development of modern science.

The term geochemistry denotes the intersection of chemistry and the Earth Sciences. It was first coined by the Swiss-German chemist Christian Friedrich Schönbein in 1838 [3]. The specialty applies to the broad theoretical framework of chemistry and the geosciences to the comprehension of the Earth, the solar system, and beyond. It became clear in modern times that such knowledge implied not only the wide-ranging understanding of our planet but also guidelines for the usage of the involved resources to improve the human condition and to safeguard life on Earth from the ill consequences of their unsound utilization [2].

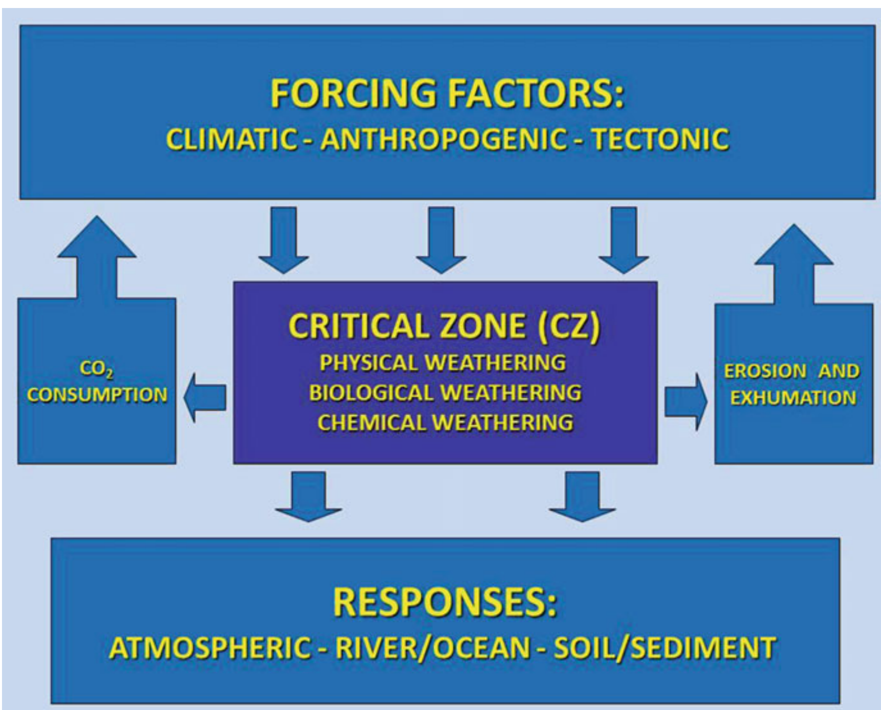
Victor Moritz Goldschmidt (1888–1947) was a major influence in the development of geochemistry in the twentieth century. Born in Switzerland but spending his career mostly in Norway and Germany, Goldschmidt has been designated as the "father of modern geochemistry." In a succession of publications in the 1920s and 1930s entitled *Geochemische Verteilungsgesetze der Elemente*, Goldschmidt laid the foundation for modern geochemistry as a discipline.

Another important figure, particularly in the history of fresh water geochemistry, is Frank Wigglesworth Clarke (1847–1931), who had begun to investigate the abundances of various elements within the Earth. His famous *Data of Geochemistry* was first published in 1908 [4] and was followed by many editions published by the US Geological Survey (USGS). Of particular interest in the field of water geochemistry was Clarke's well-known compilation of water analyses [5] which was updated in subsequent editions. Of special value in dispersing knowledge on fresh water chemistry has been the series of Water-Supply Papers, initiated in the twentieth century and published thereafter for many decades by the USGS, a role which is now mainly undertaken by dozens of scientific journals published by international editorials or scientific associations.

The core of the natural processes that determine fresh water geochemistry lies in a series of mechanisms collectively known as *weathering*. Rock weathering generates *regolith* – the mantle of in situ and transported altered material that covers landscapes across the planet and substances which are dissolved in water. There are three main types of weathering occurring at or near the Earth's surface: physical or mechanical, biological, and chemical. The attention here will be placed in the latter, but since the other types usually precede the occurrences of chemical processes – or, sometimes, act concurrently – a brief description of each of the other two follows: the disintegration of rocks with no significant chemical alteration involved in the process is known as physical (or mechanical) weathering, in which relationship with the chemical phase of the rock weathering process is close because initial chemical attack may decrease the strength of a rock, thus enabling the ensuing physical breakdown; on the other hand, living organisms, through biophysical or biochemical action, participate in material breakdown, hence contributing to the generation of regolith [6, 7].

Most minerals formed under igneous and metamorphic conditions are unstable at low temperatures and under near-surface hydrous conditions, with ample availability of oxygen and carbon dioxide. Eventually, they react to produce dissolved components (i.e., ions and molecules), mineral debris (e.g., quartz grains, rock fragments), and new mineral precipitates (e.g., oxides and hydroxides, clay minerals). The naturally occurring processes (e.g., mineral solution, hydrolysis, carbonation) that break down minerals and rocks are jointly known as chemical weathering (Fig. 1).

As specified above, the products of weathering and the formation of regolith are major participants in the chain of natural events that define



Fresh Water Geochemistry: Overview, Fig. 1 Schematic representation of the Critical Zone (*CZ*) and the associated chemical, physical and biological weathering processes, which are affected by climatic, tectonic and anthropogenic forcing over greatly different timescales. The response of the atmosphere, rivers and oceans, soils and sediments is recorded in the geologic record fresh water geochemistry. There are, however, additional sources of dissolved constituents in fresh waters. Among them, there are marine recycled salts (Box 1) that enter continents via wind-transported aerosols which are transferred to the land surface by atmospheric precipitation or as dry fallout [7]. The accumulation of water in closed basins which, through evaporation, may lead to hypersaline lakes or salt-encrusted playas (i.e., dry lake, also known as a salt flat or *sabkha*) is also an occasional source of soluble matter – and insoluble mineral phases, as well – which may be transferred elsewhere as dry material (i.e., through wind paths) of by rainfall or snowfall.

Box 1 Recycled Sea Salt and Aerosols
Sea spray generates sea salt aerosol, one of the most widely distributed natural aerosols. Aeolian paths transfer sea salt aerosols to the continents by means of atmospheric precipitation or as dry fallout. Current estimation of the total sea salt flux from ocean to atmosphere is ~3300 Tg y^{-1} (1 teragram, Tg $= 10^{12}$ g). Aerosol is a colloid of fine solid particles or liquid droplets, dispersed in air or another gas.

There are, still, other minor sources of dissolved constituents to fresh water systems, and these are the hot hydrothermal water sources, which may outcrop through fractures, in areas directly or indirectly associated with volcanic activity. These singular springs may be a source of dissolved phases in certain areas, which persists along river courses until its signature is obliterated by other geochemical processes (e.g., adsorption-desorption, mineral precipitation, dilution).

In lotic and lentic fresh water systems, the mineralization of organic matter is another process that incorporates matter to the aqueous solution. Mineralization is a biological process in which organic matter is converted into inorganic substances by microorganisms.

Global change has painfully taught humanity that its actions have reached geological proportions, so that "Anthropocene" has been coined and is widely employed to describe the time interval that frames the impact of humanity on nature. In fresh water systems, the adverse effects of human development are discernible through pollution of rivers, streams, and lakes with undesirable substances (e.g., by-products of industrial processes which are harmful to life) and by disproportionately increasing (i.e., beyond limits compatible with healthy life) the concentration of naturally occurring chemical elements or substances. Such chemical components, foreign to the natural environment, interact with it and may be included in the study of fresh water systems.

Fresh Water in Our Planet

Rivers have played a fundamental role in the historical evolution of humanity, not only supplying fresh water and being an easily accessible source of proteins but also serving as transportation pathways, useful for communication and commerce. They have frequently been used as natural borders separating regions and countries, and history shows that many wars have been fought for their control. Moreover, human beings have shown sensibility – mainly through poetry and magnificent paintings – to the aesthetical aspects that often surrounds riverine scenery. To the pre-Socratic Greek philosopher Heraclitus of Ephesus (ca 535–ca 475 BC), rivers were an irrefutable proof of the ever-present change of nature when he stated that "no man ever steps in the same river twice." In short, flowing fresh water supplied by rivers was, is, and will be a vital feature of our planet, indispensable in the foreseeable evolution of life on Earth, in general, and humankind, in particular.

Throughout the hydrological cycle, water constantly interacts with the surficial layer of the Earth. Geological materials are progressively differentiated through processes of weathering, regolith erosion, and soil and sediment development. These processes controlled by the forces of nature on a large dimension have been equated to the series of separations conducted during the course of chemical analyses [8].

The Earth has a water volume of ~1.39×10^9 km^3 of which over 96% is saline. Only 2.5% is fresh water (10.53×10^6 km^3), of which over 68% is locked up in ice and glaciers. An additional 30% is buried in the ground, as groundwater. Therefore, rivers and lakes that supply surface water for human consumption only constitute about 93.1×10^3 km^3, which is about 0.007% of total water, yet rivers are the source of most of the water people use [9, 10] (Fig. 2).

Each bar in Fig. 2 represents a *reservoir*, a subjectively defined space containing a certain mass of water. Between these reservoirs, there are transfers or fluxes of water in and out of the reservoir. The *residence time* (i.e., the average amount of time that a particle or molecule spends in a particular system) of water in a reservoir is obtained by dividing the volume of water by the flux of water into (or out of) the reservoir. Thus, the residence time of water in the ocean is ~3200–3500 year, in shallow groundwater is ~100–200 year, in lakes is ~50–100 year, in rivers is ~2–6 months, and in atmosphere is ~9–11 days.

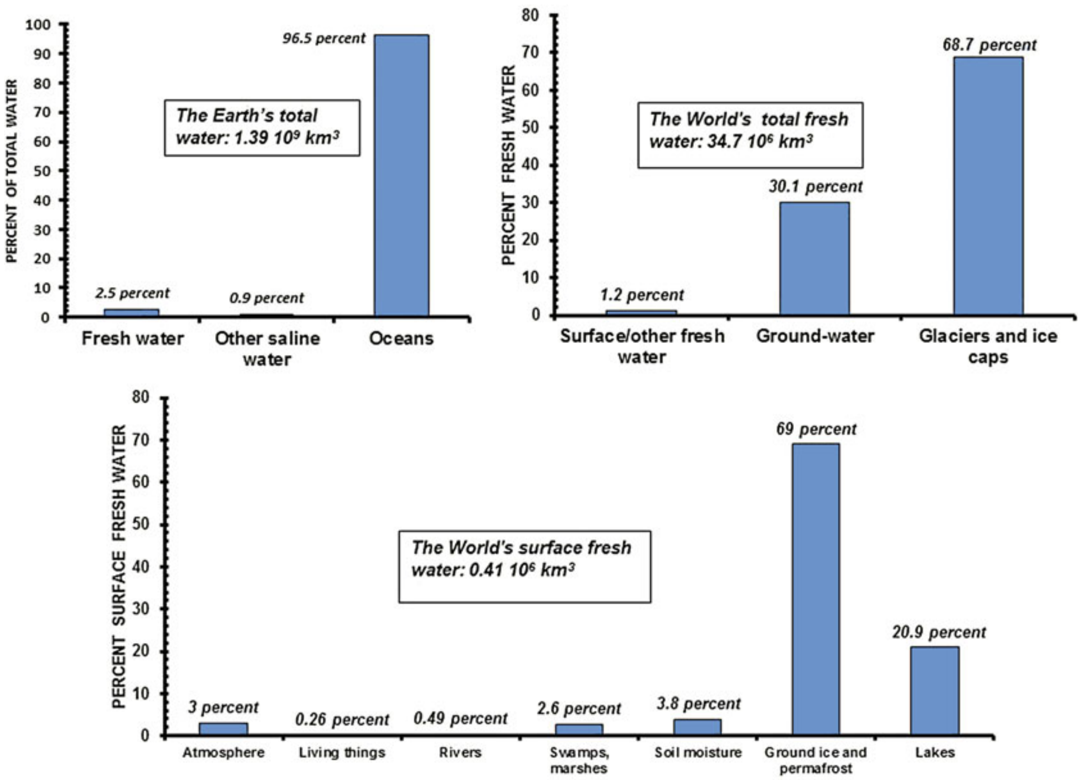
The Earth's exposed continental crust is subjected to a myriad of climates which, ultimately, are succinctly expressed in the variability of hydrologic budgets and runoffs. Runoff (R) is often used interchangeably with river discharge (D) although, strictly stated, the former is expressed as the latter, normalized by basin area (A):

$$R\ (\mathrm{mm}) = \left[Q\ \left(\mathrm{km}^3\ \mathrm{y}^{-1}\right)/A\ \left(\mathrm{km}^2\right)\right]\ 10^6$$

The dependency of R on climate and water use is clearly shown by the multivariate formula:

$$R = P - \Sigma\ (ET + S + C)$$

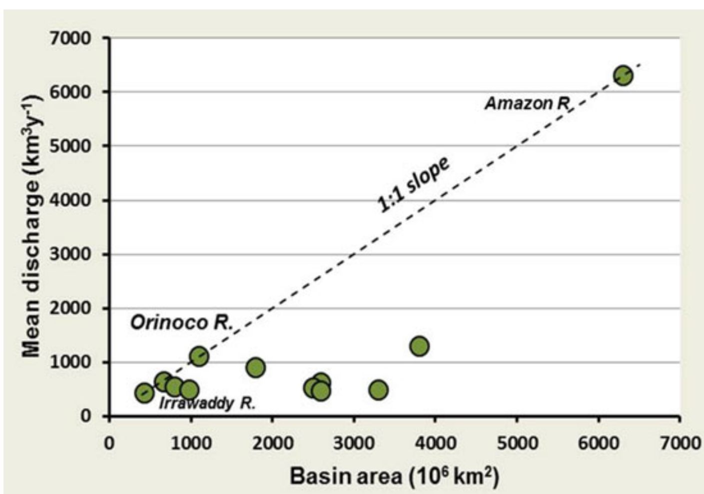
The summation of evapotranspiration (ET), water storage (S), and water consumption (C), in



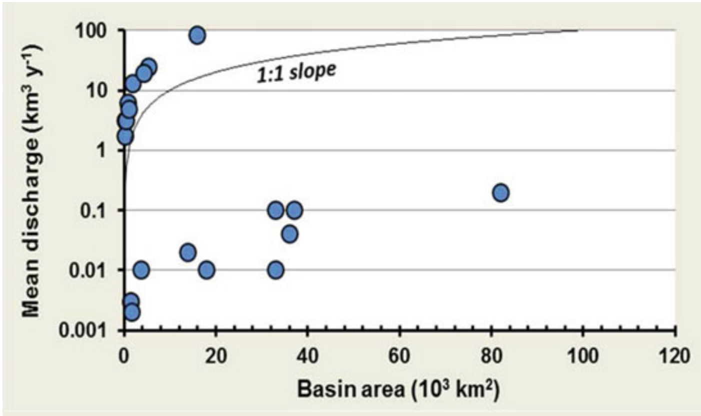
Fresh Water Geochemistry: Overview, Fig. 2 The global hydrological cycle: water on Earth is mainly stored in the oceans (*upper left*); most fresh water is locked in glaciers and ice caps (*upper right*), which also accounts for most of the fresh water existing on the surface of the planet (*lower bar graph*) proper units, is subtracted from the total atmospheric precipitation (*P*) falling over a specific drainage basin to obtain its runoff. Although there is some variability among continents, the average latitudinal distribution of global precipitation shows a clear concentration between 20°N and 25°S. On the other hand, evapotranspiration is difficult to calculate because it denotes the combination of two interconnected processes: physical evaporation (which depends on humidity, air temperature, and wind velocity) and plant transpiration (which depends on the ecosystem involved and plant types within that ecosystem). Decreasing ambient temperatures and changing vegetation define a decreasing evapotranspiration trend toward high latitudes. At any rate, all these intervening variables result in a global fresh water discharge rate, recently estimated in 36×10^9 m^3 y^{-1} [9]. Rivers with highest discharges (and largest drainage basins) are, in general, a typical tropical feature, with the Amazon, Congo, Orinoco, Changjiang, and Brahmaputra – the world's highest mean discharges – accounting for over 28 percent of the total global fresh water discharge [9]. Contrastingly, the highest runoffs (>5000 mm y^{-1}) are mostly exhibited by rivers with relatively small drainage basin areas (<1000 km^2) [9]. In general, the largest drainage basins exhibit the largest mean annual discharges.

Note (Fig. 3) that most global rivers discharging more than 400 km^3 y^{-1} fall below the 1:1 line defined by the Amazon River (top) and the Irrawaddy and the Orinoco rivers (bottom). All three rivers have the highest runoff in the set (1000 mm y^{-1}). In contrast, high runoff rivers (~4000 to ~9000 mm y^{-1}) drain tropical or temperate mountains, whereas lower runoffs (<4 mm y^{-1}) are recorded in rivers with a variety of elevations and drainage basin areas which are several orders of magnitude less than continental-size rivers (Fig. 4).

The key aspect here is that the global fresh water mass which by returning to the coastal seas closes the hydrologic cycle is a kind of permanent conveyor belt that yearly transports huge amounts of fluvial sediment and dissolved solids, both mainly the product of continental denudation. The most recent and accurate estimate of such material flux is 19 Gt y^{-1} of sediment and 3.8 Gt y^{-1} of dissolved solids [9] (1 gigaton, Gt = 10^9 t). This indicates that five times more particulate material reaches the seas than dissolved phases. It must be kept in mind that a significant proportion of both – particularly in the sediment fraction – is recycled material (i.e., products of the erosion of exposed sedimentary rocks that have already gone through the exogenous cycle or recycled sea salts, in the dissolved domain). There are, however, a number of rivers – particularly prevalent in Europe and Eurasia – where the ratio of total dissolved load (TDS) to total suspended load (TSS) is larger than 2 [9].



Fresh Water Geochemistry: Overview, Fig. 3 Among the largest rivers on Earth, none can surpass the limiting fresh water yield defined by the Amazon, the Orinoco, and the Irrawaddy rivers: 1 km^3 y^{-1} for every million km^2 of drainage area (Basic data published by Milliman and Farnsworth [9])



Fresh Water Geochemistry: Overview, Fig. 4 Greatest (plotted above the 1:1 line) and smallest runoffs (below the 1:1 line) of rivers within Milliman and Farnworth's global data base [9]. Most high-runoff rivers (>4000 mm y^{-1}) drain tropical or temperate mountains, whereas rivers with runoffs <5 mm y^{-1} drain a variety of elevations and many belong in the southern hemisphere [9]

Geology has a preeminent role in determining the concentrations in rivers draining different lithologies. The only major component which appears to remain relatively independent of lithology is dissolved SiO_2 because most river waters generally contain 100–150 mmol l^{-1} [9], indicating the relative insolubility of SiO_2 in fresh waters.

Box 2 Major Dissolved Ions

Major dissolved cations (i.e., Na^+, K^+, Ca^{2+}, and Mg^{2+}), anions (Cl^-, $SO_4{}^{2-}$, $HCO_3{}^-$, and $CO_3{}^{2-}$), and silica (SiO_2) account for the largest proportion of dissolved phases in the Earth's total fresh waters (~35 × 10^6 km^3). Fresh waters (mostly rivers, lakes, and groundwaters) have been studied with increasing attention, mainly since the nineteenth century, initially to learn if a particular water source was suitable for consumption by human beings and cattle, for its use in irrigation, in boilers, and in other minor industrial uses. In modern times there are many regions in the world where there is a water crisis, not only because of its scarcity but also for its inadequate quality. Therefore, waters in general and fresh waters in particular are

(*continued*)

Box 2 Major Dissolved Ions (continued)

studied with remarkable interest from many viewpoints: ecologists, geochemists, geographers, biologists, engineers, physicists, and many other specialists are keenly interested in probing into the most hidden details of its chemical, physical, and biological composition. In geochemistry, the foci are placed on the sources of the dissolved and particulate chemical elements, on the processes that control their concentrations, and on the partitioning between particulate and dissolved phases.

Chemical Weathering of Minerals and Rocks

Silicates dominate, by far, the proportion of rocks exposed at the Earth's land surface [11]. Chemical weathering of silicate minerals stems from disparities in the thermodynamic conditions that predominated at the time of mineral formation and the ambient conditions prevailing at the Earth's surface, when rocks are exhumed through tectonics and erosion. Moreover, the term *weathering* denotes a strong linkage with many processes associated with the hydrosphere, the atmosphere, and the biosphere. It follows that chemical weathering is caused by water – particularly acidic water – and gasses (e.g., carbon dioxide, oxygen) which attack minerals. Some ions and compounds of the original mineral are eliminated in solution, permeating the mineral residue to feed groundwater, streams, and rivers. Fine-grained solids may be washed away from the weathering site, leaving a chemically modified residue (i.e., mineralogical changes occur in the regolith to regain stability in a new environment), which is the origin of the soil layer that mantles a large proportion of the Earth's exposed crust and constitutes a substratum which is essential for the biosphere to thrive.

A key factor to understand fresh water geochemistry is the extremely diverse performance that minerals and rocks reveal when they are exposed in the Earth's surface. Goldich's [12] relative stability sequence of major rock-forming silicate minerals during weathering approaches the crystallization succession of Bowen's [13] reaction series (i.e., the least stable minerals are those that crystalize first from high-temperature magmas). An additional sign of a mineral's vulnerability to weathering is the ratio of SiO_2 to other cations (e.g., Na^+, K^+, Ca^{2+}, Mg^{2+}) in its structure. The higher the proportion of, for example, Na^+ replaceable by H^+, the more susceptible to chemical weathering the mineral will be. Hence, olivine, pyroxenes, and amphiboles (i.e., mafic minerals) may weather more promptly than plagioclase, K-feldspars, micas, and quartz (i.e., felsic minerals). This is not a strict rule, however, and exceptions are found. Chemical elements in mafic minerals have a tendency to be released early during weathering, immediately followed by elements in felsic minerals (Fig. 5).

The character and rate of weathering are influenced by climatic and/or biological conditions, along with tectonic and geomorphic factors that rule surface relief. The hydrology in the CZ exerts a strong control on the magnitude of the diffusion of chemical elements and on the level reached by material transport.

Mechanisms of Chemical Weathering

Diverse mechanisms of chemical weathering are recognizable, and various arrangements of these occur together during the breakdown of most minerals and rocks.

Dissolution/precipitation reactions: The dissolution of soluble minerals is the simplest weathering reaction. Halite dissolution is the typical reaction of a salt that dissolves in water by dissociation of ions that go into solution:

$$NaCl_{(s)} \rightarrow Na^+ + Cl^-$$

Halite is entirely or congruently dissolved in water. However, this reaction can be reversed, for example, in a saline playa subjected to intense evaporation. Through persistent evaporation in arid climates, brackish water (i.e., water too saline to be potable but less saline than seawater, with an approximate range of dissolved salts of 1000–20,000 mg l^{-1}) becomes saline water (i.e., water which has salinity similar to or greater than seawater, ~35,000 mg l^{-1}) and then – with continued evaporation – turns into brine (i.e., water significantly more saline than seawater (≫35,000 mg l^{-1})). If the concentration of both ions reaches halite's solubility product, $K_{sp} = [Na^+][Cl^-] = 37.3$ mol^2 l^{-2}, halite would begin to precipitate:

$$Na^+ + Cl^- \rightarrow NaCl_{(s)}$$

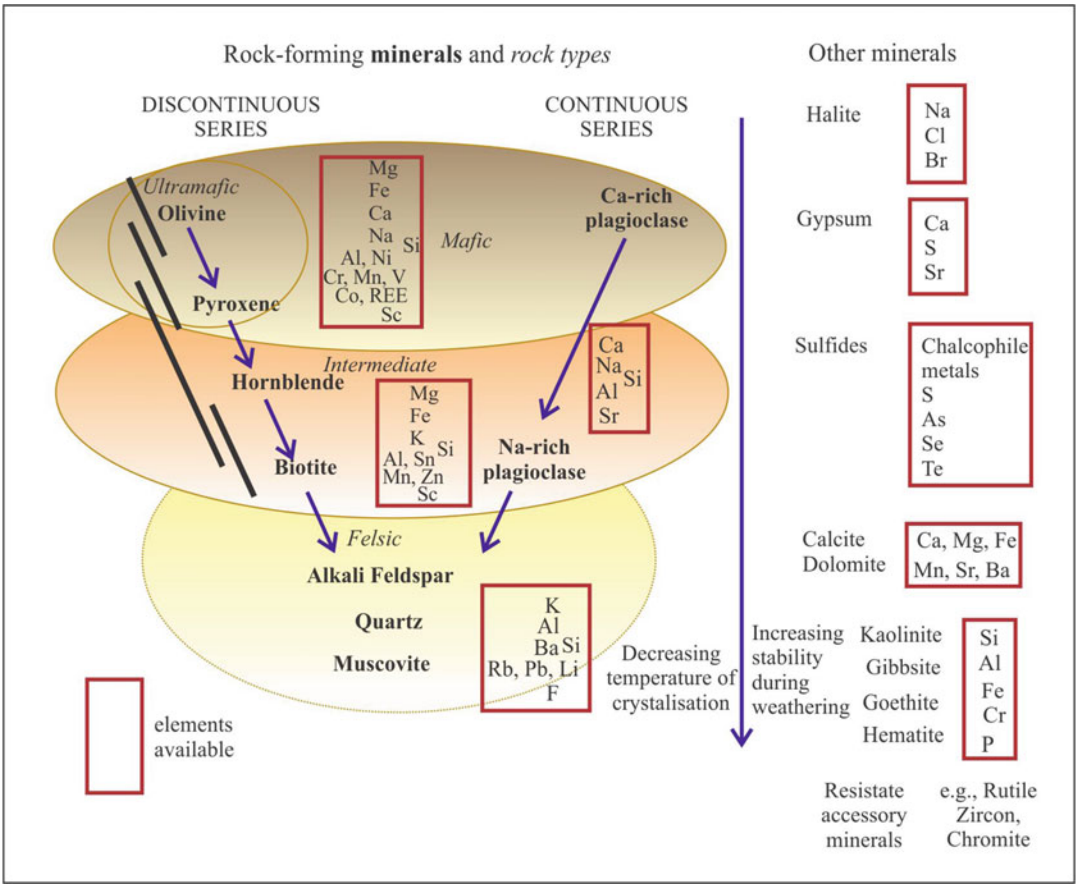
More strictly, activities (a_i) are used rather than concentrations. Therefore, the chemical equation $K_{sp} = [Na^+][Cl^-]$ can be written as (Box 3)

$$K_{sp} = a_{Cl-}\ a_{Na+}$$

Box 3 Activities

Activity a_i can be thought of as the *effective* concentration of solute i. The following equation defines the activity coefficient γ for element i:

(*continued*)



Fresh Water Geochemistry: Overview, Fig. 5 Goldich's [12] sequence of mineral susceptibility is similar to Bowen's [13] sequence of mineral crystallization from a melt. Decreasing temperature of crystallization results in increased stability during weathering [14] (© CSIRO 2008. Published by CSIRO Publishing, Collingwood, Victoria, Australia, http://www.publish.csiro.au/pid/5955.htm. Reproduced with permission)

Box 3 Activities (continued)

$$\gamma_i = a_i / m_i$$

In this expression, m_i is the molal concentration. The activity coefficient γ_i for a particular species i depends on the concentration of all the solutes present in the solution. The overall *ionic strength* of the solution is expressed as

$$I = ½\, \Sigma m_i z_i^2$$

The variable z_i is the charge of the ion concerned divided by the charge on an electron. Natural waters span a considerable range

(*continued*)

Box 3 Activities (continued)

of ionic strength: river water <0.01, seawater 0.7, and brines 1–10 mol kg^{-1}. This leads us to the Debye-Hückel equation, which works accurately for non-ideal solutions (i.e., one in which activities and concentrations are unequal) of ionic strength up to 0.01 mol kg^{-1}:

$$\log_{10} \gamma_i = -A z_i^2 I^{½}$$

A is a constant characteristic of the solvent. Although river waters typically reach ionic strengths as low as 0.002 mol kg^{-1}, it

(*continued*)

Box 3 Activities (continued)
is distinctively non-ideal. Stumm and Morgan [15] and Langmuir [16], for example, have treated extensively this particularly important physicochemical subject in connection with water chemistry.

Halite is among the most soluble salts in nature and it is not found as such (i.e., crystallized) in fresh water environments, although Cl^- and Na^+ are among the most abundant ionic species in fresh waters. This occurs because Cl^- and Na^+ are dominant ionic species in seawater which are transported as aerosols by winds back to the continents. Therefore, NaCl finds its way to fresh waters directly, through atmospheric precipitations, or indirectly, transferred by winds from endorheic playas (Box 4), supplied to rivers by outcropping salty groundwaters (i.e., a result of halite dissolution in deep aquifers), or through hydrothermal processes.

Box 4 Endorheic and Arheic
Endorheic basins are closed drainage basins (also referred to as terminal basin or as an internal drainage system) that retain water and allow no outflow to other external bodies of water, such as rivers or oceans, but converge instead into permanent or seasonal lakes or swamps, equilibrated through evaporation. Such basins are also known as endorheic or endoreic basins. An arheic area is one in which surficial drainage is almost completely lacking or where rainfall is so infrequent that all water sinks into the ground or evaporates. Such areas are also known as areic, aretic, or arhetic.

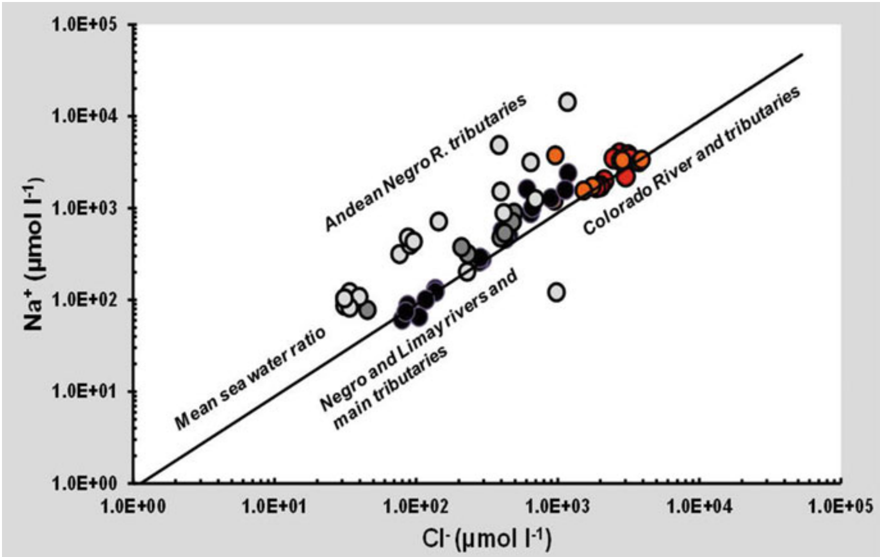
Figure 6 shows geochemical data from the Colorado and Negro rivers (Box 5). The graph (Fig. 6) shows that most data closely follow the Cl^-/Na^+ seawater ratio, thus leading to infer a clear provenance. The Negro River Andean headwaters exhibit higher Na^+ concentrations, suggesting a dissolved supply originated in the chemical weathering of Na-bearing minerals or, additionally, a likely contribution from outcropping hydrothermal solutions from volcanic terrains.

Box 5 The Paraná River and Patagonia's Colorado and Negro Rivers
Whenever adequate, data (and references) from three Argentine rivers are used throughout this chapter to illustrate different aspects of the geochemical processes occurring in fresh waters. The Paraná River is one of the Earth's largest rivers, discharging ~550 $km^3\ y^{-1}$ of fresh water; its drainage basin occupies ~$2.6 \times 10^6\ km^2$, distributed among Brazil, Paraguay, Bolivia, and Argentina. It is ~4900 km long, and the maximum elevations in the drainage basin exceed 1000 m a.s.l. The geology is complex but Mesozoic igneous rocks dominate. The ruling climate is subtropical, and the river joins (~34°S, ~58°W) the Uruguay River to integrate the Río de la Plata (~$3.17 \times 10^6\ km^2$) [18, 19]. The Colorado and Negro rivers drain the Atlantic seaboard of Argentina's Patagonia. The Colorado River (36°09′S, 70°23′W) is ~1100 km long, its drainage basin has ~22,300 km^2, and the annual water discharge is ~4.1 km^3. The Negro River (41° 01′S, 62° 47′W) is Colorado's neighbor to the south, and it is ~640 km long, its drainage basin has ~95,000 km^2, and its annual discharge is ~27.1 km^3 [18, 19]. Both have dominantly igneous Cenozoic (i.e., volcanic and intrusive rocks) Andean headwaters. The Negro drainage hosts several oligotrophic lakes of glacial origin, which the Colorado does not. Both have reservoir lakes in their drainage basins.

It is also common to find the dissolution products of gypsum ($CaSO_4.H_2O$), anhydrite ($CaSO_4$), or fluorite (CaF_2). If fluorite dissolves to equilibrium,

$$CaF_2 \leftrightarrow Ca^{2+} + 2F^-$$

The law of mass action yields



Fresh Water Geochemistry: Overview, Fig. 6 Na^{+} and Cl^{-} concentrations in the Negro/Limay, and the Colorado rivers (Argentina's Patagonia). The Colorado River and main tributaries data (*red* and *orange circles*, respectively) plot close to the mean Cl^{-}/Na^{+} sea water ratio. The Negro/Limay River system and main tributaries (*black* and *gray circles*, respectively) show a larger Na^{+} contribution (e.g., plagioclase weathering). However, most Negro's Andean tributaries (*light gray circles*), plot further away from the ratio line, implying an even larger Na^{+} weathering contribution and/or, from added hydrothermal sources (Sea water data from [17])

$$K = 10^{-10.6} = \left[Ca^{2+}\right]\left[F^{-}\right]^{2}, \text{or}$$

$$\left[F^{-}\right] = \left(10^{-10.6}/\left[Ca^{2+}\right]\right)^{1/2}$$

K is the equilibrium constant for the above reaction. This equation shows that high F^{-} waters will have low Ca^{2+} concentrations and vice versa. This is also a simple example illustrating some complexities inherent to fresh water geochemistry.

Acid hydrolysis: Continental fresh water contains dissolved chemical species which may render it acidic. A few sources determine such acidity: predominantly from the dissociation of soil zone CO_2 and, also, from the dissociation of atmospheric CO_2 in rainwater (i.e., rainwater in equilibrium with atmospheric CO_2 reaches a pH $\approx$ 5.6) to form H_2CO_3. Also important may be natural or anthropogenic SO_2, which forms H_2SO_3 and H_2SO_4 and can lead to a pH $\approx$ 4.85 in rain droplets.

The reaction of acidic weathering agents and minerals susceptible to acidic attack is known as acid hydrolysis. The weathering of calcite and/or aragonite ($CaCO_3$) is an illustrative example:

$$CaCO_{3(s)} + H_2CO_{3\,(aq)} \leftrightarrow Ca^{2+}{}_{(aq)} + 2HCO_3{}^{-}{}_{(aq)}$$

The reaction is dependent on the amount of available CO_2. The addition of CO_2 causes the increased formation of H_2CO_3, which dissolves more $CaCO_3$, displacing the equilibrium to the right hand side of the equation. In contrast, if the CO_2 flux decreases, the equilibrium moves toward the left, promoting the reverse reaction and the precipitation of calcite. This effect of CO_2 variation is a distinct example of Le Chatelier's principle. Karst topography, the Mexican *cenotes*, and the stalactites and stalagmites forming in limestone caves are all examples of the geological effect of CO_2 chemical dynamics in groundwater.

The weathering effects of CO_2 in aqueous solutions and, particularly, the interaction with $CaCO_3$ are often referred to as *carbonation.* Due to biological activity (i.e., respiration), the concentration of CO_2 in groundwater may be 20–30 times higher than in the atmosphere. This CO_2 confined within the soil structure, which in temperate climates reaches 0.1–3.5%, may rise up to 11% in tropical climates where higher temperatures promote greater biological activity and, therefore, higher CO_2 concentrations.

Another example is the congruent reaction of acid hydrolysis of magnesium-rich forsterite:

$$Mg_2SiO_{4(s)} + 4H_2CO_{3\,(aq)} \rightarrow 2Mg^{2+}{}_{(aq)} + 4HCO_3{}^{-}{}_{(aq)} + H_4SiO_{4(aq)}$$

The dissolution of silicates always produces alkalinity (Box 6). The acid (H_4SiO_4) stemming from the breakdown of olivine is weaker than the slightly stronger H_2CO_3, from which $HCO_3{}^{-}$ results

$$H_2CO_{3\,(aq)} \leftrightarrow H^+ + HCO_3{}^{-}{}_{(aq)}$$

Because carbonic acid is a diabasic acid, it dissociates in two steps:

$$4HCO_3{}^{-}{}_{(aq)} \leftrightarrow H^+ + CO_3{}^{2-}{}_{(aq)}$$

The pH of most natural waters is regulated by chemical reactions involving the carbonate system. The detailed treatment of this important system in fresh water geochemistry has been tackled by several authors (e.g., [15, 16, 20]).

Box 6 Alkalinity

Alkalinity is the capacity of water to accept protons and it is the sum effect of all bases present. The conventional definition of total alkalinity (C_B) in eq l^{-1} or meq l^{-1} is

$$C_B = HCO_3{}^{-} + 2CO_3{}^{2-} + OH^{-} - H^{+}$$

i.e., C_B equals the total equivalents of bases minus those of acids. There are other bases that contribute to total alkalinity but, generally, they are not considered in the calculation.

(*continued*)

Box 6 Alkalinity (continued)

Alkalinity is usually reported as mg l^{-1}CaCO $_3$ (or meq l^{-1} $CaCO_3$). In fresh waters with pH below 8.3, $HCO_3{}^{-}$ is commonly the only significant base. $HCO_3{}^{-}$ alkalinity in water originates from two sources: (a) from the dissolution of carbonate minerals, which, if dissolved by H_2CO_3, supplies twice as much $HCO_3{}^{-}$ alkalinity relative to H_2CO_3, as it occurs in the dissolution of dolomite:

$$CaMgCO_3 + 2H_2CO_3 \rightarrow Mg^{2+} + Ca^{2+} + 4\ HCO_3{}^{-}$$

The other source is (b) the weathering of silicates by H_2CO_3.

Most of the world's surface waters have a near neutral pH, with $HCO_3{}^{-}$ as the foremost anion [21] because of the added effect of dissolved CO_2 in soil interstitial water, the ensuing dissociation of H_2CO_3 to produce alkalinity, and the occurrence of acid hydrolysis weathering reactions.

The presence of altered mineral residues during weathering implies that incomplete dissolution is more usual than the chemical breakdown of monomer silicates which dissolve completely. Silicates are by far the most common minerals in the crust and mantle of the Earth, making up 95% of the crust and 97% of the mantle by most estimates. The mean composition of the upper crust approximates that of a granodiorite, a rock which is a phaneritic-textured felsic intrusive igneous rock similar to granite, but containing more plagioclase feldspar than orthoclase. The simplified weathering reaction for plagioclase feldspar (e.g., anorthite) might best represent average acid hydrolysis:

$$CaAlSi_2O_{8(s)} + 2H_2CO_{3\,(aq)} + H_2O \rightarrow 2HCO_3{}^{-}{}_{(aq)} + Ca2^{+}{}_{(aq)} + Al_2Si_2O_5(OH)_{4(s)}$$

The only solid product of this reaction is kaolinite, the most important member of the kaolin group of clay minerals (i.e., 1:1 layer or two-sheet structure) (e.g., [16]).

The acid hydrolysis of the Na-plagioclase albite, on the other extreme of the plagioclase series, shows

$$\begin{aligned} &2NaAlSi_2O_{8(s)} + 2H_2CO_{3\,(aq)} + 9H_2O \\ &\rightarrow 2HCO_3^{-}{}_{(aq)} + 2Na^{+}{}_{(aq)} \\ &\quad + Al_2Si_2O_5(OH)_{4(s)} + 4H_4SiO_{4(aq)} \end{aligned}$$

One significant product released to solution is H_4SiO_4. The SiO_4 tetrahedron framework is especially weak where Al has switched for Si, since the Al-O bond has a more ionic character.

Box 7 Weathering Reactions and Thermodynamics

Chemical reactions may be exothermic or, in few cases, endothermic. Such heat changes are called changes in *enthalpy* (H). There are two additional factors in the energy changes in a chemical reaction. One is *Gibbs' free energy* (G), which is the overall energy change, and the other is *entropy* (S) which is concerned with the energy changes associated with the degree of randomness of a system (e.g., a solid has a highly organized structure, whereas a gas has molecules moving in a random fashion, so that gas has high entropy and a solid has low entropy).

These energy factors are related by the well-known equation:

$$\Delta G = \Delta H - T\Delta S$$

T is temperature (in kelvins, K) and Δ means "change in." Spontaneous reactions occur when there is a decreasing value of G (i.e., ΔG is negative). A couple of factors will favor a negative value for ΔG: (a) ΔH is negative (i.e., exothermic reaction); (b) ΔS is positive (i.e., increased randomness in the reaction). Notice that if ΔH is positive (endothermic reaction), it may still be offset by a positive value of ΔS (increased randomness) to give a negative ΔG. Also, a

(*continued*)

Box 7 Weathering Reactions and Thermodynamics (continued)

negative ΔS can be offset by a negative ΔH (i.e., exothermic reaction) to give a negative ΔG (e.g., [15, 16]).

It is meaningful to cast a thermodynamic view on the above equation (Box 7). Gibbs' free energy values of formation for the above equation allow calculating

$$\Delta G^0 = 69.1 \text{ kJ}$$

The equilibrium constant for the albite weathering reaction is

$$K = [Na^+]^2[HCO_3^-]^2\left[Si(OH)_4\right]^4/[H_2CO_3]^2$$

The value of the equilibrium constant can be calculated:

$$\Delta G^0 = -2.303RT\log_{10}K$$

R is the gas constant (8.314 JK^{-1} mol^{-1}); T is the temperature in Kelvin (198.2 K for 25 °C); and 2.303 is ln to $\log_{10}$ conversion factor:

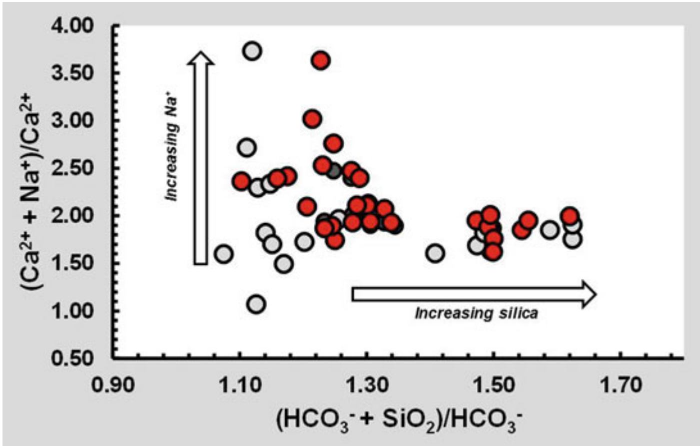
$$\log_{10} K = \Delta G0/(2.303\ R\ T)$$

$$\log K = -12.1022$$

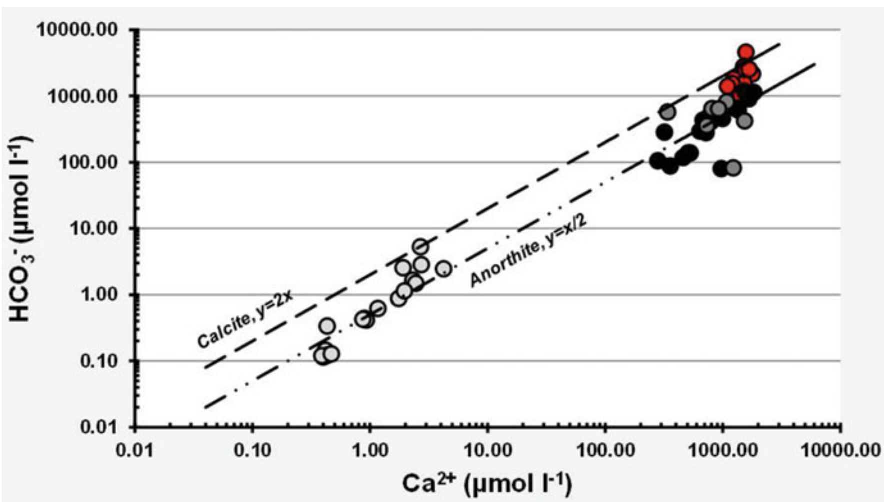
$$K = 7.9\ 10^{-13}$$

K is such a small value that a very large volume of H_2CO_3 is required to reach equilibrium. There is robust evidence, however, that the reaction takes place: kaolinite can be found in soils and sediments, and the aqueous components are found dissolved in natural waters. The explanation lies in the permanent replenish of fresh carbonated water coming into contact with the mineral surface subjected to chemical attack.

Figure 7 illustrates one aspect of the dynamics of weathering in Patagonia's Colorado and Negro rivers: solute sources may change and reactions may be more variable in certain parts of the basin



Fresh Water Geochemistry: Overview, Fig. 7 Geochemical data from Argentine Patagonia's Colorado (*red circles*), Negro (*black circles*), and Negro's Andean tributaries (*gray circles*), showing the relative increase of Na^+ and SiO_2 with respect to Ca^{2+} and alkalinity. Both increases are probably determined by a relative decrease in the significance of limestone and carbonate-bearing sedimentary rocks as a source of dissolved phases



Fresh Water Geochemistry: Overview, Fig. 8 Graph suggesting dominant lithological sources in the Colorado and Negro rivers (Argentine Patagonia). Alkalinity and Ca^{2+} concentrations in Colorado River and tributaries (*red circles*), Negro River (*black circles*), and its Andean tributaries (*gray circles*) are mostly framed by calcite and anorthite ideal dissolution lines

than in others. Clearly, Colorado River is more influenced than the Negro River by the weathering of carbonate rocks. The chemical impact of rock sources on riverine fresh water geochemistry is further illustrated in Fig. 8, where chemical data is mostly framed by the ideal composition of calcite and anorthite (i.e., the Ca end-member in the plagioclase solid-solution series). The Colorado River and its main tributaries show diverse lithological provenance for its solutes, whereas the Negro (and some headwater tributaries) are more closely associated with the plagioclase series minerals.

Reduction-oxidation (redox) reactions: The reaction

$$2FeO_{(s)} + ½O_{(g)} \rightarrow Fe_2O_3$$

has $\Delta G^0 = -240$ kJ. The negative sign means that the reaction is energetically favored and therefore is a spontaneous reaction as everybody knows by direct experience. Like Mn^{2+}, Fe^{2+} is readily dissolved by mildly acid waters during weathering, but once dissolved both elements are prone to oxidation to $Fe^{3+}$and Mn^{4+}, whose ionic potentials are in the hydrolysate range and, consequently, precipitate as minerals [e.g., goethite (FeOOH)]. This process explains the relative abundance of red soils in the Earth's tropical belt and, also, the reddish sedimentary beds found in the sedimentary column (e.g., Cretaceous), associated with extreme oxidizing environments recorded in the geological past. Iron's world river average concentration is 66 $\mu g\ l^{-1}$ [21].

Sulfides, such as pyrite (FeS_2), are frequent minerals in veins, but they are also common in mud rocks and coal deposits (e.g., [22]). The oxidation of reduced iron (Fe^{2+}) and S in pyrite causes the formation of H_2SO_4, a strong acid. This is a serious environmental hazard in materials left over after the process of separating the valuable fraction from the uneconomic portion (i.e., gangue) of an ore:

$$2FeS_{2(s)} + 2H_2O + 7O_{2\ (g)} \rightarrow 4H^+_{(aq)} + 4SO_4{}^{2-}{}_{(aq)} + 2Fe^{2+}{}_{(aq)}$$

Pyrite oxidation is followed by the oxidation of Fe (II) to Fe(III):

$$4Fe^{2+}{}_{(aq)} + O_{2(g)} + 10H_2O \rightarrow 4Fe(OH)_{3(s)} + 8H^+{}_{(aq)}$$

At the low pH values found in acid mine waters, oxidation occurs gradually. Ferric iron may react further with pyrite:

$$FeS_{2(s)} + 8H_2O + 14Fe^{3+}{}_{(aq)} \rightarrow 15Fe^{2+} + 2SO_4{}^{2-}{}_{(aq)} + 16H^+{}_{(aq)}$$

Below pH 3.5, iron oxidation is catalyzed by the bacterium *Thiobacillus thiooxidans*, and between pH 3.5 and 4.5, oxidation is catalyzed by *Metallogenium.* At pH $\gg$ 3, Fe(III) precipitates as goethite (FeOOH).

Reduced Fe-bearing silicate minerals (e.g., pyroxenes, amphiboles, and Fe-rich olivine) may also undergo oxidation, as it occurs with fayalite:

$$Fe_2SiO_{4(s)} + ½O_{2\ (g)} + 5H_2O \rightarrow 2Fe\ (OH)_{3(s)} + 4H_4SiO_4$$

Colloidal $Fe(OH)_3$ dehydrates to generate a variety of oxides, e.g., Fe_2O_3 (hematite) and FeOOH (goethite).

Weathering Intensity and Rate

Clearly, weathering is the foremost process that maintains a permanent supply of dissolved substances to the Earth's water mass. Continued weathering of its upper crust has occurred ever since the planet had water and a liquid ocean (i.e., ~4.5 10^9 y), and rock weathering has been –and still is – largely responsible for the ocean chemical composition (i.e., world's oceans have a mean salinity of 35 g kg^{-1}, but seawater is not uniformly saline throughout the world; the vast majority of seawater fluctuates between 31 and 38 g kg^{-1}). Ocean salinity was not only determined by the permanent supply of dissolved substances through eons of continental wear away, but it was also the result of outgassing (i.e., HCl and other gasses) from the Earth's interior via submarine volcanoes and hydrothermal vents.

The specialized literature states that ocean salinity has been stable for billions of years, as a consequence of a chemical/tectonic system which removes as much salt as it is deposited. In other words, chemical stability in the ocean does not mean that the dissolved input from the continents has experienced little variability throughout the geological column. Quite the contrary, continental weathering intensity and rate has changed markedly over geological time inasmuch climate and tectonics are dynamic forcing factors (Fig. 1). Hence, it is important to learn about present-day weathering intensity and rate so that it would be possible to probe into past weathering dynamics and, further, into fresh water variability over time. Basically, weathering intensity alludes to the degree of decay of minerals and rocks at a certain point in time, whereas the rate of such decay ideally refers to the amount of change per unit time. These terms may be linked inasmuch a high degree of weathering intensity usually indicates a relatively swift alteration rate.

As weathering advances and the more mobile elements are washed away, the chemical composition of regolith will change accordingly. This is the reason why the study of weathering intensity in regolith is often used as proxy to probe into fresh water geochemistry. There are two methodologies that lead to the assessment of such change: (a) working out the ratio of highly resistant minerals (e.g., quartz, zircon) with others which are more easily removed and (b) following the so-called absolute approach which consists of comparing the chemical composition of the parent material with that of the weathered debris (e.g., [23, 24]).

Determination of weathering intensity with relative methods: Uncertainty in the composition of the parent material (i.e., "fresh" rock) leads to the calculation of the ratio of more stable and less stable oxides, expressing the result as an index. Bland and Rolls [23] supply an example:

$$WR_{(h)} = \begin{array}{l}(\text{zircon} + \text{tourmaline}) \\ \times/(\text{amphiboles} + \text{pyroxenes})\end{array}$$

$$WR_{(l)} = \text{quartz}/\text{feldspars}$$

These weathering ratios (*WR*) aim to unveil the situation in the heavy (*h*) and light (*l*) mineral fractions.

Most methodologies used in the relative domain are based on the assumption that Al_2O_3 is largely immobilized during weathering. The most popular index is known as the "chemical index of alteration" or *CIA*:

$$CIA = [Al_2O_3/(CaO^* + Na_2O + K_2O + Al_2O_3)] \times 100$$

CaO^* represents CaO adjusted for apatite and Ca-bearing carbonates using P_2O_5 and CO_2 in the correction procedure [25]. *CIA* values of 45–55 indicate nearly no weathering (i.e., the Earth's average upper crust has a *CIA* of 47). Products of intense weathering (e.g., kaolinite, gibbsite) have *CIA* values of ~100, whereas other clay minerals (e.g., smectite, illite) have values of about 70–80. Primary minerals have much lower *CIA* values (e.g., 50 for plagioclase, 0–20 for amphiboles and pyroxene). It arises then that the *CIA* value of a bulk detrital sample will vary substantially depending on the proportions of primary minerals and clay minerals in it. Figure 9a shows mean CIA values determined in Paraná River TSS samples. The trend suggests a relationship with discharge. Mountainous (i.e., mostly Andean) tributaries transport coarser sediment, with somewhat lower CIAs (Fig. 9b) [26].

There are several other relative approaches, such as the *chemical index of weathering* or *CIW* [27]:

$$CIW = [Al_2O_3/(CaO + Na_2O + Al_2O_3)]100$$

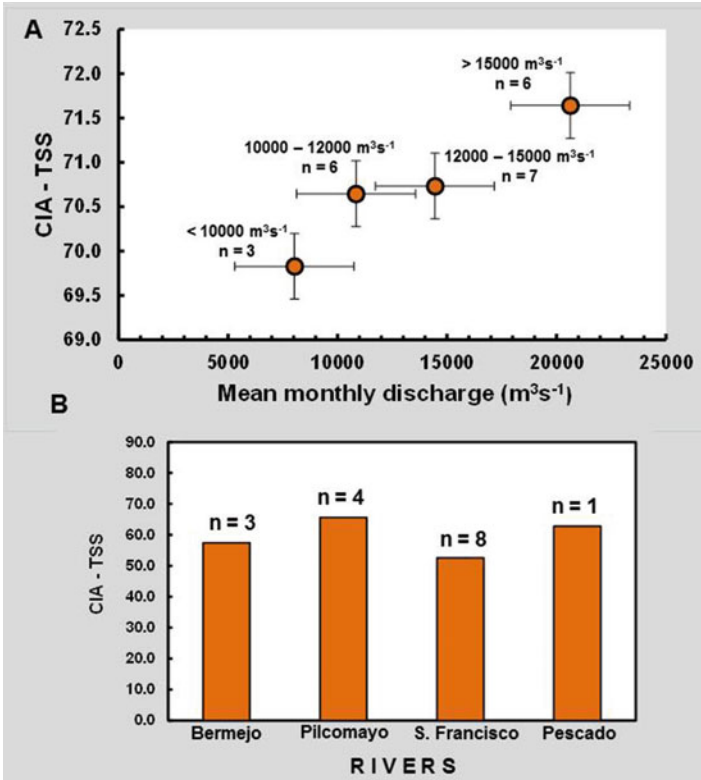
Like the *CIA*, the *CIW* value increases as the degree of weathering increases as well. This index has proved particularly useful in weathering scenarios where K_2O has not totally been removed from the regolith (e.g., diagenetic *illitization*).

Another popular index is the *weathering index of Parker* or *WIP* [28], which is based on the proportions of alkali and alkaline earth metals:

$$\begin{aligned} WIP = {} & [(2Na_2O/0.35) + (MgO/0.9) \\ & +2K_2O/0.25) + (CaO/0.7)] \end{aligned}$$

WIP values are usually between ~100 and 0, with the least weathered rocks having the highest values. The assumption that all CaO in the examined regolith is held in silicate minerals is implied. Parker [29] included the susceptibility to weathering of the elements involved in the equation by including bond strength values (as denominators) as a measure of the energy required to break the cation-to-oxygen bonds of the respective oxides.

Weathering indices have often been devised or adapted to attain a clearer view on the breakdown of specific rock types. Such is the case of the *mafic*



Fresh Water Geochemistry: Overview, Fig. 9 Values for the chemical index of alteration (*CIA*) in total suspended solids from the Paraná River system. (**a**) Plot of mean CIA values and typical error bars as a function of river discharges. The samples were collected in the middle stretch (~600 km upstream the mouth). (**b**) Average CIA values determined in TSS samples from Paraná's mountainous tributaries [26]

index of alteration or *MIA*, which is an index proposed to extend the equation of the *CIA* to include the mafic elements (i.e., Fe and Mg), turning the index into a more amenable tool to investigate weathering in basalts [30]. In an oxidizing alteration environment, Fe is considered an immobile element along with Al_2O_3, and the equation for *MIA* is

$$MIA_{(O)} = \left[\left(Al_2O_3 + Fe_2O_{3(T)}/(CaO^* + MgO + Na_2O + K_2O + Al_2O_3 + Fe_2O_{3(T)}\right)\times\right]100$$

When the environment of alteration is reducing and Fe is leached along with Mg, total Fe is considered a mobile element, and the *MIA* calculation becomes

$$MIA_{(R)} = [Al_2O_3/(CaO^* + MgO + Na_2O + K_2O + Al_2O_3 + Fe_2O_{3(T)})] \times 100$$

As it happens with the *CIA*, increasing index values always represent a more altered rock, and a value of 100 specifies complete removal of mobile elements and their transfer to the aqueous domain. As it is the case with the *CIA*, the molar CaO is corrected for the presence of carbonate and apatite, considering only the silicate-bound Ca (CaO^*).

Relative elemental losses attributable to rock weathering can also be analyzed from the dissolved realm. Tardy [31] calculated the molecular ratio from the concentration of different dissolved species determined in surface waters draining granitic and gneissic terrains:

$$Re = \left[3\left(Na^+ + 3(K^+) + 2\left(Ca^{2+}\right) - (SiO_2)\times\right]/\left[0.5(Na^+) + 0.5\,(K^+) + \left(Ca^{2+}\right)\right]$$

The coefficients used in the above equation depend on the bedrock's major primary minerals and correspond to an average granitic composition with feldspars and micas. It is assumed that if $Re \approx 0$, gibbsite is basically formed; if $Re \approx 2$, kaolinite is the dominant solid product of rock alteration; and if $Re \approx 4$, the products are mainly smectites [32].

It may be concluded, then, that the indices that can be computed with a significant variety of methodologies basically reflect the probabilities of elemental mobility during weathering processes, expanding the insight on the likely sources which control fresh water geochemistry. Depetris et al. [24] have described several procedures used to appraise weathering intensity through the chemical analyses of regolith or sediments.

Determination of weathering intensity with absolute methods: The assumption behind any absolute method is that a computable association can be established between an unaltered rock and the overlying weathered material. In order to calculate the losses of different chemical components in a weathered profile, it is necessary to select an insoluble oxide which will be the *benchmark* in the analysis. As with relative methods, Al_2O_3 is mainly used as it is relatively insoluble at pH values down to 5.5, such as found in rainwater. The ratio between the Al_2O_3 content of the bedrock and of the regolith is used to calculate the losses in the other compounds. Examples of this ingenious procedure, originally formulated by S. S. Goldich [12], have been treated in detail in several books (e.g., [23, 24, 33, 34]).

Weathering rate: The original observations made by Goldich [12] showed that field observations on the weathering sequence of igneous rocks were the opposite to Bowen's reaction series that ordered minerals in the crystallization sequence of magmas. This qualitative approach was significantly enriched by thermodynamic and kinetic facts which expanded the insight on the susceptibility to weathering of minerals and rocks.

The weathering rate R (mol m^{-2} s^{-1}) is understood as a chemical flux, and for a primary silicate mineral, it is usually defined by the relationship:

$$R = \Delta M / \mathrm{S}\, t$$

ΔM (mol) is the mass change due to weathering, S (m^2) is the total surface area of the reacting mineral, and t (s) is the duration of the reaction [35]. Alternatively, the rate calculation can be based on the mass of solutes produced during weathering:

$$\Delta M_{j,\,\mathrm{solute}} = \Delta c_j \mathrm{V}_s$$

This methodology usually involves comparing changes between initial and final solute concentrations Δc_j (M) in a known volume of water V_s (L).

The mass changes due to weathering (ΔM) can be tackled through several approaches, like bulk changes in regolith, small-scale changes in mineral and rock compositions, changes based on solute compositions, characterization of fluid transport, and weathering based on solutes in soils and in groundwater [34]. Of particular significance for this chapter is the weathering rate assessed through the flux of surface water solutes:

$$\begin{aligned} Q_{weathering} = Q_{watershed\ output} - Q_{precipitation} \\ - Q_{anthropogenic} \pm Q_{biology} \\ \pm Q_{exchange} \end{aligned}$$

The expression shows that weathering constituents connected with watershed discharge must be adjusted to other potential sources or sinks within the watershed, such as atmospheric and anthropogenic inputs, biological contributions, and the impact of ion exchange processes in soils. Generally, in pristine watersheds the weathering flux is the difference between the watershed output and the precipitation input, because biologic and exchange reactions are presumed to be at a steady state.

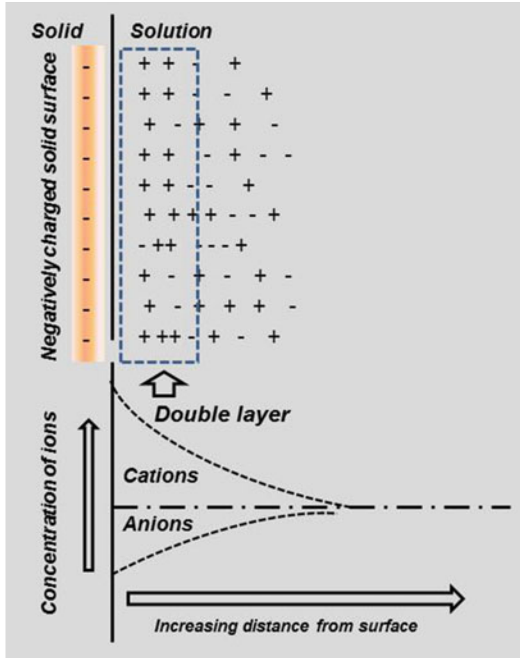
Measuring solute fluxes in surface fresh water discharge is an indirect method to estimate chemical weathering rates. Surface water dissolved substances represent discharges from other weathering environments that are spatially and momentarily integrated by the watershed flux. Variations in solute concentrations are supported by a mixture of several end-member sources, for example, a groundwater component, a component representing waters from regolith, and an organic component comprising soil-derived near-surface runoff. The interpretation of the solute discharge flux is further obscured because the relative proportions of the mentioned components fluctuate seasonally. In a mountainous drainage basin, for instance, hill-slope waters dominate during the rainy season, groundwater rules at lessened flow during the dry season, and storm events appear to control the flux of the organic component. Therefore, the interpretation of solute discharge fluxes and the ensuing calculation of weathering rates should be undertaken with caution, even in well-researched watersheds. Despite some uncertainty surrounding the calculation of weathering rates in river basins of widely variable sizes, TDS and SiO_2 fluxes and yields have proved valuable in assessing, at a global scale, the relative significance of physical vs. chemical weathering rates [9].

Exchangeable Ions

All clay minerals exhibit ion exchange behavior to some degree [36, 37], as also do natural organic compounds and colloidal oxide-hydroxides (i.e., also collectively referred to as oxyhydroxides). This is a process that takes place wherever small particles (i.e., less than 1 or 2 μm) are in contact with aqueous solutions, for example, in regolith and soil, in fluvial and marine sediments, in aquifers, etc.

The properties of colloidal suspensions result from electrical charges on the surface of the particles. Such electrical charges (i.e., mostly colloidal) are linked to their small size and subsequent large specific surface area. The atoms on the surface of solid particles are moderately bonded because they are not surrounded by ions of the opposite charge. When a small particle such as clay (e.g., vermiculite, smectite) is suspended in water, some of the interlayer cations pass into solution, developing a negatively charged silicate structure surrounded by a disperse cloud of cations (Fig. 10). Such suspensions may be stable for long periods of time but are flocculated if the ionic strength of the aqueous solution (e.g., in an estuary) increases markedly. Clays, feldspar, quartz, and manganese oxides are negatively charged except in significantly acidic environments (i.e., pH < 2.0) where they acquire a positive charge. The iron oxide-hydroxides can have both negative and positive charges in the 5–9 pH range. In contrast, particulate aluminum oxide-hydroxides exhibit positive charges except at pH > 9.

Cations adsorbed on small particles surfaces which are in contact with dilute electrolyte solutions (e.g., river water) are exchangeable for ions



Fresh Water Geochemistry: Overview, Fig. 10 Simple outline of the electrical double layer in a submicron particle. The excess negative charge on the particle surface is balanced by an excess concentration of positive charges (i.e., cations) in solution near the particle-solution interface

in the solution. Such ion exchange reactions can attain equilibrium that causes water-particle suspensions to respond to chemical changes in the solution in agreement with Le Châtelier's principle. This capability of small particles to adsorb exchangeable ions can be measured (and therefore defined) by the *cation exchange capacity* (CEC), which is a procedure that involves the measurement of uptake and release of NH^{4+} from 1 M ammonium acetate ($NH_4CH_3CO_2$) solution at pH 7.0. CEC is not a precise or fundamental quantity because it varies as a function of pH and, also, as a function of the ions occupying the exchange sites in the particle.

Exchange equilibrium between monovalent (or two divalent) cations can be represented by the mass action equation:

$$A - \text{clay} + B^+ = B - \text{clay} + A^+$$

Typical CEC in clays are vermiculites, 120–200 meq/100 g; smectites, 80–150 meq/100 g; illites, 10–40 meq/100 g; kaolinite, 1–10 meq/100 g; and chlorite, <10 meq/100 g.

Sayler and Mangelsdorf [38] studied the cation exchange characteristics of Amazon River suspended sediment in order to determine the supply of exchangeable cations to Amazon's geochemical fluxes. The range of exchangeable cation compositions was equally narrow in the river and in seawater. In river water, the exchangeable cation complement (equivalent basis, exclusive of H^+) was 80% Ca^{2+}, 17% Mg^{2+}, and 3% Na^+ plus K^+. In seawater Na^+ and Mg^{2+} are about equal (38%) while Ca^{2+} ~ 15% and K^+ ~ 9%. On reaction with seawater, river suspended sediment took up an amount of Na^+ equal to about 1/3 of the dissolved river load, as well as amounts corresponding to 15–20% of the dissolved fluvial K^+ and Mg^{2+}.

Adsorption

Adsorption is the adhesion of atoms, ions, or molecules to the surface of a solid or, in a more general sense, the buildup of solutes in the surrounding area of a solid-solution interface. The adsorption mechanisms (i.e., surface-based processes) can be divided into *electrostatic adsorption*, where the ions in the aqueous solution are pulled toward a surface of the opposite electrical charge (i.e., the previously discussed CEC is an example); *physical adsorption*, where the attraction to the solid surface is caused by moderately weak van der Waals forces; and *chemical adsorption*, where chemical bonding develops between the solute molecule and one or more atoms exposed on the surface of the solid.

Adsorption is among the most important chemical processes influencing the movement of contaminants in water, particularly in groundwater. It is, as well, an important effect on mineral dissolution rates [35, 39], but the attention here is mostly directed to the role of adsorption phenomena in the presence of trace elements in fresh water. It is a known fact that the concentrations of trace elements in natural waters are far below the values that would be predicted for saturation with respect to a solid phase. The most general reason for the low concentrations of trace elements in fresh waters is their adsorption onto solid phases such as colloidal particles of iron or manganese oxide-hydroxides.

Adsorption is usually described through *isotherms*. The term stems from the fact that the measurements were made at constant temperature. In the procedure, the amount of solute (i.e., adsorbate) on the surface (i.e., adsorbent) is determined as a function of its concentration. The quantity adsorbed is nearly always normalized by the mass of the adsorbent to allow comparison of different materials.

The simplest adsorption isotherm is the linear distribution coefficient or linear K_d:

$$K_d = \mathrm{m}_{i(ads)}/\mathrm{m}_{i(soln)}$$

The variable $\mathrm{m}_{i(\mathrm{ads})}$ is the concentration of the chemical species adsorbed on the solid phase (i.e., usually moles kg^{-1} of solid), and $\mathrm{m}_{i(\mathrm{soln})}$ is the concentration of the species in solution (i.e., moles l^{-1}).

The value of a specific K_d is a function of the properties of the adsorbent and the composition of the adsorbate; both must be measured experimentally in each system under study, and the result cannot be extrapolated from one system to another.

There are several available isotherms, applicable to diverse systems: Langmuir, Freundlich, BET, and Kisliuk, among several others (e.g., [15, 16, 20]). The mathematical methodology used in every case spans from comparatively simple empirical equations to complex mechanistic models of interaction at the adsorbate-adsorbent interface. The scientific literature offers many examples of studies performed with the purpose of gaining insight on the complexities of adsorption processes and on their important environmental implications [40–42].

Organic Matter in Fresh Water Systems

Besides water and inorganic dissolved and particulate phases, rivers transport organic matter (OM) which may end up, eventually, in the coastal sea. Such organic matter may be autochthonous, if it is the result of the biological production that occurs within the water body (e.g., photosynthesized algae, bacteria, copepods) or, conversely, may be allochthonous, if it is organic debris from eroded soils, dead microorganisms, and even tree leaves and other remains from land plants. As a result of such processes, all fresh waters contain dissolved organic molecules, organic colloids, and particulate organic matter. Before the late 1960s, geochemists were inclined to ignore organic components in fresh waters because of analytical difficulties and theoretical complexities involved in the subject. Soon afterward (i.e., early 1970s), scientific interest began to expand [43, 44] largely nurtured by the rising awareness on climate change and the availability of new, more evolved analytical instrumentation. It is now recognized that organic substances in fresh waters – besides the role played in the global carbon cycle – are important in weathering processes, in light attenuation and photochemical reactions, in diagenetic processes, and in the transport of trace metals.

Organic matter is considered dissolved if it passes through a 0.45 μm cellulose acetate membrane filter. Such filtrate is customarily designed as *dissolved organic carbon* (DOC) or *dissolved organic matter* (DOM). The material retained by the filter is known as *particulate organic carbon* (POC). The most commonly reported analytical parameters are *total organic carbon* (TOC) and DOC. Although straightforward, this definition is arbitrary because many submicron organic colloids and some microorganisms can pass through 0.45 μm membrane filters and, also, because the effective pore size of 0.45 μm cellulose filters is reduced by partial clogging during the filtration of a water sample.

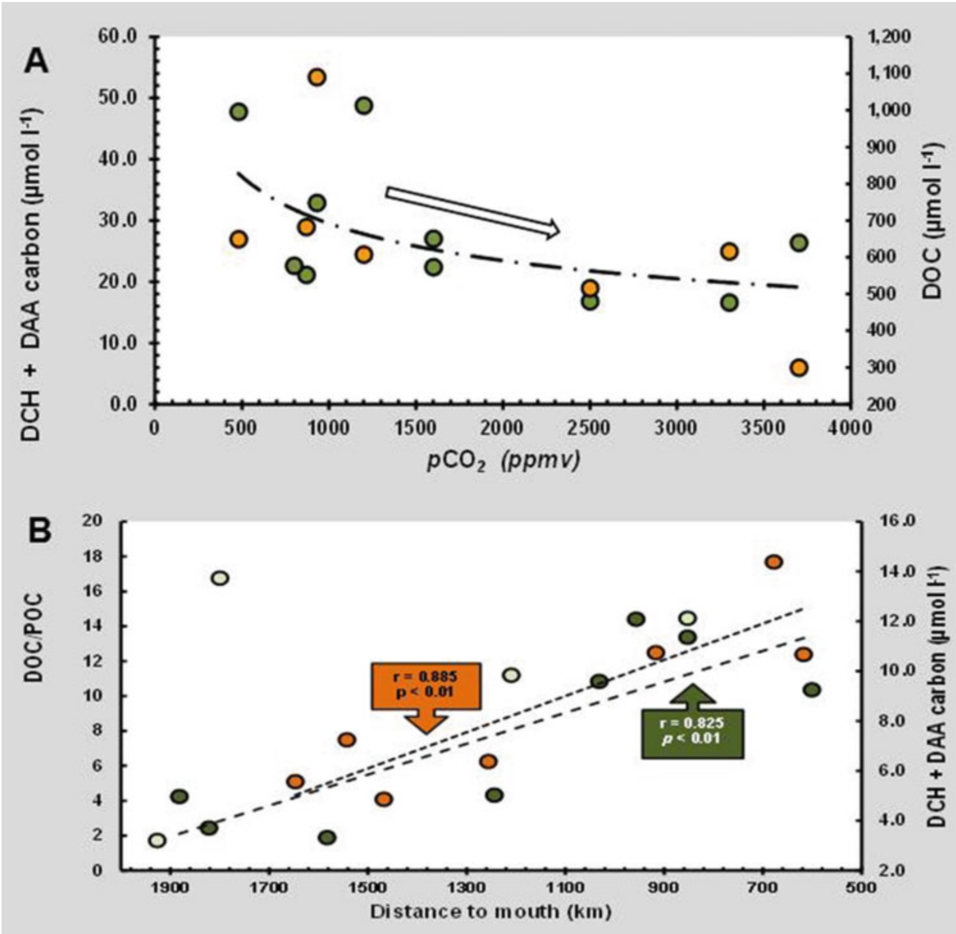
TOC and DOC: The great impact that climate change awareness directed toward the understanding of the global carbon cycle determined that several attempts were made in the scientific community to attain a reasonably accurate estimate of riverine carbon fluxes. In over 20 years, the estimates almost doubled: from a TOC flux ~16.5 Tmol y^{-1} in 1975 [45] to ~31.5 Tmol y^{-1} in 1996 [46] (1 Tmol = 10^{12} mol). Currently, 28–31 Tmol y^{-1} is assumed as the likely variability of the annual riverine TOC flux from the continents to the oceans, which corresponds to a global TOC mean concentration of 730–880 μmol l^{-1}. On a global scale, the average concentrations of DOC fluctuate between ~400 and ~480 μmol l^{-1} (~4.4–5.8 mg l^{-1}).

Six distinct fractions make up the river-transported dissolved organic matter: hydrophobic acids, bases, and neutral compounds and hydrophilic acid, bases, and neutral compounds. In average fresh waters, more than 80% of DOC is distributed between hydrophobic acids and hydrophilic acids in a 2:1 ratio. Less than 20% of DOC is evenly distributed between hydrophilic bases and the two neutral bases. Besides, in average fresh waters, fulvic acids and humic acids account for ~60% of DOC, roughly distributed in a 3:1 ratio [47].

On the basis of available global riverine biogeochemical data, total mean concentration of dissolved hydrolysable sugars is ~2.6 μmol l^{-1}, with hexoses (glucose, galactose, rhamnose, mannose, and fucose) more abundant than pentoses (xylose, arabinose, ribose, and lyxose). The sum of the mean concentration of hydrolysable sugars accounts for ~3.0% of the organic carbon in dissolved organic matter [47].

Similarly, the average fresh water contains ~1.3 μmol l^{-1} of total hydrolysable amino acids. Jointly, they account for 1.8% of the organic carbon in dissolved organic matter and ~18% of its nitrogen. The most plentiful class of lignin-derived phenols are vanillyl phenols. As a group, the lignin-derived phenols account for 0.6% of the organic carbon determined in fresh water DOM [47]. Using the available mean values, the addition of sugars, amino acids, and lignin-derived phenols accounts for less than 6% of the dissolved organic carbon measured in fresh waters [47].

TOC, sugars, and amino acids were measured in the Paraná River. DOC concentrations fluctuated between ~500 and 850 μmol l^{-1}, higher than the global average. Figure 11a shows the variability (i.e., in a hydrological year) of dissolved sugars plus amino acids, and DOC, as a function of the CO_2 partial pressure (pCO_2). Both variables show a decreasing trend as pCO_2 increases, thus suggesting that at least part of the decreasing trend may be attributable to increased biological



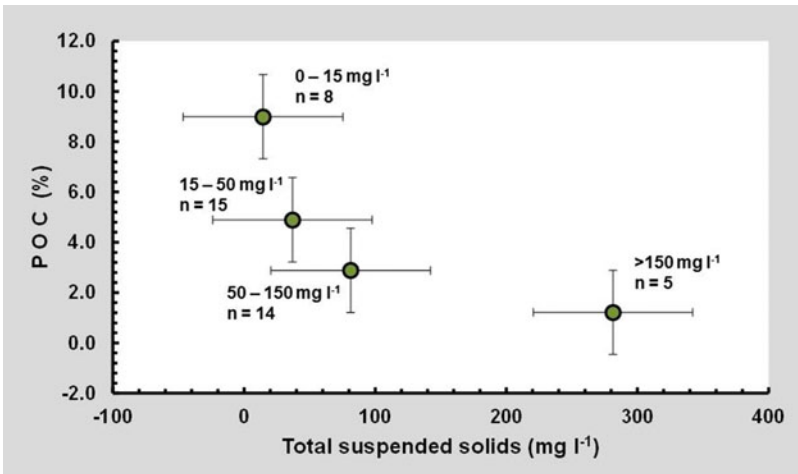
Fresh Water Geochemistry: Overview, Fig. 11 (**a**) Variability of dissolved sugars + amino acids carbon and DOC, as a function of CO_2 partial pressure (pCO_2). The samples were collected in the Paraná River, throughout a hydrological year, about 600 km upstream from the mouth. *Green circles* correspond to dissolved sugars + amino acids carbon; *orange circles* to DOC. (**b**) Joint downriver variability of dissolved sugars + amino acids carbon, and DOC/POC atomic ratio in the Paraná River. *Dark green circles* correspond to sugars + amino acids carbon measured in the main channel and *light green circles* to samples from tributaries; *orange circles* correspond to DOC/POC ratios (The data were collected within the SCOPE/UNEP International Carbon Project (e.g., [48]))

consumption (i.e., respiration). Figure 11b, on the other hand, shows that both sugars plus amino acids and the DOC/POC ratio increase toward the river mouth, a characteristic which may be connected with Paraná's extensive flood valley (i.e., the dynamic water exchange between lotic and lentic environments). In the middle Paraná River, mean dissolved hydrolysable sugar concentrations were ~12.5 µmol l^{-1} and mean dissolved hydrolysable amino acids ~15.8 µmol l^{-1}; jointly they accounted for ~5% of DOC [48].

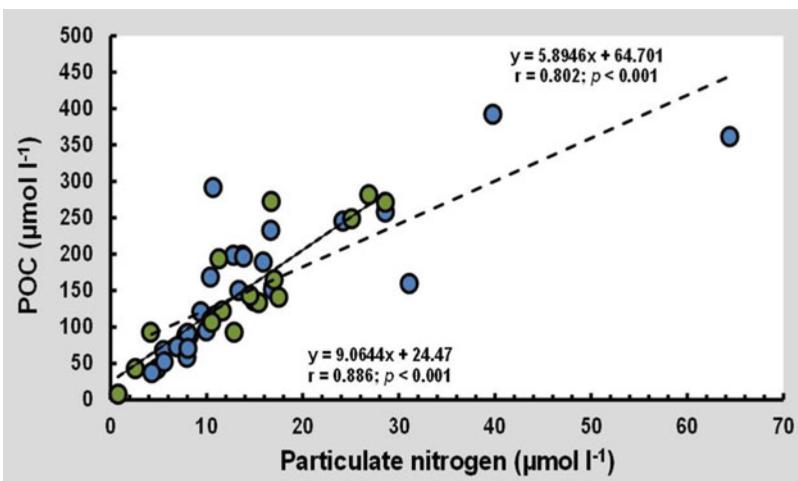
Particulate organic matter (POM): Soils hold the largest carbon pool of all terrestrial ecosystems [49], a significant part of which is POM, eventually eroded and transported – mainly via fresh water systems – to the world's oceans. POM is subjected to different processes in fresh water environments, which include mineralization, disaggregation, and sedimentation; such processes determine POM's role, behavior, and fate in aquatic ecosystems. POM can be divided into living (e.g., bacteria and plankton) and nonliving (i.e., detritus of broadly variable size) fractions. Generally, most of the living part is promptly recycled, along with a marginal detrital fraction.

POM (or POC) in fresh water (i.e., rivers and lakes) is often expressed as a percentage of total suspended sediment (TSS) and as concentrations. It has been known for many years that there is an inverse relationship between TSS and the percentage of POC (e.g., [50]): the larger the TSS concentration, the smaller the relative contribution of POC to total TSS. Figure 12 shows an example for the middle Paraná River.

The carbon to nitrogen (C/N) ratio is an adequate indicator of the provenance of organic matter. In general, a C/N > 20 can be taken as an indicator of vascular plant material, those between 15 and 8 seem to indicate terrigenous sources (i.e., soil humus), and lower values are thought to represent planktonic sources. The Paraná River exhibited a TSS C/N ratio that fluctuated between 9.8 and 11.5, roughly following the prevailing hydrological situation [48]. Figure 13 exhibits in more detail the variability of C/N ratios in the middle Paraná's TSS. The data collected during low waters (water discharges ≈16000 $m^3 s^{-1}$ or lower) showed a significant regression slope ~9, indicating a mostly allochthonous organic matter dominated by soil-derived detritus. In contrast, when discharge distinctly exceeded the mean (i.e., there was ENSO-induced exceptional flooding



Fresh Water Geochemistry: Overview, Fig. 12 Variability of POC concentrations (% of TSS) and TSS in the Paraná River, about 600 km upstream from the mouth. Mean concentrations (*green circles*) and typical error bars (The data were collected within the SCOPE/UNEP International Carbon Project (e.g., [48]))



Fresh Water Geochemistry: Overview, Fig. 13 POC and PN variability in the Paraná River, about 600 km upstream from mouth. Data divided into samples collected with discharges lower than mean Q (~16,000 $m^3 s^{-1}$, *green circles*), and samples collected with discharges higher than mean Q (~17,000–50,700 $m^3 s^{-1}$, *blue circles*) (The data were collected within SCOPE/UNEP International Carbon Project (e.g., [48])) during the sampling period [18]), the regression slope for the corresponding data decreased to ~5.9, thus suggesting that the organic fraction transported in suspension had become dominantly autochthonous (i.e., algae).

The source of POC can also be identified nowadays by a number of powerful instrumental techniques, like the determination of stable isotopes with mass spectrometry or the organic fingerprint by means of gas chromatography/mass spectrometry. An example is the study performed several years ago by Onstad et al. [51], identifying POM sources in rivers from the continental USA. The work undertook the study of the elemental, stable carbon isotope and lignin-phenol compositions of suspended particulate organic matter collected from rivers draining South Central USA. The atomic C/N ratios (~11 ± 2) were similar to those reported worldwide for riverine POM. The corresponding stable isotope data ($\delta^{13}C$ values varied from −18.5‰ to −26.4‰) was consistent with the C3 and C4 plant distribution in the drainage basin; substantially degraded lignin-phenol indicated input from angiosperm-rich plant materials. In short, the obtained results signified that highly degraded organic matter is a key constituent of fine-grained POM transported by rivers of the Central USA.

The most recent estimate on the world's total suspended sediment (TSS) flux to the oceans is 19 Gt y^{-1}[9] (1 Gt = 10^9 t). Likewise, the latest estimate for the TOC flux to the coastal seas is 14.4 Tmol y^{-1} [46]. Therefore, the flux of organic particles is only a small fraction of the total mass of continental debris poured yearly by rivers into the coastal seas. Such POC flux can be translated into a mean concentration of ~330–400 μmol l^{-1} (~3.9–4.4 mg l^{-1}). Based on a limited number of estimates, the global mean DOC/POC ratio is ~1.20 [47].

Nutrients in Rivers and Lakes

Nutrients are chemical elements vital to the development of plant and animal life. In rivers and lakes, nutrients are required for the growth of algae that form the foundation of an intricate food web sustaining the entire aquatic ecosystem. The most common nutrients in rivers and lakes are N and P. Si also plays a significant role in primary biological production.

Although eutrophication is the natural process of organic matter enrichment of lakes and rivers through the stimulation of nutrients, human activity has dramatically increased its rate in many water bodies, altering the associated biological and physical characteristics. Excessive nutrients not only affect fresh water quality but also may adversely impact on health in humans and livestock. Phosphorous is not toxic to human adults in moderate concentrations, but high levels of NO_3^- in drinking water (e.g., >10 mg l^{-1}) can adversely affect livestock or human infants. An undesirable consequence of nutrient overabundance is excessive algal growth, such as some cyanobacteria (i.e., blue-green algae), which produce toxins that may affect the liver and the nervous system.

Background levels of N and P in fresh waters are generally quite low. Nitrogen is present in rather small quantities in igneous rocks and in the oceans. The greater portion of the amount on the planet occurs uncombined in the atmosphere (78.09% by volume). In fresh waters, it may occur in several forms, depending on the level of oxidation. It may be present as NH_3 or as NH_4^+. In NO_2^- it has the form N^{3+}, and NO_3^- is the most common species of N in fresh waters, in the completely oxidized state (N^{5+}).

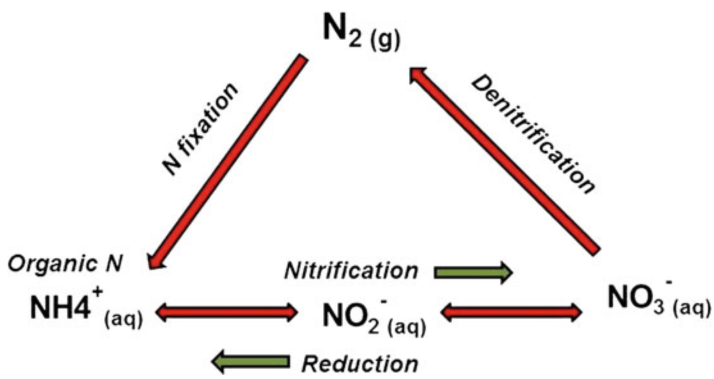
The N cycle is the biogeochemical cycle which converts N into various chemical forms as it flows through the atmosphere and terrestrial and marine ecosystems. The conversion of N can be completed through both biological and physical processes. Particularly important in the N cycle are fixation, ammonification, nitrification, and denitrification (e.g., [52]): all N that is available to biota was originally derived from N fixation, either by lightning or by free-living and symbiotic microbes associated with leguminosae. However, N fixation contributes only to about 12% of the N that is annually assimilated by land plants. The remaining N must be obtained from internal recycling and from decay of dead tissue in regoliths. Microorganisms mediate, under aerobic conditions, the nitrification of NH_4^+ (and NH_3)

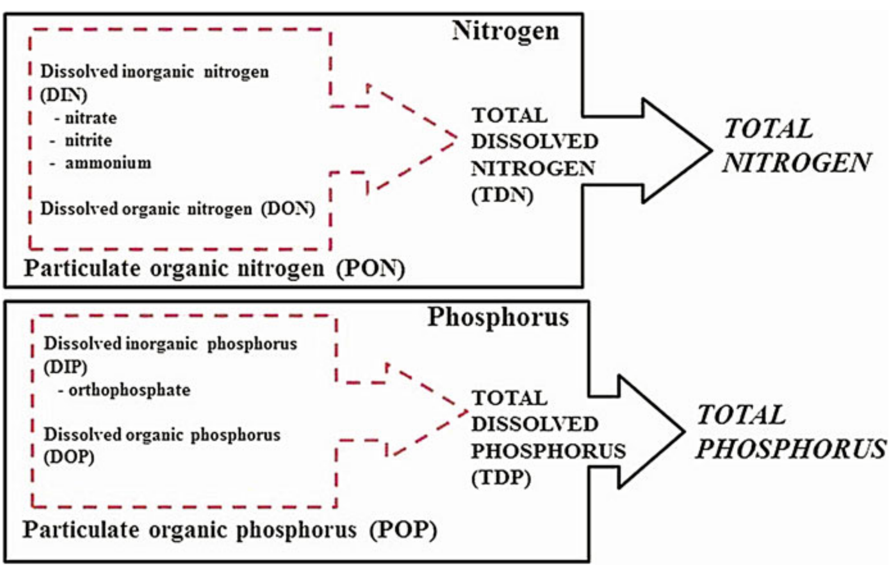
to NO_2^- and NO_3^-. Organic compounds can cause the reduction of NO_3^- to N_2 (denitrification). The most important biologically mediated N conversion processes are schematically presented in Fig. 14.

Nitrate concentrations in fresh waters are usually less than 0.6 mg l^{-1}; in streams and rivers, natural N sources include eroded regolith and soils, leaves, grasses, and other organic debris from the riparian vegetation. The extensive use of fertilizers and insufficiently treated sewage are, undoubtedly, the most important consequence of human activities. The outline in Fig. 15 shows the speciation of N and P in fresh waters.

The weathering of apatite [$Ca_{10}(PO_4)_6(OH,F,Cl)_2$] releases calcium phosphate, which is soluble to some degree in CO_2-supersaturated water. This is the main natural source of P, along with other less abundant phosphates. In contrast with the global cycles of C and N, the global cycle of P is exceptional because it does not have a significant gaseous component. Concentrations of P in world's average river water and river TSS are 25 and 1150 μg g^{-1}, respectively [17]. Therefore, background PO_4^{3-} in fresh water is usually less than 0.1 mg l^{-1}; eroded soil and sediments are the primary natural sources of P. In many areas, the flux of P in rivers is significantly higher than the

Fresh Water Geochemistry: Overview, Fig. 14 Schematic diagram showing the main biologically mediated N conversion processes. NO and NO_2 may occur as intermediaries, both in nitrification and denitrification (Modified from [15])





Fresh Water Geochemistry: Overview, Fig. 15 Schematic diagram showing main chemical species in dissolved and particulate, organic and inorganic N and P

Fresh Water Geochemistry: Overview, Table 1 Nutrients (means) and runoff in Patagonian rivers. Data in $\mu mol\ l^{-1}$

River	Runoff ($mm\ y^{-1}$)	DOC	$NO_3^- - N$	$PO_3^{3-} - P$	SiO_2	N:P
Colorado	151	99.91	2.28	1.87	183.08	1.22
Negro	251	99.91	2.07	5.17	199.72	0.40
Chubut	32	174.84	11.57	2.71	199.72	4.26
Deseado	11	174.84	6.85	6.13	316.22	1.12
Chico	65	632.75	9.00	12.59	94.87	0.71
Santa Cruz	1445	66.61	9.00	2.94	28.29	3.06
Coyle	9	424.61	0.79	7.75	39.94	0.10
Gallegos	2089	557.82	47.12	4.84	249.65	9.73
*Mean**		334.11	27.96	4.19	162.00	6.22

*Runoff-weighted mean

natural background as a result of pollution and fertilizer runoff.

The ratio of N to P specifies which nutrient will possibly limit the growth of algae (i.e., primary biological productivity). C, N, and P occur in algal tissue in a surprisingly coherent ratio of atomic weights of 106:16:1 known as the Redfield ratio [53]. This ratio is named after the American oceanographer Alfred C. Redfield, who described this empirically developed stoichiometric ratio, which he described in an article published in 1934. Redfield analyzed thousands of samples of marine biomass and found that globally the elemental composition of marine organic matter (dead and living) was remarkably constant worldwide. The stoichiometric ratios of C, N, and P persist relatively constant from both the coastal and open ocean regions, as a result of the incorporation of these elements in photosynthesis and growth:

$$106CO_2 + 16NO_3^- + HPO_4^{2-} + 122\ H_2O + 18H^+ \rightarrow (CH_2O)_{106}(NH_3)_{16}(H_3PO_4) + 138O_2$$

Based on this natural ratio and on abundant bioassay experimentation with lake algae, it is generally accepted that when N:P ratio decreases below 16:1, algae will have a limited N supply per unit of P and, accordingly, experience nitrogen constraint, while ratios above 16:1 indicate P limitation. At N:P ratios oscillating between 10:1 and 20:1, a combined limitation by both nutrients is likely to occur.

The lakes in Patagonia owe their remarkable oligotrophy to the limiting power of N [54]. This is in opposition to what is often seen in many other world lakes, which productivity is basically controlled by P.

The biogeochemical typology of Patagonian main water courses was studied in order to determine the factors controlling the continental fluxes of C and nutrients. Patagonia's rivers are partially responsible for the noticeable biological productivity of Patagonia's coastal region. Data in Table 1 [55] suggest that N is also a limiting factor in the biological productivity of Patagonian rivers. The significantly higher N:P ratio of the Gallegos River is likely caused by the low-grade coal-bearing beds and abundant mudstones exposed in its drainage basin. Silica appears to be adequately furnished in some rivers (e.g., the Colorado, the Negro, the Chubut, the Deseado, and the Gallegos) and poorly provided in others (e.g., the Chico, the Santa Cruz, and the Coyle).

Minor and Trace Elements

In igneous petrology and in high-temperature geochemistry, a trace element is one whose concentration is less than 0.1% of a rock's total mass. In natural waters, however, trace elements are generally characterized by concentrations lower than

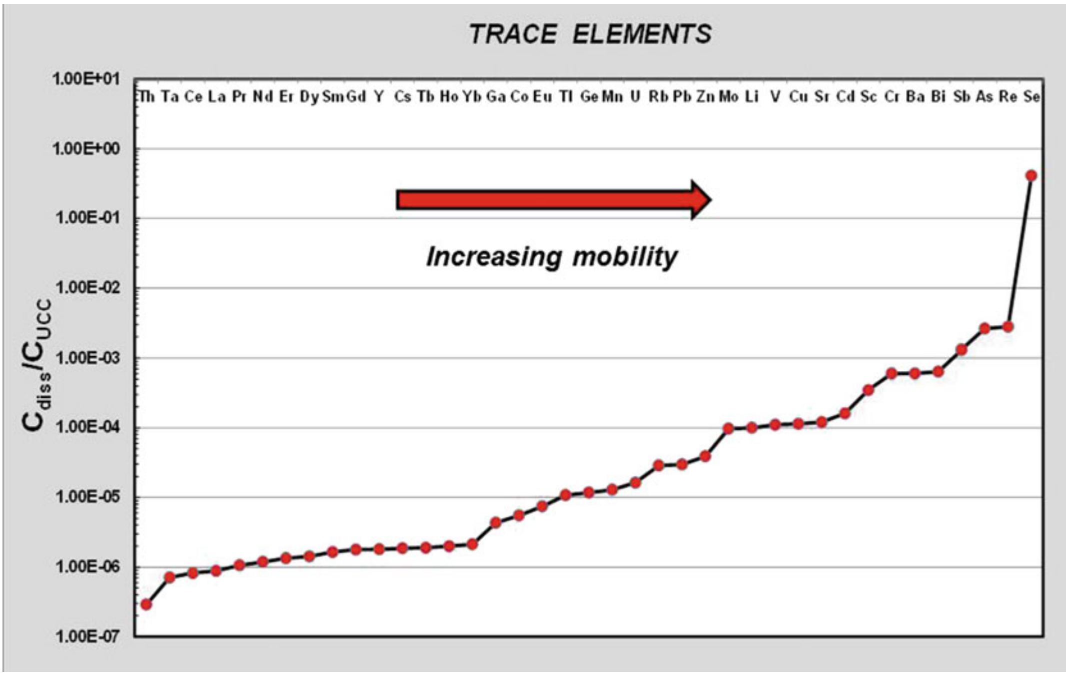
1 mg l^{-1} (i.e., or 1 ppm). Major elements in rocks may be in trace concentrations in natural waters (e.g., Fe, Al, and Ti), and, contrastingly, some elements in trace amounts in rocks are major constituents in fresh waters (i.e., C, Cl) [56]. This guides to what W. M. White [56] considers as the best definition of a trace element: "an element whose activity obeys Henry's Law in the system of interest." This denotes sufficiently dilute concentrations that interactions between atoms of the element under consideration are sufficiently rare that its behavior is independent of its concentration (Box 8). Trace element concentrations in fresh waters span over ten orders of magnitude, similarly to the range of abundance in the Earth's crust.

Box 8 Concentration Units of Trace Elements
Trace element concentrations are frequently reported in weight units as parts per million (ppm), parts per billion (ppb), or parts per trillion (ppt), which correspond to milli-, micro-, and nanograms per kilogram. Usually used molarity units are milli-, micro-, or nanomol per liter (i.e., mM, μM, and nM).

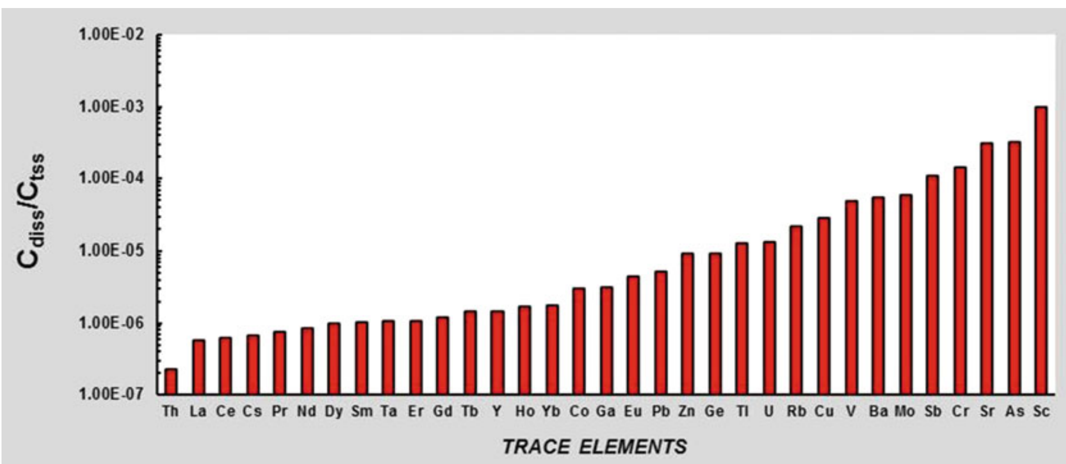
In fresh water systems, trace element concentrations not only depend on their crustal abundance but also on their mobility during the weathering process of minerals and rocks. Other mechanisms may be locally significant, such as hydrothermal contributions or the infrequent fallout of volcanic ashes into lakes. An interesting example in connection with their mobility is supplied by one of the rare earth elements (REE): europium (Eu). Under reducing conditions, such as those existing within the mantle or lower crust, Eu may exist in the divalent state, Eu^{2+}. This causes an increase in ionic radius (i.e., about 17%) making it basically equal to Sr^{2+} (ionic radius = 1.26 Å for a coordination number = 8). Therefore, Eu substitutes freely in place of Sr in feldspars, particularly in anorthite (i.e., Ca-plagioclase), steering to a distinguishing geochemical behavior compared to the other REE. Plagioclase is among the first minerals to weather, and hence, Ca, Sr, and Eu are delivered to the aqueous environment during the early stages of chemical weathering. Eu, in very low concentrations (e.g., ~10–30 ng l^{-1}), does not remain for a long time in dissolved state due to the peculiar behavior that REE exhibit during weathering and transport and is rapidly adsorbed onto exchange surfaces. This important aspect of REE dynamics will be treated ahead, but the example is relevant to show the transfer of trace elements from minerals to fresh waters and some of the involved complexities. At this point it is worth mentioning that, in general, trace elements are far more fractionated by weathering and transport processes than major elements.

Gaillardet et al. [21] proposed an approximate classification of trace element mobility in fresh waters. The first group involves the highly mobile elements (i.e., mobility close to or greater than that of Na): Cl, C, S, Re, Cd, B, Se, As, Sb, Mo, and Sr. The second group includes those elements considered moderately mobile, with a mobility ~10 times less than Na: U, Os, Li, W, Mn, Ba, Cu, Ra, Rb, Co, and Ni. The third group comprises the nonmobile elements, with mobility 10–100 times less than that of Na: REE, Zn, Cr, Y, V, Ge, Th, Pb, Cs, Be, Ga, Fe, and Hf. Finally, there is a group of the most immobile trace elements, which are more than 100 times less mobile than Na: Nb, Ti, Zr, Al, and Ta. This mobility scheme [21] arises from contrasting the dissolved concentrations in rivers with the mean continental abundance of trace elements [17]. A similar exercise was prepared in Fig. 16, but comparing Paraná River dissolved trace elements data with the upper continental crust composition [17].

The diagram, ordered by increasing mobility, shows elements such as Se, Re, As, and Sb as the most mobile and Th and the REE as the least mobile (Box 9). There are some differences regarding the mobility scheme of Gaillardet et al. [21], which are more obvious when both the dissolved and the particulate phases (i.e., TSS) are used in a sort of partition coefficient



Fresh Water Geochemistry: Overview, Fig. 16 UCC-normalized diagram of dissolved trace element concentrations in the middle Paraná River. The graph shows increasing mobility (from *left* to *right*) of trace elements during weathering and transport processes (UCC data from [17]; Paraná River data from [26])



Fresh Water Geochemistry: Overview, Fig. 17 Mean dissolved trace element concentration/mean TSS trace element concentration ratio in the middle Paraná River (i.e., about 600 km upstream from mouth) (Data from [26])

diagram (K_d) for the Paraná River (Fig. 17). The minor variances with the global scheme [21] are attributable to changes in the lithology as well as to other intervening factors (e.g., climate, relief). At any rate, the dominance of some mobile elements is evident in Paraná's material (e.g., As, Sr, Sb) as well as the immobility of the REE, for example.

Box 9 Filtration: A Key Procedure in Trace Element Fresh Water Geochemistry
Dissolved and particulate phases in fresh water samples are separated for subsequent analysis by means of vacuum or pressure filtration. Major dissolved components are usually analyzed after filtration with 0.45 μm pore size cellulose acetate membrane filters; trace elements require a finer pore size (0.22 μm) because there is ample evidence that they tend to associate with submicron particles which are not retained by ~1/2 μm pore size filters. Moreover, recent research has shown that dissolved REE concentrations are significantly reduced by ultrafiltration (from 3 to 100 kDa, 1 kDa (Dalton) = 5.0 10^{-3} μm), thus implying that there is an organic/inorganic colloidal phase (especially in waters with low pH), which can pass through the fine pores (i.e., 0.2 μm) of usual filtration membranes ([21] and references therein).

Table 2 [57] shows a set of heavy metals (mostly transition elements [58]) whose total concentrations (i.e., particulate + dissolved) were determined in the middle Paraná River during a major flooding event (i.e., the ENSO-triggered flood of 1982–1983) [58]. The decreasing mean concentration sequence for the considered time period was Mn > Zn > Cu > Ni > Cr > Pb > As > Cd. In most metals, over 70% of the total concentration was accounted for by the particulate phases, being Cd (~30%) and Pb (~44%) the metals with the largest relative dissolved fraction. Correlation analysis showed that the total concentration of Ni was significantly ($p < 0.05$) correlated with Cu, Zn, and As, as also was Cd, with Zn and Pb. The increased discharge recorded during the extraordinary flood (i.e., the river reached flood peaks about four times as large its mean discharge, i.e., ~16,000 $m^3\ s^{-1}$) did not show a discernible (and significant) relationship with total metal concentrations.

The attention now will switch to a special group of trace elements, the REE. Following La (lanthanum, Z = 57), electrons begin to occupy the 4f orbitals, thus shaping the 14 metals identified as lanthanides or REE. The chemical properties of all REE are remarkably similar, including a stable trivalent state.

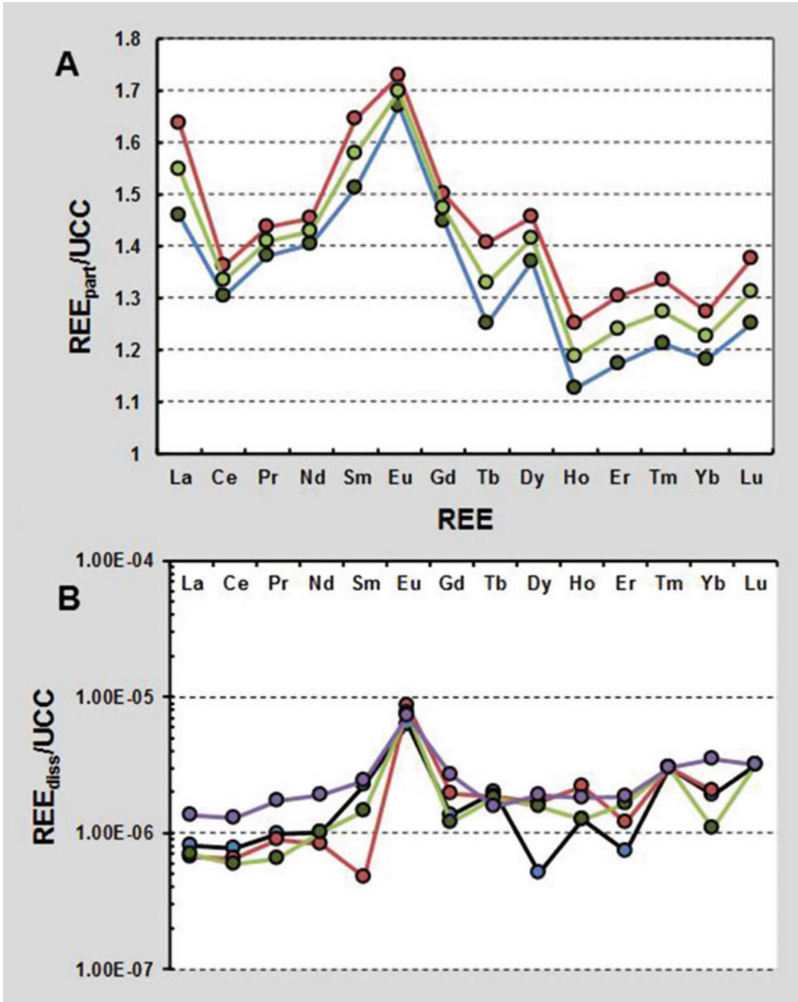
There is a steady decrease of ionic radius from La^{3+} to Lu^{3+}, known as the "lanthanide contraction." Consequently, the *light rare earths* (LREE, La to Sm) are incompatible elements (Box 10), whereas the *heavy rare earths* (HREE, Gd to Lu), owing to their smaller ionic radii, are more easily accommodated in the crystal structure of some rock-forming minerals. The REE, though nearly identical in other chemical properties, range continuously in behavior from incompatible to selectively compatible.

Box 10 Incompatible and Compatible Elements
In geochemistry and igneous petrology, an incompatible element is one that is unfitting in size and/or charge to the cation sites of the minerals in which it is included. Such elements are concentrated during the fractional crystallization and the generation of magma by the partial melting of the Earth's mantle and crust. There are two groups of incompatible elements that have difficulty entering the solid phase: One known as LILE, or *large-ion lithophile elements*, includes elements having large ionic radius, such as K, Rb, Cs, Sr, and Ba. The other group, called HFSE, or *high field strength elements*, includes elements of large ionic valences (or high charges), such as Zr, Nb, Hf, REE, Th, U, and Ta. In contrast, compatible elements are depleted in the crust and enriched in the mantle, with Ni and Ti as typical examples. In general, the mantle is implied when an element is referred to as being "compatible" without specifying the rock.

Figure 18 shows the concentrations of dissolved and particulate (i.e., in the TSS) in the middle Paraná River. Figure 18a exhibits the UCC-normalized extended diagram, often referred to

Fresh Water Geochemistry: Overview, Table 2 Total heavy metal concentrations determined in the Paraná River (about 600 km upstream from the mouth) during the 1982–83 ENSO flood

Date	Cr_{Total}			Mn_{Total}			Ni_{Total}			Cu_{Total}			Zn_{Total}			As_{Total}			Cd_{Total}			Pb_{Total}		
	Conc.	± *s.d.*	% part.	Conc.	± *s.d.*	% part.	Conc.	± *s.d.*	% part.	Conc.	± *s.d.*	% part.	Conc.	± *s.d.*	% part.	Conc.	± *s.d.*	% part.	Conc.	± *s.d.*	% part.	Conc.	± *s.d.*	% part.
08.06.82	6.6	0.2	91	68.7	1	87	8	0.6	>99	5.7	0.6	70	32.9	0.2	>99	7.7	0.3	>99	0.31	0.03	42	7.9	0.5	4
09.07.82	7.5	0.2	88	94.7	1.5	96	9	0.5	>99	6.4	0.3	92	73.2	1.5	90	n.d.	n.d.	n.d.	0.77	0.1	91	8.1	0.2	96
09.23.82	4.9	0.2	90	77.9	1	92	5.3	0.4	>99	n.d.	n.d	n.d	26.2	0.5	>99	3.3	0.1	>99	0.74	0.08	73	0.9	0.1	67
10.07.82	19.1	0.2	97	42.5	0.3	82	10.4	0.8	56	2.6	0.1	54	41.1	3.7	29	0.5	0.1	>99	1.27	0.01	98	n.d	n.d.	n.d
10.27.82	32.3	0.7	96	93.8	3.1	99	0.8	0.2	>99	4.9	0.2	67	26.1	0.7	74	1.3	0.2	62	0.08	0.01	>99	1.5	0.1	80
11.24.82	8.5	0.3	93	96.1	1.4	98	6	1.2	75	43.6	0.6	92	28.2	1	80	3.1	0.4	77	0.04	0.005	>99	n.d.	n.d.	n.d.
12.15.82	6.1	0.4	92	12	0.2	>99	10	0.8	39	n.d.	n.d.	n.d.	22.7	0.7	76	2.6	0.4	77	0.23	0.03	>99	7.2	0.9	56
12.28.82	5.6	0.2	95	25	0.6	72	4.1	0.3	85	n.d.	n.d.	n.d.	15.9	0.3	58	6.3	0.3	90	0.33	0.04	>99	11.4	0.3	68
02.03.83	4.2	0.1	95	16.3	0.4	40	12.5	0.7	>99	4.8	0.3	52	30.5	1.3	56	7.5	0.4	93	0.13	0.03	54	8.3	0.4	14
02.23.83	6.2	0.1	97	30.7	0.7	46	3.1	0.8	45	3.3	0.2	21	22.8	1.6	49	4.5	0.4	89	0.08	0.01	38	6	0.4	45
03.11.83	2.9	0.1	93	144	1.8	>99	6.5	0.5	83	6.6	0.2	67	61.7	2.9	81	4.8	0.5	88	0.67	0.01	94	12.6	0.7	56
04.14.83	31.2	0.4	99	45.2	0.2	86	7.3	1	25	8.7	0.2	74	22.3	0.6	57	2.7	0.03	>99	0.9	0.01	6	9.5	0.4	11
05.03.83	8.4	0.4	71	55	0.7	>99	6.5	0.4	63	10.2	0.3	64	79.3	1.4	>99	2.5	0.1	>99	0.65	0.07	46	9.9	1.1	42
05.03.83	5.3	0.2	>99	54.7	0.7	80	7	0.5	>99	25.1	1	70	34	0.4	>99	7.1	0.3	>99	0.4	0.05	>99	8.4	2.1	15
08.23.83	2.2	0.2	45	58.2	0.9	95	16.2	1.1	>99	11	0.2	85	64.2	1.5	63	6.6	0.2	>99	0.53	0.07	85	10.7	0.3	51
09.21.83	27	0.2	48	112.1	0.3	86	7	0.1	57	3.9	0.1	38	29.5	0.2	84	5.6	0.3	89	0.11	0.01	45	1.8	0.04	83
10.26.83	5	0.2	96	59.4	0.8	94	14.3	0.9	73	21	0.2	89	45.4	3.2	95	12.9	1	93	0.3	0.04	90	1.7	0.2	71
12.06.83	3.6	0.2	83	101.8	0.5	93	12.5	0.5	74	30.5	0.2	92	49.9	1.3	76	8.2	0.5	91	0.8	0.09	80	1.4	0.1	79
02.15.84	8.5	0.3	89	80	1	91	11	0.8	67	41	0.4	92	50.3	0.8	54	7.4	0.3	93	0.1	n.d.	<1	1.6	0.1	81
04.08.84	2.8	0.1	79	29.7	0.9	>99	15.7	0.3	76	20.5	0.2	>99	66.1	2.5	54	4.8	0.1	88	1.1	0.02	84	13.1	1	60
05.30.84	5.7	0.4	>99	113.3	4.5	>99	3.9	0.2	>99	2.5	0.1	>99	18.7	0.2	>99	8.5	0.3	>99	0.08	0.02	>99	5.2	0.2	60
08.01.84	0.9	0.1	>99	53.7	1.2	31	6.8	0.6	43	7.9	0.4	39	27.3	1.6	24	2.9	0.3	83	0.26	0.02	15	2.4	0.1	75



Fresh Water Geochemistry: Overview, Fig. 18 (**a**) UCC-normalized extended REE diagram for Paraná's TSS. The samples were collected in the middle stretch (i.e., ~600 km upstream from mouth). Notice Eu/Eu* and Ce anomalies. (**b**) UCC-normalized dissolved REE concentrations determined in the Paraná River middle stretch (i.e., ~600 km upstream from mouth). Notice logarithmic y-axis, and outstanding Eu/Eu* (Paraná River data from [26])

as spider diagram or spidergram. The diagram shows significant Ce and Eu anomalies and, also, a fractionation between LREE and HREE, which is expressed in a La_N/Yb_N mean ratio $\approx$ 11.6 (i.e., most post-Archean sedimentary rocks have somewhat uniform REE patterns with $La_N/Yb_N < 15$; the subscript N denotes chondrite normalization [59]).

Under oxidizing conditions, Ce^{3+} may be oxidized to Ce^{4+} steering to a decrease in ionic radius (i.e., about 15%). This oxidation is accompanied by a decrease in solubility and the precipitation of CeO_2, a process which explains the negative anomaly frequently found in sediments.

The Eu anomaly is related to the process described at the beginning of this section. The anomaly is correctly expressed as [59]

$$Eu/Eu^* = Eu_N/(Sm_N.Gd_N)^{0.5}$$

Again, the subscript N denotes chondrite normalization, although sometimes other standards can be used for normalization (e.g., mid-ocean ridge basalt or MORB). In Paraná's TSS, the anomaly is positive ($Eu/Eu^* \approx 0.73$) thus reflecting the significance of plagioclase in the TSS source materials. In the dissolved realm, REE concentrations are several orders of magnitude (10^{-5} to 10^{-7}) lower than the

REE concentrations in the Earth's crust (Fig. 18b), and the europium anomaly appears as persistent.

The complicated and multifaceted combination of several factors determines the mobility of trace elements. There are several such factors which are worth mentioning: (a) the relative susceptibility of minerals and rocks to weathering; (b) the supply to the aqueous system of non-weathering sources, such as anthropogenic, atmospheric, and volcanic (i.e., hydrothermal); (c) the relative ease with which the elements are coprecipitated, adsorbed, etc.; and (d) the capability of trace elements to be complexed by ultrafine colloidal particles.

Isotopes in Fresh Waters

Isotopes are atoms of the same element that have the same number of protons and electrons but have a different number of neutrons. The difference in the number of neutrons between the various isotopes of a single element means that the different isotopes have dissimilar masses. For example, deuterium, a hydrogen isotope, has one neutron and one proton, and it is denoted as ^{2}H or D. In short, the isotopes of a particular element have the same atomic number but different atomic weights.

Some isotopes have nuclei that do not decay naturally to other isotopes. Such isotopes are known as *stable*. In contrast, there are unstable isotopes (i.e., *radioactive*) which have nuclei that spontaneously decay over time to form other isotopes. There is a third kind of isotopes which are formed by radioactive decay but reach a stable condition and do not decay themselves. These isotopes are called *radiogenic.*

Isotopes have proved to be extremely valuable research tools in both high-temperature and low-temperature geochemistry. Stable isotopes are generally used to recognize different water sources or processes that have influenced water during its historical evolution (e.g., after its evaporation in the open ocean). Radioactive isotopes are used predominantly to measure age. Radiogenic isotopes are less extensively used in the study of fresh waters; they are mostly used in the determination of the source (or sources) of specific elements [60, 61].

Stable isotopes: The atomic number of a specific element determines its chemical properties, so that there is a similarity among the chemical behaviors of the various isotopes of the same element. The minor noticeable differences result solely from disparities in mass. The differences in mass cause isotopic fractionation in nature, particularly among lighter elements, because a minor difference in mass is a significant fraction of the total atomic mass. *Fractionation* occurs, for example, when the isotopic ratios in specific phases differ from one another (e.g., the ratio ^{1}H/^{2}H in rain is different from the ratio in the oceans) [62].

Stable isotopic compositions are normally reported as delta (δ) values in parts per thousand (denoted as ‰ or per mil) enrichments or depletions relative to an internationally agreed standard of known composition. With oxygen, for example, δ^{18}O is defined by

$$\delta^{18}\mathrm{O} = \left\{\left[\left({}^{18}\mathrm{O}/{}^{16}\mathrm{O}\right)_{\mathrm{sample}} - \left({}^{18}\mathrm{O}/{}^{16}\mathrm{O}\right)_{\mathrm{standard}}\right] / \left({}^{18}\mathrm{O}/{}^{16}\mathrm{O}\right)_{\mathrm{standard}}\right\} 1000$$

For δ^{18}O, a widely used reference standard is V-SMOW, an acronym for *Vienna Standard Mean Ocean Water*. Another standard (V-PDB, which is a belemnite – i.e., an extinct Jurassic-Cretaceous mollusk – from the Pee Dee Formation of South Carolina, USA) is used to measure ^{18}O in carbonates.

The important stable isotopes of oxygen (i.e., ^{16}O and ^{18}O) and the stable isotopes of hydrogen (i.e., ^{1}H and ^{2}H or D) are extensively used jointly to determine sources of water.

The $H_2{}^{18}O$ molecule is 11% heavier than $H_2{}^{16}O$, which makes it slightly more difficult to evaporate. The vapor in equilibrium with water will be somewhat undersupplied in $H_2{}^{18}O$ and HDO (i.e., $^1H^2H^{16}O$) relative to the coexisting liquid water. Because atmospheric water vapor is produced by evaporation of seawater, it is slightly impoverished in these heavier molecules. Moreover, precipitation of rain makes the remaining vapor even more depleted. Atmospheric precipitation becomes progressively more deprived in $H_2{}^{18}O$ and $HD^{16}O$ with increasing latitude and with increasing distance from the ocean, over the continents. The difference between vapor and liquid will remain about 10‰ and so make the rain increasingly lighter in $\delta^{18}O$. This process, in which material is progressively removed from a system without subsequent re-equilibration, is called Rayleigh fractionation. The δD and $\delta^{18}O$ values in precipitation – and, therefore, in fresh waters – generally plot close to a straight line, known as the *global meteoric water line* (GMWL) [62]:

$$\delta D = 8\,\delta^{18}O + 10$$

The slope and intercept of the *local meteoric water line* (LMWL) for rain from a specific drainage basin can be different from the GMWL. For that reason, the *deuterium excess* (*d*) has been defined to describe these different water lines:

$$d = \delta D - 8\,\delta^{18}O$$

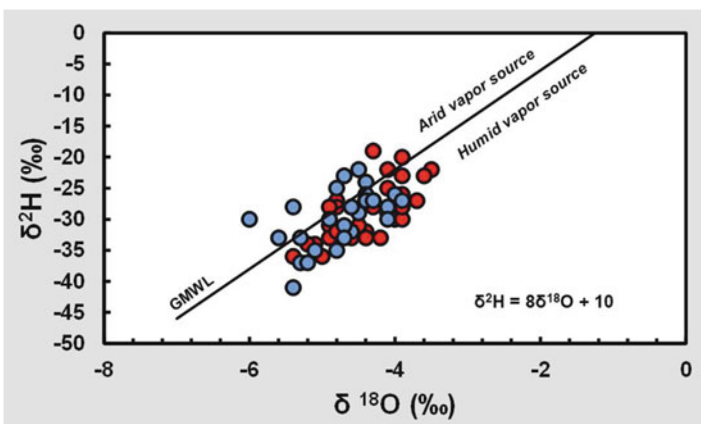
Evaporation from open water surfaces (e.g., lakes, reservoirs, ponds), or mixing with evaporated water, plots below the meteoric water line, usually with a slope between 2 and 5.

Stable isotopic data from the Paraná River was used to illustrate the variability verifiable in a large river with a wide-ranging floodplain, periodically subjected to inundation (Fig. 19). The Paraná is one of the world's large rivers which are significantly influenced by periodic climatic events (e.g., the ENSO), and stable isotopes have contributed to disclose a significant part of the intervening hydrological processes [63, 64].

Paraná's floodplain is profusely dissected by channels, oxbows, and shallow ponds, which store a large volume of water during the seasonal high flow stage for variable periods of time. During the low stage, the water stored for some time in the floodplain reenters the river's main stem, modifying to some extent its chemical composition. Figure 19 also shows that a large proportion of the samples plot below the GMWL, suggesting that there is a larger water contribution from humid vapor sources than from arid vapor sources.

Carbon has two stable isotopes, ^{13}C (1.1%) and ^{12}C (98.9%). Similarly to other stable isotopes, results are reported as $\delta^{13}C$. There are several possible natural sources for the carbon stable isotopes found in fresh waters: (a) the dissolution of calcite, aragonite, or dolomite, which delivers to fresh waters relatively heavy carbon isotopes; (b) in contrast, the oxidation of organic matter supplies relatively light isotopic carbon; and (c) the transport of $CO_{2\,(g)}$ from soil atmosphere, which also introduces relatively light carbon.

Carbon is an element that mediates among the hydrosphere, the lithosphere, the biosphere, and the atmosphere. The circulation through all these compartments triggers fractionation which can be advantageously used in the identification of sources and paths. In general, ^{13}C is used to identify sources of carbon and is particularly valuable for discriminating carbon derived from organic



Fresh Water Geochemistry: Overview, Fig. 19 Lower Paraná River stable isotopic data [63], separated in below the mean (*red circles*) and above the mean discharges. Notice that the majority of the points fall below the global mean water line (*GMWL*), suggesting a dominating tropical vapor source matter (i.e., isotopically light) from carbon derived from carbonate minerals (i.e., isotopically heavy) (e.g., [65]).

An example on the use of $\delta^{13}C$ in fresh waters was produced by Brunet et al. [66], studying the signature of different dissolved inorganic carbon (DIC) sources in the rivers of Patagonia. The study allowed distinguishing a set comprised by Patagonia's largest rivers (i.e., the Colorado, the Negro, the Chubut, and the Santa Cruz), which have large glacial lakes and/or reservoirs in their drainage basins and, therefore, exhibit high $\delta^{13}C_{DIC}$ (i.e., −3 to −7) due to significant CO_2 degassing to the atmosphere. The other set clusters smaller rivers (i.e., the Deseado, the Chico, the Coyle, and the Gallegos) where organic carbon oxidation appears to result in a negative $\delta^{13}C_{DIC}$ signature (i.e., −7 to −13) for the dissolved riverine carbon pool.

Nitrogen has two stable isotopes, ^{14}N and ^{15}N. Average abundance of ^{15}N in air (0.366%) is very constant (i.e., atmospheric N contains about 1 atom of the stable ^{15}N per 273 atoms of ^{14}N), and, hence, air is used as the standard ($\delta^{15}N$ = 0‰) for reporting d ^{15}N values. Fractionation occurs along the pathway of several biochemical reactions which include N; most terrestrial materials have $d^{15}N$ compositions between −20‰ and +30‰. Many plants fix nitrogen and organisms cycle this N into the soil. Additional sources of nitrogen include fertilizers produced from atmospheric N with compositions of 0 ± 3‰ and animal manure with $\delta^{15}N$ values generally in the range of +10‰ to +25‰ [67]. Rock sources of N are considered negligible. $\delta^{15}N$ has been used to identify sources – chiefly anthropogenic sources – of NO_3^- in natural waters, sometimes in combination with $\delta^{18}O$.

Sulfur has four stable isotopes: ^{32}S (95.02%), ^{33}S (0.75%), ^{34}S (4.21%), and ^{36}S (0.02%) [62]. Stable isotopic compositions are reported as ratios of $^{34}S/^{32}S$ in ‰ relative to the standard VCDT (Vienna Canyon Diablo Troilite). The terrestrial range spans from +50‰ to −50‰, with some infrequent values falling outside. The $\delta\ ^{34}S$ of the ocean is currently about +20‰, but has changed significantly in geological history. Variations in the $\delta^{34}S$ values are caused by two kinds of processes: (a) reduction of SO_4^{2-} to S^{2-} by bacteria which results in an increase in the ^{34}S of the residual SO_4^{2-} and (b) several types of exchange reactions which result in ^{34}S being concentrated in the compound with the highest oxidation state of S (e.g., [62, 68]). In the study of fresh waters, $\delta^{34}S$ can be used to distinguish different sources of sulfur, but the technique is not widely used in hydrochemical research.

There are many examples in the specialized literature on the use of stable isotopes in geochemical research (e.g., [68]), and the list of stable isotopes available for their use in geochemical applications is expanding, in coherence with the development of new, more evolved instrumentation. Lithium is a suitable example.

The large mass differences (~15%) of its two stable isotopes, ^{6}Li (7.5%) and ^{7}Li (92.5%), produce large isotopic fractionation in terrestrial systems, from −20‰ to +40‰, and appear to be strongly fractionated in low-temperature systems [69]. Like the other stable isotopes, variations in Li isotopes are reported as departures from an international standard and expressed in delta notation:

$$\delta^7\mathrm{Li}\ (‰) = \left[\left({}^7\mathrm{Li}/{}^6\mathrm{Li}\right)_{\mathrm{sample}}/\left({}^7\mathrm{Li}/{}^6\mathrm{Li}\right)_{\mathrm{L-SVEC}} - 1\right] 1000$$

The used standard is the US National Institute of Standards and Technology (NIST) highly purified Li_2CO_3 reference material NIST L-SVEC, which has a $^{7}Li/^{6}Li$ = 12.02 ± 0.03. Positive values of δ^7Li reflect heavier isotopic ratios, in keeping with other isotope systems [69]. Concerning natural waters, δ^7Li in lakes fluctuates between ~+32‰ and +13‰, between ~+38‰ and +5‰ in river waters, and between ~+40‰ and −4‰ in marine pore fluids. Seawater fluctuates narrowly between ~+32‰ and +30‰ [69]. A recent study in the Mackenzie River drainage basin showed that ^{7}Li is enriched in the dissolved load and that δ^7Li can fluctuate by 20‰ within the large river basin. The δ^7Li of the particulate load replicates that of the bedrock and ranges between −2‰ and +3‰ [70].

Radioactive isotopes: Tritium (^{3}H or T) is a radiogenic and radioactive isotope of H with a half-life of 12.4 year. It is produced naturally in the atmosphere by the interaction of cosmic rays with N and O. However, the most important source is the testing of thermonuclear weapons which took place between 1952 and 1969. It is used as a tracer to evaluate timescales for the mixing and flow of waters (i.e., in the upper layers of the ocean, in lakes, and in groundwater) because it is deemed to be relatively conservative geochemically, and it is ideally appropriate for studying processes that occur on a timescale of less than 100 years [62].

Like ^{3}H, radiocarbon (^{14}C) is generated in the atmosphere by the interaction of cosmic rays with N, and, also, it was supplied in large amounts during the atmospheric testing of nuclear weapons. However, ^{14}C has a much larger half-life (i.e., 5730 year) than ^{3}H, making it a convenient tool for dating waters as old as 50,000 year. The most important reaction between slow cosmic-ray neutrons and the nucleus of stable ^{14}N is

$$ {}^{1}{}_{0}n + {}^{14}{}_{7}\mathrm{N} \rightarrow {}^{14}{}_{6}\mathrm{C} + {}^{1}{}_{1}\mathrm{H} $$

In the equation, ${}^{1}{}_{0}n$ is the neutron and ${}^{1}{}_{1}$H is the proton that is emitted by the product nucleus. The ^{14}C produced in the atmosphere is carried down to the Earth's surface by atmospheric precipitation and is assimilated in the biomass or conveyed into lakes, oceans, and groundwater. Radiocarbon decays to ^{14}N, so once segregated from the atmosphere, the amount of ^{14}C decreases with time:

$$ \left({}^{14}\mathrm{C}\right)_t = \left({}^{14}\mathrm{C}\right)_0 e^{-kt} $$

The amount of ^{14}C present at time t is $(^{14}\mathrm{C})_t$. $(^{14}\mathrm{C})_0$ is the amount present at $t = 0$, and k is the decay constant. The half-life $t_{1/2}$ with k is expressed by the equation

$$ t_{1/2} = \ln 2/k $$

Besides its use as a tool to determine water age, ^{14}C can be used to approximate groundwater flow velocity [60].

The decay of U and Th to Pb is one of the most commonly used mineral dating techniques in the Earth Sciences. Various intermediate products in the decay pattern have been used in water studies. One of them, ^{222}Rn, is produced by the decay of ^{226}Ra, which is itself a decay product of ^{238}U. Radon is a gas and is transported by eddy diffusion from the source into the overlying water. Its radioactive isotope ^{222}Rn has a half-life of 3.8 days so that its concentration would decrease in accordance with the distance from the source; if the diffusion rate is slow, ^{222}Rn will be found close to the source; if it is fast, ^{222}Rn will be transported to greater distances from the source before it vanishes by decay.

In recent years, ^{222}Rn has been extensively used to study groundwater discharge, particularly in coastal regions, where fresh water flow to the ocean is identified as submarine groundwater discharge (SGD) (e.g., [71]), currently evaluated worldwide as a very significant source of water and dissolved material [72].

Radium has no stable – or nearly stable – isotopes, and, therefore, a standard atomic weight cannot be provided. Radium has 33 known isotopes from ^{202}Ra to ^{234}Ra, also products of the decay chain of ^{238}U (i.e., often referred to as the radium series). With a longer half-life than Rn (1600 year), ^{226}Ra is extensively used to identify areas with significant SGD (e.g., [73]), sometimes in conjunction with shorter-lived Ra isotopes (e.g., ^{223}Ra and ^{224}Ra) and, also, with ^{222}Rn (e.g., [74]).

Radiogenic isotopes: There is a group of elements that exhibit variations in their isotopic composition, which result from radioactive decay occurring within minerals over geologic timescales. Such isotopic variations supply natural fingerprints of rock-water interactions and have been used extensively in the investigation of weathering processes and hydrological dynamics. It must be added here that before the extended use in studies on weathering and hydrology, each of these isotopic systems was employed profusely in geochronology and petrology.

The most broadly employed isotopic systems in the Earth Sciences are Rb-Sr, U-Pb, and K-Ar. The latter is not directly relevant in most studies of rock-water interaction because Ar is a noble gas that may be mixed with atmospheric argon, thus limiting its usefulness in most weathering application.

$^{87}Sr/^{86}Sr$, $^{206}Pb/^{207}Pb$, $^{208}Pb/^{207}Pb$, $^{234}U/^{238}U$, $^{143}Nd/^{144}Nd$, and $^{187}Os/^{186}Os$ are among the isotopic systems that have been extensively used in weathering studies, and the pertinent scientific literature is remarkably abundant (e.g., [75] and references therein), probably approaching 400 by now. $^{87}Sr/^{86}Sr$ is undoubtedly the isotopic system most widely used as an isotopic tracer because of the large variability in isotopic composition and the importance attached to source tracing studies and to the understanding of the cycling of Ca, the analog element. These systems have been used, as well, to trace dissolved phases in fresh waters (i.e., rivers, springs, and small streams) [75].

Box 11 CHUR Model Ages and the Epsilon (ε) Notation

The CHUR (i.e., Chondritic Uniform Reservoir) model ages presuppose that the Earth's primeval mantle had the same isotopic composition as the average chondritic meteorite at the formation of the Earth, which is taken to be 4.6 Gy. For Nd isotopes, CHUR represents the composition of the bulk Earth. A model age calculated relative to CHUR, therefore, is the time in the geologic past at which the sample suite separated from the mantle reservoir and acquired a different Sm/Nd ratio.

Isotope ratios are only strictly comparable if the samples are of the same age. The epsilon value (ε) is a measure of the deviation of a sample or sample suite from the expected value in a uniform reservoir and may be used as a normalizing parameter for samples of different age.

$$\varepsilon_{Nd}(0) = \left\{\left[(^{143}Nd/^{144}Nd)_{sample\,(t)}/(^{143}Nd/^{144}Nd)_{CHUR\,(t)}\right] - 1\right\} 10000$$

This is called the epsilon notation whereby one epsilon unit represents a one part per 10,000 deviations from the CHUR composition.

An illustrative example [76] has been used to identify the likely material sources in the Paraná River by means of the isotopic signatures exhibited by $^{87}Sr/^{86}Sr$, $^{143}Nd/^{144}Nd$, and $\varepsilon_{Nd}(0)$ (Box 11). The data obtained for the Paraná's middle stretch is presented in Table 3 [76]. The corresponding isotopic data, along with data from the Uruguay River, are plotted in Fig. 20.

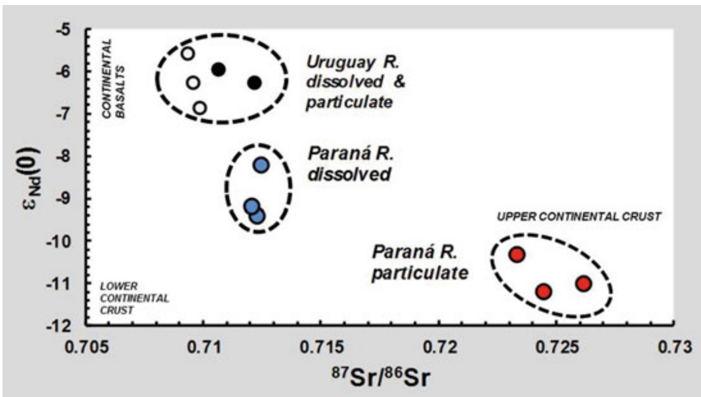
The Andean tributaries of the Paraná River (i.e., the Bermejo and Pilcomayo rivers) supply most of the sediment load (~60–70%) transported by Paraná's middle and lower stretches. However, most of the water (~63%) that reaches the outflow is supplied by the upper Paraná River [18], with headwaters in tropical Brazil. The region is significantly mantled by Jurassic-Cretaceous Paraná tholeiitic basalts. Figure 20 suggests that the difference in provenance for the particulate and dissolved phases is reflected by the isotopes, with the $^{87}Sr/^{86}Sr$ signature more radiogenic (i.e., upper crustal) in the TSS than in the dissolved fraction which shows, instead, a signal closer to that of continental basalts. In contrast, the Uruguay River [76] shows a significantly uniform signature for both particulate and dissolved phases, thus reflecting the predominance of continental tholeiitic basalts (i.e., Serra Geral Fm) as a prevailing source.

Mechanisms Controlling Fresh Water Geochemistry

In 1970, R.J. Gibbs [77] proposed a simple model for the mechanisms controlling world water chemistry. Gibbs' scheme included three major systems that would explain most of the variability accounted for by the Earth's surface waters. The mechanisms are atmospheric precipitation (i.e., Na^+ and Cl^- as major ions), rock (weathering) dominance (i.e., Ca^{2+} and HCO_3^- as major ions), and evaporation-fractional crystallization processes (i.e., again, Na^+ and Cl^- as major ions). The precipitation control is one end-member of a series, whereas rock dominance is the other end-member; precipitation-supplied water defines the dilution operating all along the series. The third major mechanism that controls

Fresh Water Geochemistry: Overview, Table 3 Dissolved and particulate Sr and Nd isotopes in the Paraná River, 420 km upstream the mouth

Phase	Date	$^{87}Sr/^{86}Sr$	±2σ	$^{143}Nd/^{144}Nd$	±2σ	ε_{Nd} (0)
diss.	Dec. 1993	0.712485	0.000009	0.512217	0.000009	−8.2
diss.	Dec. 1993	n.d.	n.d.	0.512018	0.000013	−12.1
diss.	Sept. 1994	0.712292	0.000006	0.512157	0.00001	−9.4
diss.	Sept. 1994	0.712056	0.000006	0.512165	0.000012	−9.2
part.	Dec. 1993	0.723338	0.00001	0.512108	0.000009	−10.3
part.	Dec. 1993	0.724472	0.000006	n.d.	n.d.	n.d.
part.	Sept. 1994	0.72617	0.000008	0.512061	0.000009	−11.2
part.	Sept. 1994	n.d.	n.d.	0.512073	0.000009	−11



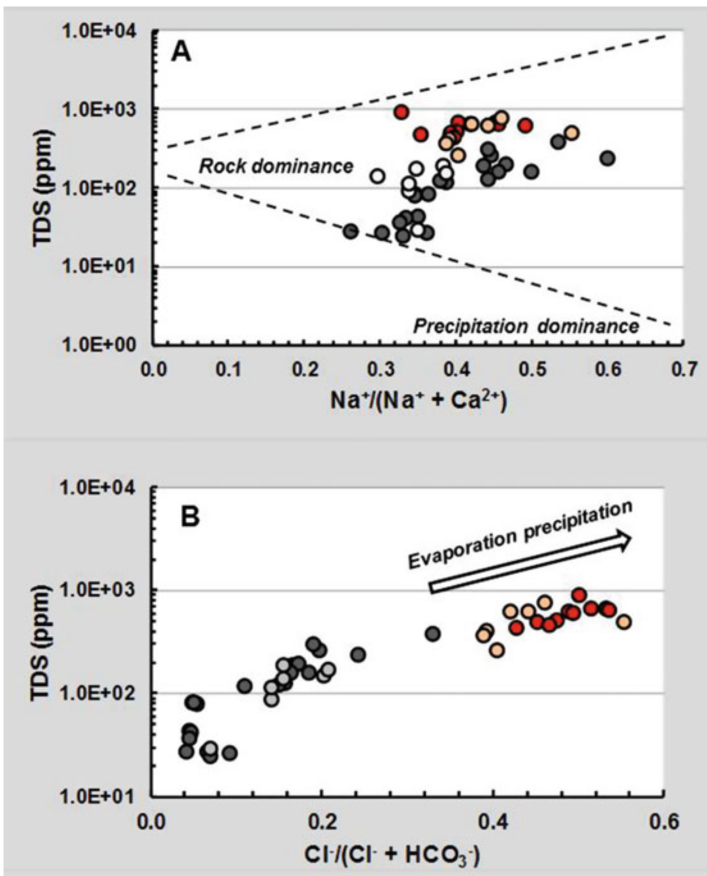
Fresh Water Geochemistry: Overview, Fig. 20 Radiogenic isotopes determined in dissolved and particulate phases from the Paraná and Uruguay rivers. $^{87}Sr/^{86}Sr$ and $\varepsilon_{Nd}(0)$ suggest differing sources for both phases in the Paraná, whereas they indicate a common source (i.e., continental basalts) in the case of the Uruguay River (Paraná and Uruguay rivers' basic data from [76])

the chemical composition of the Earth's surface waters is identified as the evaporation-crystallization process, which is also a continuum extending from the Ca-rich, medium salinity (i.e., fresh water) end-member cluster to the contrasting, Na-rich (i.e., saline waters and brines) end-member in the other extreme of the series. Gibbs collated these three mechanisms in a widely known boomerang-shaped scheme, which was (and still is) widely cited in scientific journals and reproduced in books (e.g., [33]).

Gibbs [77] developed the model based on over 130 water analyses of major lakes and rivers from around the world. Some departures from the original scheme were pointed out by several researchers, and, hence, a modified version of the original model was published over 20 years later [78]. Gibbs' approach basically focused on two dominant factors: rock dominance (i.e., a proxy for weathering) and climate, fluctuating between two extremes (i.e., water deficit and overabundance). The other natural factors often mentioned in the literature as intervening in the control of water chemistry (i.e., biota and relief) frequently show some degree of association with climate and geology. It follows, however, that biota and relief always play a role in defining water chemistry although it may be subdued or amplified in specific scenarios. Milliman and Farnsworth [9] have probed into the control of TDS at the global level and have arrived at the conclusion that bedrock lithology is the governing factor in the character and quantity of total dissolved delivery, but climate (i.e., mostly atmospheric precipitation) controls the rate of chemical weathering and, hence, water chemistry.

Following Gibbs' scheme, Fig. 21 shows the spreading of fresh water samples collected in Patagonia's Negro and Colorado rivers and their main upper tributaries [79]. The Negro River is Ca-dominated and, therefore, appears more controlled by weathering than the Colorado, which is more Na-dominated and hence, placed somewhere along the evaporation-precipitation series (Fig. 21a). Moreover, the Negro River and main tributaries are clearly dominated by alkalinity, most likely the product of silicate weathering and dissolution of carbonates (Fig. 21b) [79].On the other hand, the Colorado and tributaries show an incipient trend toward the Cl-rich extreme (Fig. 21b).

As rivers flow toward their respective mouths, the character and concentration of their TDS load change at often unpredictable rates. The reason for a particular chemical change is, sometimes, obvious (e.g., the main stem receives one or more tributaries with a different chemical signature). In other instances, the cause for the chemical change is not evident and it may obey to other factors. Figure 22 shows the chemical change in two different rivers, represented by the $Na^+/(Na^+ + Ca^{2+})$ ratio as a function of their distances to their corresponding mouths. Patagonia's Colorado and Negro rivers were sampled twice, in the downriver direction. The first sampling took place in May 1972 and the second in January 1973. Both neighboring rivers are almost parallel and flow eastward across northern Patagonia, from the Andes down to the SW Atlantic Ocean. Figure 22a shows for the Colorado River a progressive Ca^{2+} enrichment toward the mouth determined during the May 1972 sampling. The other downflow sampling (e.g., January 1973) did not result in a regression line significantly different from zero.

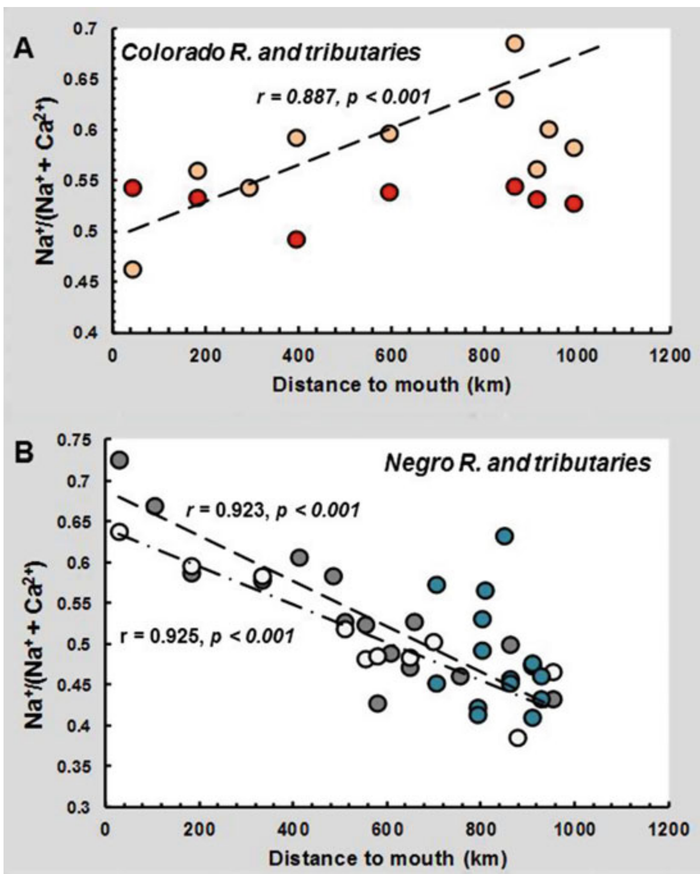


Fresh Water Geochemistry: Overview, Fig. 21 Colorado River and tributaries (*red* and *pink circles*, respectively) and Negro River and tributaries (*dark grey* and *light grey circles*, respectively) plotted in a Gibbs' diagram [77, 78] (Argentina's Patagonia). Data show clear rock dominance for the Negro and tributaries, and a tendency toward the evaporation-precipitation realm in the case of the Colorado and its main tributaries

The contribution of increased subsurface flow, probably triggered by rainfall, is most likely the factor determining the Ca^{2+} concentration increase registered during the first field excursion.

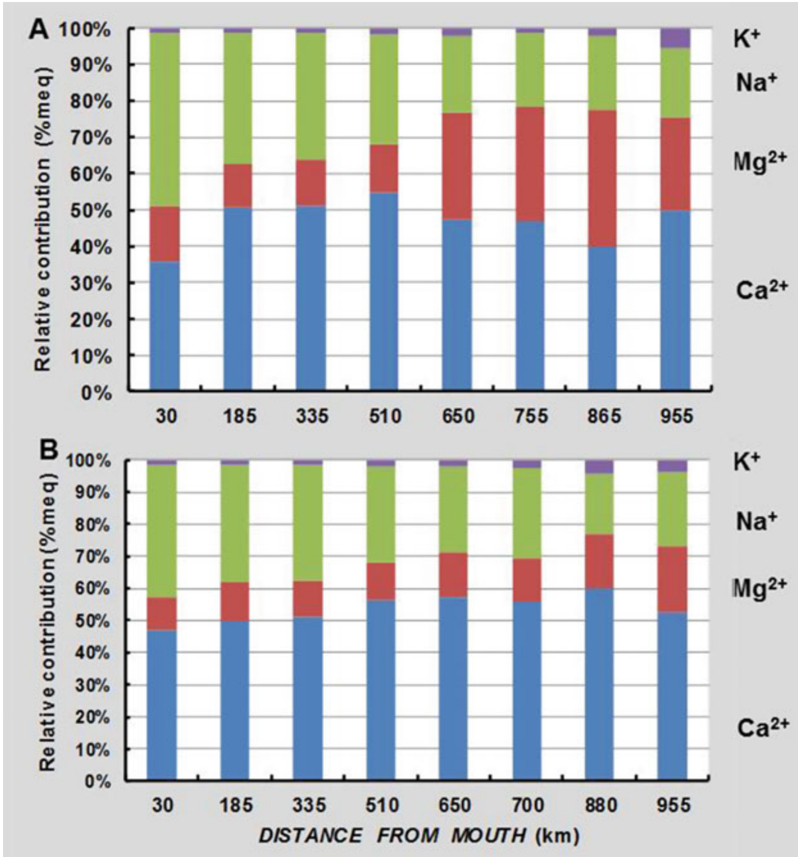
In the Negro River, there is a contrasting setting, with a marked relative increase in Na^{+} concentration, apparently similar in both sampling exercises (Fig. 22b). In this case, the reason for the change rests most probably in outcropping salty groundwater, in sufficient volume and concentration to change the river's chemical character. The ratio value increases ~21% in 400 km of downriver flow.

Although the mechanism controlling fresh water chemistry in the Negro River appears relatively coherent in both sampling instances (Fig. 21a, b), there are aspects which Gibbs' model does not consider and may be significant at the time of analyzing the processes that control water chemistry. Figure 23 shows, again, the variability of the main ions as a function of the distance to the river's mouth. In this opportunity, the data set has been expanded, considering the four main ions ($\Sigma Z^{+} = Na^{+} + K^{+} + 2Mg^{2+} + 2Ca^{2+}$, in % meq l^{-1}). The comparison of the two bar graphs shows that in the upper stretches, Mg^{2+} was relatively more abundant in the first sampling trip (May 1972) than in the second (May 1973). Both alkaline earths, Ca^{2+} and Mg^{2+}, showed a decreasing trend toward the mouth, whereas Na^{+} accounted for a larger relative proportion of ΣZ^{+} in both downriver sampling exercises. The role of K^{+} was equally subordinate in both instances, probably due to the consistent role of adsorption phenomena (Fig. 23). Once more, this approach suggests the contribution of groundwater, but this possibility, as well as the occurrence of mineral precipitation or adsorption onto colloids, should be validated through the use of modeling and isotopes.



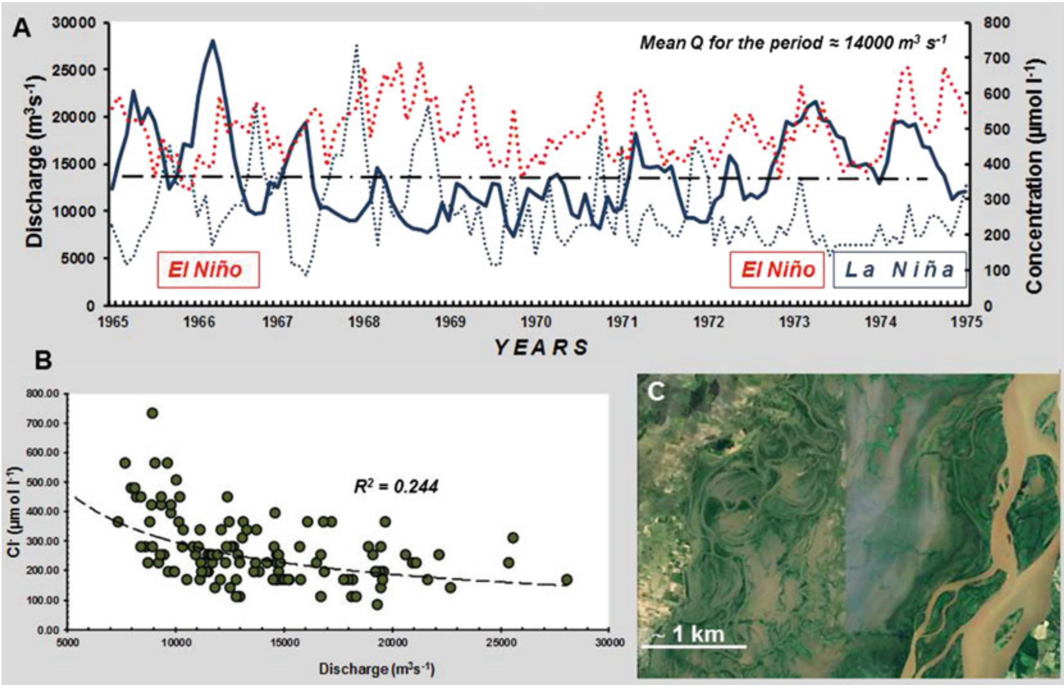
Fresh Water Geochemistry: Overview, Fig. 22 Divergent evolution of $Na^{+}/(Na^{+} + Ca^{2+})$ as a function of the distance to the mouth in (**a**) the Colorado River (*red circles*) and main tributaries (*pink circles*), and (**b**) in the Negro River (*dark gray circles*), main tributaries (*light grey circles*), and Andean tributaries (*blue circles*); Argentine Patagonia (**b**). The decreasing significance of Ca^{2+} in the Negro system is interpreted as the growing contribution of groundwater [79]

Fresh Water Geochemistry: Overview, Fig. 23 The bar graphs show the relative main cations contributions (ΣZ^+) in the Negro River (Argentine Patagonia) – during low waters – as a function of the distance to the mouth. (**a**) In May 1972 (mean monthly Q, ~235 $m^3\ s^{-1}$) and (**b**) in September 1973 (mean monthly Q, 395 $m^3\ s^{-1}$). The Negro River long-term mean annual discharge is ~860 $m^3\ s^{-1}$ [18, 79]



Another aspect that may distort the picture depicting the factors controlling water chemistry in fluvial systems is the occurrence of extensive, gently sloped floodplains, totally carved by secondary channels, oxbows, and ponds. Such is the case of the middle and lower Paraná River (also of the Amazon, the Orinoco, and other large world rivers), which have a long (~900 km long) and wide (~40–70 km) floodplain that holds a myriad of relatively shallow water bodies, framed by abundant riparian vegetation (i.e., the well-known *varzea* in the Amazon River). The inundation or drainage of this extensive territory is driven by the Paraná's seasonal hydrological stage. Superimposed on this harmonic discharge succession, there are ENSO-triggered high- and low-water events that occur with 3–5-year periodicity (e.g., [18, 19]). These water level permanent fluctuations determine over-the-bank flooding in the valley (particularly during El Niño floods). In contrast, drainage occurs during low waters, frequently in coherence with La Niña events. The significant volume of water (probably as much as 10% of the Paraná's annual discharge) thus stored in the valley for variable periods of time is – while isolated in the valley – subjected to several physical/biogeochemical or chemical processes (e.g., evapotranspiration, ion exchange, adsorption-desorption, organic matter mineralization, mineral dissolution-precipitation, etc.) which modify the chemical character of the accumulated water mass. Thus, during low stages (i.e., below-the-mean water gage level), the chemically modified water flows out of the flood valley and reenters the main stem, altering to some extent the chemical signature of the water volume delivered by the headwaters.

Figure 24 shows some elements to assist in the understanding of the above described process. Figure 24a shows the discharge variability in the



Fresh Water Geochemistry: Overview, Fig. 24 (**a**) Variability of discharge, Cl^- (*blue dotted line*) and alkalinity (*red dotted line*) concentrations in the middle Paraná River for the period (1965–1975); the occurrence of El Niño and La Niña events are indicated. (**b**) The same data plotted in a scatter diagram of Cl^- against discharge showing a poorly correlated nonlinear relationship. (**c**) Satellite image of the middle Paraná River, ~600 km upstream the mouth, showing the main channel (*right*) and the profusion of lotic and lentic environments in the riparian zone

middle Paraná River (~600 km upstream the mouth) for a 10-year period (1965–1975) along with two chemical variables: Cl^- (i.e., a conservative element) and alkalinity. Both show a poor coherency with water discharge, as demonstrated in Fig. 24b, where discharge accounts only for 24% of the Cl^- variance in a nonlinear relationship. Figure 24c shows a portion of a Google map image, where the relative importance of small lotic and lentic water bodies in the valley can be evaluated.

Acknowledging the significance of Gibbs' typology but, at the same time, emphasizing that his scheme of world water chemistry was leaving unexplained ~20% of the Earth's waters, M. Meybeck [80] tentatively proposed a more elaborate model, seeking to include rivers that were left out in Gibbs' proposal (Table 4). In must be kept in mind that the relative significance of different water types depends on the global representativeness of the used database, and there are still many small, remote, or insufficiently analyzed rivers which are not represented in global data sets of pristine or sub pristine rivers and tributaries. Although obvious, it must be said that human pollution has complicated further the appearance of a conclusive interpretation on the mechanisms controlling the Earth's fresh water resources.

Future Directions: Modeling, Instrumentation, and Sustainability

Modeling: A wide number of applications in low-temperature geochemistry are found in the literature since the introduction of computer modeling in the Earth Sciences. The scope of such endeavor in the field of low-temperature geochemistry has embraced from research in the essential processes of water-rock interactions to the ameliorated management of industrial and hazardous wastes (e.g.,

Fresh Water Geochemistry: Overview, Table 4 Occurrence of water types: dominant chemical sources and control factors (Modified from [21])

Water type	SZ^+ (meq l^{-1})	Rainfall Input (%)	NaCl (%) *(nonrain)*	$CaSO_4$ (%) *(nonrain)*	$CaCO_3$ (%)	$MgCO_3$ (%)	Na_2CO_3 (%)	$MgSO_4$ (%)	Na_2SO_4 + $MgCl_2$ (%)	Dominant control factors
Extremely dilute	< 0.185	20			78	1		1		*Vegetation/rain/ silicates*
Very dilute	0.185 – 0.375	12			87	0.5		0.5		*Rain/silicates*
Dilute	0.375 – 0.75	6			84	1	8	1		*Silicates/rain*
Medium dilute	0.75 – 1.5	4			85	1	8.5	1	0.5	*Silicates/rain*
Medium mineralized	1.5 – 3.0	2			94	2.5	0.5	0.5	0.5	*Carbonates/silicates*
Mineralized	3.0 – 6.0		4.5	2	84	2	0.5	0.5	6.5	*Carbonates*
Highly mineralized	6.0 – 12.0		15	24	42	2	1.5	4.5	11	*Evaporites/pyrite/ carbonates*
Subsaline	12.0 – 24.0		45.5	26		0.5	8	6.5	13.5	*Evaporation/ evaporites*
Saline	> 24.0		48.5	25				6	20.5	*Evaporation/ evaporites/pyrite*

[20]). In such environments, generally understood as those in the temperature range of 0–100 °C and close to atmospheric pressure, intricate hydro-biogeochemical reactions join together in an assortment of interconnected processes that affect nature and human beings in a two-way avenue. Geochemical models are sophisticated tools that allow describing multicomponent, multiphase chemical reactions in a sufficiently transparent mode so as to reveal the main driving forces. The main processes that they must deal with are mineral dissolution and precipitation, aqueous inorganic and organic speciation and complexation, ion exchange, solute adsorption and desorption, redox processes, reaction during fluid flow, etc.

A comprehensive description of all the computer codes available to assist in the undertaking of problem-solving in fresh water geochemistry is beyond the purpose of this chapter. The treatment is limited to mentioning the importance of software packages developed, for example, by the USGS [81], which offers a wide variety of tools, particularly in the field of groundwater applications, as aids in the interpretation of water quality data. Worth mentioning are SOLMINEQ.GW, PHREEQC, and WATEQ4F all of which can be downloaded free of charge at the USGS web page. Also popular among low-temperature geochemists is the set of five programs known as The Geochemist's Workbench™. This code performs speciation, mass transfer, reaction-path calculations, isotopic calculations, sorption and independent redox calculations, and temperature dependence for 0–300 °C.

The US Environmental Protection Agency (US EPA) supports MINTEQA2 [82], which is an equilibrium speciation model that can be used in the calculation of the equilibrium composition of dilute aqueous solutions in natural aqueous systems or in the laboratory. The model, which has a Windows version, is convenient for calculating the equilibrium mass distribution among dissolved species, adsorbed species, and multiple solid phases under a variety of conditions, including a gas phase with constant partial pressures.

D.K. Nordstrom [39] has analyzed in detail the modeling of low-temperature geochemical processes, pointing to advantages and drawbacks of the available computer codes, most of them in the public domain. Future efforts should be aimed toward developing normalized test cases, confronting a wide variety of processes against which code performances can be matched and verified. Other fields open for future improvement are more evolved comparisons between analytical and computed speciation to attain accuracy estimates of aqueous speciation, the development of routine techniques for estimating uncertainties in model calculations, and more exhaustive analyses of fine-grained minerals which are reactive phases in geochemical systems [39].

Instrumentation: Analytical instrumentation is another field offering possibilities of future expansion, which ultimately would result in increased analytical capacity, and more precise and exact geochemical data. Inductively coupled plasma-mass spectrometry (ICP-MS) is a proper example of a technique which, during the last decade, has enlarged significantly the analytical capability of many geochemical laboratories around the world. More data is being reported in the ever-increasing publication rate of specialized books and scientific journals.

Mass spectrometry is another technique that offers great advantages in the forthcoming use of isotopes in fresh water geochemistry. Radiogenic, radioactive, and stable isotopes have proven to be an important and powerful tool in the investigation of many aspects of hydrology and weathering. Altogether, the use of Nd, Hf, and Os isotopes in weathering and hydrology can still grow, and much additional research will be needed to expand the knowledge on the behavior of these systems and, furthermore, to establish their usefulness in routine investigations. Increasingly, the combined use of several isotopic systems will enlarge our comprehension in a number of hydrological and geochemical applications.

Fresh water biogeochemistry is another field which has gained from new and more evolved instrumentation. Organic geochemistry facilities provide capability to develop scientific understanding of organic compounds in soils, sediments, and waters. Investigations cover themes such as organic metabolic paths in fresh water bodies, pollution studies and environmental forensics, terrestrial and marine paleoenvironmental analysis, climate change, and several other similar fields. In terms of instrumentation, different types of chromatographic analysis coupled with mass spectrometry have proved to be powerful tools – and will continue to do so – in probing into the complex realm of biogenic markers.

Sustainability: Water is a limited and irreplaceable resource that is central to human well-being and to all life on Earth. It is only renewable if understood as a thoughtfully administered vital resource. More than 1.7 billion people live in river basins where supply reduction through use surpasses natural recharge, a tendency that – according to the UN – will see two-thirds of the world's population living in water-stressed countries by 2025. Water can pose a serious challenge to sustainable development. It is decisive for socioeconomic growth and healthy ecosystems and for human survival itself. In a more focused perspective, it is vital for stopping the global problem of disease and improving the health, productivity, and welfare of populations. If properly managed, water can play a crucial role in supporting the resilience of social, economic, and environmental systems, rapidly exposed to sudden and unpredictable changes. Water also plays a vital role in the adaptation to climate change, operating as the crucial link between the climate system, the human society, and the environment.

Through scientific knowledge of a resource, the main stepping stone to its sustainable use may be reached. This is where and when fresh water geochemistry will continue to prove its value, as an indispensable aid to secure safe fresh water for future generations. More and better scientific resources, worldwide, will undoubtedly result in a more accurate map of the road to sustainability.

Bibliography

Primary Literature

1. Montanarella L, Panagos P (2015) Policy relevance of Critical Zone Science. Land Use Policy 49:86–91
2. White WM (2017) Geochemistry. In: White WM (ed) Encyclopedia of geochemistry. Springer, Heidelberg, pp 1–10
3. Reinhardt C (2008) Chemical sciences in the 20th century: bridging boundaries. Wiley, New York
4. Clarke FW (1908) Data of geochemistry. Bulletin, vol 330. US Geological Survey, Washington, DC
5. Clarke FW (1914) Water analyses from the laboratory of the United States Geological Survey, Water Supply Paper 364. US Geological Survey, Washington, DC
6. Field J, Little D (2009) Regolith and biota. In: Scott KM, Pain CF (eds) Regolith science. CSIRO Publishing/Springer, Collingwood/Dordrecht, pp 175–217
7. Allen PA (1997) Earth surface processes. Blackwell Science, Oxford
8. Rankama K, Sahama TG (1950) Geochemistry. University of Chicago Press, Chicago
9. Milliman JD, Farnsworth KL (2011) River discharge to the coastal ocean. A global synthesis. Cambridge University Press, Cambridge
10. Shiklomanov IA, Rodda JC (eds) (2003) World water resources at the beginning of the 21st century. Cambridge University Press, Cambridge
11. Garrels RM, Mackenzie FT (1971) Evolution of sedimentary rocks. W. W. Norton, New York
12. Goldich SS (1938) A study in rock weathering. J Geol 46:17–58
13. Bowen NL (1928) Evolution of the igneous rocks. Princeton University Press, Princeton
14. McQueen KG (2009) Regolith geochemistry. In: Scot KM, Pain CF (eds) Regolith science. CSIRO Publishing/Springer, Collingwood/Dordrecht, pp 175–217
15. Stumm W, Morgan JJ (1996) Aquatic chemistry, 3rd edn. Wiley-Interscience, New York
16. Langmuir D (1997) Aqueous environmental geochemistry. Prentice Hall, Upper Saddle River
17. Li Y-H (2000) A compendium of geochemistry. From solar nebula to the human brain. Princeton University Press, Princeton
18. Pasquini AI, Depetris PJ (2007) Discharge trends and flow dynamics of South American rivers draining the southern Atlantic seaboard: an overview. J Hydrol 333:385–399
19. Depetris PJ, Pasquini AI (2008) Riverine flow and lake level variability in southern South America. EOS Trans Am Geophys Union 89(28):254–255
20. Drever JI (1997) The geochemistry of natural waters, 3rd edn. Prentice Hall, Upper Saddle River
21. Gaillardet J, Viers J, Dupré B (2005) Trace elements in river waters. In: Drever JI (ed) Surface and ground water, weathering, and soils. Elsevier, Amsterdam, pp 225–272
22. Potter PE, Maynard JB, Depetris PJ (2005) Mud & mudstones. Introduction and overview. Springer, Heidelberg
23. Bland W, Rolls D (1998) Weathering. An introduction to the scientific principles. Arnold, London
24. Depetris PJ, Pasquini AI, Lecomte KL (2014) Weathering and the riverine denudation of continents. Springer, Dordrecht
25. Fedo CM, Nesbitt HW, Young GM (1995) Unraveling the effects of potassium metasomatism in sedimentary rocks and paleosols, with implications for paleoweathering conditions and provenance. Geology 23:921–924

26. Depetris PJ, Pasquini AI (2007) The geochemistry of the Paraná River: an overview. In: Iriondo MH, Paggi JC, Parma MJ (eds) The middle Paraná River: limnology of a subtropical wetland. Springer, Berlin
27. Harnois L (1988) The CIW index: a new chemical index of weathering. Sediment Geol 55(3–4):319–322
28. Hamadan J, Burnham CP (1996) The contribution of nutrients from parent material in three deeply weathered soils of peninsula Malaysia. Geoderma 74:219–233
29. Parker A (1970) An index for weathering of silicate rocks. Geol Mag 107:501–505
30. Babechuk MG, Widdowson M, Kamber BS (2014) Quantifying chemical weathering intensity and trace element release from two contrasting basalt profiles, Deccan Traps, India. Chem Geol 363:56–75
31. Tardy Y (1971) Characterization of the principal weathering types by the geochemistry of waters from some European and African crystalline massifs. Chem Geol 7:253–271
32. Boeglin JL, Probst JL (1998) Physical and chemical weathering rates and CO_2 consumption in tropical lateritic environment: the upper Niger basin. Chem Geol 148:137–156
33. Faure G (1991) Principles and applications of geochemistry. Prentice Hall, Upper Saddle River
34. Krauskopf KB, Bird DK (1995) Introduction to geochemistry, 3rd edn. McGraw-Hill, New York
35. White AF (2005) Natural weathering rates of silicate rocks. In: Drever JI (ed) Surface and ground water, weathering, and soils. Elsevier, Amsterdam, pp 133–168
36. Velde B (1992) Introduction to clay minerals. Chemistry, origins, uses and environmental significance. Springer, Heidelberg
37. Meunier A (2005) Clays. Springer, Heidelberg
38. Sayler FL, Mangelsdorf PC (1979) Cation-exchange characteristics of Amazon River suspended sediment and its reaction with seawater. Geochim Cosmochim Acta 43(5):767–779
39. Nordstrom DK (2005) Modeling low-temperature geochemical processes. In: Drever JI (ed) Surface and ground water, weathering, and soils. Elsevier, Amsterdam, pp 37–72
40. Dali-yousef M, Ouddane B, Derriche Z (2006) Adsorption of zinc on natural sediments of Tafna River (Algeria). J Hazard Mater 137(3):1263–1270
41. Ito A, Otake T, Shin K-C, Ariffin KS, Yeoh F-Y, Sato T (2017) Geochemical signatures and processes in a stream contaminated by heavy mineral processing near Ipoh city, Malaysia. Appl Geochem 82:89–101
42. Szynkiewicz A, Borrok DM (2016) Isotope variations of dissolved Zn in the Rio Grande watershed, USA: the role of adsorption on Zn isotope composition. Earth Planet Sci Lett 433:293–302
43. Hood DW (ed) (1970) Symposium on organic matter in natural waters. University of Alaska, Institute of Marine Sciences, College
44. Bolin B, Degens ET, Kempe S, Ketner P (1979) The global carbon cycle. SCOPE 13. Wiley, Chichester
45. Garrels RM, Mackenzie FT, Hunt C (1975) Chemical cycles and the global environment. Kaufmann, Los Altos
46. Ludwig W, Probst JL, Kempe S (1996) Predicting the oceanic input of organic carbon by continental erosion. Glob Biogeochem Cycles 10:23–41
47. Perdue EM, Ritchie JD (2005) Dissolved organic matter in freshwaters. In: Drever JI (ed) Surface and ground water, weathering, and soils. Elsevier, Amsterdam, pp 273–318
48. Depetris PJ, Kempe S (1993) Carbon dynamics and sources in the Paraná River. Limnol Oceanogr 38(2):382–395
49. (2007) http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr_topic1.pdf
50. Ittekkot V, Laane RWPM (1991) Fate of riverine particulate matter. In: Degens ET et al (eds) Biogeochemistry of major world rivers. SCOPE, vol 42. Wiley, New York, pp 233–243
51. Onstad GD, Canfield DE, Quay PD, Hedges JI (2000) Sources of particulate organic matter in rivers from continental USA: lignin phenol and stable carbon isotope compositions. Geochim Cosmochim Acta 64(2):3539–3546
52. Boyer EW, Howarth R (eds) (2002) The nitrogen cycle at regional to global scales. Springer, Heidelberg
53. Redfield AC, Ketchum BH, Richards FA (1963) The influence of organisms on the composition of seawater. In: Hikk MN (ed) The sea, vol 2. Wiley, New York
54. Díaz M, Pedrozo F, Reynolds C, Temporetti P (2007) Chemical composition and the nitrogen-regulated trophic state of Patagonian lakes. Limnologica 37:17–27
55. Depetris PJ, Gaiero DM, Probst JL, Hartmann J, Kempe S (2005) Biogeochemical output and typology of rivers draining Patagonia's Atlantic seaboard. J Coast Res 21(4):835–844
56. White WM (2017) Trace elements. In: White WM (ed) Encyclopedia of geochemistry. Springer, Heidelberg, pp 1–2
57. Pasquini AI, Depetris PJ (2012) Hydrochemical considerations and heavy metal variability in the middle Paraná River. Environ Earth Sci 65:525–534
58. Canil D (2017) Transition elements. In: White WM (ed) Encyclopedia of geochemistry. Springer, Heidelberg, pp 1–4
59. McLennan SM (1989) Rare earth elements in sedimentary rocks: influence of provenance and sedimentary processes. In: Lipin BR, McKay GA (eds) Geochemistry and mineralogy of rare earth elements. Mineralogical Society of America, Washington, DC, pp 169–200
60. Faure G (1986) Principles of isotope geology, 2nd edn. Wiley, New York
61. White WM (2015) Isotope geochemistry. Wiley-Blackwell, New York
62. Kendall C, Doctor DH (2005) Stable isotope applications in hydrologic studies. In: Drever JI (ed) Surface and ground water, weathering, and soils. Elsevier, Amsterdam, pp 319–364
63. Panarello HO, Dapeña C (2009) Large scale meteorological phenomena, ENSO and ITCZ, define the Paraná River isotope composition. J Hydrol 365:105–112

64. Pasquini AI, Depetris PJ (2010) ENSO-triggered exceptional flooding in the Paraná River: where is the excess water coming from? J Hydrol 383:186–194
65. Mook WG (2005) Introduction to isotope hydrology: stable and radioactive isotopes of hydrogen, carbon, and oxygen. CRC Press, New York
66. Brunet F, Gaiero DM, Probst JL, Depetris PJ, Gauthier Lafaye F, Stille P (2005) $\delta^{13}C$ tracing of dissolved inorganic carbon sources in Patagonian rivers (Argentina). Hydrol Process 19:3321–3344
67. Kendall C (1998) Tracing nitrogen and cycling in catchments. In: Kendall C, McDonell JJ (eds) Isotope tracers in catchment hydrology. Elsevier, Amsterdam, pp 519–576
68. Hoefs J (2009) Stable isotope geochemistry. Springer, Berlin/Heidelberg
69. Tang YJ, Hong-Fu Z, Ji-Feng Y (2007) Review of the lithium isotope system as a geochemical tracer. Int Geol Rev 49:374–388
70. Millot R, Vigier N, Gaillardet J (2010) Behaviour of lithium and its isotopes during weathering in the Mackenzie Basin, Canada. Geochim Cosmochim Acta 74:3897–3912
71. Burnett WC, Dulaiova H (2006) Radon as a tracer of submarine groundwater discharge into a boat basin in Donnalucata, Sicily. Cont Shelf Res 26(7):862–873
72. Kwon EY, Kim G, Primeau F, Moore WS, Cho HM, DeVries T, Sarmiento JL, Charette MA, Cho YK (2014) Global estimate of submarine groundwater discharge based on an observationally constrained radium isotope model. Geophys Res Lett 41. https://doi.org/10.1002/2014GL061574
73. Moore WS (2003) Sources and fluxes of submarine groundwater discharge delineated by radium isotopes. Biogeochemistry 66:75–93
74. Schlüter M (2002) Fluid flow in continental margin sediments. In: Wefer G, Billet D, Hebbeln D, Jorgensen BB, Schlüter M, Van Weering T (eds) Ocean margin system. Springer, Heidelberg, pp 205–217
75. Blum JD, Erel Y (2005) Radiogenic isotopes in weathering and hydrology. In: Drever JI (ed) Surface and ground water, weathering, and soils. Elsevier, Amsterdam, pp 365–392
76. Henry F, Probst JL, Thouron D, Depetris PJ, Garçon V (1996) Nd-Sr isotopic compositions of dissolved and particulate material transported by the Paraná and Uruguay rivers during high (December 1993) and low (September 1994) water periods. Sci Géol Bull 49:89–100
77. Gibbs RJ (1970) Mechanisms controlling world water chemistry. Science 170:1088–1090
78. Gibbs RJ (1992) A reply to the comment of Eilers et al. Limnol Oceanogr 37(6):1338–1339
79. Depetris PJ (1980) Hydrochemical aspects of the Negro River, Patagonia, Argentina. Earth Surf Proc 5:181–186
80. Meybeck M (2005) Global occurrence of major elements in rivers. In: Drever JI (ed) Surface and ground water, weathering, and soils. Elsevier, Amsterdam, pp 207–223
81. https://water.usgs.gov/software/
82. https://www.epa.gov/exposure-assessment-models/minteqa2

Books and Reviews

Albarède F (2003) Geochemistry. An introduction. Cambridge University Press, Cambridge
Allen PA (1997) Earth surface processes. Blackwell Science, Oxford
Barceló D, Kostianoy AG (eds) (1980) The handbook of environmental chemistry. Springer, Heidelberg
Berkowitz B, Dror I, Yaron B (2014) Contaminant geochemistry. Springer, Heidelberg
Boyd CE (2015) Water quality. Springer, Heidelberg
Christensen ER, Li A (2014) Physical and chemical processes in the aquatic environment. Wiley, New York
Clark I (2015) Groundwater geochemistry and isotopes. CRC Press, New York
Killops SD, Killops VJ (1993) An introduction to organic geochemistry. Longman S & T, Burnt Mill
Krauskopf KB, Bird DK (1995) Introduction to geochemistry, 3rd edn. McGraw-Hill International Editions, New York
Merkel BJ, Nordstrom DK, Planer-Friedrich B (eds) (2008) Groundwater geochemistry. Springer, Heidelberg
Osadchyy V, Nabyvanets B, Linnik P, Osadcha N, Nabyvanets Y (2016) Processes determining surface water chemistry. Springer, Heidelberg
Otonello G (1997) Principles of geochemistry. Columbia University Press, New York
Stumm W (1992) Chemistry of the solid-water interface: processes at the mineral-water and particle-water interface in natural systems. Wiley, New York
Van Loon G, Duffy SJ (2005) Environmental chemistry: a global perspective. Oxford University Press, Oxford
White WM (2013) Geochemistry. Wiley-Blackwell, New York
White WM (2015) Isotope geochemistry. Wiley, New York

Groundwater Impacts of Radioactive Wastes and Associated Environmental Modeling Assessment

Rui Ma[1], Chunmiao Zheng[2,3] and Chongxuan Liu[3,4]
[1]School of Environmental Studies, China University of Geosciences, Wuhan, China
[2]Department of Geological Sciences, University of Alabama, Tuscaloosa, AL, USA
[3]Southern University of Sciences and Technology, Shenzheng, China
[4]Pacific Northwest National Laboratory, Richland, WA, USA

Article Outline

Glossary

Absorption The process in which a dissolved substance is incorporated into the interior of a solid grain.

Adsorption The adhesion of a chemical species onto the solid surface.

Contaminant fate and transport The ultimate state of contaminants and the processes by which the contaminants migrate through the subsurface.

Groundwater Water that exists in liquid form beneath the land surface, filling the cracks, voids, and pore spaces in earth materials. The subsurface strata that store and transmit groundwater are referred to as aquifers.

Mass transfer The interaction and exchange of solutes in mobile state with those in immobile state through either physical or chemical processes.

Radioactive waste A waste product that contains radioactive material. The majority of radioactive waste is "low-level" waste, which has low levels of radioactivity per unit of mass or volume. Depending on the type and nature of radioactive wastes, it could take hours to thousands of years to diminish their radioactivity.

Reactive transport model The mathematical model that couples hydrogeological, geochemical, and biological processes to simulate and predict the contaminant fate and transport in the subsurface.

Sorption/desorption Sorption is a general term used to describe both adsorption and absorption by which a dissolved substance is attached to the surface or incorporated into the interior of a solid grain. The reverse process from the sorbed phase to the dissolved phase is referred to as desorption.

Definition of the Subject

Ever since the dawn of the nuclear age, especially since the 1970s with the work related to exploration of Yucca Mountain as a potential nuclear waste repository [1] and remediation of former nuclear fuel processing sites such as Hanford [2], there have been significant public concerns over groundwater impacts of radioactive wastes. This is because groundwater is a vital water resource with tremendous values to public water supplies and ecological lives, and because groundwater provides a pathway for potential spread of radioactive contaminants, posing significant risks to human health and ecological systems. Investigation of groundwater impacts of radioactive wastes requires understanding hydrogeological,

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_203

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media, LLC, part of Springer Nature 2018, https://doi.org/10.1007/978-1-4939-2493-6_203-3

geochemical, and biological processes that control the fate and transport of radionuclide contaminants. It will also require developing numerical simulators to integrate these processes for future projection under both natural and engineered remediation scenarios. While some sporadic information is available at specific sites, few attempts have been made to integrate scattered information into a coherent framework to answer relevant questions such as what are the most common sources of radioactive contaminants in groundwater and how they are transported and transformed in the aquifer.

Introduction

Radioactive waste contamination in soil and groundwater poses long-term risks to human health and environment. Public awareness on this subject has steadily increased worldwide since the inception of the nuclear age. Significant efforts have been made to understand the fate and transport of radionuclide wastes in environments, and a number of research and remediation programs have been established to remediate contaminated sites in the United States.

The radionuclide fate and transport in groundwater is controlled by hydrologic, microbiologic, and geochemical processes that operate in the subsurface environment. These processes and their coupling control contaminant migration and persistence, and efficiency of remediation technologies. The understanding of the fate and transport of radionuclide contaminants can improve our ability to forecast contaminant destination and select cost-effective remediation technologies such as mobilization, immobilization, or in-ground degradation.

Reactive transport models are important tools to systematically integrate physical, chemical, and biological processes and data that are critical to understand and predict the fate and transport of radionuclide contaminants in the subsurface. Properly designed and calibrated models can describe the interactions of competing processes at a range of spatial and time scales, and hence are helpful for optimizing field operations and designing monitoring systems for remediating contaminated sites.

This article is intended to provide a review of the major sources of radioactive wastes and their impacts on groundwater contamination, to discuss the major biogeochemical processes that control the transport and fate of radionuclide contaminants in groundwater, and briefly describe the evolution of mathematical models designed to simulate and assess the transport and transformation of radionuclides in groundwater.

Overview on Groundwater Impacts of Radioactive Wastes

Source of Radioactive Wastes

Radioactive wastes were generated primarily from the production of nuclear fuels for the weapons program and electricity energy, development and operation of commercial power reactors, nuclear weapons tests, fuel reprocessing, waste storage and disposal activities, and nuclear accidents. According to the statistics by Ahearne [3], the combined volume of all radioactive wastes (excluding that in the soil and water) from both the government and commercial sources in the United States (US) is about 5.5 million m^3 and the total radioactivity from all anthropogenic sources is about 31 billion Ci. Over the past decades, the uranium mining for production of nuclear fuels have resulted in a large volume of mine and mill tailings, which contain all of the naturally occurring radioactive elements. This has left a legacy of environmental pollution across the countries in the world, such as former Soviet Union, US, Germany, France, Eastern European countries [4].

Contamination of Groundwater Caused by Radioactive Waste Release

Radionuclide contaminants have been detected in subsurface sediments and groundwater as a result of intentional and accidental release of radionuclide-containing wastes during storage, processing, and disposal of nuclear materials [5–7] as well as the leaching of uranium mill tailings [8, 9]. About 30–80 million m^3 of soils and 1,800–4,700 million m^3

of water have been contaminated by radionuclides in US [10]. Among them, most contamination occurred at US Department of Energy (DOE) sites used for nuclear weapons production, where totally over 5700 individual contaminant plumes have been detected in subsurface [11]. At DOE sites, the radionuclides (e.g., uranium, cesium, strontium, thorium, and tritium) have been normally co-disposed with chlorinated solvents (e.g., perchloroethylene (PCE) and trichloroethylene (TCE)) and metals (e.g., lead, chromium, and mercury) and thus have been often identified together with these contaminants in groundwater [5]. According to DOE estimation, approximately 38 million m^3 of groundwater was contaminated from the wastes generated from uranium mill processing or mill tailings [12]. Among the radionuclide contaminants, actinide elements uranium, neptunium, and plutonium are the most problematic [13] because of their long decay half-lives in subsurface sediments and groundwater.

Groundwater Contamination by Radionuclides at Well-Known Sites

Thousands of sites around the world are currently contaminated with radionuclides [7]. In particular, spectacular examples of such sites can be found in the former Soviet Union and US. Major contaminated areas are often located at or near facilities that reprocessed nuclear fuels from production reactors [6]. In the US, the major sites are at Hanford in Washington state, Savannah River in South Carolina, and Oak Ridge in Tennessee [5, 6]. The contamination at Oak Ridge was caused by underground injection of cesium and strontium wastes, and at Savannah River resulted from the release of mixed fission product solutions into streams and seepage basins, and at Hanford originated from the discharge of mixed fission wastes into soils and surface ponds [6].

Among the contaminated sites at US, the Hanford site has been dubbed “the dirtiest place on Earth” [14]. The contamination at the Hanford site, which was divided into three areas namely 100 Area, 200 East and West Areas, and 300 Area, has mainly occurred at locations of nuclear fuel fabrication, fuel irradiation, strategic radionuclide separation, and waste storage and disposal for plutonium production [5]. About 67 of the 149 single shell tanks used for storing nuclear wastes were suspected to have released over 1.9 million L of tank waste to the vadose zone [15]. The tank leaks, combined with discharge of liquid waste through ponds, trenches, pipelines, and cribs, caused a large quantity of radioactivity (through the year 2000: two million Ci) and 100,000 to 300,000 t of toxic chemicals residing in the vadose zone [2]. Wastes at the site have migrated through the vadose zone, resulting in groundwater contamination including nitrate, chromium, tritium, uranium, strontium-90, technetium-99, iodine, carbon tetrachloride, and others [16–18]. For example, waste disposal in 300 Area resulted in a groundwater plume of uranium (VI) with an area of 0.4 ~ 0.5 km^2 that exceeded the drinking water standard of 30 μg/L up to present [19].

For over 30 years, trillions of liters of acidic plating wastes containing high levels of uranium and nitric acid were generated at the Y-12 Facility, Oak Ridge, Tennessee and were discarded into unlined S-3 Ponds. The wastes were neutralized and denitrified and the area was capped and converted to a parking lot in 1984 [20]. Despite these treatments, radionuclide contaminants continued to migrate from the source along geologic strike and dip to groundwater [21, 22].

The Savannah River Site was established in the early 1950s to produce nuclear materials, primarily tritium and plutonium-239 for nuclear weapon purposes, but also plutonium-238 and various transplutonium radionuclides for medical, industrial, and scientific applications. The production was ended in 1988 [23, 24]. The operations at the Savannah River Site have resulted in the migration of radionuclides (e.g., uranium, cesium, radium, thorium, and tritium) into groundwater at various locations, predominantly in the central areas of the site [24, 25].

Many abandoned mine processing sites were also contaminated with radionuclides during uranium mining activities. Contaminated groundwater at the former uranium mill site located at Naturita, Colorado, which processed uranium and vanadium ores intermittently from 1930s to 1958, occurred in the thin alluvial deposits of the

San Miguel River floodplain. High concentrations of uranium were measured in the groundwater below and downgradient of the former tailings pile [26]. The uranium at the Old Rifle UMTRA field experiment site in western Colorado in the aquifer originated at mill tailings (now removed), percolated through a 4-m thick vadose zone to the water table, and was transported laterally through the aquifer via groundwater flow with maximum uranium concentrations 300 $\mu g\ L^{-1}$ [27].

Controlling Biogeochemical Processes

The fate and transport of radionuclides in groundwater is controlled by geochemical and biological processes, in addition to the hydrological ones. This section discusses four biogeochemical processes that have been identified to have major effects on the fate and transport of radionuclide contaminants in groundwater.

Adsorption/Desorption

The adsorption to sediment surfaces is a major geochemical process controlling the mobility of radionuclide contaminants in oxic groundwater. This process is regulated by interfacial chemistry of the prevailing mineral surfaces [9, 26, 28–32]. For example, spectroscopic characterization, laboratory transport experiments, and numerical simulations have revealed that adsorption is a primary process controlling uranium transport in the groundwater at Hanford 300 Area [33, 34], at a U(VI) mill located in Naturita, Colorado [26], at the Oak Ridge site in Tennessee [35], and at the Old Rifle UMTRA site in western Colorado [27]. Many experimental studies have also demonstrated that plutonium can be adsorbed onto a variety of minerals and mineral assemblages [36–38].

Sorption of actinides, particularly plutonium, onto submicrometer-sized colloids increases their mobility in groundwater [39]. Certain actinides can be stabilized in natural waters through the formation of actinide pseudo-colloids, in which the actinide sorbs onto aquatic colloids. This process alone can increase the actinide concentrations by many orders of magnitude over the values expected from solubility calculations [24, 39–41]. Colloid-facilitated transport is likely one of major mechanisms for long-distance transport of actinides in groundwater [39].

The kinetic adsorption/desorption behavior of radionuclide contaminants has often been observed in uranium(VI)-contaminated sediments with controlling mechanisms generally not well understood [32, 34, 42–47]. The kinetic uranium (VI) release from the Hanford 300Area sediments was found to result from diffusional mass transfer from intragrain, intracoating, and intragrain aggregate regions based on microscopic and spectroscopic characterizations of the sediments [48, 49].

Aqueous Complexation

Uranium(VI) can complex with various ligands such as carbonate in groundwater to form aqueous complexes. The aqueous complexation process stabilizes uranium in aqueous phase and decreases its tendency to bind to mineral surfaces [50]. Fox et al. [51] and Dong et al. [52] found that calcium can have a significant impact on the aqueous speciation of uranium(VI) under neutral to mild alkaline pH conditions through formation of ternary uranium(VI)-calcium-carbonate aqueous species. The aqueous and surface uranium (VI) complexation is sensitive to important groundwater chemical composition including pH, carbonate, and calcium concentrations [9, 26, 51]. Consequently, the hydrogeochemical conditions in groundwater at field sites have a great impact on radionuclide-contaminant fate and transport. The groundwater redox conditions have also been found to impact the mobility of other selected radionuclides [53].

Precipitation/Dissolution

Radionuclide fate and transport is also affected by processes, such as precipitation/dissolution reactions [46, 54, 55], coprecipitation with other minerals [56], and microbially induced mineralization [22]. The actinide precipitation/dissolution may occur in intragrain regions where thermodynamic conditions for precipitation/dissolution reactions may significantly different from bulk solution [46, 54].

Precipitation and co-precipitation processes play an important role in uranium stabilization under both reducing and oxidizing conditions [33, 57–59]. Under oxidizing conditions, radionuclides can react with carbonate and phosphate to form carbonate and phosphate minerals (e.g., autunite, $Ca(UO_2)_2(PO_4)_2$). Under reducing conditions, anaerobic bacteria can reduce uranium (VI) to uranium(IV) to form poorly soluble uraninite [22, 27, 35, 58]. Radium can be sorbed and coprecipitated with Fe–Mn oxyhydroxides, gypsum, barite [4, 60], and amorphous silica [61]. The stability of coprecipitated radionuclides is controlled by the host mineral solubility and stability. For example, microbial reduction of iron minerals can indirectly contribute to radium-226 release in groundwater [60].

All the processes mentioned above can act cooperatively or sequentially to control the transport behavior of radionuclide contaminants in aquifers. For example, a study by Catalano et al. [33] using the sediments collected from the Hanford 300 Area revealed that uranium coprecipitated with calcite in shallow vadoze zone sediments, formed metatorbernite ($Cu(UO_2PO_4)_2 \cdot 8H_2O$) that coexisted with adsorbed species at intermediate depths in the vadose zone, and occurred predominantly as adsorbed onto phyllosilicates in the deeper vadose zone and groundwater.

Bioreduction

Bioreduction has been proposed as a remediation approach to immobilize redox-sensitive radionuclides in subsurface environments. Radionuclides can be reduced by various dissimilatory bacteria including metal-reducing bacteria and sulfate-reducing bacteria that use radionuclides as terminal electron acceptors [62, 63]. Microbial reduction of uranium(VI) to insoluble uranium(IV) by the injection of ethanol has been demonstrated at DOE Environmental Remediation Sciences Program (ERSP) Field Research Center in Oak Ridge, Tennessee [64] and Old Rifle UMTRA field site in western Colorado. The results from Wu et al. [22] demonstrated that aqueous uranium concentrations below the USEPA maximum contaminant level ($<0.126\ \mu M$) can be achieved in situ, that bioreduced/immobilized uranium is stable under anaerobic conditions, and that infiltration of dissolved oxygen into the bioreduced area promotes spatially variable oxidation of uranium (IV) and mobilization of uranium(VI). Studies at Old Rifle UMTRA site demonstrated that the immobilization of uranium (VI) in groundwater can be achieved by iron-reducing bacteria stimulated by acetate amendment at the field scale [27, 65].

However, one major unresolved question in terms of bioreduction as a viable remediation technology is the long-term stability of bioreduced radionuclides such as biogenic uranium(IV). Zhong et al. [66] demonstrated that biogenic U(IV) readily oxidizes once groundwater environment returns to oxic condition. Wan et al. [67] presented evidence that bioreduced uranium(IV) was reoxidized under reducing conditions because carbonate accumulation promotes the formation of highly stable carbonato-uranium(VI) complexes under neutral to slightly alkaline conditions. Further researches are needed to enable the bioreduction as a remediation technology.

Environmental Modeling Assessment

Numerical modeling can help scientists and engineers understand and predict the radioactive contaminant fate and transport in the subsurface. Specifically, the modeling can be used to integrate conceptual understanding into a consistent and numerical framework, to test hypotheses on physical and chemical processes under field relevant conditions, to plan for field experiments under uncertainties, to interpret and analyze the field experimental data, and finally to help with the design of remedial alternatives. In an iterative and complementary way, field experiments and modeling activities can work together to enable us to gain new insights and to improve our predictive capabilities on contaminant fate and transport. This section discusses two major types of approaches for modeling radionuclide fate and transport in the subsurface.

Isotherm-Based Transport Modeling

This type of models was based on the simplification that a sorption isotherm involving a single distribution coefficient (Kd) can be used to describe the sorption equilibrium of radioactive contaminants [12]. The Kd-based sorption isotherm can be directly incorporated into a hydrological transport equations to simulate the coupled effects of advection, dispersion, and sorption processes [68].

As discussed before, however, the radionuclide adsorption/desorption is controlled by complex geochemical and microbiological conditions at field and a single Kd value is usually not adequate to represent the sorption behavior over a wide range of geochemical conditions. This is because Kd values are sensitive to geochemical conditions and vary as geochemical conditions change. For example, the Kd value for uranium(VI) can vary by 5 orders of magnitude over the pH range from 6 to 9 and by 4 orders of magnitude at pH 8 as the partial pressure of CO_2 gas increases from its value in air to 0.01 atm [9, 26].

In contaminants plumes where the groundwater compositions change spatially and temporally, the Kd values could also have complex spatial patterns and evolve temporally along with the transport processes. Consequently significant errors and uncertainties may be introduced in reactive transport simulations if a constant-Kd modeling approach is used at sites where groundwater chemistry varies temporally and spatially [9, 26, 28–31].

Multicomponent Reactive Transport Modeling

Over the last two decades, a number of models that couple advective-dispersive-diffusive transport processes with "full" geochemistry, including pH, redox-state, sediment/rock–water interactions have been developed, such as PHT3D [69], MIN3P [70], and PHAST [71]. In these coupled models, the solute transport and chemical reactions are rigorously simulated often in three-dimensional groundwater flow systems. The mechanistic treatment of chemical reactions in the coupled multicomponent, multispecies mass transport has obvious advantages over the empirical isotherm-based transport models since the models more realistically account for complex geochemical processes.

In contrast to the constant-Kd modeling approach, the coupled models normally incorporate surface complexation reactions into solute transport models through the mass action equations describing the equlibria between aqueous chemical species and species formed at mineral surfaces to account for the adsorption/desorption [9]. Surface complexation models (SCM) can account for variations in chemical conditions and aqueous speciation, and thus can describe spatial and temporal changes of radionuclide adsorption [26].

There are two major approaches for applying the SCM concept to natural subsurface systems: the Component Additivity (CA) and Generalized Composite (GC) approaches [9]. The CA approach utilizes documented SCMs for well-characterized surfaces and detailed sediment characterization of the study site to determine the quantity of each reactive surface in sediments and then assembles an SCM for the sediment from its basic components [9, 26]. In the GC approach, adsorption is assumed to occur on "generic" surface sites that represent average properties of the sediment surfaces because the surface of the mineral assemblage is considered too complex to be quantified in terms of the contributions of individual phases to adsorption. Adsorption can be described by mass laws written with "generic" surface functional groups, with the stoichiometry and formation constants for each mass law determined by fitting experimental data for the mineral assemblage as a whole [9, 72].

The SCM approach has not been commonly used to describe adsorption in field-scale reactive transport modeling studies because of a poor understanding of the thermodynamics of surface complex formation in natural systems and the lack of field data. However, there is a growing application of SCM in multi-component reactive transport modeling to simulate the radionucide fate and transport in aquifers due to recent advances in computer codes and availability of extensively geochemical characterization data in some filed sites, such as uranium reactive transport models developed at the Naturita site, Colorado [26], the Hanford site [34, 47, 73], the Oak Ridge site [35], and the Old Rifle site [27]. In addition to the models that consider equilibrium-based surface complexation reactions, models that couple kinetic mass transfer processes with surface complexation reactions have also been developed to simulate the kinetic adsorption and desorption behavior of radionuclides. For example, Liu et al. [34, 47] proposed a multirate SCM to consider the effects of diffusive mass transfer on U(VI) adsorption/desorption processes and the approach was evaluated at the Hanford site [73, 74].

The multicomponent reactive transport model can also take into account for other biogeochemical reactions including aqueous complexation, precipitation/dissolution of radionuclide-containing minerals and bioreduction (e.g., references 35, 75). Thus, the multicomponent reactive transport model can better describe the transport and fate of radionuclide contaminants in the aquifer. However, obtaining sufficient and accurate field data for properly parameterizing a field-scale multicomponent reactive transport model will remain a major challenge.

Future Directions

Future research should be aimed at improving fundamental understanding of radionuclide transport processes in heterogeneous subsurface media and facilitating transfer of knowledge and insights gained from laboratory experiment to field application.

Mass Transfer Processes of Radionuclide Contaminants in Heterogeneous Media

The variability in the physical and chemical properties of subsurface media, such as hydraulic conductivity, porosity, grain size, and reactive surface area, may vary by several orders of magnitude within an aquifer. The heterogeneity of physical and chemical properties causes spatial variations in groundwater flow velocity, reaction rate, and speciation. These variations may be associated with a range of different predominating phenomena such as preferential flow and contaminant migration pathways, hydraulically inaccessible zones into which solutes may only diffuse, and mineral grains by which solutes are selectively sorbed (e.g., references 76–82). The individual phenomena in turn contribute to a variable extent to the spreading (dispersion) of chemical contamination, emerge and disappear at different time scales, and lead to variable reaction types and rates during solute mass transport.

As a result of physical and chemical heterogeneity, the mass transfer could occur in the subsurface at multiple scales. The release of contaminants including cesium-137, chromium (VI), strontium-90, and uranium from contaminated sediments at the Hanford site to groundwater is found to be controlled by mass transfer in laboratory experiments [46, 55, 83–85]. Mass transfer limitations also occur at increasingly larger (macroscopic) scales in the field (e.g., references 54, 86) such as between coarse and fine-textured zones in a given facies, or between different geologic formations and facies, such as highly conductive and less conductive aquifer layers [87]. Understanding multiscale mass transfer processes and their implications to contaminant migration and remediation at the field scale is at the forefront of reactive transport science and is a critical need for the remediation of contaminated sites [87].

Upscaling of Radionuclide Transport from Laboratory to Field Scales

Laboratory experiments provide important information on parameters and key insights for contaminant transport processes. Even complex reactive transport models may be reasonably well constrained by measured data at the laboratory scale, where a large number of measurements and observations are available. However, this is generally not the case for field-scale problems where reactive transport may be affected or controlled by strongly chemically and/or physically heterogeneous conditions (e.g., references 74, 88–91). The heterogeneity at different scales significantly influences and drastically complicates the upscaling of solute transport and its analysis and prediction at larger scales (e.g., references 86, 92–95). Consequently, important differences can exist between the experimental conditions under which a laboratory model was developed and calibrated, and those present in the field, including the ratios of reaction and transport time scales and variability in reactant properties and distribution.

Without appropriate consideration of the upscaling problem, reliable prediction of reactive transport processes is impossible. Therefore, scale dependence and associated upscaling of transport parameters for porous media have been actively studied for two decades. For the most part, however, the existing research has focused on the development of theoretical or empirical upscaling methods in relatively simple systems [96–98] and on the investigation of the scale dependence of different transport parameters such as diffusion coefficients, geochemical reaction rates, sorption coefficients, and retardation factors [99] and other factors (e.g., reference 100) at various scales ranging from pore scale to column experiments, and to field tracer tests. However, very few studies have been reported on upscaling of transport processes of radionuclide contaminants in groundwater [74]. Thus, the unsolved problems is how to systematically upscale the reactive transport process and parameters for reactive transport models involving complex, physically and chemically heterogeneous systems across multiscales.

Bibliography

Primary Literature

1. Bodvarsson GS, Boyle W, Patterson R, Williams D (1999) Overview of scientific investigations at Yucca Mountain – the potential repository for high-level nuclear waste. J Contam Hydrol 38:3–24
2. Gephart RE (2003) Hanford: a conversation about nuclear waste and cleanup. Battelle Press, Columbus
3. Ahearne JF (1997) Radioactive waste: the size of the problem. Phys Today 50(6):24–29
4. Abdelouas A (2006) Uranium mill tailings: geochemistry, mineralogy, and environmental impact. Elements 2:335–341
5. Riley RG, Zachara JM, Wobber FJ (1992) Chemical contaminants on DOE lands and selection of contaminant mixtures for subsurface science research, DOE/ER-0547T. U.S. Department of Energy, Washington, DC
6. Bradley DJ, Frank CW, Mikerin Y (1996) Nuclear contamination from weapons complexes in the former Soviet Union and the United States. Phys Today 49:40–45
7. Whicker FW, Shaw G, Voigt G, Holm E (1999) Radioactive contamination: state of the science and its application to predictive models. Environ Pollut 100:133–149
8. Abdelouas A, Lutze W, Nuttall HE (1999) Uranium contamination in the subsurface; Characterization and remediation. Rev Mineral Geochem 38:433–473
9. Davis JA, Meece DE, Kohler M, Curtis GP (2004) Approaches to surface complexation modeling of uranium (VI) adsorption on aquifer sediments. Geochim Cosmochim Acta 68(18):3621–3641
10. Ewing RC (2004) Environmental impact of the nuclear fuel cycle. In: Gieré R, Stille P (eds) Energy, waste and the environment: a geochemical perspective. Geological society special publication, vol 236. The Geological Society, London, pp 7–23
11. Lee D, Walton MR, Megio JL (2005) Biological and chemical interactions with U(VI) during anaerobic enrichment in the presence of iron oxide coated quartz. Water Res 39:4363–4374
12. Zhu C, Anderson G (2002) Environmental applications of geochemical modeling. Cambridge University Press, London
13. Renshaw J, Butchins LJC, Livens FR, May I, Charnock JM, Lloyd JR (2005) Bioreduction of uranium: environmental implications of a pentavalent intermediate. Environ Sci Technol 39:5657–5660
14. Fishlock D (1994) The dirtiest place on earth. New Sci 1913:34–37
15. Zachara JM, Serne J, Freshley M, Mann F, Anderson F, Wood M, Jones T, Myers D (2007) Geochemical processes controlling migration of tank wastes in Hanford's vadose zone. Vadose Zone J 6:985–1003
16. Gee GM, Oostrom M, Freshley MD, Rockhold ML, Zachara JM (2007) Hanford site vadose zone studies: an overview. Vadose Zone J 6:899–905
17. Um W, Serne RJ, Brown CF, Last GV (2007) U (VI) adsorption on aquifer sediments at the Hanford site. J Contam Hydrol 93:255–269
18. Hartman MJ, Morasch LF, Webber WD (2007) Hanford site groundwater monitoring for fiscal year 2006. Richland, Washington, Pacific Northwest National Laboratory, Richland
19. Hartman MJ, Webber WD, Fluor Hanford, Inc (2008) Hanford site groundwater monitoring for fiscal year 2007.DOE/RL-2008-01, Revision 0. Pacific Northwest National Laboratory, Richland
20. Phillips H, Watson DB, Roh Y (2007) Uranium deposition in a weathered fractured saprolite/shale. Environ Sci Technol 41:7653–7660
21. Wu W, Carley J, Fienen M, Mehlhorn T, Lowe K, Nyman J, Luo J, Gentile ME, Rajan R, Wagner D, Hickey RF, Gu B, Watson D, Cirpka O, Kitanidis P, Jardine J, Criddle CS (2006) Pilot-scale in situ bioremediation of uranium in a highly contaminated aquifer. 1. Conditioning of a treatment zone. Environ Sci Technol 40(12):3978–3985
22. Wu W, Carley J, Luo J, Ginder-Vogel MA, Cardenas E, Leigh MB, Hwang C, Kelly SD, Ruan C, Wu L, Nostrand JV, Gentry T, Lowe K, Mehlhorn TL, Caroll S, Luo W, Fields MW, Gu B, Watson D, Kemner K, Marsh T, Tiedje J, Zhou J,

Fendorf S, Kitanidis PK, Jardine PM, Criddle C (2007) In situ bioreduction of uranium(VI) to submicromolar levels and reoxidation by dissolved oxygen. Environ Sci Technol 41:5716–5723
23. Carlton WH (1997) Assessment of neptunium, americium, and curium in the Savannah River Site Environment. Westinghouse Savannah River Co., Aiken, WSRC-TR-97-00266
24. Dai M, Kelley JM, Buesseler KO (2002) Sources and migration of plutonium in groundwater at the Savannah River site. Environ Sci Technol 36:3690–3699
25. Westinghouse Savannah River Co (1998) The Savannah River Site's groundwater monitoring program: third quarter 1997. U.S. Department of Energy, Washington, DC, ESH-EMS-970490
26. Curtis GP, Davis JA, Naftz DL (2006) Simulation of reactive transport of uranium (VI) in groundwater with variable chemical conditions. Water Resour Res 42:W04404. https://doi.org/10.1029/2005WR003979
27. Yabusaki SB, Fang Y, Long PE, Resch CT, Peacock AD, Komlos J, Jaffed PR, Morrison SJ, Dayvault RD, White DC, Anderson RT (2007) Uranium removal from groundwater via in situ biostimulation: field-scale modeling of transport and biological processes. J Contam Hydrol 93:216–235
28. Read D, Ross D, Sims RJ (1998) The migration of uranium through Clashach sandstone: the role of low molecular weight organics in enhancing radionuclide transport. J Contam Hydrol 35:235–248
29. Bethke CM, Brady PV (2000) How the Kd approach undermines group water cleanup. Ground Water 38(3):435–443
30. Glynn PD (2003) Modeling Np and Pu transport with a surface complexation model and spatially variant sorption capacities; implications for reactive transport modeling and performance assessments of nuclear waste disposal sites; reactive transport modeling in the geosciences. Comput Geosci 29(3):331–349
31. Zhu C (2003) A case against Kd-based transport models: natural attenuation at a mill tailings site; reactive transport modeling in the geosciences. Comput Geosci 29(3):351–359
32. Bond DL, Davis JA, Zachara JM (2008) Uranium (VI) release from contaminated vadose zone sediments: estimation of potential contributions from dissolution and desorption, Chapter 14. In: Barnett MO, Kent DB (eds) Adsorption of metals to Geomedia II. Elsevier, Amsterdam, pp 375–416
33. Catalano JG, Mckinley JP, Zachara JM, Heald SM, Smith SC, Brown GE (2006) Changes in uranium speciation through a depth sequence of contaminated Hanford sediments. Environ Sci Technol 40(8):2517–2524
34. Liu C, Zachara JM, Qafoku NP, Wang Z (2008) Scale-dependent desorption of uranium from contaminated subsurface sediments. Water Resour Res 44:W08413. https://doi.org/10.1029/2007WR006478
35. Luo J, Weber F, Cirpka OA, Wu W, Nyman JL, Carley J, Jardine PM, Criddle CS, Kitanidis PK (2007) Modeling in-situ uranium(VI) bioreduction by sulfate-reducing bacteria. J Contam Hydrol 92:129–148
36. Keeney-Kennicutt WL, Morse JW (1985) The redox chemistry of $Pu(V)O_2^+$ interaction with common mineral surfaces in dilute solutions and seawater. Geochim Cosmochim Acta 49(12):2577–2588
37. Sanchez AL, Murray JW, Sibley TH (1985) The adsorption of plutonium on goethite. Geochim Cosmochim Acta 49:2297–2307
38. Duff MC, Hunter DB, Triay IR, Bertsch PM, Reed DT, Sutton SR, Shea-McCarthy G, Kitten J, Eng P, Chipera SJ, Vaniman DT (1999) Mineral associations and average oxidation states of Sorbed Pu on Tuff. Environ Sci Technol 33:2163–2169
39. Novikov P, Kalmykov SN, Utsunomiya S, Ewing RC, Horreard F, Merkulov A, Clark SB, Tkachev VV, Myasoedov BF (2006) Colloid transport of plutonium in the far-field of the Mayak Production Association, Russia. Science 314:638–641
40. Kim JI (1993) The chemical behavior of transuranium elements and barrier functions in natural aquifer systems. Mater Res Soc Symp Proc 294:3–21
41. Kim JI (1994) Actinide colloids in natural aquifer systems. Mater Res Soc Bull 19:47–53
42. Braithwaite A, Livens FR, Richardson S, Howe MT (1997) Kinetically controlled release of uranium from soils. Eur J Soil Sci 48:661–673
43. Barnett MO, Jardine PM, Brooks SC, Selim HM (2000) Adsorption and transport of uranium(VI) in subsurface media. Soil Sci Soc Am J 64:908–917
44. Baik MH, Cho WJ, Hahn PS (2004) Sorption of U (VI) onto granite surfaces: a kinetic approach. J Radioanal Nucl Chem 260:495–502
45. Qafoku NP, Zachara JM, Liu C, Gassman PL, Qafoku OS, Smith SC (2005) Kinetic desorption and sorption of U(VI) during reactive transport in a contaminated Hanford sediment. Environ Sci Technol 39:3157–3165
46. Liu C, Zachara JM, Yantansee W, Majors PD, McKinley JP (2006) Microscopic reactive diffusion of uranium in the contaminated sediments at Hanford, USA. Water Resour Res 42:W12420. https://doi.org/10.1029/2006WR005031
47. Liu C, Shi S, Zachara JM (2009) Kinetics of uranium (VI) desorption from contaminated sediments: effect of geochemical conditions and model evaluation. Environ Sci Technol 43(17):6560–6566
48. Arai Y, Marcus MA, Tamura N, Davis JA, Zachara JM (2007) Spectroscopic evidence for uranium bearing precipitates in vadose zone sediments at the Hanford 300-area site. Environ Sci Technol 41:4633–4639
49. Stubbs JE, Veblen LA, Elbert DC, Zachara JM, Davis JA, Veblen DR (2009) Newly recognized hosts for uranium in the Hanford site vadose zone. Geochim Cosmochim Acta 73(6):1563–1576
50. Waite TD, Davis JA, Payne TE, Waychunas GA, Xu N (1994) Uranium (VI) adsorption to ferrihydrite: application of a surface complexation model. Geochim Cosmochim Acta 58(24):5465–5478

51. Fox PM, Davis JA, Zachara JM (2006) The effect of calcium on aqueous uranium(VI) speciation and adsorption to ferrihydrite and quartz. Geochim Cosmochim Acta 70:1379–1387
52. Dong W, Ball WP, Liu C, Wang Z, Stone AT, Bai J, Zachara JM (2005) Influence of calcite and dissolved calcium on uranium(VI) sorption to a Hanford subsurface sediment. Environ Sci Technol 39:7949–7955
53. Hu QH, Zavarin M, Rose TP (2008) Effect of reducing groundwater on the retardation of redox-sensitive radionuclides. Geochem Trans 9:12. https://doi.org/10.1186/1467-4866-9-12
54. Liu C, Zachara JM, Qafoku OS, McKinley JP, Heald SM, Wang Z (2004) Dissolution of uranyl microprecipitates in subsurface sediments at Hanford Site, WA. Geochim Cosmochim Acta 68:4519–4537
55. McKinley JP, Zachara JM, Liu C, Heald SM (2006) Microscale controls on the fate of contaminant uranium in the vadose zone, Hanford site, Washington. Geochim Cosmochim Acta 70:1873–1887
56. Gómez P, Garralón A, Buil B, Turrero MJ, Sánchez L, de la Cruz B (2006) Modeling of geochemical processes related to uranium mobilization in the groundwater of a uranium mine. Sci Total Environ 366:295–309
57. Abdelouas A, Lutze W, Nuttall E (1998) Chemical reactions of uranium in ground water at a mill tailings site. J Contam Hydrol 34:343–361
58. Abdelouas A, Lutze W, Gong W, Nuttall EH, Strietelmeier BA, Travis BJ (2000) Biological reduction of uranium in groundwater and subsurface soil. Sci Total Environ 250:21–35
59. Ohnuki T, Kozai N, Samadfam M, Yasuda R, Yamamoto S, Narumi K, Naramoto H, Murakami T (2004) The formation of autunite ($Ca(UO_2)_2(PO_4)_2 \cdot nH_2O$) within the leached layer of dissolving apatite: incorporation mechanism of uranium by apatite. Chem Geol 211:1–14
60. Martin AJ, Crusius J, Jay McNee J, Yanful EK (2003) The mobility of radium-226 and trace metals in pre-oxidized subaqueous uranium mill tailings. Appl Geochem 18:1095–1110
61. Landa ER (2004) Uranium mill tailings: nuclear waste and natural laboratory for geochemical and radioecological investigations. J Environ Radioact 77:1–27
62. Lovely DR, Coates JD (1997) Bioremediation of metal contamination. Curr Opin Biotechnol 8:285–289
63. Liu C, Gorby YA, Zachara JM, Fredrickson JK, Brown CF (2002) Reduction kinetics of Fe(III), Co(III), U(VI), Cr(VI), and Tc(VII) in cultures of dissimilatory metal-reducing Bacteria. Biotechnol Bioeng 80(6):637–649
64. Wu W, Carley J, Gentry T, Ginder-Vogel MA, Fienen M, Mehlhorn T, Yan H, Carroll S, Nyman J, Luo J, Gentile ME, Fields MW, Hickey RF, Watson D, Cirpka OA, Fendorf S, Zhou J, Kitanidis P, Jardine PM, Criddle CS (2006) Pilot-scale in situ bioremediation of uranium in a highly contaminated aquifer. 2: U(VI) reduction and geochemical control of U(VI) bioavailability. Environ Sci Technol 40:3986–3995
65. Anderson RT, Vrionis HA, Ortiz-Bernad I, Resch CT, Long PE, Dayvault R, Karp K, Marutzky S, Metzler DR, Peacock A, White DC, Lowe M, Lovley DR (2003) Stimulating the in situ activity of *Geobacter* species to remove uranium from the groundwater of a uranium-contaminated aquifer. Appl Environ Microb 69(10):5884–5891
66. Zhong L, Liu C, Zachara JM, Kennedy DW, Szecsody JE, Wood BD (2005) Oxidative remobilization of biogenic uranium (IV) precipitates: effects of Iron (II) and pH. J Environ Qual 34(5):1763–1771
67. Wan J, Tokunaga TK, Brodie E, Wang Z, Zheng Z, Herman D, Hazen T, Firestone MK, Sutton SR (2005) Reoxidation of bioreduced uranium under reducing conditions. Environ Sci Technol 39:6162–6169
68. Zheng C, Wang PP (1999) MT3DMS, A modular three-dimensional multi-species transport model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems; documentation and user's guide. U.S. Army Engineer Research and Development Center Contract Report SERDP-99-1, Vicksburg, 202 pp. http://hydro.geo.ua.edu/mt3d/
69. Prommer H, Barry DA, Zheng C (2003) MODFLOW/MT3DMS based reactive multicomponent transport modeling. Ground Water 41(2):247–257
70. Mayer KU, Frind EO, Blowes DW (2002) Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions. Water Resour Res 38:1174. https://doi.org/10.1029/2001WR000862
71. Parkhurst DL, Kipp KL, Engesgaard P, Charlton SC (2004) PHAST – a program for simulating groundwater flow, solute transport and multicomponent geochemical reactions. USGS Tech Methods 6-A8:154 pp
72. Davis JA, Payne TE, Waite TD (2002) Simulating the pH and pCO_2 dependence of uranium(VI) adsorption by a weathered schist with surface complexation models. In: Geochemistry of soil radionuclides. Soil Science Society of America Inc., Madison, pp 61–68
73. Ma R, Zheng C, Prommer H, Greskowiak J, Liu C, Zachara J, Rockhold M (2010) A field-scale reactive transport model for U(VI) migration influenced by coupled multirate mass transfer and surface complexation reactions. Water Resour Res 46:W05509. https://doi.org/10.1029/2009WR008168
74. Greskowiak J, Prommer H, Liu C, Post VEA, Ma R, Zheng C, Zachara JM (2010) Comparison of parameter sensitivities between a laboratory and field scale model of uranium transport in a dual domain, distributed-rate reactive system. Water Resour Res 46:W09509. https://doi.org/10.1029/2009WR008781
75. Fang Y, Yabusaki SB, Morrison SJ, Amonette JP, Long PE (2009) Multicomponent reactive transport

modeling of uranium bioremediation field experiments. Geochim Cosmochim Acta 73:6029–6051
76. Gelhar LW (1986) Stochastic subsurface hydrology – from theory to applications. Water Resour Res 22(9):135S–145S
77. Dagan G (1989) Flow and transport in porous formations. Springer, New York
78. Barber LB (1994) Sorption of chlorobenzenes to Cape Cod aquifer sediments. Environ Sci Technol 28:890–897
79. Friedly JC, Davis JA, Kent DB (1995) Modeling hexavalent chromium reduction in groundwater in field-scale transport and laboratory batch experiments. Water Resour Res 31:2783–2794
80. Kleineidam S, Rugner H, Grathwohl P (1999) Impact of grain scale heterogeneity on slow sorption kinetics. Environ Toxicol Chem 18:1673–1678
81. Allen-King RM, Divine DP, Robin MJL, Alldredge JRG (2006) Spatial distributions of perchloroethylene reactive transport parameters in the Borden Aquifer. Water Resour Res 42. https://doi.org/10.1029/2005WR003977
82. Descourvières C, Hartog N, Patterson BM, Oldham C, Prommer H (2010) Geochemical controls on sediment reactivity and buffering processes in a heterogeneous aquifer. Appl Geochem 25:261–275
83. Liu C, Zachara JM, Smith SC, McKinley JP, Ainsworth CC (2003) Desorption kinetics of radiocesium from the subsurface sediments at Hanford site, USA. Geochim Cosmochim Acta 67:2893–2912
84. McKinley JP, Zachara JM, Smith SC, Liu C (2007) Cation exchange reactions controlling desorption of $^{90}Sr^{2+}$ from coarse-grained contaminated sediments from the Hanford formation, Washington. Geochim Cosmochim Acta 71(2):305–325
85. Zachara JM, Ainsworth CC, Brown GE Jr, Catalano JG, McKinley JP, Qafoku O, Smith SC, Szecsody JE, Traina SJ, Warner JA (2004) Chromium speciation and mobility in a high level nuclear waste vadose zone plume. Geochim Cosmochim Acta 68(1):13–30
86. Zheng C, Gorelick SM (2003) Analysis of solute transport in flow fields influenced by preferential flow paths at the decimeter scale. Ground Water 41(2):142–155
87. Zachara J, Freshley M, Andersen G, DePaolo D, Fredrickson J, Haggerty R, Kent D, Konopka A, Lichtner P, Liu C, McKinley J, Rockhold M, Rubin Y, Szecsody J, Versteeg R, Ward A, Williams B, Zheng C (2007) Integrated Field-Scale Subsurface Research Challenge, Multi-Scale Mass Transfer Processes Controlling Natural Attenuation and Engineered Remediation: an IFC Focused on Hanford's 300 Area Uranium Plume. Proposal to the U.S. Department of Energy Office of Biological and Environmental Research LAB 06-16 – Environmental Remediation Science Program. Pacific Northwest National Laboratory, Richland, Washington
88. Morrison SJ, Tripathi VS, Spangler RR (1995) Coupled reaction/transport of a chemical barrier for controlling U(VI) contamination in groundwater. J Contam Hydrol 17:347–363
89. Zhu C, Hu FQ, Burden DS (2001) Multi-component reactive transport modeling of natural attenuation of an acid groundwater plume at a uranium mill tailings site. J Contam Hydrol 52:85–108
90. Bain JG, Mayer KU, Blowes DW, Frind EO, Molson JWH, Kahnt R, Jenk U (2001) Modeling the closure-related geochemical evolution of groundwater at a former uranium mine. J Contam Hydrol 52:109–135
91. Yabusaki SB, Fang Y, Waichler SR (2008) Building conceptual models of field-scale uranium reactive transport in a dynamic vadose zone-aquifer-river system. Water Resour Res 44:W12403. https://doi.org/10.1029/2007WR006617
92. Feehley CE, Zheng C, Molz FJ (2000) A dual-domain mass transfer approach for modeling solute mass transfer in heterogeneous porous media, application to the MADE site. Water Resour Res 36:2501–2515
93. Haggerty R, Harvey CF, von Schwerin CF, Meigs LC (2004) What controls the apparent timescale of solute mass transfer in aquifers and soils? A comparison of experimental results. Water Resour Res 40:W01510. https://doi.org/10.1029/2002WR001716
94. Gorelick SM, Liu G, Zheng C (2005) Quantifying mass transfer in permeable media containing conductive dendritic networks. Geophys Res Lett 32:L18402. https://doi.org/10.1029/2005GL023512
95. Zheng C, Bianchi M, Gorelick SM (2011) Lessons learned from 25 years of research at the MADE site. Ground Water 49(5): 649-662
96. Seeboonruang U, Ginn TR (2006) Upscaling heterogeneity in aquifer reactivity via exposure-time concept: forward model. J Contam Hydrol 84: 127–154
97. Fernàndez-Garcia D, Llerar-Meza G, Gómez-Hernández JJ (2009) Upscaling transport with mass transfer models: mean behavior and propagation of uncertainty. Water Resour Res 45:W10411. https://doi.org/10.1029/2009WR007764
98. Heβe F, Radu FA, Thullner M, Attinger S (2009) Upscaling of the advection–diffusion–reaction equation with Monod reaction. Adv Water Resour 32:1336–1351
99. Deng H, Dai Z, Wolfsberg A, Lu Z, Ye M, Reimus P (2010) Upscaling of reactive mass transport in fractured rocks with multimodal reactive mineral facies. Water Resour Res 46: W06501. https://doi.org/10.1029/2009WR008363
100. Wang F, Bright J (2004) Scale effect and calibration of contaminant transport models. Ground Water 42(5):760–766

Groundwater Salinity Due to Urban Growth

José Joel Carrillo-Rivera[1] and Samira Ouysse[2]
[1]Instituto de Geografía, UNAM CU, Coyoacán, Mexico
[2]Cadi Ayyad University, Marrakech, Morocco

Article Outline

Glossary

Aquifer unit It is a geological formation, part of a formation, or a number of formations that yield water substantially and with adequate quality for the expected usage

Basin It is often referred to as the drainage basin or watershed where rainfall is gathered with a common discharge outlet; it is considered to have no additional inflow or outflow

Flow systems They are manifested by the presence of groundwater flows with contrasting hierarchy (local, intermediate, and regional), in which their components may be clearly defined from field evidence in conjunction with modeling of groundwater hydraulics, geochemistry, geomorphology, isotopes, and associated soil and vegetation cover

Groundwater vulnerability to contamination It is the tendency or likelihood for a contaminant to reach a specified position in the groundwater system after introduction at some location above the uppermost aquifer unit

Hydraulic conductivity It is the rate of water that is mobilized through a unit section under a unit hydraulic gradient; such value is a function of its degree of saturation, attaining its maximum at 100% saturation, and is also a function of water density and viscosity

Local flow system It is the flow that is travelling a short distance and depth, recharging and discharging in the same valley

Regional flow system It is recharged at the highest elevation and then travels deeply to the basement depth and discharges at the lowest plain in the region

Urban sprawl It is a multifaceted concept, which includes the spreading of dwelling construction outward of a city and its suburbs to its outskirts of low density and becoming an auto-dependant development on rural land, high segregation of uses (e.g., stores and residential), and various design features that encourage car dependency

Vertical flow It is groundwater traveling downward (as in a recharge area) or upward (as in a discharge area). This upward flow component occurs largely during intensive abstraction which induces deep groundwater ascent

Water balance It, as applied to a watershed (surface-basin), is referred to the application of a lumped parameter approach that lacks consideration of spatial and temporal variations of precipitation, evaporation, transpiration, recharge, runoff, and change in storage as well as the processes involved

Water quality change It is the chemical evolution of water quality with abstraction time in a borehole

Wellhead protection area It ideally is the entire groundwater inflow area for the well; often, this inflow area is too large to be managed (or

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_439

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_439-3

defined) effectively, so a smaller location around the well is often chosen and delineated by flow lines basically incorporating horizontal flow; no vertical flow upward is included

Definition of the Subject and Its Importance

The population of the world is increasing at different rates according to the period of time and geographical position; however, at present, the number of cities that have a population larger than one million is estimated to be more than 336; while in the 1950s, 83 cities were in that range. Such urban development has been seen in Mexico City, which has more than 18.6 million of inhabitants, this number could be larger if the adjoining urban sprawls are included (nearly 28 million). Cities have often solved their water supply requirements by means of groundwater abstraction. In ancient times, most of those megacities were founded where surface waters and springs were accessible. Later, when the number of inhabitants and economical activities increased, shallow dug wells were the solution to cope with the increasing water needs. Afterward, the population and economical activities of cities have seen a large increase; and by 1900s, many cities have experienced a rapid growth that has led to their actual size. Due to this excessive growth, groundwater was regarded as a major alternative.

Initially, during the second half of the ninetieth century, groundwater abstraction in Mexico City was made on the surrounding plains by means of boreholes that provided artesian yields with good water quality and sufficient quantities to comply with existing local requirements. In view of the extensive and continuous increase of water needs, limited supply has led to additional boreholes construction. Later, by the end of the first half of the twentieth century, many boreholes were no longer flowing artesian waters, so the acquisition of the turbine pump started a new abstraction period. At that time, the prevailing view was that water quality would basically remain constant with abstraction time due to the use of a hydrological conceptual model known as the "water balance," applied in Mexico since the middle of the 1960s. This conceptual model included an important assumption based on the presence of purely radial horizontal flow to the abstraction borehole as published by Theis [23], a condition that was supposed to prevail in existing abstraction boreholes.

Initially, boreholes were drilled to shallow depths in the vicinity of 100 m, reaching few meters of the whole aquifer thickness. At present, deeper boreholes may reach 400 m but are still far from tapping the full thickness of the aquifer, which is estimated to be of more than 3000 m.

Originally, the deterioration of groundwater quality was solely expected to arise from spillage on the soil surface of anthropogenic and industrial substances (i.e., gasoline leaks, leachates from rubbish pits, improper management of sewage water, etc.), not from substances naturally occurring in the groundwater. The common prevailing wellhead protection area was considered to be the only source of contamination responsible of the quality change of the supplied groundwater. This misconception was overcome by the presence of vertical ascending flow from abstraction boreholes.

Vertical component of groundwater flow has been overlooked as well as its meaning in producing related environmental impacts [7]. Vertical ascending components of flow become paramount when viewed in the context that usually the screen of the abstraction borehole is crossing less than one tenth of the full aquifer thickness. Below this abstraction depth, groundwater flow systems with contrasted water quality travel in the vicinity of local flows that were initially captured by the boreholes at shallow depth. Regional flows with both high temperature and salinity travel at depth and were represented by the Peñón de los Baños spring (44 °C; Cl, 650 mg/L; Li, 1550 μg/L) reported since pre-European arrival times. Regional flows are being slowly but steadily induced to extraction wells, changing the initial water quality. Additionally, there is an increase in the disappearance of springs due to extracting water from boreholes at the spring site. The recognition of the presence of such flow systems is very important. As population and economic activities increase water consumption, the induced vertical flow rate will increase, which in turn will noticeably change the obtained water quality over time.

Introduction

Groundwater has provided feasible alternatives to water needs for many growing cities worldwide. The definition of groundwater functioning is essential when looking for water accessibility in terms of quantity and quality in both time and space. Such knowledge could assist in providing an adequate definition of water sources to reduce the risk of salinity increase of abstracted groundwater due to the growing water needs along with the urban and industrial growth. Establishment of adequate groundwater management policies needs to be in accordance with the understanding of the functioning of existing flow systems and the groundwater environmental manifestation due to human intervention.

The understanding of groundwater functioning is most important in regions where there is an increase in water needs and where shortages in the water supply often affect available infrastructure. The limiting conditions for appropriate development of big cities in many regions could be solved under land-use planning; this refers to the method used by the public sector to influence future spatial distribution of people and economic activities at various scales. Land-use planning processes that are undertaken at local, provincial, and regional levels require covering a broad range of issues, such as housing, commercial and other nonresidential facilities, roads and transportation, schools and utilities, and farming. What is meant to be considered proper land-use planning implies that the proposed land use is in agreement with what the land "attitude" and its environment could provide without affecting its sustainability.

Ideally, a socioeconomic and environmental development program should be assisted by an adequate political decision-making to include the definition of the groundwater flow systems. Required actions to increase development involve a responsible and meticulous knowledge of the nature of the territory, not only surface elements but also groundwater availability and its connection to neighboring surface drainage basins. The common water availability analysis carried out for additional development in Mexico has been related to the *water balance* since late 1960s. This usually involves hydrological factors that are difficult to obtain through the application of the scientific method. This is due to the (1) simplification of natural processes; (2) use of estimates of variables which are not directly measured, such as runoff, precipitation, evapotranspiration, groundwater abstraction, infiltration, and groundwater recharge; (3) presumption of instant infiltration of local precipitation that is reflected in an ascending water table; and (4) the lack of consideration of water quality response. Additionally, prevailing geological conditions of Mexico, and in particular those related to the location of Mexico City, implies the presence of aquifer units with thicknesses beyond 3000 m that extend over 100,000 km^2. Therefore, most surface drainage basins have an underground hydraulic connection with near or faraway surface drainage basins [24]. A starting point with most computed water balance carried out on a surface drainage basin generally fails to represent a consistent scenario which involves the groundwater functioning as a connected system, and as a result, any groundwater abstraction program would fail to consider the changing of water quality over time.

Abstraction response in terms of water quality and its relevant control before the water reaches the borehole pump may be defined applying the *Tóthian Groundwater Flow Systems* (TGFS), methodology that permits water quality control by *understanding before acting* [6, 8]. An approach that has been continuously applied since the 1960s needs to be included in a law ad hoc as to direct any groundwater-related activity to this method; the aim is to understand and, therefore, control related processes responsible for observed mineralization increase before this effect is produced and observed in an abstraction borehole. Although economical and more environmentally friendly, this method was not given adequate interest as compared to the other black box models which often create additional undesirable effects.

From the perspective of safeguarding obtained water quality, the concept of "borehole protection area" (see http://www.state.nj.us/dep/njgs/whpaguide.pdf) in complex hydrogeological systems lacks the required control of particular needed measures of protection to preserve obtained water quality and to recommend appropriate monitoring programs for borehole production. Under purely horizontal flow conditions, the sources of water quality variations are solely expected to arrive from the surface and not from beneath the tapped aquifer unit. The concept of groundwater vulnerability to contamination has limited value in protecting existing groundwater supply if every alteration of water quality is solely expected to be of exogenic origin, rather than natural.

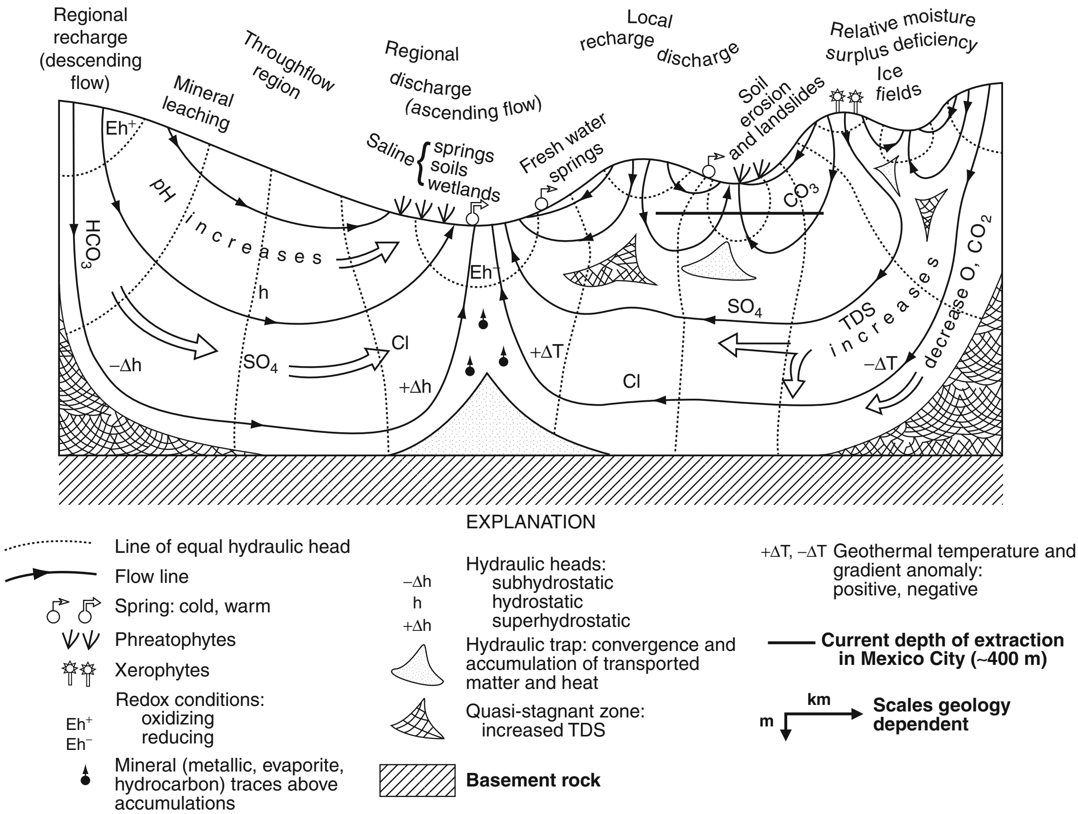
A different approach is therefore needed to define expected changes in the obtained groundwater quality over time. A better understanding of groundwater functioning will facilitate the process. Additionally, the inclusion of socioeconomic and political variables has an equal importance in preventing effects generating extracted water quality deterioration if abstraction and water requirements are properly planned in agreement with the groundwater flow regime. A constant increase in abstraction yield requires an interdisciplinary understanding to assist decision-makers in reaching a better assessment on how concepts such as a specific land-use planning may change water quality and availability. A particular policy of land-use planning on a territory with certain water availability could develop a sustainable use according to the knowledge of water sources functioning and their natural limits. Conversely, without a clear perception of the functioning of water sources, decision-makers are very limited in projecting the outcome of their decisions regarding the steady increase of related cities requirements. In Ref. [7], associated environmental impacts due to a lack of attention to the functioning of groundwater flow systems in Mexico are identified.

In both semiarid and temperate regions, groundwater sources have become crucial to maintain social and economical developments [12, 14, 17]. Fresh groundwater is available in many areas as recently infiltrated rainfall manifested in local groundwater flow systems; therefore, it is common sense to improve strategies for its adequate management since it usually remains the lower-cost option for drinking water. This low-cost option is favored not only financially but also environmentally, especially when water transfers among neighboring basins or surface water sources have become polluted. Another alternative often used is to treat the extracted groundwater. However, such a procedure has proved to be inadequate in the case of the control of the continuous inducement of deep saline waters in Mexico City which surpass local treatment methods and capacity designed to cope with the expected strict levels of salinity. One important issue related to this inefficient rate is linked to the changing of water quality with abstraction time. The quality of the inflow water to the plant will hardly meet design specifications, and consequently, treatment plants may fail to meet drinking water production standards.

Groundwater Flow Systems Hierarchy

Groundwater flows in 3-D from its recharge to its discharge area (i.e., river, wetland, lake, spring, sea) travelling through paths of different lengths and depths differentiating local- to regional-scale flow systems [25]. The presence of flow systems with contrasting hierarchy may be clearly defined from field evidence in conjunction with modeling of groundwater hydraulics, geochemistry, geomorphology, edaphology, and natural vegetation response. Contrasting geomorphological characteristics give groundwater a distinctive flow system with particular physical functioning [25], resulting in specific flow directions: vertical downward in recharge areas, horizontal in transit areas, and vertical upward in discharge areas (Fig. 1). Groundwater flow systems are constantly linked; therefore, it is paramount to understand their functioning to manage them adequately in agreement with their environmental conditions to reduce the impacts of human water usage and land-use changes.

Regions have distinct physical and chemical characteristics for different flow systems, especially



Groundwater Salinity Due to Urban Growth, Fig. 1 Gravity-driven flow systems developed in a sedimentary drainage basin defining local, intermediate, and regional flow systems (Adapted from [26]). The location of the groundwater abstraction for Mexico City is indicated by the depth of boreholes as related to the full aquifer unit

when thick and extensive aquifer units are present. Usually, the changing in temperature (cold in recharge areas, high in the discharge area) and chemical conditions (i.e., water with high dissolved oxygen, low pH, and low salinity in recharge areas; and low dissolved oxygen, high pH, and high salinity in discharge areas) are key indicators of the existing flow systems. Groundwater flow systems travel in the prevailing geological framework structure along the horizontal and vertical plans. Local flows travel short distances at shallow depth, while those that initiate their traveling path in a basin and discharge in a neighboring basin constitute intermediate flows including several local flows (Fig. 1), all of which may be contained by a regional flow. This flow is limited by a basement rock which consists of a regional geological unit with a low hydraulic conductivity; underneath it there are practically no flow conditions (i.e., Cambrian or Precambrian igneous rock). A sketch of the location of Mexico City as related to the prevailing flow systems is represented in Fig. 1 in view of the current depth of abstraction boreholes.

In Mexico, as a result of the thickness and extent of geological formations that constitute the aquifer unit, as well as the variability of climate characteristics, regional flows travel several hundreds of kilometers, often flowing from a temperate climate region to one with arid to semiarid characteristics, constituting the basic source of urban and economical development as the city of San Luis Potosí [6]. Such development has been in progress without proper consideration being given to water availability not only in quantity terms but also regarding its quality.

Importance of Groundwater for Urban Development

The Mexico City Case

A large groundwater storage capacity becomes important when urban development condition exceeds surface water availability. However, the awareness of groundwater functioning is necessary to be understood and incorporated in any development plan of megacities as Mexico City (and other major cities worldwide), which is characterized by thick and extensive aquifer units, to avoid inducing vertical inflows of deep groundwater flow systems that might deteriorate the quality of the originally obtained groundwater supply.

From its establishment, water needed for the development of Mexico City was originally provided by groundwater sources (springs). As time advanced, both development and population increased, and as a result water supply needs were mainly overcome with groundwater usage. Abstraction rate had a relatively stable increase from 1900 to the 1940s (from ≈2 to ≈4.5 m^3/s, respectively); however, starting from 1940s, a rampant increase in water needs reached a claimed water supply in excess of 60 m^3/s by the end of the twentieth century, where about 76% was obtained from groundwater sources (Mexico basin ≈67.0% and Lerma basin ≈9.4%) and 23.6% from surface waters (Mexico basin ≈ 2.3% and Cutzamala basin ≈21.3%). This steady abstraction has produced depletion in the water-table level from artesian conditions to a depth of more than 50 m; a consequent response of the flow system to such increase in groundwater extraction was perceived in the changing of water quality.

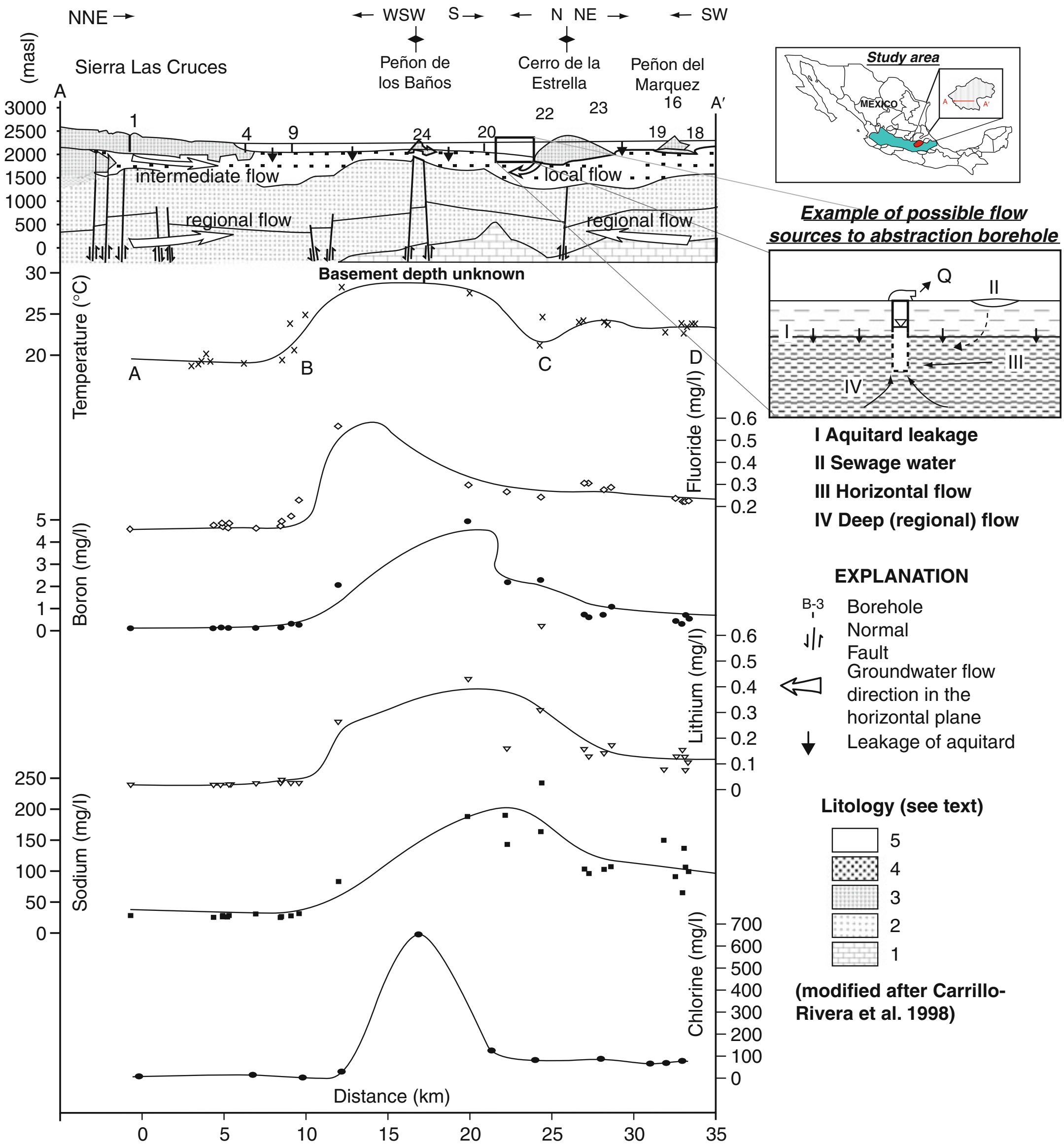
Groundwater Quality Response

Mexico City is located on a plain within a natural closed basin where its 80% of water supply was provided by groundwater since the beginning of the twenty-first century. The groundwater supply has generally been of good quality for almost a century, from the late 1840s when the first borehole was constructed to the early 1970s when the extracted water quality drastically changed. Initially, groundwater abstraction was obtained mainly through artesian boreholes at depths of about 100 m. As technology advanced, construction of deeper boreholes was feasible, and the turbine pump was employed to solve the no overflowing conditions; deep boreholes went down to about 400 m, which is their current depth beneath the foundations of the city. Reported groundwater salinity problems in the northwest of Mexico City started by 1976, with an increase of about 300% in salinity; for example, Cl, Fe-Mn, and TDS increased from 32 mg/L, 0.1 mg/L, and 230 mg/L, respectively, to 55 mg/L, 0.45 mg/L, and 750 mg/L, respectively.

An understanding of the changing of the chemical groundwater quality as a result of the continuous increase of abstraction in the Mexico basin (Fig. 2) may be reached when considering that thick aquifer units and recharge conditions are directly responsible for the presence of contrasting groundwater flows traveling at different depths with different path lengths. Intensive abstraction with partially penetrated boreholes also contributes to water quality changes through vertical ascent.

Geological Framework

Geologically, Mexico City is located on a plain in a closed basin situated above a graben structure developed during the Oligocene, where a thick succession of volcanic and lacustrine materials was deposited. During the Oligocene, the basin drained to the south. This drainage outlet was closed during the Pleistocene from a series of volcanic activities [10] that formed the Chichinautzin Sierra to the south. The closure was part of an intense regional volcanic activity that formed the Trans-Mexican Volcanic Belt (TMVB), which now covers Central Mexico from east to west with a length of about 1100 km and a width of 400 km. The Mexico basin is located at the center of TMVB (Fig. 2). The extrusive events have occurred since 700,000 years before present until the Quaternary; during the later stage, volcanism



Groundwater Salinity Due to Urban Growth, Fig. 2 Groundwater quality evolution along a section located in the south of the Mexico basin. Inset: Location of the basin in the trans-Mexican Volcanic Belt

allowed a substantial heterogeneous layer of ash, interbedded with extensive alluvial and lacustrine deposits, to accumulate in the various lakes formed after the closure of the basin. The full volcanic succession consists of a Middle Tertiary volcanic unit that includes clastic material as well as basalt, rhyolite, and andesite rock units with a total thickness of 3000 m. These units are partially covered by Lower Pliocene deposits characterized by lacustrine and pyroclastic material with a total thickness of 600 m. These rocks are overlaid by Plio-Quaternary rocks consisting mainly of basaltic andesite, andesite, and pyroclastic material that dominate the highlands and are interstratified with contemporaneous alluvial and lacustrine deposits; their total thickness on the plain is over 1000 m. Recent volcanic episodes and regional tectonic events have produced an intensive vertical fault structure identified to be crossing beneath the plain [10].

These volcanic rocks overlie a regional Cretaceous sequence of calcareous deposits that consist of limestone, sandstone, and shale outcropping beyond the southern and northern basin limits; this deepest and oldest identified geological unit of 1500 m of total thickness is at the center of the basin. Deep drilling encountered the upper limit of this unit at 1581 m below ground surface. Karstic and fractured features are reported based on partial and total loss of drilling fluid and the drop of about 8 m of the drilling tools [28].

Recent deposits cover the entire plain, comprising some 600 m of fluvial and alluvial deposits. A lens of fine material (microfossils, volcanic ash, and, to a minor extent, lacustrine clay interbedded with sand, silt, and occasionally gravel) is included within these deposits; it outcrops in Lake Texcoco, underlying the rest of the plain except at the margins; in general, these deposits contain [19] high Na (up to 9200 ppm) and Mn (up to 1920 ppm). Total thickness increases gradually from the edge of the plain to Lake Texcoco, where it attains approximately 300 m. This unit forms a main aquitard saturated with water of heterogeneous quality, which locally acts as a confining or semi-confining unit to the aquifer material beneath. This aquifer unit includes the extrusive and sedimentary rocks that extend regionally over Central Mexico with a thickness of more than 4000 m, and it lays above a basement rock that has not been physically described.

The Mexico basin is one of the largest of series of closed basins located in the Mexican Trans-volcanic Belt [28]. The basin has a drainage extent of about 9600 km^2; however, the actual part covered by the city and its extended area is estimated to 2400 km^2. An abrupt relief of the surrounding mountains, with altitudes in excess of 5000 m amsl, slopes toward the flat-lying center of the basin to an altitude of approximately 2230 m amsl.

Climate Framework

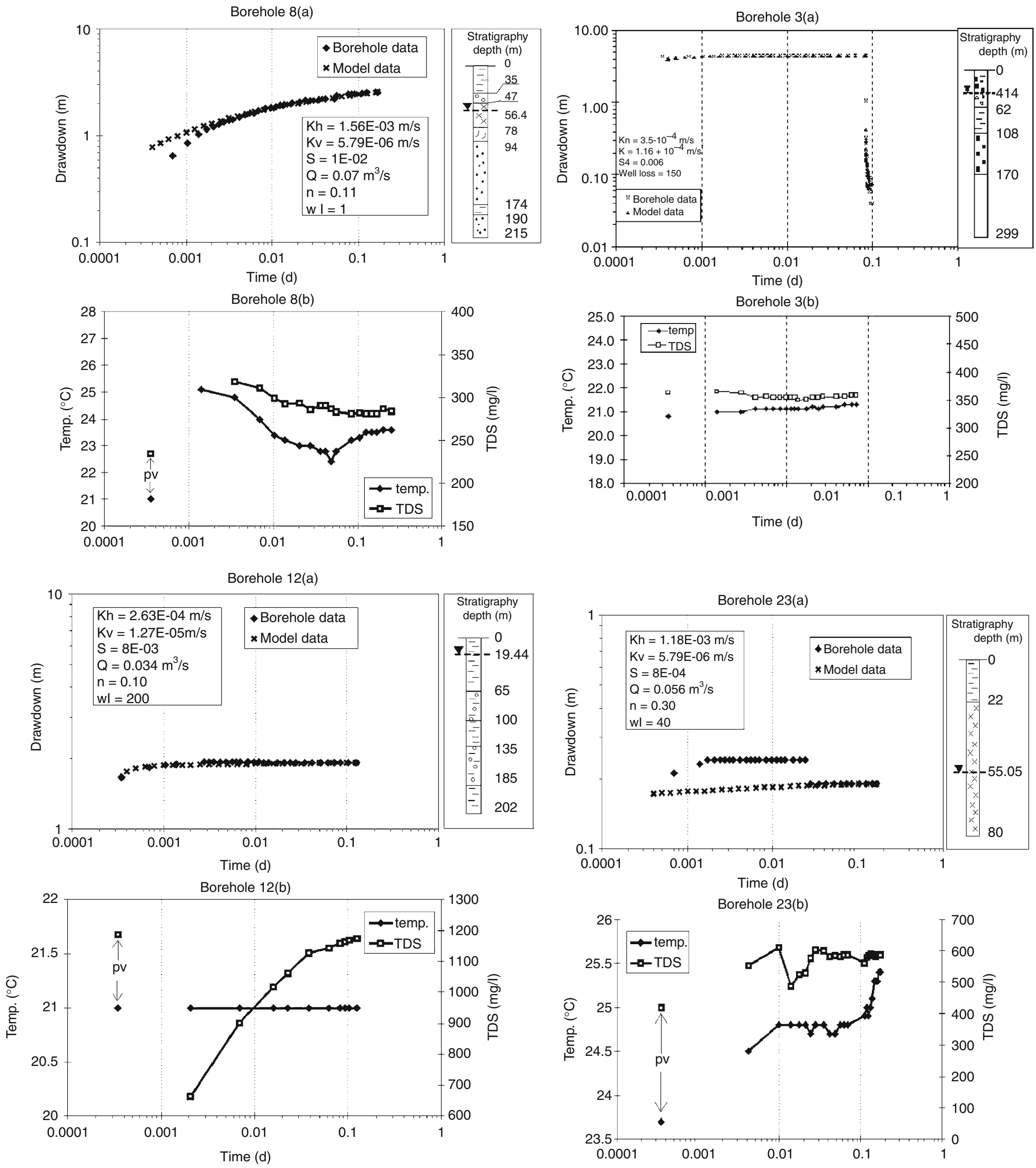
The present climate conditions where Mexico City is located are those of subtropical but with a significant variation in the mean annual precipitation across the basin. Precipitation ranges from 1800 mm in the southern Sierra Chichinautzin, 1100 mm in the west (Sierra Las Cruces), and 1200 mm in Sierra Nevada to 600 mm toward the remnant Lake Texcoco toward the center of the basin. The rainy season is mainly in the summer months. The mean annual temperature is from 12° to 14 °C in the sierras of Chichinautzin and Las Cruces and 15 °C in Lake Texcoco site. The potential annual evapotranspiration is around 1400 mm. Rainfall produces negligible surface runoff on the permeable rocks of the Sierra Chichinautzin to the south of the city. In contrast, substantial runoff is generated in Sierra La Cruces to the west of the city, a condition that has caused severe historical flooding.

Water Quality Response to Groundwater Abstraction

Mexico City reportedly abstracts more than 50 m^3/s of groundwater from its basin to supply different water needs. An action plan for positive environmental response to groundwater abstraction should rely on an understanding of the flow systems functioning to assist in a sustainable planning of groundwater management by defining which particular flow is more vulnerable to intensive local abstraction in time and space.

Four field cases of the Mexico City basin [8] were examined; the interference of particular nearby boreholes was not incorporated. The first case presented by borehole 3 in Fig. 3 shows abstraction (and recovery test) data that illustrate the local confining characteristics of the geological environment manifested as the classical drawdown-time response; the evidence of the horizontal flow is reflected on the relatively steady values of TDS and temperature of obtained water with abstraction time.

The case of borehole 8 shows a response traditionally considered to be the only possibility of chemical water quality change in this basin, portraying the acknowledged leaky effect from the semi-confining unit located above the level of groundwater abstraction. There is an inflow of cold water and withdrawal from this unit



Groundwater Salinity Due to Urban Growth, Fig. 3 Different responses to groundwater abstraction: Case 1 (borehole 3) confined conditions, case 2 (borehole 8) semi-confined conditions, case 3 (borehole 23) ascending groundwater flow, and case 4 (borehole 12) leakage effect

producing water with lower temperature. As abstraction progresses, inflow water is expected to reach observed previous values (pv) when the borehole was stopped to initiate the test, at about 21 °C. The TDS decrease shows a change in the chemical characteristics of inflow water due to leakage from the semi-confining unit.

Drawdown-time data in borehole 23 would appear to also suggest a standard leaky response where the drawdown rate is diminished along the duration of the test; however, the increase of groundwater temperature may imply that water is induced through vertical ascent of intermediate or regional flows traveling beneath.

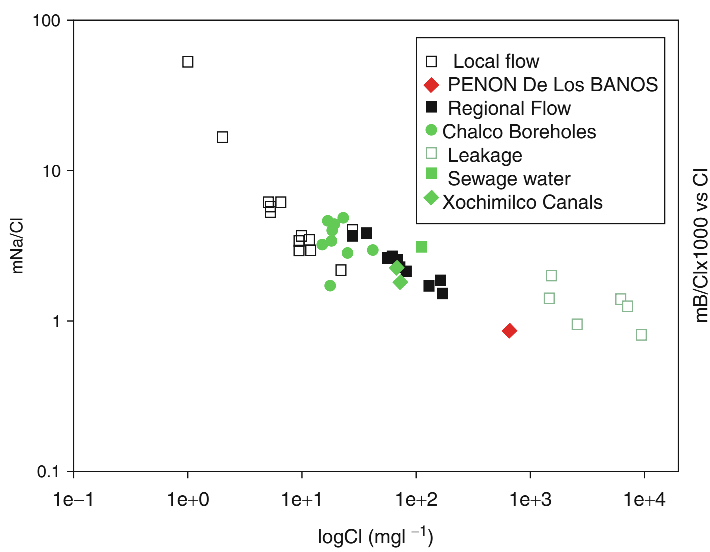
The response of the increase in TDS and constant temperature of borehole 12 suggests that obtained water is derived from standard semi-confined conditions, where aquitard material above and below the tapped aquifer strata provides a steady response in terms of salinity values. Water is derived from the same shallow depth, keeping temperature at a steady value of 21 °C.

The overall response of each of these boreholes in terms of obtained water quality results from the combined effect of (1) hydrogeological characteristics of tested aquifer units (i.e., lack or presence of confined conditions, geological structure, Kv/Kh ratio), (2) discharge yield, (3) effect of drawdown on the local hydraulic potential (i.e., relative thickness among flow systems of different densities), (4) travelling path length of any flow system to the level of abstraction in the particular borehole, (5) borehole construction design as related to the stratigraphic units, and (6) position of the water-table. All factors should be further studied to better plan actions to improve the quality of obtained water and to assess possible borehole techniques to better manage flow control. Suggestions on techniques to be applied are proposed in Ref. [8].

As suggested in Fig. 3, the response of borehole 23 is implied to be general to the prevailing long-term abstraction yield. The increase over time of groundwater abstraction in the basin of Mexico has reached far more than 50 m^3/s, with individual boreholes extracting an average of 60 L/s. The effect of the partial penetration of production boreholes is that of radial flow in 3-D; relative shallow abstraction levels in the Mexico City plain generate an ascending flow from deeper levels. This can be depicted in the presence of the inflow of water of contrasting chemical quality as well as of higher temperature (Fig. 2). This condition appears when considering the graph of Fig. 4, where mNa/Cl-log Cl suggests that extracted shallow water (local flow) is evolving to water representing the former thermal spring of Peñón de los Baños (end-member). The inducement of regional flow with different quality into the supplied waters increases the concentration of obtained salinity in the water. The vertical components of ascending flow are key factors of observed water quality changes. Understanding water quality evolution could enhance better control processes of regional flow inducement and assist in predicting where and how waters of more desirable quality might be obtained as suggested in Ref. [6].

Vertical Flow Control to Prevent Water Quality Deterioration

The identification and understanding of abstracted groundwater quality changes resulting from contrasting flow systems supplying a borehole is



Groundwater Salinity Due to Urban Growth, Fig. 4 Relation of the content of mNa/Cl and Cl for selected boreholes tapping local flows and water mixture representing local and regional end-member of former spring of Peñón de los Baños one tool for possible control of obtained water quality [6, 8]. These authors interpreted resulting field tests carried out in partially penetrating boreholes tapping a > 2000-m fractured volcanic (rhyolitic) rock sequence and proposed a technique to induce the production of the required quality water to a particular abstraction borehole. The procedure involves the definition of a particular response of flow systems mixture under prevailing local abstraction and hydrogeological variables, such as vertical and horizontal hydraulic conductivities, porosity, hydraulic head response between cold and warm flows, and rate and time of abstraction. The response is basically defined through long-term aquifer tests that include measurement of standard drawdown-time (s-t) response plus electrical conductivity (EC) and temperature at the various t intervals of the duration of the test. Additional collection of water samples for chemical and isotopic analyses was carried out to support obtained results.

Available information on vertical flow in volcanic terrain with large thickness [8] is important in understanding the factors affecting groundwater flow distribution and obtained water quality in the vicinity of the abstraction borehole. An abstraction yield will present a particular water quality, depending on the construction design of the borehole, lithology, associated flow systems, and their physico-chemical components. Obtaining a specific water quality is difficult to implement; this difficulty may be overcome by understanding the control of obtained water quality related to tests carried out in each particular borehole. Methods include an initial step-drawdown test, followed by a constant abstraction-rate aquifer test; both in which temperature and EC measurements must be conducted. The response of the borehole could be further defined and understood when full sampling and water analysis are carried out to define major, minor, and trace elements as well as stable isotopes to differentiate the various components of inflowing groundwater flow systems.

Future Directions

Understanding and defining the groundwater flow systems before planning for their abstraction is a basic rule that requires further attention in Mexico, especially when it comes to resolving responses related to groundwater quality deterioration and yield. The perception of groundwater behavior achieved in the second half of the twentieth century was obtained mainly based on the concept of shallow, thin aquifer units. This acquired knowledge provides a basis to define groundwater response and propose agreeable solutions for thick aquifer units.

Such powerful techniques have been applied in growing cities as Mexico City, where the increasing need for water has produced changes in the flow system regime. The continuous development of an interdisciplinary approach as the flow system theory [25] is useful for defining the response and acquiring solutions to the diversity of parameters involved in extracting water of the desired quality. The application of this theory is a powerful tool as it provides answers that are in agreement among the diverse disciplines involved.

Increase in population and economic activities in large cities relies on developing local groundwater flow systems that are often hydraulically linked to regional flow systems when the lithological sequence has a thickness greater than 1000 m. An interdisciplinary understanding must be reached where groundwater flows of different chemical compositions and temperatures travel in a stratified mode. A 3-D modeling is useful for a clear understanding of the problem and introducing solution for thick aquifer units where regional flow maybe induced. Usually, current analysis carried out is made solely on the horizontal scale; however, vertical components of flow need to be acknowledged. Aquifer tests, as well as other common hydrogeological tools, may define groundwater response and its functioning by incorporating the observation of the quality and temperature of extracted groundwater.

Changes in the obtained groundwater quality due to an increase of abstraction rate are reflected by the chemistry and temperature of water flowing from a different level to that of the targeted aquifer unit. Water quality changes could become key factors for density variation identification during water abstraction, indicating the presence of semi-confining conditions through the rate of change of drawdown over time (s-t) to that of inflow of water from regional systems at depth.

The growth of cities located above thick aquifer units (i.e., >1000 m), accompanied with increasing water needs, makes the identification of local, intermediate, and regional groundwater flow systems essential. Two main issues arise in the analysis of thick regional aquifers in Central Mexico: (1) the position of basement rock, which in turn implies the coverage of the aquifer unit involved, and (2) effects due to groundwater chemistry change, such as water density variations, affecting actual hydraulic heads, and flow direction. These variables become more important in thick aquifer units.

Bibliography

Primary Literature

1. AIC (1995) El Agua y la Ciudad de México. Academia de la Investigación Científica, Academia Nacional de Ingeniería, Academia Nacional de Medicina, National Academy of Sciences (through the National Research Council), 364 pp
2. Ángeles-Serrano G, Perevochtchicova M, Carrillo-Rivera JJ (2008) Posibles controles hidrogeológicos de impacto ambiental por la extracción de agua subterránea en Xochimilco, México. J Lat Am Geogr 7(I):39–56
3. Bouwer H (1978) Groundwater hydrology, McGraw-hill series in water resources and environmental engineering. McGraw-Hill, Sydney. 480 pp
4. Carrillo-Rivera JJ (1998) Monitoring of exploited aquifers resulting in subsidence, example: Mexico City. In: Van Lanen HAJ (ed) Monitoring for groundwater management in (semi-)arid regions. Studies and reports in hydrology, vol 57. UNESCO, Paris, pp 151–165
5. Carrillo-Rivera JJ, Angeles-Serrano G, Hernández GG, Hergt T (2002) Estudio de hidrología subterránea sobre el área de Xochimilco, Distrito Federal. In: Programa rector de manejo del área natural protegida, ejidos de Xochimilco y San Gregorio Atlapulco. Secretaría del Medio Ambiente del Gobierno del Distrito Federal, Distrito Federal
6. Carrillo-Rivera JJ, Cardona A, Edmunds WM (2002) Use of abstraction regime and knowledge of hydrogeological conditions to control high-fluoride concentration in abstracted groundwater: San Luis Potosi basin, Mexico. J Hydrol 261:24–47
7. Carrillo-Rivera JJ, Cardona A, Huizar-Alvarez R, Graniel E (2008) Response of the interaction between groundwater and other components of the environment in Mexico. Environ Geol 2:303–319
8. Carrillo-Rivera JJ, Cardona A (2008) Groundwater flow system response in thick aquifer units: theory and practice in Mexico. Selected papers, XXXIII-IAH internacional congress, Zacatecas, México, vol 12. Asociación Internacional de Hidrogeólogos, Editorial Balkema, Taylor & Francis, Leiden, pp 25–46
9. CAVM (Comisión De Aguas Del Valle De México) (1966) Datos del Valle de México, Período 1959–1963. Bol Mec Suelos 4
10. De Cserna Z, De la Fuente DM, Palacios NM, Triay L, Mitre SLM, Mota PR (1988) Estructura geológica, gravimétrica, sismicidad y relaciones neotectónicas regionales de la Cuenca de México, Institute of Geology, UNAM, México. Boletín 104:71pp. (4 maps)
11. Edmunds WM, Carrillo-Rivera JJ, Cardona A (2002) Geochemical evolution of groundwater beneath Mexico City. J Hydrol 258:1–24
12. Foster S, Garduño H, Evans R, Olson D, Yuan T, Zhang W, Han Z (2004) Quaternary aquifer of the North China plain - assessing and achieving groundwater resource sustainability. Hydrogeol J 12:81–93
13. Freeze A, Cherry J (1979) Groundwater. Prentice, London. 490 pp
14. Graniel-Castro D, Morris LB, Carrillo-Rivera JJ (1999) Effects of urbanization on groundwater resources of Mérida, Yucatan. Environ Geol 37(4):303–312
15. Herrera I, Alberro J, León JL, Chen B (1974) Análisis de los asentamientos para construcción de los lagos del Plan Texcoco. UNAM, Instituto de Ingeniería, Internal Report No. 340
16. Huizar-Alvarez R, Carrillo-Rivera JJ, Angeles-Serrano G, Hergt T, Cardona A (2004) Chemical response to groundwater extraction southeast of México City. Hydrogeol J 12:436–450
17. Llamas MR (1999) La inserción de las aguas subterráneas en los sistemas de gestión integrada. Bol Geol Min 110(4):253–370
18. Marsal RJ, Graue R (1969) The subsoil of lake Texcoco. In: Carrillo N (ed) The subsidence of Mexico City and Lake Texcoco project. Secr. de Hacienda y Crédito Publico, México, pp 167–203
19. Mesri G, Rokhsar A, Bohor BF (1975) Composition and compressibility of typical samples of Mexico City clays. Geotechnique 25:527–554
20. Ortega GA, Cherry JA, Rudolph DL (1993) Large-scale aquitard consolidation near Mexico City. Ground Water 31:708–718
21. Rivera A, Ledoux E (1991) Non-linear modelling of groundwater flow and total subsidence in the Mexico City aquifer-aquitard system. In: Proceedings of the fourth international symposium on land subsidence, Houston, May 1991, IAHS Publication No. 200, pp 45–58
22. Rudolph DL, Cherry JA, Farvolden RN (1991) Field investigations and solute transport in a lacustrine aquitard near Mexico City. Water Resour Res 27:2187–2201
23. Theis CV (1935) The relation between the lowering of the piezometric surface and rate and duration of discharge of a well using ground water storage. Trans Am Geophys Union 16:519–524
24. Tóth J (1995) Hydraulic continuity in large sedimentary basins. Hydrogeol J 3:4–16

25. Tóth J (1998) Groundwater as a geological agent: an overview of the causes, processes, and manifestations. Hydrogeol J 7:1–14
26. Tóth J (2008) From the artesian paradigm to basin hydraulics. Eötvös Loránd University, Institute of Geography and Earth Sciences, Budapest
27. Vázquez-Sánchez E (1995) Hidrogeología del Acuífero de la Ciudad de México. MSc thesis en groundwater. UACPyP, Posgrado en Geofísica, Instituto de Geofísica, UNAM
28. Vázquez-Sanchez E, Jaimes-Palomera R (1989) Geología de la Cuenca de México. Geofis Int 28(2):133–190

Books and Reviews

Cardona A, Carrillo-Rivera JJ, Huizar-Alvarez R, Graniel-Castro E (2004) Salinization in coastal aquifers of arid zones: an example from Santo Domingo, Baja California Sur, Mexico. Environ Geol 45(3):350–366

Cardona A, Carrillo-Rivera JJ, Castro-Larragoitia GJ, Graniel-Castro EH (2008) Combined use of indicators to evaluate waste-water contamination to local flow systems in semi-arid regions: San Luis Potosi, Mexico. Selected papers XXXIII-IAH internacional congress, Zacatecas, México, vol 12, Asociación Internacional Hidrogeólogos, Ed Balkema, Taylor & Francis, Leiden, pp 85–104

Carrillo-Rivera JJ (1988) The Sierra Madre occidental. In: The geology of North America, hydrogeology, vol 0-2. The Geological Society of North America, Boulder, pp 87–88

Carrillo-Rivera JJ (1993) The hydrogeology of the San Luis Potosi area, Mexico. Ground Water J Assoc Ground Water Sci Eng (Abstracts Students Section, Marzo-Abril, p 330)

Carrillo-Rivera JJ (1998) Monitoring of exploited aquifers resulting in subsidence, example: Mexico City. In: Monitoring for groundwater management in (semi-) arid regions. Studies and reports in hydrology, vol 57. UNESCO, Paris, pp 151–166

Carrillo-Rivera JJ (2000) Application of the groundwater-balance equation to indicate interbasin and vertical flow in two semi-arid drainage basins, Mexico. Hydrogeol J 8(5):503–520

Carrillo-Rivera JJ (2002) Aquifer evaluation. In: Moore JE (ed) Field hydrogeology. Lewis, Boca Raton, pp 79–87. ISBN:1-56670-587-8. (Chap 6)

Carrillo-Rivera JJ (2003) Lack of a conceptual system view of groundwater resources in Mexico, editor's message. Hydrogeol J 11(5):519–520

Carrillo-Rivera JJ (2004) Congreso conjunto: XXXIII asociación internacional de hidrogeológos y 7° asociación Latinoamericana de hidrología subterránea para el desarrollo, Zacatecas, Zac. Méx. Rev Invest Geogr 55:175–177

Carrillo-Rivera JJ, Adrián Ortega G (2008) Special number of environmental geology. In: XXXIII international hydrogeologic congress on groundwater flow understanding from local to regional scales, vol 55. Springer, Berlin, pp 235–464. ISSN: 0943-0105

Carrillo-Rivera JJ, Cardona A (2002) Capítulo 14. In: Michael Price (ed) Agua subterránea. Chapman and Hall. Published by Limusa, Noriega Editores

Carrillo-Rivera JJ, Cardona A (2008) Groundwater flow system response in thick aquifer units: theory and practice in Mexico. Selected papers, XXXIII-IAH internacional congress, Zacatecas, México, vol 12. Asociación Internacional de Hidrogeólogos. Editorial Balkema, Taylor & Francis, Leiden, pp 25–46

Carrillo-Rivera JJ, Ortega Guerrero MA (eds) (2008) Groundwater flow understanding: from local to regional scales Selected papers, XXXIII-IAH international congress, Zacatecas, México, vol 12. Asociación Internacional de Hidrogeólogos Editorial Balkema, Taylor & Francis, Leiden, p 186. ISBN-13: 978 0 203 94579 7

Carrillo-Rivera JJ, Clark DI, Fritz P (1992) Investigating recharge of shallow and paleo-groundwater in the Villa de Reyes Basin, SLP, Mexico, with environmental isotopes. Appl Hydrogeol Off J Int Assoc Hydrogeol 4:35–48

Carrillo-Rivera JJ, Cardona A, Moss D (1996) Importance of the vertical component of groundwater flow: a hydrochemical approach in the valley of San Luis Potosí, Mexico. J Hydrol 185:23–44

Carrillo-Rivera JJ, Cardona A, Hergt T (2001) Inducción de agua termal profunda a zonas someras: Aguascalientes, México. Rev Latinoam Hidrol 1(1):41–53

Carrillo-Rivera JJ, Varsányi I, Kovács LÓ, Cardona A (2007) Tracing groundwater flow systems with hydrogeochemistry in contrasting geological environments. Water Air Soil Pollut 184:77–103

Carrillo-Rivera JJ, Cardona A, Edmunds WM (2007) Groundwater flow functioning in arid zones with thick volcanic aquifer units: North-Central Mexico. In: international symposium on advances in isotope hydrology and its role in sustainable water resources management (HIS-OIEA) Proc, 21–25 May 2007, vol 1, Vienna, pp 199–211

Facundo-Castillo JR, Carrillo-Rivera JJ, Antigüedad-Auzmendi I, González Hernández P, Leláes-Diaz R, Hernández-Diaz R, Cáceres-Govea D, Hernández-Santana JR, Suárez-Muñoz M, Melán-Rodriguez C, Rodríguez-Piña M (2008) Chemical and geological control of spring water in eastern Guaniguanico mountain range, Pinar del Rio, Cuba. Environ Geol 55(2):247–267

Fagundo J, Carrillo-Rivera JJ, Antigüedad I, González P, Peláez R, Suárez M, Melián C, Hernández R, Cáceres D (2005) Caracterización hidrogeoquímica del sistema de flujo local-regional de la Sierra del Rosario (Cuba). Rev Latinoam Hidrol 5:75–90

Gutiérrez de MacGregor MT, Carrillo-Rivera JJ, Valdez-Quijada R (1997) Impact of Mexican and USA policies in urban growth and natural resources in the northern border of Mexico, vol 15. Latin American Studies, Nihon-Burajiru Chou Kyokai, Tokio, pp 49–62

Huizar-Alvarez R, Hernández-García G, Carrillo-Rivera JJ (2009) Simulation of the effects from the groundwater flow on the hydrological balance of the Tecocomulco lagoon, Central Mexico. Open Environ Sci J 3:1–13

Huizar-Álvarez R, Hernández GG, Carrillo-Martinez M, Carrillo-Rivera JJ, Hergt T, Ángeles-Serrano G (2003) Geologic structure and groundwater flow in the Pachuca-Zumpango sub-basin, central Mexico. Environ Geol 43:385–399

Martinez SE, Carrillo-Rivera JJ (2006) Socio-economic constraints of groundwater in capital La Rioja, Argentina. Environ Geol 49(6):875–886. https://doi.org/10.1007/s00254-006-0183-7

Segovia N, Taméz E, Peña P, Carrillo-Rivera JJ, Acosta E, Armienta MA, Iturbe JL (1999) Groundwater flow system in the valley of Toluca, Mexico: an assay of natural radionuclide specific activities. Appl Radiat Isot 50:589–598

Van-Lanen AJH, Carrillo-Rivera JJ (1998) Framework for groundwater monitoring in (semi-) arid regions. In: Monitoring for groundwater management in (semi-) arid regions. Studies and reports in hydrology, vol 57. UNESCO, Paris, pp 7–20

Mine Water Inrush

Qiang Wu[1], Shuning Dong[2], Bo Li[3] and Wanfang Zhou[4]
[1]National Engineering Research Center of Coal Mine Water Hazard Controlling, China, China University of Mining and Technology, Beijing, China
[2]Xian Branch, China Coal Technology and Engineering Group Corp., Shannxi, China
[3]Key Laboratory of Karst Environment and Geohazard Prevention, Guizhou University, Guiyang, China
[4]ZeoEnvironmental, LLC, Knoxville, TN, USA

Article Outline

Glossary

Caving zone The completely collapsed part of the overlying rock in response to mining. The rock has the characteristics of irregularity, bulking property, and poor density. The caving zone is not effective in preventing water from entering the mining area.

Draft zone of confined water Refers to the upward penetration height of water from the confined aquifer into the geologic barrier along fissures or fractures in the floor.

Floor rock pressure failure zone Due to mining-induced stresses, the floor rock is continuously destroyed, resulting in a rock failure zone. Hydraulic conductivity increases significantly in the altered rock formation.

Goaf water Water in old excavated roadways, abandoned mines, and caves. The water pressure is generally large. Once exposed, the goaf water can suddenly gush into active mining areas, resulting in mass destruction. Water gushing may be accompanied with emission of harmful gases. The gushing water duration is typically short.

Mine water inrush classification A system of classifying water inrushes into different types based on the characteristics of mine water inrush, water inrush pathway, water inrush mode, and hazard form.

Mine water inrush An incident in which a large amount of water suddenly poured into mines through water-conducting pathways from aquifers, water-rich caves, goafs, and other strong water-bearing geological bodies.

Water-filling channel The pathways of water flowing into mines, such as faults, rock pores, fissures, conduits, caves, and collapse columns.

Water-filling source Sources of water flowing into mines. They mainly include atmospheric precipitation, surface water, groundwater, and goaf water.

Water-filling strength The amount of groundwater flowing into mines. This parameter is used to reflect complexity of mine hydrogeological conditions.

Definition of Mine Water Inrush and Its Importance

Mine water inrush is a phenomenon in which a great volume of groundwater suddenly gushes into underground workings when water-bearing media, such as high-pressure confined aquifers, water-flowing fractures, water-rich karst caves, and goafs, are exposed during tunneling or mining. Mine water inrush generally occurs dramatically and can submerge tunnels in a short period of time, jeopardize mine production, and cause casualties. As the largest coal-producing country

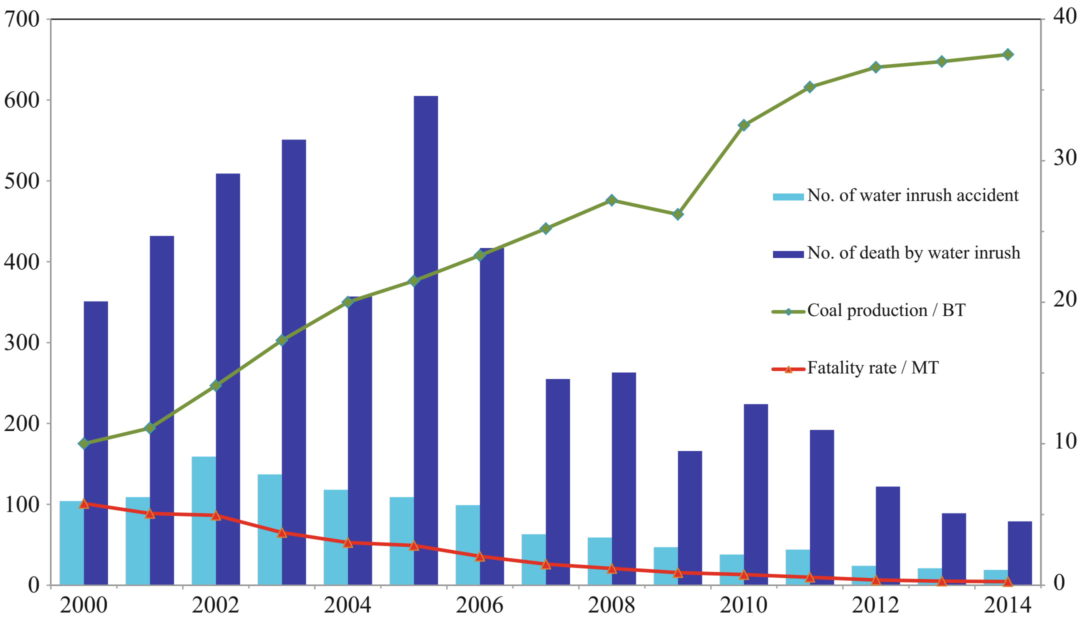
J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_967

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_967-1

in the world, China has abundant coal resources which encompass a vast geographical distribution. Furthermore, China is a land formed by multiple tectonic plates that were spliced through numerous tectonic movements, showing complex hydrogeological conditions of ore deposits. Therefore, China is one of the countries that is most seriously impacted by mine water inrush in the world. According to the statistics of China's State Administration of Work Safety, mine water inrush has become the second most serious disaster after gas explosion in serious and major coal mine accidents. At present, more than 2.5×10^{10} tons of coal reserves are threatened with water inrush, mainly in developed industrial areas, such as North China, East China, and Southern China where coal reserves account for about 70% of the national coal resource. Figure 1 shows the number of water-related hazards and casualties in coal mines of China between 2000 and 2014. In recent years, with the progress of science and technology, mining equipment and mining technology in coal mine production and construction have been greatly improved. Furthermore, the number of accidents and casualties caused by mine water inrush has been declined in general, but the casualties and property losses induced by accidents are still serious. Therefore, it is practically significant to clearly recognize the current situation of mine water inrush, study countermeasures for preventing water inrush, and take effective engineering controls. Knowledge on water inrush ensures safety production of coal mines, frees coal reserves from threats of water disaster, and guarantees sustainable and stable development of society and economy.

Introduction

Mine water inrush studies suggest that different types of water inrush generally call for corresponding prevention and control technologies. Therefore, it is necessary to classify water inrush according to distinct characteristics of mine water inrush. Due to diversity and complexity of mine water inrush, the systematic, overall, and comprehensive classifications of mine water inrush are rarely reported at present [1–4]. In fact, because scientific classification of mine water inrush is a huge systematic classification



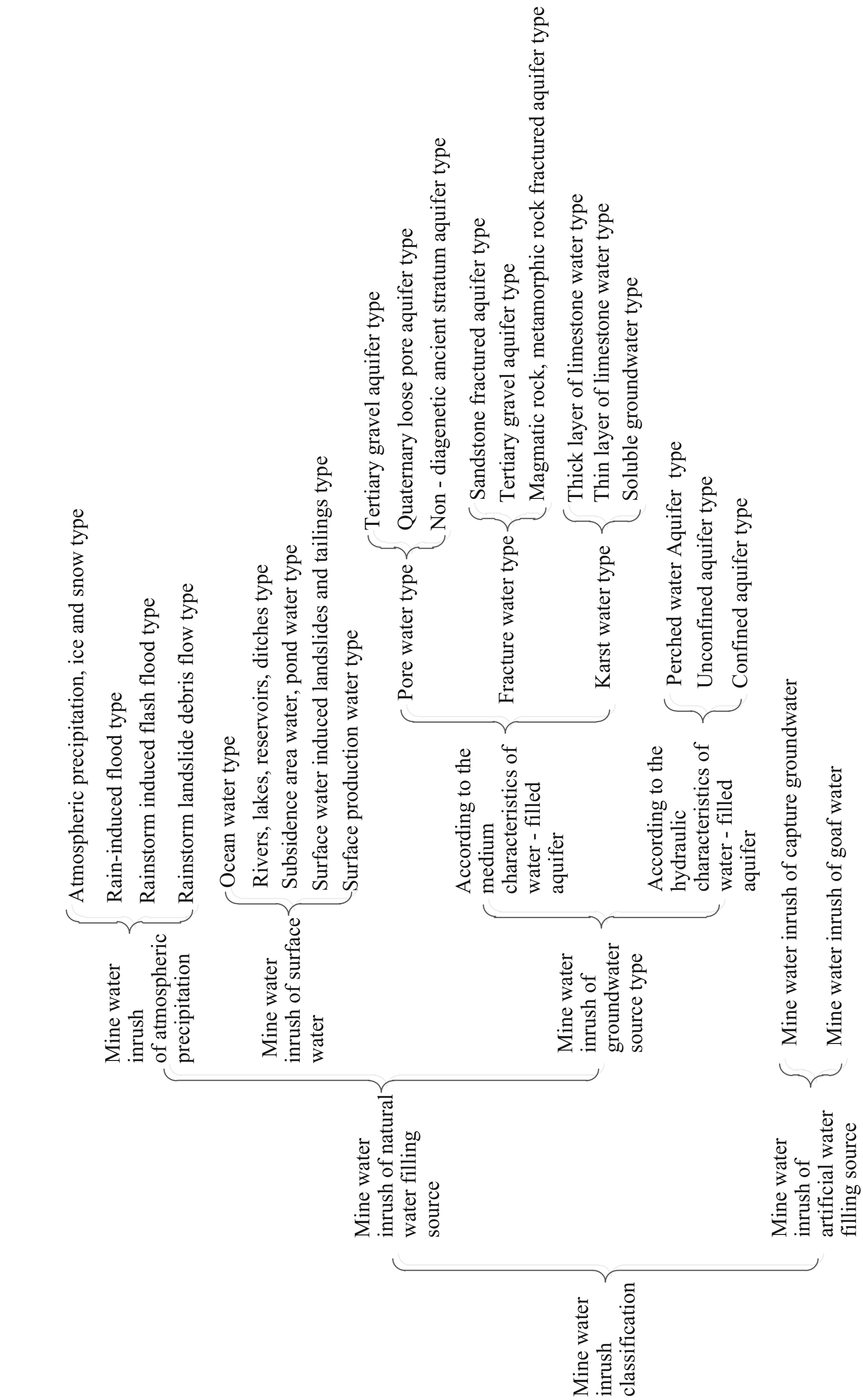
Mine Water Inrush, Fig. 1 Number of water-related hazards and casualties in coal mines of China between 2000 and 2014 project, it needs to develop qualitative, quantitative, and combined classification methods. Furthermore, it requires a large amount of field data and cases of water inrush as well as supports of relevant scientific theories [5, 6]. Therefore, the classification of mine water inrush has important theoretical and practical value for basic theory research, investigation and exploration, evaluation and prediction, design of detection equipment, prevention and control technologies, and comprehensive utilization of mine water [7–9].

Principles for Classification of Mine Water Inrush

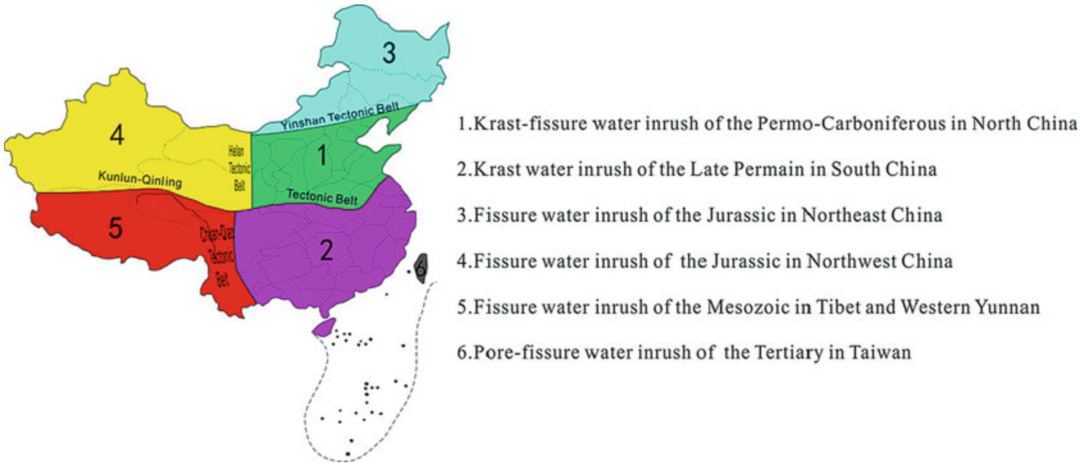
The division bases need to be constructed prior to classification of mine water inrush. According to case histories of water inrush and specific characteristics, water-filling source, water-filling channel, water-filling strength, harm forms, economic loss and casualties, and time-varying characteristics are used as the bases for classifying water inrush [10–12]. Water-filling source refers to all of the water sources that exist in and show hydraulic connections with ore bodies and surrounding rock strata that can cause continuous mine water inrush (gush) during the mining process. Water-filling channel is the path for these water sources entering into pits. Harm forms indicate pit water inrush (gush) showing characteristics, such as abnormal temperature and corrosivity. In addition, economic losses and casualties measure the magnitude of economic losses and the number of casualties directly resulting from mine water inrush, respectively. Time-varying characteristics show the temporal relationship between mine water inrush (gush) and the progress of mining engineering [13–15].

Main Types of Mine Water Inrush

1. Classifying mine water inrush in accordance with water-filling source. Based on the nature of water-filling source, water inrush can be divided into water inrush of natural and artificial sources (Fig. 2). The natural water-filling source includes the mine water inrush directly recharged by atmospheric precipitation, mine water inrush of surface water-filling sources (large-scale surface water bodies, such as seas, lakes, rivers, pools, bogs, and reservoirs), and water inrush of groundwater source [16, 17]. Among them, the groundwater-source-type water inrush can be divided into unconsolidated pore water-filling source type, bedrock fissure water-filling source type, and karst water-filling source type in soluble karst rocks through medium characteristics of water-filling aquifers [18, 19]. According to the classification criteria, China's coal mine groundwater inrush types can be obtained from Fig. 3. According to hydraulic characteristics of water-filling aquifers, the groundwater-source-type water inrush can be divided into perched water-filling source type, phreatic water-filling source type, and confined water-filling source type. In addition, the artificial source water inrush consists of mine water disasters of groundwater captured source type and goaf water source type [20, 21].
2. The classification can also be carried out according to the location and contact relationship between minable seams and water-filling aquifers [22–24]. In accordance with the relative locations of minable seams and water-filling aquifers, water inrush can be classified as water inrush from coal roof, coal floor, and periphery water-filling source. Furthermore, based on the contact relationship between minable seams and water-filling strata, water inrush can be further divided into six types with direct and indirect roof water-filling source, direct and indirect floor water-filling source, and direct and indirect surrounding water-filling source [25, 26].
3. Water disasters can also be classified in light of the water-filling channel. Water-filling channel can be divided into natural and artificial water-filling source passages [27, 28]. The natural water-filling source passages include water inrush with point karst collapse column passage, linear fracture (fissure) zone passage, narrow strip concealed outcrop passage, plain fracture network (thinning area of partial plain aquifuge), and earthquake-induced passage [29]. The artificial water-filling source passages are divided into those with passages in roof caving fractured zone, roof cut caving fractured zone and roof pump caving zone, floor rock pressure failure zone, floor draft zone of



Mine Water Inrush, Fig. 2 Mine water inrush type classification according to their sources



Mine Water Inrush, Fig. 3 The types of coal mine water inrush in China

confined water, ground karst collapse zone, and poorly sealed boreholes [30–32].

4. Mine water inrush can be classified according to the harm forms into normal-temperature, moderate- to high-temperature, or corrosive water disasters [33].
5. One can divide mine water inrush in accordance with economic losses and causalities. In accordance with causalities or direct economic losses, mine water disasters can be divided into extremely large, very large, large, and general ones [34, 35].
6. Classification can be carried out according to time-varying characteristics as well. Based on time-varying characteristics, water inrush is divided into instant, hysteretic, skipping, or gradually varying ones [36, 37].

Main Characteristics of Different Types of Mine Water Disasters

1. Mine water inrush directly recharged by atmospheric precipitation. Atmospheric precipitation is the main supply source of groundwater, and all ore deposits filled with water are directly or indirectly related to atmospheric precipitation [38, 39]. The source of atmospheric precipitation described here indicates the only water source for direct water filling of ore deposits. There is a synchronous correlation or delayed correlation between the disaster time and the precipitation time. Moreover, catastrophability is related to precipitation and rainfalls and is generally proportional to rainfalls [40–43].
2. Water inrush of surface water. For ore deposits close to large-scale surface water bodies, such as seas, lakes, rivers, reservoirs, and pools, it is critical to clarify the influences of surface water under natural conditions and after mining on ore deposit mining. This is a key process of hydrogeological exploration in mining areas and hydrogeological work in mines [44]. Surface water is generally large in volumes. Once surface water forms hydraulic connection with mining activities and influencing ranges, the catastrophability is likely to rise [45, 46].
3. Water inrush of groundwater. This type of>disaster is complex. According to medium characteristics of water-filling aquifers, the water inrush can be divided into those with sources of pore water-filling unconsolidated sediments, bedrock fissure water filling, and karstic water filling in soluble rocks. Furthermore, perched water, phreatic water, and confined aquifer source water disasters are included in this type in accordance with hydraulic characteristics of water-filling aquifers [47, 48]. From the perspectives of water-bearing media and hydraulic characteristics, confined aquifers of karstic water show strong water abundance in general. Therefore, once such aquifers are connected due to mining activities and influencing ranges, the water inrush of karstic water source generally demonstrates the greatest catastrophability [49].

4. Water inrush of captured water. Because of mining, the groundwater depression cone constantly extends, and thereby mining activities strongly transform natural groundwater flow fields in mining areas. The new supply water source obtained in artificial groundwater flow fields is known as captured water source. The captured water source includes spring water in groundwater drainage areas, surface water (seas, lakes, and rivers), the neighboring aquifers in one side of drainage areas in groundwater flow zone in mining areas, and groundwater in the adjacent hydrogeological units. Therefore, the catastrophability is generally proportional to water abundance of the supply source [50–53].
5. Water inrush of goaf water. Because part of goaf remained open after mining, the goaf becomes filled with water in the late stage. If the edge of water bodies is mined, water in the goaf can suddenly gush into underground mines, causing mine water inrush accidents. According to statistics, this type of inrush has the largest number and strongest catastrophability in serious mine water disasters [54, 55]. Such a type of mine water inrush is unexpected with large amounts of water inflow, causing great damages. The goaf water is often acid and shows high concentration of hydrogen sulfide gas. However, due to the limited size of the water-storing space, the water flow may last for a short duration, and the water can be easily drained.
6. Mine water inrush of roof water. Mine water inrush of roof source occurs when mining activities and influencing ranges (caving and fissured zone and water-flowing structure) affect the aquifers overlying the ore body [56]. The catastrophability is directly related to the water abundance and connectivity of the overlying water-filling aquifers. Greater catastrophability results from stronger water abundance and connectivity of aquifers in the influencing ranges of mining activities [57].
7. Mine water inrush of floor water. Mine water inrush of floor source is triggered when the mining activities and influencing ranges (zone destroyed by mine pressure and water-flowing structures) affect aquifers underlying the ore body [58]. Similar to the water inrush of roof source, the catastrophability of floor water inrush is directly correlated with water abundance in and connectivity to the underlying aquifers. If aquifers in the influencing ranges of mining activities show remarkable water abundance and connectivity, the catastrophability of mine inrush is strong [59, 60].
8. Mine water inrush of periphery water. Such disasters result from mining activities and influencing ranges affecting the aquifers around the ore bodies. The water-filling sources can be direct or indirect. The catastrophability has a proportional relation with water abundance of surrounding water-filling aquifers and connectivity of fissures [61–63]. In general, direct water-filling sources in above three types of mine water inrush refer to the water source directly contacted with mined ore or water source which can be affected by roof water-flowing caving zone or floor rock pressure failure zone and thereby contacted with mined ore bodies. The indirect water-filling source indicates the water-filling source that enters mines by passing through water-resisting rocks through certain water-flowing structures or via leakage. It mainly distributes around mined ore body but does not directly contact with ore bodies or locates outside normal caving zone or zone destroyed by mine pressure [64].
9. Mine water inrush through natural water-filling passage. While mining ore body, various paths for water-filling source entering pits are referred to as water-flowing passages. Moreover, the mine water disaster caused by water gushing into pits through non-artificial water-flowing passages is called water inrush of natural water-flowing passage [65–68]. The characteristics are described as follows:
   1. As for water inrush through karst collapse column passages, the hydraulic relation of groundwater in coal series strata and different water-filling aquifers can be connected by karst collapse column passages, thus increasing catastrophability of such mine water disasters [69, 70].
   2. Water inrush through passages in linear fracture (fissure concentrated) zone mainly takes place in fault concentrated zone, fault

intersection point, fault convergence, or fault tip. The passages link the close hydraulic connection between water-filling rock strata, thus causing mine water disasters [71].

3. In view of water inrush of narrow strip concealed outcrop passage, according to the practical experience in China, the quaternary pore aquifer group is very likely hydraulically connected to the coal series and water-filling aquifer group of the thick carbonate formations at the narrow strip concealed outcrops. As a result, water disaster happens through narrow strip concealed outcrop passages, resulting in strong catastrophability.
4. Water inrush through passage in plain fracture networks (thinning area of local plain aquifuge). In the northern area of North China-type coalfield, stresses have been released through rock fracturing in the brittle water-resisting strata under multistage tectonic stresses in the geological history. Therefore, concentrated cracks and joints in different directions are present in the water-resisting strata. These fractures and joints form plain extended fracture networks with a planar distribution on the whole. With the increases of groundwater head difference in the upper and lower water-filling aquifer groups, such fracture networks form vertical water exchange in a plain leaky form and cause water disasters of plain fracture network passages [72].
5. Water inrush through earthquake-induced passage. When strong earthquakes occurred, fractures in different scales were formed near the epicenter by coupling of cyclic tension and compression of seismic forces with shear. Mine water inrush disaster occurs when fractures near the coal seams develop and connect with surrounding aquifers [73].

10. Mine water inrush through artificial water-flowing passages. The mine water inrush caused by water gushing into pits through artificial water-flowing passages is known as inrush of artificial water-flowing passage. Such type includes water inrush with passages in the following media [74–75]:
(1) Roof caving fractured zone
(2) Roof cut caving and fractured zone
(3) Roof pump caving fractured zone
(4) Zone destroyed by mine pressure of floor
(5) Floor draft zone of confined water
(6) Ground karst collapse zone
(7) Borehole with poor sealing quality

Subtypes (1), (2), and (3) are similar and caused by the fact that upper aquifers are connected due to roof rock damages triggered by mining activities. The difference is that caving fractured zone in (1) is mainly developed in horizontal or gently tilted strata, while caving zone in (3) is mainly developed in steep dip strata. Furthermore, cut caving fractured zone in (2) is formed in thick and extremely thick strata of sandstone or coarse sandstone with a large modulus of elasticity on the roof of coal seams. Caving does not happen in limited mining ranges. When caving occurs, it takes place in a large range, damaging roofs or floors. Similarly, subtypes (4) and (5) are similar. They are mine water disasters resulting from the fact that the lower aquifers are connected because of floor rock damages induced by mining activities. The differences lie in that zone damaged by mine pressure in (4) is formed in strata closely neighboring the lower ore bodies, while draft zone of confined groundwater in (5) develops in the top of lower aquifers of ore bodies. In large-scale water pumping and dewatering practices of karst water-filling deposits, surface karst collapse is well developed in mining areas and surrounding areas. These collapses allow surface water and atmospheric precipitation to be filled into mines to form subtype (6). When roadways or working face is advanced to boreholes with poor sealing quality, groundwater in roof and floor water-filling aquifers of coal seams gush into tunneling face via these boreholes, thus leading to (7).

11. Normal-temperature, moderate- to high-temperature, and corrosive water inrush. The normal-temperature water disaster refers to

water inrush in the normal temperature range of local groundwater. Under the effects of abnormal geotherm, the water disaster in which the temperature of water inrush is higher than the normal water temperature is known as moderate- to high-temperature water disaster. Corrosive water disaster means that the source of water inrush is corrosive to mining machinery equipment, drainage equipment, and roadways.

12. Instant, hysteretic, skipping, and gradually varied water inrush. Instant water disaster refers to the water inrush occurring in the working face of mines, while hysteretic water disaster refers to that appearing in the goaf behind the working face. With the gradual growth of mining depth and large-scale mining of under-group coal seams, the crustal stress and water pressure of the mining environment also increase. Hysteretic water disasters induced by different passages including faults, fracture concentrated zones, or karst collapse columns occur more frequently in recent years. Skipping water disaster refers to water inrush during which the inrush amount constantly changes with time while gradually varied water disaster is water inrushes during which the inrush amount gradually increases or decreases [76, 77].

Typical Prediction Methods of Mine Water Inrush

Due to differences in geological condition and enrichment law of coal resources in various regions, mine water inrush does not occur in many coal-producing countries such as the United States, Canada, Australia, and India. As a matter of fact, water disasters only take place in some countries including former Soviet Union, Poland, Hungary, Yugoslavia, and China. Therefore, countries suffering from mine water disasters are pioneers in studying the prevention of the disasters which has lasted for over 100 years [78].

The study of mine water inrush begins with Europe, in the 1940s. Weg Frence from Hungary proposed the concept of relative water-resisting layers aiming at the karst water inrush from coal floors [79]. The researcher suggested that water inrush is correlated with the thickness of water-resisting aquifuge and the water pressure of aquifer, and the water inrush process is restricted by the ratio of the thickness of equivalent water-resisting aquifuge to the value of the water pressure. The ratio was called relative water-resisting aquifuge. He also suggested that no water inrush occurred when the relative thickness of the aquifuge was greater than 1.5 m/atm in the mining process. Eighty percent to 88% of the water inrush occurred where the thickness of relative water-resisting aquifuge was less than this value. Thus, many countries with confined water mining have learned from the concept and suggested that water inrush will not occur if the thickness of the aquifuge is greater than 2 m/atm. From the 1960s to 1970s, the Hungary Mining Technology Committee included the thickness of relative water-resisting aquifuge into *mining safety codes* and provided explanations on different mining conditions [80]. Scholars in many countries such as former Soviet Union and Yugoslavia also began to study the effect of the relative water-resisting aquifuge during this period involving the influence of stress change caused by the goaf on the thickness of relative water-resisting layer and relationship between water flows and rock structures. The Xi'an Exploration Group of China Coal Research Institute began to investigate mine water inrush in the 1960s and proposed applying the water inrush coefficient as the criterion to predict whether water will burst out from floors or not.

From the 1970s to the late 1980s, scholars specialized in rock mechanics in numerous countries investigated the failure mechanism of floors while studying the stability of pillars. For example, C.F. Santos and Z.T. Bieniawski analyzed the load-bearing capacity of floors by introducing the critical energy release point based on the improved Hoek-Brown's rock strength criterion [82]. Li Baiying proposed "down three zone" theory, which indicates that there are three zones in the mining of coal floors just like mining overlying strata including floor failure zone, intact rock stratum, and the confined water-rising zone [83]. Wang Zuoyu, Liu Hongquan, and others put forward the theory of "in situ tension and failure at zero position" [84], which points out that the mine and water pressures exert a combination effect on working face, and the influence range on coal seams can be divided into three sections: advanced pressure compressing, pressure-releasing expansion, and post-mining pressure compressing and stable stages. Since the 1990s, the methods for studying mine water inrush have been diversified with the development of science and technology. For instance, the Bureau of Mines in the United States analyzed the stability of roadways under multiple environments by using finite element model. V.F. Bense investigated the deformation mechanism of materials around fault zones and the spatial distribution of water conductivity by applying the digital image analysis [85]. Academician Qian Minggao of China University of Mining and Technology established the KS theory of rocks in mining floors according to the stratified structure characteristic of coal floors. The theory indicates that there is a stratum with the largest load-bearing capacity in the coal floor below the mining failure zone and above the aquifer, which is called "key strata" [86]. The Institute of Geology, Chinese Academy of Sciences proposed the theory of "high-permeability passage," which states that whether water inrush from floors occurs depends on the existence of water inrush passages. It is divided into two cases. On the one hand, there are water inrush passages connecting with the water source in the hydrogeological structure of floors. On the other hand, there are no natural water inrush passages in coal floors, but new penetrated passages with high permeability are formed due to the deformation and damages to weak sections along the rock structures of floors under the effects of engineering stress, crustal stress, and underground water [87]. The Xi'an Exploration Group of China Coal Research Institute in the 1990s proposed the theory of coupling of rock mechanics and hydraulics. The theory holds that floor water inrush is the result of rock, water (floor pressure water), and stress (mining stress and ground stress) interaction [88]. The mining pressure makes the floor aquiclude produce a certain depth of the permeable fracture, reducing the strength of the rock, weakening the aquiclude performance, and resulting in the redistribution of the seepage field. When the confined water along the water rupture zone to further rise, rocks are softened due to water and continue to expand the cracks. The interaction between the water and the surrounding rocks reduces the minimum principal stress of the rock mass. When the minimum principal stress is less than the confined water pressure, fracturing expansion takes place, and water inrush will occur [89, 90].

In summary, the development history of mine water inrush theory reflects the process of continuous understanding of mine water inrush disaster, which is a process from practice to theory and from theory to practice. Apart from the above research on water inrush mechanism, some excellent achievements have also been made on predicting and forecasting mine water inrush [91–96].

Water Inrush Coefficient Method

The water inrush coefficient refers to the water pressure that can be borne by unit water-resisting aquiclude, which is expressed by the following equation [81]:

$$T = P/M$$

where T is the water inrush coefficient, in unit of MPa/m; M is the thickness of aquiclude floor, in unit of m; and P is the water pressure on floor aquiclude, in unit of MPa. When the water pressure exceeds the critical water inrush coefficient, water inrush possibly occurs. Based on statistics of actual water inrushes, the critical water inrush coefficients are 0.10 MPa/m for normal segments and 0.06 MPa/m for complicated tectonic segments.

Vulnerability Index Method

There are numerous factors causing water inrush from coal floors, and the water inrush process exhibits a complex nonlinear dynamic characteristic. Owing to the traditional water inrush, coefficient method only considers water pressure and aquiclude thickness, which fails to fully describe the water inrush from coal floors – a nonlinear dynamic phenomenon with multiple complex mechanisms controlled by multiple factors [97]. Thus, academician and professor

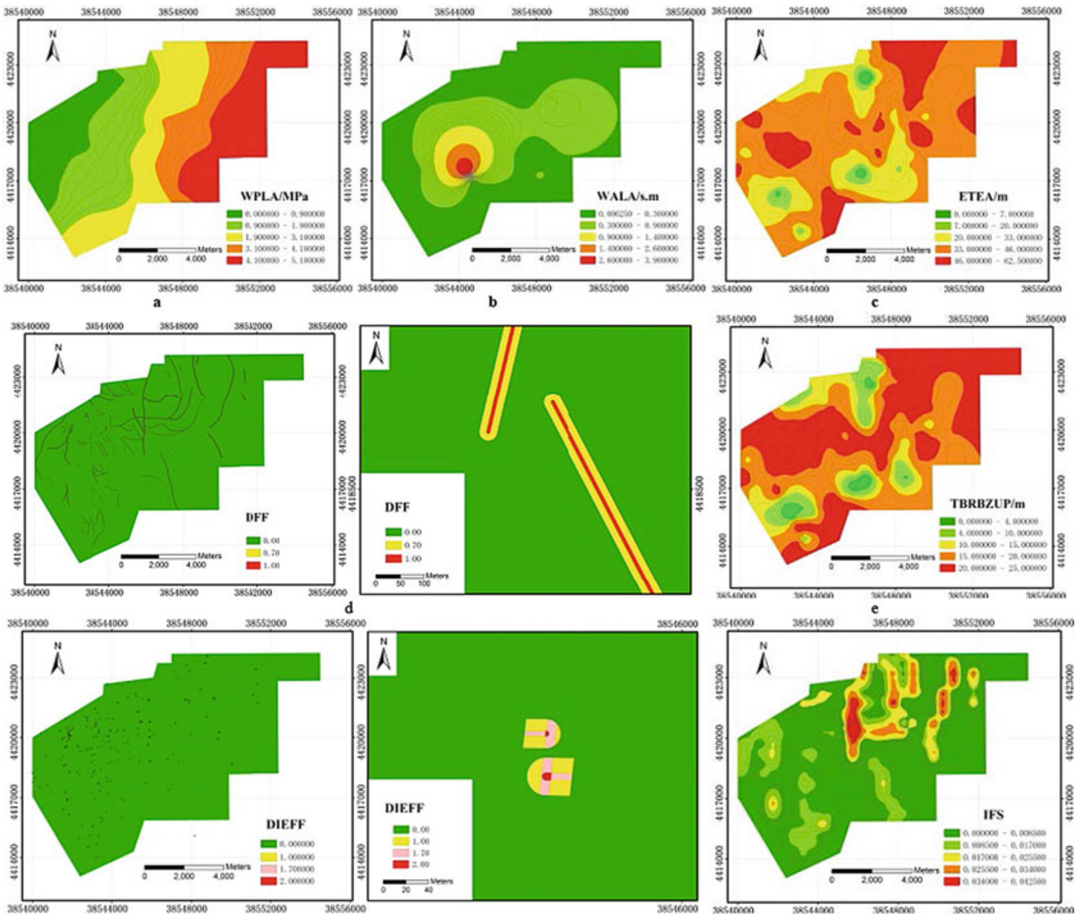
Wu Qiang of China University of Mining and Technology (Beijing) has been engaged in this challenging task in the last 20 years. He introduced multi-source information integration theory in the 1990s and proposed vulnerability index method for forecasting water inrush risk from coal floors in the 1990s [98–100]. The method can reasonably reflect the water inrush process from coal floors and forecast water inrush from floors by considering the interaction among various influence factors of water inrush from floors and the relative weights of factors. Vulnerability index method is used for predicting water inrush from coal floors by coupling the mathematical model with the geographic information system (GIS) with functions in analyzing and processing spatial information. The mathematical model determines the weight coefficients of multiple control factors of water inrush. The "vulnerability index method" reflects the complicated mechanisms in water inrush processes and can predict the water inrushes reasonably well. The specific steps are described as follows:

1. Collect data about the mine under study. Based on the analysis of mine water inrush mechanism in the mine, the main controlling factors of water inrush from coal floors are determined through three aspects – water abundance of aquifers, water-resisting layer of floor, and geological structure. Typical factors include equivalent thickness of effective aquiclude, thickness of brittle rock under broken zone by underground pressure, distribution of faults and folds, water abundance of aquifer, and water pressure of aquifer.
2. The geological and hydrogeological data are used to establish the attribute database and develop the thematic map for each factor. The collected data including drilling data, coordinates of water monitoring points, and quantitative values are inputted into GIS to develop the thematic maps of the main controlling factors by using functions of GIS including data storage, spatial data analysis and processing, and result display. The maps reflect the spatial characteristics of the factors. Figure 4 shows an example.
3. The weights of main controlling factors can be calculated by using either constant weight mode or variable weight model. The variable weight model considers the relative importance of controlling factors and the influence of changes of index values of controlling factors on the comprehensive evaluation value in different evaluation units. The role is realized by adjusting weights in response to the change of index values. Evaluation results are more in line with the actual situations. The relevant theory of the variable weight model can be obtained from references [100].
4. The superposition process is conducted to compose information storage layers of relevant factors into one information storage layer, so that the produced information storage layer contains the information of all relevant factors. The evaluation model of water inrush vulnerability is established to calculate the vulnerability indexes of each evaluation cell, producing a zoning map for the water inrush risk. Figure 5 shows an example. The accuracy of forecasting and evaluation results is tested by using locations of known water inrush points or locations of water inrush after mining.

Three-Map and Two-Prediction Method

The three maps in this method employ the following maps: zoning map for water abundance of water-filling aquifers in coal seam roofs, safety zoning map for roof caving, and comprehensive zoning map for water gushing (inrush) conditions of roofs [101–102]. According to the method, the necessary and sufficient condition for the water filling in coal seam roofs is that water-flowing fractured zones in roofs formed by coal mining connect overlying direct water-filling aquifers, and the aquifers have large water abundance in the caving zone of the mining face. In zones with low water abundance, serious water disasters are not likely to happen even if water-flowing fractured zones connect aquifers. Severe water disasters only take place in zones with large water abundance under the condition that water-flowing fractured zones connect aquifers.

The two predictions in this method consist of groundwater flux rate prediction for the working

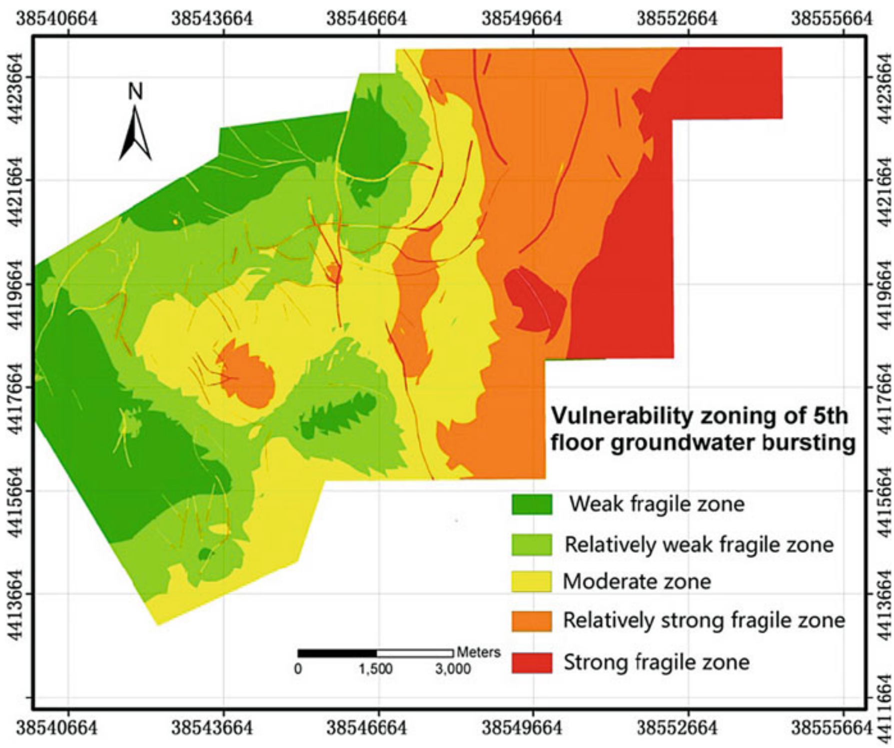


Mine Water Inrush, Fig. 4 Thematic maps of main factors influencing the floor water bursting. (**a** – water pressure of limestone aquifer; **b** – water abundance of limestone aquifer; **c** – equivalent thickness of effective aquiclude; **d** – distribution of faults and folds; e – thickness of brittle rock under broken zone by underground pressure; **f** – distribution of intersection and endpoint of faults and folds; **g** – index of fault scale)

face prior to mining and groundwater flow rate prediction for the working face during mining.

The method can favorably predict water disasters in coal seam roofs, and the concrete steps are as follows:

1. Determine water abundance of water-filling aquifers in coal seam roofs. On the basis of considering the main controlling factors influencing water abundance of water-filling aquifers in roofs, weights of the factors are solved according to the controlling effects of each factor on the water abundance. Then, the zoning map for water abundance of water-filling aquifers in roofs is developed through overlay analysis of multi-source geo-information. Figure 6 shows an example.
2. Develop the zoning map for roof caving. The developing height of caving and fractured zone after mining is calculated by using the empirical formula in *Coal Mine Safety Regulations*. Afterward, the safety zoning map for roof caving is plotted according to the distance from the height of the above zone to the water-resisting layer between coal seam roof and aquifer floor. Figure 7 shows an example.
3. Through the overlay analysis of the generated zoning map for water abundance of water-filling aquifers in roofs and safety zoning map for roof caving, the comprehensive zoning map



Mine Water Inrush, Fig. 5 Risk zoning map of water inrush

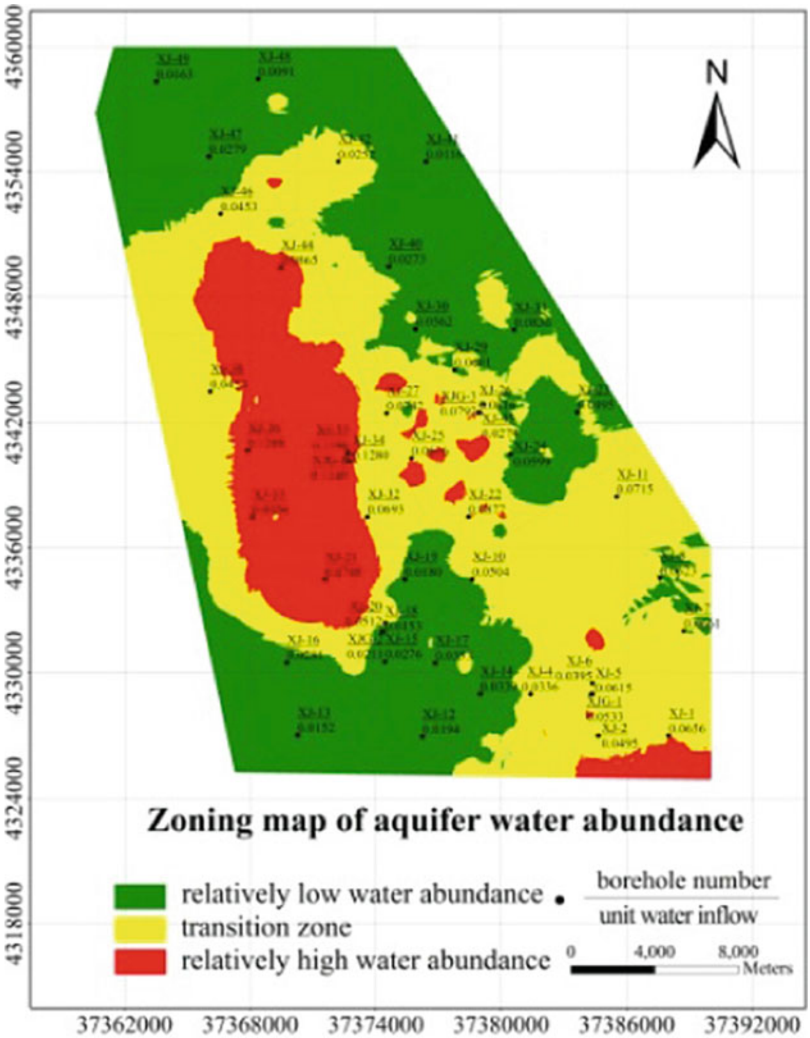
for water gushing (inrush) conditions of roofs is developed. Figure 8 shows an example.

4. The "two predictions" refer to the mine water inrush quantity predictions in natural and man-made conditions. The natural conditions mean that the geological and hydrogeological conditions are not disturbed by mining activities, while the man-made conditions mean that the geological and hydrogeological conditions are disturbed by mining or transformed by some engineering measures such as groundwater drainage and reinforcement of the fractured strata by grouting. The two predictions are based on mine-specific conceptual site model, three-dimensional groundwater flow field, calibration of model parameters, and stress distributions during mining. Figure 9 shows the predicted water flow rate in a mining face in natural conditions, whereas Fig. 10 shows the predicted water flow rate in the same mining face in man-made conditions.

Five-Map and Two-Coefficient Method

Five-map and two-coefficient method is another coal floor water hazard evaluation method. The five maps consist of contour map of coal floor protection layer failure depth, contour map of coal floor protection layer thickness, contour map of water head above coal floor, contour map of effective protection layer thickness, and evaluation map of mining above confined aquifer. The two coefficients are water inrush coefficient and the coefficient of mining above the confined aquifer. Water inrush coefficient is the ratio of "effective protection layer thickness" to water head value. Coefficient of mining above the confined aquifer is the index that represents hydraulic pressure resistance per meter of stratum. This method can favorably predict coal seam floor water disasters, and the detailed steps are as follows:

1. Develop contour map of coal floor protection layer failure depth. In the process of mining, coal floor is fractured to a certain depth, also

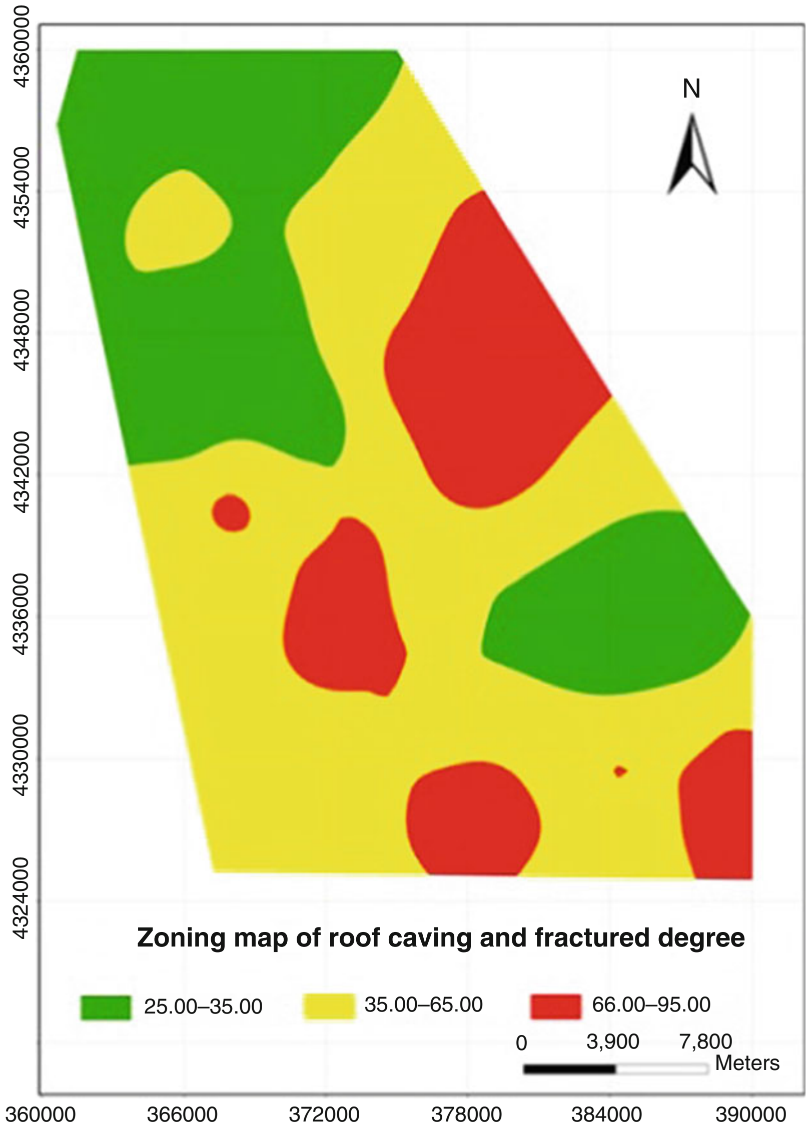


Mine Water Inrush, Fig. 6 Zoning map of aquifer water abundance

referred to as "floor failure depth," which induces water flow in the rock stratum. The floor failure depth can be obtained with experiments and empirical calculations.

2. Develop contour map of coal floor protection layer thickness. The rock stratum between coal floor and aquiclude roof is named as "protection layer." It is the barrier that stops confined water from flowing into the mining space. Its thickness and variation pattern can be obtained from exploratory boreholes.
3. Develop contour map of water head above coal floor. The contour map is based on groundwater monitoring data. The confined water heads are different for different floor elevations.
4. Develop contour map of effective protection layer thickness. The effective protection layer thickness is calculated by subtracting the floor failure depth from the coal floor protection layer thickness. It is a portion of the protection layer that possesses real ability of water head pressure resistance and safety protection function.
5. Finally, develop the evaluation map of mining above confined aquifer based on the existence and thickness of effective protection layer in accordance with the comprehensive analysis on "two-coefficient and three-class evaluation." The evaluation map shows the relative water inrush risks of mining above the confined aquifer.

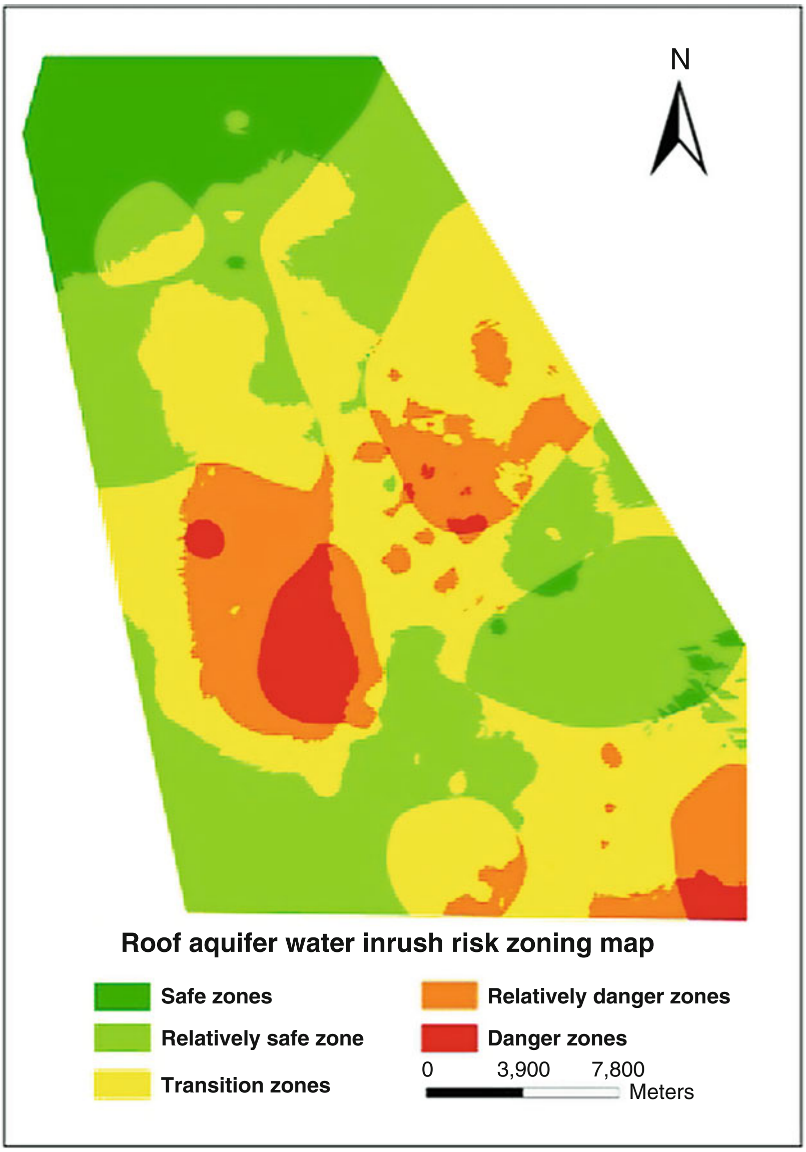
Mine Water Inrush, Fig. 7 Zoning map of roof caving and fractured degree



Future Directions

The prevention and control technology for mine water inrush has recently witnessed a great progress and development. Because of the diverse coalfield hydrogeological conditions and complex coalfield geological structures, the mine water inrush is concealed, and the factors and mechanisms are different, especially with the large-scale mining of deep mineral resources. Prevention and control of mine water inrush is still austere, and a lot of work needs to be done [103]. Many problems are to be solved and challenges are to be overcome. The following summarizes the development trends of future prevention and control technologies for mine water disasters [104–106]:

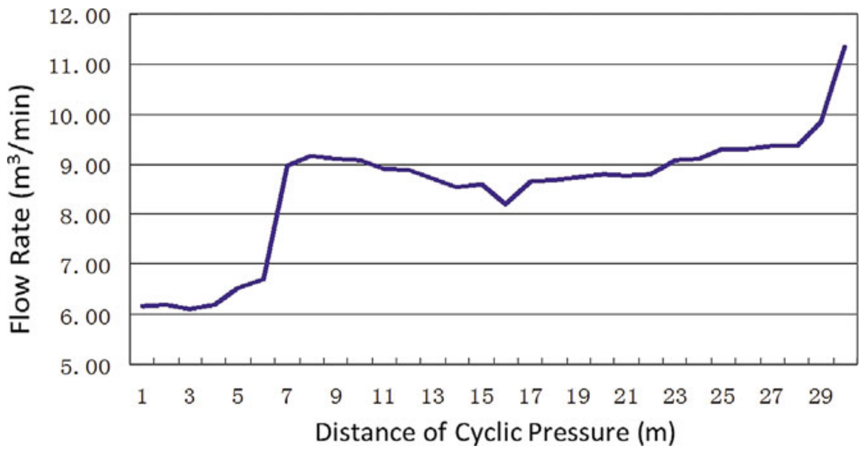
1. A detailed, intelligent, and geologically guaranteed system for safety and efficient mining is expected to be developed. Coal mine geological security system should include two parts – the production geological security subsystem and safety geological protection



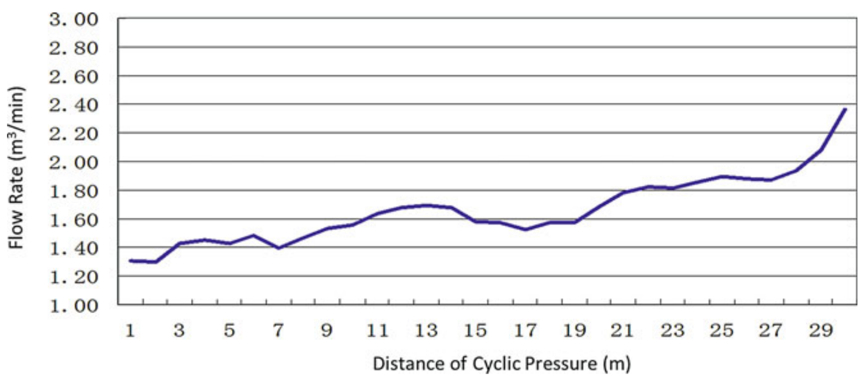
Mine Water Inrush, Fig. 8 Risk zoning map of roof aquifer water

subsystem. Regardless of mine production or mine safety, the basic geologically guaranteed system is a prerequisite. The basic geological guarantee is a large systematic engineering and involves many aspects [107, 108]. It covers survey and prospecting of hydrogeological conditions, forecasting of water regimes and disasters and high-precision and detailed detection, and locating of technology and equipment for geological structures (particularly concealed disaster-causing structures) and water abundance of water-filling aquifers. Apart from these, it also includes drilling and construction technology and equipment for rapid and efficient control of mine water disasters and emergency rescue and dynamic monitoring and early warning of hydrogeological conditions for mine water-filling and mining effects. These are basic issues guaranteeing the safety production of coal mines and need to be analyzed using the theories and methods for large-scale system engineering.

2. Studies should be conducted on the basic theories and techniques for prevention and control



Mine Water Inrush, Fig. 9 Dynamic curve of inflow in mining face in natural state



Mine Water Inrush, Fig. 10 Dynamic curve of inflow in mining face in artificial modification state

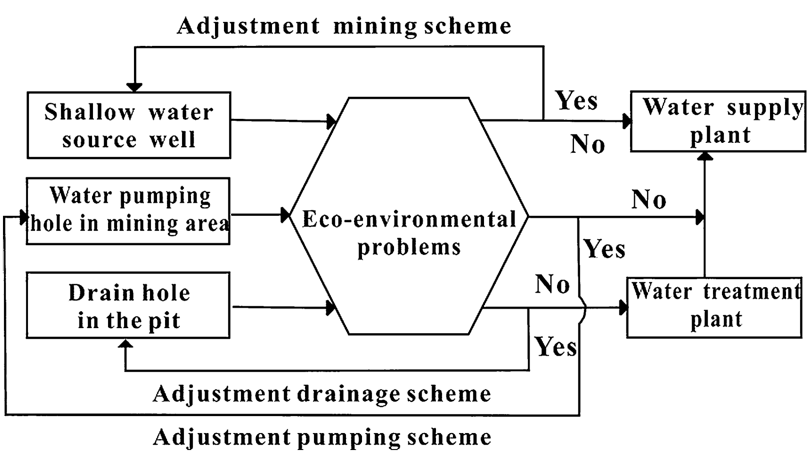
of mine water disasters in the mining of deep coal seams. As the shallow and upper-group coal resources are depleted, it is inevitable to enhance the exploitation of deep coal resources [109, 110]. In response to such situation, lots of work need to be done, for example, the deep and lower coal seam occurrence rule, the characteristics and distribution of concentrated high ground stress, high water pressure, high water temperature, and high gas pressure, as well as evolution characteristics of deep rock mechanical behavior in the "four-high" environment; geophysical and hydrophysical applications to supplementing exploration; hydrogeological conditions in supplementary exploration; in-depth analysis of the recharge, runoff, and drainage characteristics of deep karst water; dynamic evolution law of mining-induced fractures; and dynamic mechanism and triggering condition of deep mine water inrush. In addition, the economically and technically feasible detailed detection and locating techniques for deep concealed geological structures, the chain reaction effect of the coupled cross-feed of various geological disasters in coal mining, and the forecasting theory, prevention technology, monitoring, and warning approach for water inrush need to be investigated.

3. The theoretical and technical research on coal-water couplet-resource mine construction and the development of key technologies for water-controlled coal mining are supposed to be enhanced. Water-abundant mining areas are generally faced with conflicts of coal mining

(dewatering), water supply, and eco-environment protection. The coal-water couplet-resource mine construction and the technology development for water-controlled coal mining are effective approaches for solving the above conflicts. The key technologies for these methods include:

1. For mines where dewatering is feasible, the three-in-one optimization and combination method that consists of mine drainage, water supply, and eco-environment protection can be applied (Fig. 11). The drainage measures can be implemented on the ground, underground, or a combination of both. It is advisable to use a sewer system with a cleanup shunt, which can greatly reduce the cost of mine water [111].
2. For mines where dewatering is not feasible, the three-in-one method also can be used. The measures of groundwater control may include grouting reinforcement of coal seam floor and reconstruction of aquifer, blocking water channels by grouting, and change in coal mining method such as optimization of the mining process by filling as well as room pillar mining method. Where the quaternary strong aquifer is present, implementation of local slight blasting helps to inhibit the development of fracture zone height. The use of one-time mining or caving coal mining technology locally should be restricted because they tend to produce great damage to the roof and floor of the coal seam. Evaluation of mining suitability study should be performed to determine the non-suitable mining areas. In some cases, cautionary measures should be taken such as pre-cutting the groundwater flow to the mine by establishing ground drains and pre-draining strong groundwater runoff zones. While it is desirable to minimize the mine water inflow, efforts should be made to maximize the use of the mine drainage after treatment. Through the effective control and utilization of mine water, groundwater resources are protected in the mine area, and the groundwater levels are controlled to avoid deterioration of the mining ecosystems and geological environment.
3. The five-in-one optimization and combination method that consists of mine water control, treatment, utilization, recharge, and eco-environment protection can be used in mines with unique recharge conditions [112]. Under the scenarios in which effective measures are taken to prevent and control water for safe production, mine water treatment should be performed to maximize the use of mine water in underground production and groundwater supply process. The remaining mine water is processed and recharged back into the underground to achieve the goal of sustainable coal resources development and overall water resources and ecological environment protection.

Mine Water Inrush, Fig. 11 The sketch map of drainage – water supply – eco-friendly



4. Organizational coordination. As the optimization system of "three-in-one" or "five-in-one" involves three separate different departments of mine, water, and environmental protection, cooperation is essential between them.

4. The forecasting and evaluation theory and field measurement technology for super-high water-conducting fractured zone of roofs and extra-deep mine pressure disturbed zone of floors in integrated mechanized mining with large mining height and high one-time-take-whole caving are also a future research direction. From the beginning of the twentieth century, most of China's large coal mines have adopted integrated mechanized mining methods with large mining height and high one-time-take-whole caving. These coal mining methods created high productivity and efficient mines and have also brought new hydrogeological problems. The height of the roof caving fractured zone and depth of the floor rock pressure failure zone are much greater than those from the traditional stratified mining. The roof and floor filling water sources that are not in the disturbance zone of the traditional mining may become parts of the disturbance zone in response to the new mining methods, which poses a threat to the safety of coal mines [113, 114]. The abnormally disturbed zones in coal roofs and floors induced by the powerful mechanized mining increase the probability of connecting to aquifers in roofs and floors, goaf water, or surface water and lead to a higher risk of mine water inrush.
5. In the study on the drilling and geophysical integration in technology for exploitation water of advanced drilling, exploitation water of advanced drilling of underground mining faces generally includes drilling and geophysical exploration, each having its own advantages and disadvantages.

   The advantages of drilling are intuitive and clear. As long as the drilling intercepts the water body, one can prove its exact location and thus implement drainage to achieve the purpose of advanced exploration. However, the shortcomings are limited control range, existence of exploration blind spots, long duration, and large investment. Interception of no water only means that there is no water content in the location of the borehole. Multiple boreholes must be designed to determine the distribution of the water content in the entire range of the excavation work surface [115, 116]. On the other hand, the shortcomings of geophysical methods are nonunique interpretation, decrease in detection accuracy with the increase of the detection range (volumetric effect), and strong influences on the detection accuracy by the environment. The advantage of geophysical methods is that these methods can quickly and cost-effectively carry out a comprehensive survey of the entire mining face. When the detection range is small, the accuracy is better [117, 118]. In view of the above advantages and disadvantages, the geophysical methods are typically used first, which is followed by drilling to verify the geophysical results. A true integration of drilling and geophysical exploration has not been achieved.
6. A monitoring and early-warning technology system for mine water disasters with functions of monitoring mining-induced deformation and potential water inrush is expected to be developed in the future. The occurrence of water inrush cannot happen unless the following three conditions are met, i.e., water-filling source, water-filling channel, and water-filling strength. Water inrush will not occur if any of these conditions are not met. The microseismic monitoring technology can perform real-time, planar, high-precision positioning and three-dimensional display and analysis of the mining process of the aqueduct to determine the channel type, channel space position, and deformation scale, but the microseismic monitoring technology cannot detect and determine if the channel is filled with water or not. The micro-earthquake itself cannot monitor and provide early warning of water inrush potential, because water inrush will not occur if there is no water supply, although the mining deformation may be great [119, 120]. In addition to microseismic monitoring technology, the monitoring technology based on fiber grating

communication can monitor mining-induced deformation and water pressure in real time; however, the multiparameter monitoring system can merely conduct point monitoring but fails to monitor the whole working face [121]. Therefore, developing the monitoring and early-warning technology system for mine water disasters is another important research direction in the future. The system is expected to perform concurrent monitoring of two parameters – mining-induced deformation and water-filling sources.

Bibliography

1. Wu Q, Zhao SQ, Li JS, YJ F, Yin SX, Liu SQ (2011) The preparation background and the main points of rule of mine prevention and cure water disaster. J China Coal Soc 36(1):70–74
2. Wu Q, Cui FP, Zhao SQ, Liu SQ, Zeng YF, YW G (2013) Type classification and main characteristics of mine water disasters. J China Coal Soc 38(4): 561–565
3. Zhao YS (1994) Rock fluid mechanics in mine. China Coal Industry Publishing House, Beijing
4. Wu Q, Jin Y (1995) Decision system of water control in mines of north china type coal fields. Coal Industry Publishing House of China, Beijing
5. State Administration of Work Safety Supervision (2009) Provisions for mine water control. China Coal Industry Publishing House, Beijing
6. Wu Q (2009) Interpretation of provisions for mine water control. China University of Mining and Technology Press, Xuzhou
7. State Administration of Work Safety Supervision (2011) Coal mine safety regulations. China Coal Industry Publishing House, Beijing
8. Wu Q (2001) Supervision of coal mine safety (mine water disaster and investigation). China University of Mining and Technology Press, Xuzhou
9. DJ X, Shao DS, Nie JW, Chen HJ (2013) Advances in the mechanism of groundwater gushing from bed separation of hard overlying strata induced by repeated coal mining. Sci Technol Rev 31(23):75–79
10. Wu Q, Li ZY (2002) Mine water prevention and control. China University of Mining and Technology Press, Xuzhou
11. Wang JM (2000) Measurement and physical analogue on water inrush from coal floor induced by progressive intrusion of artesian water into protective aquiclude. Chin J Geotech Eng 21(9):546–549
12. Zhu SY, YJ J, Zhao ZZ, Liu DQ (2009) Field measurement study on deformation and destruction of “three-soft” coal seam floor of Chaohua coal mine. Chin J Geotech Eng 31(4):639–642
13. Wu Q, Liu JT, Zhong YP, Yin ZR, Li JM, Hong YQ, Ye GJ, Tong YD, Dong DL (2002) The numeric simulations of water-bursting time-effect for faults in Zhaogezhuang coal mine, Kailuan, China. J China Coal Soc 27(5):511–516
14. Yin SX, Wu W, Li YJ (2008) Study on Karst collapse column and water inrush in North China coalfield. China Coal Industry Publishing House, Beijing
15. Sui WH, Cai GT, Dong QH (2007) Experimental research on critical percolation gradient of quicksand across overburden fissures due to coal mining near unconsolidated soil layers. Chin J Rock Mech Eng 26(10):2084–2091
16. Han SQ, Fan LM, Yang BG (1992) Some hydrogeological and engineering-geological problems concerning development of north Shaanxi Jurassic coalfield. Coal Geol China 4(1):49–52
17. State Bureau of Coal Industry (2010) Buildings, water, railway and main shaft and this coal pillar and press coal mining regulations. China Coal Industry Publishing House, Beijing
18. Wu Q, Xu H, Pang W (2008) GIS and ANN coupling model: an innovative approach to evaluate vulnerability of karst water inrush in coalmines of north China. Environ Geol 54(5):937–943
19. Li B, Wu Q, Chen YL, Wang GL, Xu L (2016) Multifactor evaluation for the water yield properties of coal floor aquifers, based on GIS. Water Pract Technol 11(1):75–85
20. Li JK (1990) Prevention and control of mine karst water. China Coal Industry Publishing House, Beijing
21. Luo LP, Peng SP (2005) Mechanism study of water-inrush hazard of floor strata in mining on confined aquifer. J China Coal Soc 30(4):459–462
22. Wu Q, Chen H, Liu SQ (2010) Methodology and application on size-limited structure predictions with ANN based on loop overlapping theory: a case study of Lingzi coal mine in Zibo. J China Coal Soc 35(3):449–453
23. Wu Q, Zhang ZL, Zhang SY (2007) A new practical methodology of the coal floor water inrush evaluating II – the vulnerable index method. J China Coal Soc 32(11):1121–1126
24. Yin GZ, Li X, Han PB, Li MH, Li WP, Deng BZ (2016) Experimental study on overburden strata fracture evolution law in three dimensional mine induced stress conditions. J China Coal Soc 41(2):406–413
25. Zhao TC (2006) Integrated control technology for Ordovician limestone water in North China. China Coal Industry Publishing House, Beijing
26. Wu Q, Fan ZL, Liu SQ (2011) Water-richness evaluation method of water-filled aquifer based on the principle of information fusion with GIS: water-richness index method. J China Coal Soc 36(7): 1124–1128
27. Sui WH, Dong QH (2008) Variation of pore water pressure and its precursor significance for quicksand disasters due to mining near unconsolidated formations. Chin J Rock Mech Eng 27(9):1908–1916

28. Li B, Chen YL (2016) Risk assessment of coal floor water inrush from underlying aquifers based on GRA–AHP and its application. Geotech Geol Eng 34(1):143–154
29. Wu Q, Zhang ZL, Ma JF (2007) A new practical methodology of the coal floor water bursting evaluating I: the master controlling index system construction. J China Coal Soc 32(1):42–47
30. Liu WT, Zhang WQ, Li JX (2000) An evaluation of the safety of floor water-irruption using analytic hierarchy process and fuzzy synthesis methods. J China Coal Soc 25(3):278–282
31. Li B, Wu Q (2015) An analysis of parameters sensitivity for vulnerability assessment of groundwater inrush during mining from underlying aquifers based on variable weight model. J Min Saf Eng 32(6):911–917
32. Wu Q, Liu Y, Liu D, Zhou W (2011) Prediction of floor water inrush: the application of GIS-based AHP vulnerable index method to Donghuantuo coal mine, China. Rock Mech Rock Eng 44(5):591–600
33. Jia MK (2005) A new way of genetic classification on roof falling of bolt supporting roadway. J China Coal Soc 30(5):568–570
34. Qian MG, Liu TC (1991) Mining pressure and strata control. China Coal Industry Publishing House, Xuzhou
35. Wang CX, Wang HM (2004) Thinking about theories and practices on mine water control. Coal Geol Explor 32(11):100–103
36. Liu TQ (1986) Safe extraction of near soft layer underlying a thick loose aquifer. Coal Sci Technol 13(2):14–18
37. Yang YG, Yuan JF, Chen SZ (2006) R/S analysis and its application to the forecast of mine inflows. J China Univ Min Technol 16(4):425–428
38. Liu SC, Liu XM, Jiang ZH, Xing T, Chen MZ (2009) Research on electrical prediction for evaluating water conducting fracture zones in coal seam floor. Chin J Rock Mech Eng 28(2):348–355
39. Qiao W, Li WP, Sun RH, Li XQ, Hu G (2001) Formation mechanism of dynamic impact failure zone of super dynamic water inrush in coal mine. Chin J Geotech Eng 33(1):1726–1733
40. YC X, Liu SQ (2011) Study on method to set safety coal and rock pillar for fully mechanized top coal caving mining under body. Coal Sci Technol 39(11):1–4
41. Wu Q, Zhou W (2009) Prediction of groundwater inrush into coal mines from aquifers underlying the coal seams in China: application of vulnerability index method to Zhangcun Coal Mine, China. Environ Geol 57(5):1187–1195
42. Yang TH, Shi WH, Li SC (2016) State of the art and trends of water-inrush mechanism of nonlinear flow in fractured rock mass. J China Coal Soc 41(7):1598–1609
43. Coal Science Research Institute Beijing Mining (1981) Surface movement of coal mine and law of overburden rock fracture and its application. Coal Industry Publishing House, Beijing
44. Wang JM, DH Y (2010) Simulation of water hazards caused by burst of water cells formed. Chin J Geotech Eng 32(2):231–237
45. Qiao W, Huang Y, Zb Y, Guo W, Zhou DK (2014) Formation and prevention of water inrush from roof bed separation with full-mechanized caving mining of ultra-thick coal seam. Chin J Rock Mech Eng 33(10):2076–2084
46. Bense VF, Van D, Berg EH (2003) Deformation mechanisms and hydraulic properties of fault zones in unconsolidated sediments. Hydrogeol J 11(3):319–332
47. Fan LM, Ma XD, Jiang H (2016) Risk evaluation on water and sand inrush in ecologically fragile coal mine. J China Coal Soc 41(3):531–536
48. Zhu WB, Wang XZ, Kong X, Liu WT (2009) Study of mechanism of stope water inrush caused by water accumulation in overburden separation areas. Chin J Rock Mech Eng 28(2):306–311
49. JL X, Zhu WB, Wang XZ (2011) Mechanism and criteria of crushing sand near loosening sand stone aquifer. J Min Saf Eng 8(3):333–339
50. Gao YF (1996) Rock movement, "four-zone" model and dynamic displacement analysis. China Coal Society 21(1):23–25
51. Li ZK, Hu J (2007) A study on working face roof dynamic separation water harness. Coal Geol China 19(2):35–38
52. Xie XH, BY S, Gao YF, Duan XB (2005) Numerical study on water inrush above a confined aquifer in coal mining using hydro-fracturing. Chin J Rock Mech Eng 24(6):987–993
53. Yang TH, Tang CA, Tan ZH (2007) State of the art of inrush models in rock mass failure and developing trend for predication and forecast of ground water inrush. Chin J Rock Mech Eng 26(2):268–277
54. Wu Q, Zhu B, Li JM, Hong YQ, Qian ZJ (2008) Numerical simulation of lagging water-inrush mechanism of rock roads near fault zone. J China Univ Min Technol 37(6):780–785
55. Wen-hong LV (2014) Measure and simulation for development height of water conducted crack zone in overburden roof. J Xi'an Univ Sci Technol 03:309–313
56. Liang YP, Wen GC (2000) Comprehensive analysis method of "three-zone" classification on mine roof strata. Coal Sci Technol 28(5):39–42
57. XJ H, Li WP, Cao DT (2012) Index of multiple factors and expected height of fully mechanized water flowing fractured zone. J China Coal Soc 37(4):613–620
58. Wu Q, Liu YZ, Luo LH, Liu SQ, Sun WJ, Zeng YF (2015) Quantitative evaluation and prediction of water inrush vulnerability from aquifers overlying coal seams in Donghuantuo coal mine, China. Environ Earth Sci 74(2):1429–1437
59. Zhang Y, Pang YH (2010) Water-inrush mechanical model based on a theory of coupled stress-seepage. J China Univ Min Technol 39(5):659–664
60. Zhang ZL, Gao YF, Wu Q, Wei SM (2013) Discussion on the technical system of solid prevention and control on mining flooding. J China Coal Soc 38(3):378–383

61. Liu ZJ, YQ H (2007) Solid-liquid coupling study on water inrush through faults in coal mining above confined aquifer. J China Coal Soc 32(10):1046–1050
62. YC X, Li JB (2014) "Pore-fractured lifting type" mechanical model for floor water inrush of the grouting enforcement working face. J China Univ Min Technol 43(1):49–55
63. Shi LQ, Han J (2004) Floor water-inrush mechanism and prediction. China University of Mining and Technology Press, Xuzhou
64. Wu Q, Zhou YJ, Liu JT, Zhong YP, Yin ZR, JM LI, Hong YQ, Zhou RG (2003) The mechanical experiment study on lag mechanism of water-bursting of fault under coal seam. J China Coal Soc 28(6):561–565
65. Zhang JC, Zhang YZ, Liu TQ (1997) Rock mass permeability and coal mine water inrush. Geological Publishing House, Beijing
66. Feng MM, Mao XB, Bai HB, Wang P (2009) Experimental research on fracture evolution law of water-resisting strata in coal seam floor above aquifer. Chin J Rock Mech Eng 28(2):336–341
67. ZM X (2011) Mining-induced floor failure and the model, precursor and prevention of confined water inrush with high pressure in deep mining. J China Coal Soc 36(8):1422–1424
68. Zhao YS, Hu YQ (2004) Theory and technology of coal mine on the confined aquifer. China Coal Industry Publishing House, Beijing
69. Zhou W (1997) Paleocollapse structure as a passageway for groundwater flow and contaminant transport. Environ Geol 32(4):251–257
70. Luo LP, Peng SP (2005) Mechanism study on water-inrush hazard of floor strata in mining on confined aquifer. J China Coal Soc 30(4):439–462
71. Wu Q, Pang W, Dai YC, Yu J (2006) Vulnerability forecasting model based on coupling technique of GIS and ANN in floor ground water bursting. J China Coal Soc 31(3):314–319
72. Moutsopoulos KN, Papaspyros J, Tsihrintzis VA (2009) Experimental investigation of inertial flow processes in porous media. J Hydrol 374(3–4):242–254
73. Jiang QM (2009) Coal floor strata failure depth test of working face at big mining depth. Coal Geol Explor 37(4):30–33
74. Zhu SY, YJ J, Zhao ZZ (2009) Field measurement study on deformation and destruction of "three-soft" coal seam floor of Chaohua coal mine. J Geotech Eng 31(4):639–642
75. ZM X, Sun YJ, Dong QH, Zhu ZK (2012) Closing mechanism of mining-induced fracture in coal mine aquifuge and its application. J Min Saf Eng 29(5):613–618
76. Yin SX, Wu Q, Wang SX (2004) Studies on characters and forming mechanism of karstic collapse columns at mine area of north China. Chin J Rock Mech Eng 23(1):120–123
77. WK B, Mao XB (2009) Research on effect of fault dip on fault activation and water inrush of coal floor. Chin J Rock Mech Eng 28(2):386–394
78. Hosseini N, Oraee K, Shahriar K, Goshtasbi K (2013) Studying the stress redistribution around the longwall mining panel using passive seismic velocity tomography and geostatistical estimation. Arab J Geosci 6(5):1407–1416
79. Qian MG, Miao XX, Xu JL (2000) Key strata theory in ground control. China University of Mining and Technology Press, Beijing
80. Wu Q, Zhao SQ, Sun WJ, Cui FP, Wu C (2013) Classification of the hydrogeological type of coal mine and analysis of its characteristics in China. J China Coal Soc 38(6):901–905
81. Guan ET (2012) Origin of water bursting coefficient and process of modification. Coal Geol China 24(2):30–32
82. Wang YH, Shen W (1996) Prevention and control of China coal mine flooding. China Coal Industry Publishing House, Beijing
83. Santos CF, Bieniawski ZT (1967) Floor design in underground coalmines. Rock Mech Rock Eng 22(4):249–271
84. Li BY (1999) "Down Three Zones" in the prediction of the water inrush from coalbed floor aquifer-theory, development and application. J Shandong Inst Min Technol (Nat Sci) 18(4):11–18
85. Bense VF, Van den Berg EH, Van Balen RT (2003) Deformation mechanisms and hydraulic properties of fault zones in unconsolidated sediments; the Roer Valley Rift System, The Netherlands. Hydrogeol J 11(3):319–332
86. Qian MG, Miao XX, JL X (1996) Theoretical study of key stratum in ground control. J China Coal Soc 21(3):225–230
87. Wang ZY, Liu HQ (1993) Mining in confined aquifer. China Coal Industry Publishing House, Beijing
88. Shi LQ (2009) Summary of research on mechanism of water inrush from seam floor. J Shandong Univ Sci Technol 28(3):17–23
89. Zhou W, Li GY (2001) Geological barrier– a natural rock stratum for preventing confined karst water from flowing into mines in North China. Environ Geol 40(8):1003–1009
90. Xie GX, Chang JC, Hua XZ (2007) Influence of mining velocity on mechanical characteristics of surrounding rock in fully mechanized top-coal caving face. Chin J Geotech Eng 29(7):963–967
91. Nelson EP, Kullman AJ, Gardner MH, Batzle M (1999) Fault-fracture networks and related fluid flow and sealing, brushy canyon formation, West Texas. Faults Subsurf Fluid Flow Shallow Crust 21(3):69–81
92. Motyka J, Pulido-Bosch A (1985) Karstic phenomena in calcareous-dolomitic rocks and their influence over the inrushes of water in lead-zinc mines in Olkusz region (south of Poland). Int J Mine Water 4(2):1–11
93. He MC, Xie HP, Peng SP, Jiang YD (2005) Study on rock mechanics in deep mining engineering. Chin J Rock Mech Eng 24(16):2803–2813
94. Noghabai K (1999) Discrete versus smeared versus element-embedded crack models on ring problem. J Eng Mech 125(6):307–314

95. Kuznetsov SV, Trofimov VA (2002) Hydrodynamic effect of coal seam compression. J Min Sci 39(3):205–212
96. Wolkersdorfer C, Bowell R (2005) Contemporary reviews of mine water studies in Europe. Mine Water Environ 24(1):2–37
97. Wu Q, Wang M, Wu X (2004) Investigations of groundwater bursting into coal mine seam floors from fault zones. Int J Rock Mech Min Sci 41(4):557–571
98. Wu Q, Yang L, Zhu B (2009) Application of vulnerability index method in coal floor water bursting evaluation in Zhaogezhuang coalmine. Coal Geol China 21(6):40–44
99. Wu Q, Wang JH, Liu DH, Cui FP, Liu SQ (2009) A new practical methodology of the coal floor water bursting evaluating IV:the application of AHP vulnerable index method based on GIS. J China Coal Soc 34(2):231–238
100. Wu Q, Li B, Liu SQ (2013) Vulnerability assessment of coal floor groundwater bursting based on zoning variable weight model: a case study in the typical mining region of Kailuan. J China Coal Soc 38(9):1516–1521
101. Wu Q, Xu K, Zhang W (2016) Further research on “three maps-two predictions” method for prediction on coal seam roof water bursting risk. J China Coal Soc 41(6):1341–1347
102. Wu Q, Huang XL, Dong D, Yin ZR, Li JM, Hong YQ, Zhang HJ (2000) “Three maps-two predictions” method to evaluate water bursting conditions on roof coal. J China Coal Soc 25(1):62–67
103. Wu Q (2014) Progress, problems and prospects of prevention and control technology of mine water and reutilization in China. J China Coal Soc 39(5):795–805
104. Qiang W, Li D (2006) Management of karst water resources in mining area: dewatering in mines and demand for water supply in the Dongshan mine of Taiyuan, Shanxi Province, North China. Environ Geol 50(8):1107–1117
105. Wu Q, Zhou WF, Li D, Di ZQ, Miao Y (2006) Management of karst water resources in mining area: dewatering in mines and demand for water supply in the Dongshan mine of Taiyuan, Shanxi Province, North China. Environ Geol 50(8):1107–1117
106. Zhang HR, Zhang WQ, Wen XL, Wang C (2000) Design of continuous observation work on mining failure feature in the floor of mine and its practice. Min Res Dev 20(4):1–4
107. Li HT, Lei BL, Yang GM (2005) Development progress and prospect of China mine geological safety guarantee system. Coal Geol Explor 33(Suppl):9–13
108. Zhang BL, Guo WJ, Zhang XG, Shen BT, Zhang T, Kong H (2016) Development and application of analogue testing system for floor confined water rise in coal mining. J China Coal Soc 41(8):2057–2062
109. Dong SN (2010) Some key scientific problems on water hazards frequently happened in China’s coal mines. J China Coal Soc 35(1):66–71
110. Chmitt DR, Zoback MD (1989) Poroelastic effects in the determination of the maximum horizontal principal stress in hydraulic fracturing tests-a proposed breakdown equation employing a modified effective stress relation for tensile failure. Int J Rock Mech Min Sci Geomech Abstr 26(6):499–506
111. Wu Q, Dong DL, Shi ZH, Xiong W, Sun WD, Ye GJ, Li SW, Liu JT (1999) Study on the optimized combination of drainage – water supply – eco – environmental trinity in North China coalfield. Sci China Earth Sci 29(6):567–573
112. Wu Q, Wang ZQ, Guo ZK, Zhang DY, Chen YJ, Zhao PF, Chen Y, Wang TJ, Liu C, Wang Y (2010) A research on an optimized five-in-one combination of mine water control, treatment, utilization, backfilling and environment friendly treatment. China Coal 36(2):109–112
113. Duan HF, Jiang ZQ, Wang YD, Shao MX, Zhao LJ, Zhu QL (2012) Rock expansion boundary antipermeability strength and its application in the coal mine floor water inrush evaluation. J Min Saf Eng 29(1):95–99
114. ST G, Jiang BY, Wang CQ, Li NN (2013) Simulation research on overburden failure and lead abutment pressure distribution of fully-mechanized sublevel caving face. Min Res Develop 33(3):13–14
115. Tian J, Han G, YJ W, Wu M (2006) Status of pilot probing technology for blind heading. Coal Sci Technol 34(8):17–20
116. Qi CS (2005) Underground engineering geological prediction technology tunnel construction. Tunnel Constr 25(3):9–12
117. Liu SD, Guo LQ (2007) MSP technology and the application of detecting small geologic structure in coalmine lane way. J Anhui Univ Sci Technol (Nat Sci Ed) 27(S1):22–25
118. Liu SC, Yue JH, Liu ZX (2005) Hydrogeophysical exploration technique and application in coal seams. China University of Mining and Technology Press, Xuzhou
119. Yan XL, Chen XH, Yan XY (2011) Analysis on microseismic law when fully mechanized coalface passed through fault. J China Coal Soc 36(S1):83–87
120. Zheng C, Yang TH, Yu QL, Zhang PH, Liu HL, Zhang SJ (2012) Stability evaluation of excavated rock mass in mines based on microseismic monitoring. J China Coal Soc 37(S2):280–286
121. Zhang PS, Liu SD, Wu RX, Cao Y (2009) Dynamic detection of overburden deformation and failure in mining workface by 3D resistivity method. Chin J Rock Mech Eng 28(9):1870–1875

Mineral and Thermal Waters

Adam Porowski
Stable Isotope Laboratory, Institute of Geological Sciences Polish Academy of Sciences (INGPAN), Warszawa, Poland

Article Outline

Glossary

Codex Alimentarius Collection of internationally recognized standards, codes of practice, guidelines, definitions, and recommendations relating to food, food production and processing, and food safety, including standards for bottled drinking waters and natural mineral waters. The Codex Alimentarius is developed and maintained by the Codex Alimentarius Commission established in November 1961 by the Food and Agriculture Organization of the United Nations (FAO). The World Health Organization (WHO) joined the Commission in June 1962. The first session of the Commission was held in Rome in October 1963.

Curative water Groundwater which meets the standards of mineral water (i.e., at least 1 g/kg dissolved solid) or which contains less than 1 g/kg of dissolved solids but having concentration of pharmacologically active compound above the lower limit, including temperature above 20 °C (i.e., thermal water).

Geothermal energy resources Total amount of thermal energy (heat) accumulated in the Earth's crust down to given depth, referred to particular area for which the calculations are made and for mean annual temperature at the Earth's surface.

Geothermal energy Is a heat energy generated and stored in the Earth.

Geothermal system Consists of three elements: (i) heat source, (ii) reservoir (i.e., aquifer), and (iii) fluid – i.e., thermal water (or steam) as a heat carrier; it comprises the entire hydrogeologically connected area with the recharge zone, reservoir aquifer, and outflow zone; the geothermal aquifer plays a role of geothermal reservoir.

Lindal diagram Diagram that illustrates the possible utilization of thermal waters with respect to their temperature.

McKelvy diagram Diagram that utilizes the resources and reserves to assess the amount of any minerals on the Earth in the light of two distinct parameters: degree of certainty and the profitability (technically, these parameters are used as axes of the diagram). Such diagram can be used also to assess the resources of geothermal energy.

Mineral water May be defined in different ways depending on the field of application, i.e., hydrogeology, balneology, or bottled water industry. Traditionally term "mineral water" was referred to specific groundwater used for curative purposes which characterized higher than average concentrations of dissolved solids or gases or higher temperature. The formal definition of mineral water has been proposed for the first time by L. Grünhut in 1911 during the International Balneological Congress in Nauheim, Germany. It was established that this term can be referred only to the groundwater which contains at least 1 g/kg of dissolved solids. Nowadays, the value of 1 g/L of total

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_978

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_978-1

dissolved solids is commonly accepted border between fresh and mineral waters also from hydrogeological point of view.

Thermal water Also called as geothermal water, are defined as groundwater with a temperature of at least 20 °C at the outflow.

Definition of the Subject

From the beginning of time, water has been essential resource for survival of man. Except freshwaters used for drinking and household purposes, the mineral and thermal waters used by man from centuries have their special place in the development of societies, nations, countries, and civilizations.

Mineral and thermal waters from hydrogeological point of view are groundwaters. The groundwater is defined as the free water (i.e., able to gravitational flow) present beneath Earth's surface in pores and fissures of rock formations. Groundwater constitutes about 1.7% of the world's total water, which accounts for the volume of around 23.4×10^6 km^3. More than 50% of that volume constitutes groundwater which occurs in strata below 1 km depth where water is mostly saline and hot. The water is an extraordinary solvent that dissolves more substances than any other liquid in nature. That is why groundwaters through interaction with rocks as they travel through underground environment gradually acquire their unique physicochemical characteristics, mineral content, and temperature. Part of these waters, depending on the specific chemical features, temperature, and place and form of occurrence at the surface, have been used by people for drinking, bathing, healing, therapy, or religious rites for thousands of years. The use of these resources for heating, personal hygiene, and curative and recreational purposes is deeply integrated in the history of our civilization. Archeological findings show that mineral and thermal waters have been used for bathing since the Bronze Age (about 12,000–3,000 years BC). Many hot springs have been used in connection with religious rites in Egypt and by the Jews of the Middle East; the Greeks, Turks, and Romans were famous for their spa development and use from Persia to England. The term "mineral water" traditionally is connected with balneology and is referred to groundwaters used for curative purposes in which total dissolved solids are higher than that of waters commonly used for drinking or household purposes at particular area. "Balneology" relates to warm baths for healing purposes with natural thermal waters, different temperatures, chemical composition, or viscosity (pure liquid or mud). The first formal definition of "mineral water" proposed by L. Grünhut in 1911 during the International Balneological Congress in Nauheim, Germany, indicates that this term should be referred exclusively to groundwaters with total dissolved solids of 1 g/kg solution. Groundwaters which contain less than 1 g/kg of dissolved solids but having concentration of pharmacologically active compound (e.g., F^-, Br^-, I^-, HBO_2, CO_2, Fe, H_2SiO_3, etc.) above the medically (balneologically) accepted lower limit can be counted among mineral curative waters as the so-called specific waters. The temperature is one of these pharmacologically active compounds of mineral curative waters. The groundwaters with a temperature of at least 20 °C at the outflow are unanimously called thermal (or geothermal) waters.

The use of mineral and thermal waters developed and evolved over time: from curative and healing bathing or drinking at the place of their occurrence and origin to the emergence of two very important sectors of economy, namely, (i) bottling water industry and (ii) geothermics – science/industry pertaining to the exploration and utilization of geothermal heat for direct use and electric power generation. The uniqueness of the use of mineral and thermal waters from the beginnings to these days is characterized by a firmly established belief in the curative powers of mineral and thermal water springs in the countries in which they occur. Nowadays this is corroborated by a growing number of balneological healing centers, health resorts, and spas, which are well established in the system of public health care in many countries of the world. On the other hand, the bottling water markets positioned bottled waters as very good, healthy, and safe alternative for less healthy packaged beverages as carbonated soft drinks or fruit juice drinks.

Mineral and thermal waters still play an important role in social, cultural, and industrial development of many countries providing the human health, food, and renewable energy security.

Introduction

Mineral and thermal waters constitute a special category of groundwater which has been used by man for drinking, bathing, healing, therapy, or religious rites for thousands of years. From the beginning to today, people have believed in the curative powers of mineral and thermal water springs in the places and countries in which they occur. That is why throughout the centuries, thermal waters have always occupied a significant position in the history and development of nations, countries, civilizations, and societal relations. Nowadays utilization of mineral and thermal waters falls into three main areas which constitute important sectors of the economy, namely, (i) balneology – relates to the most traditional use of mineral and thermal water, balneotherapy, health and wellness industry, and warm baths for healing purposes with natural thermal waters of different chemical composition or viscosity (pure liquid water or mud); (ii) water bottling industry, relates to specially distinguished and regulated water-food market and to consumption of natural mineral waters, springwater, table waters, etc., packed and sold for consumers; and (iii) geothermics, geothermal energy sector connected with exploration and utilization of geothermal heat produced by thermal water for direct use and electric power generation.

This chapter is focused on the explanation of the key issues of mineral and thermal water occurrence, formation in the Earth's hydrosphere, and their exclusive socioeconomic role in modern societies resulting directly from the ways they are used and applied. The first chapter provides basic information on the distribution and movement of water in Earth's hydrosphere to show the position and relative volume of the groundwaters: groundwaters are direct source of mineral and thermal waters used by man. The second chapter deals with terminology and definitions of mineral and thermal waters and explains mutual relationships between these two water types. Confusion of terminology is particularly evident in relation to mineral waters: they are understood differently in different areas of usage, e.g., balneology vs. water bottling industry. The next three chapters are focused on the application and utilization of mineral and thermal waters in three main areas: balneology, geothermal energy, and water bottling industry. Balneology and balneotherapy refer to the oldest and the most traditional use of mineral and thermal waters and were described with special focus on historical aspects.

Geothermal energy is perhaps the least understood of all alternative energy resources among the general public and even among workers in other energy fields. In the twentieth century, high-temperature water resources have been used for the production of electricity, while medium- and low-temperature resources are used for domestic heating, from individual houses to whole communities, as well as for industrial and agricultural purposes. It usually comes as a surprise to learn that the use of thermal waters as geothermal heat carriers saves the equivalent of more than two million tonnes of oil per year. Various ways of utilization of thermal water heat are shown in light of the global use of geothermal energy.

The water bottling industry is a key sector in the economies of many countries. According to the newest analytical reports, the global bottled water market was valued at approximately USD 170.0 billion in 2014 and is expected to reach approximately USD 280.0 billion by 2020. In terms of volume, global bottled water market stood at around 290.0 billion liters in 2014. Marketed bottled mineral waters are treated as food, have their own definitions, and must comply with special safety regulations and standards which are shown in detail in the last chapter together with the newest data concerning consumption of bottled water in the world.

Distribution of Water in Earth's Hydrosphere: Sources of Mineral and Thermal Waters

Water – transparent fluid which forms oceans, streams, rivers, lakes, ice caps, and rains – is essential to life on our planet. About 71% of the Earth's surface is covered with water (Fig. 1).

About 96.5% of the Earth's water is stored in oceans. The remaining 3.5% is stored in glaciers and ice caps, groundwater, and soil moisture. More than 97% of the Earth's water is salt water, and only 2.5% accounts for freshwater. More than 75% of total water in land areas is locked in glacial ice or is saline. Only a small percentage (≤1%) of the world's total water supply is actually available to humans as freshwater for drinking and other purposes like industry, transportation, heating and cooling, etc. (Fig. 2).

Although the mass of water on Earth remains quite constant over time, its distribution among various reservoirs is variable depending on a wide range of climatic factors. This process of continuous movement (cycling) of water masses between Earth's lithosphere, atmosphere, hydrosphere, and biosphere is called the hydrological cycle (Fig. 3).

The processes that drive the movement of water from one reservoir to another are evaporation, condensation, precipitation, deposition, runoff, infiltration, sublimation, transpiration, melting, and groundwater flow.

G**roundwater** constitutes about 1.7% of the world's total water, which accounts for a volume of around 23.4×10^6 km^3 [1]. About 50% of this volume constitutes groundwater which occurs in strata below 1 km deep where water is usually saline and hot. Water has a distinguishing feature of dissolving a wider range of substances than any other liquid. That is why the variable chemistry of groundwater is an index of its interaction with geological environment, residence time, movement, and storage.

Groundwater is the source of mineral and thermal water.

Mineral and Thermal Waters: Terminology and Mutual Relations

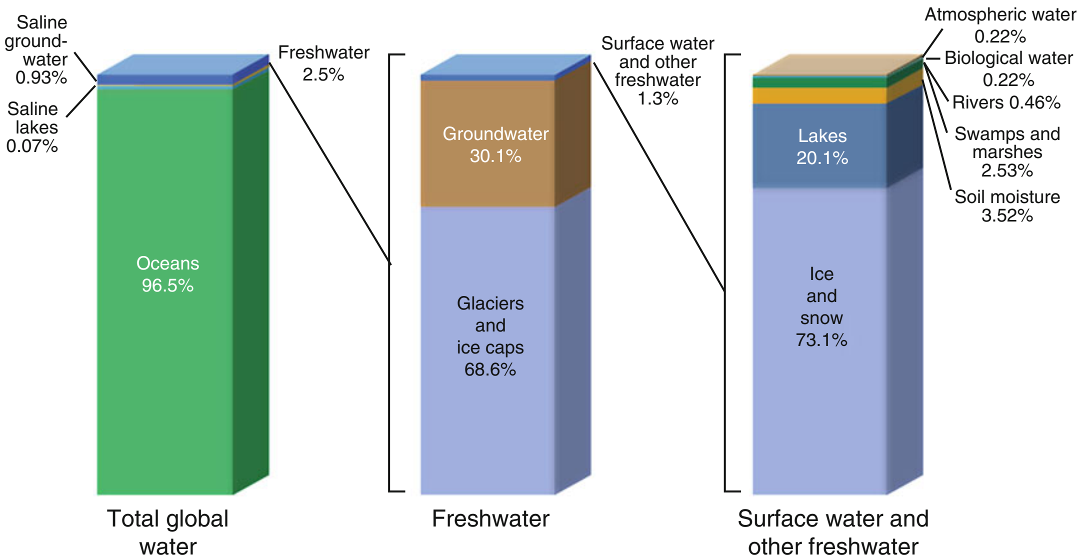
The term "mineral and thermal waters" is used widely in the literature and traditionally combines two kinds of waters, namely, mineral water and thermal water. Why is such combination justified? And how are these kinds of waters defined?

The term "mineral water" has not been precisely defined or understood for a long period of time. Today it is usually colloquially used to mean bottled water in contrast to tap water. The confusion with unambiguous definition of mineral water stems from different understanding of this term in different fields of its application and use. There are three main areas which refer to mineral water and describe this term in slightly different ways, namely, (i) balneological industry, (ii) water bottling industry, and (iii) the area of groundwater research – hydrogeology and hydrogeochemistry.

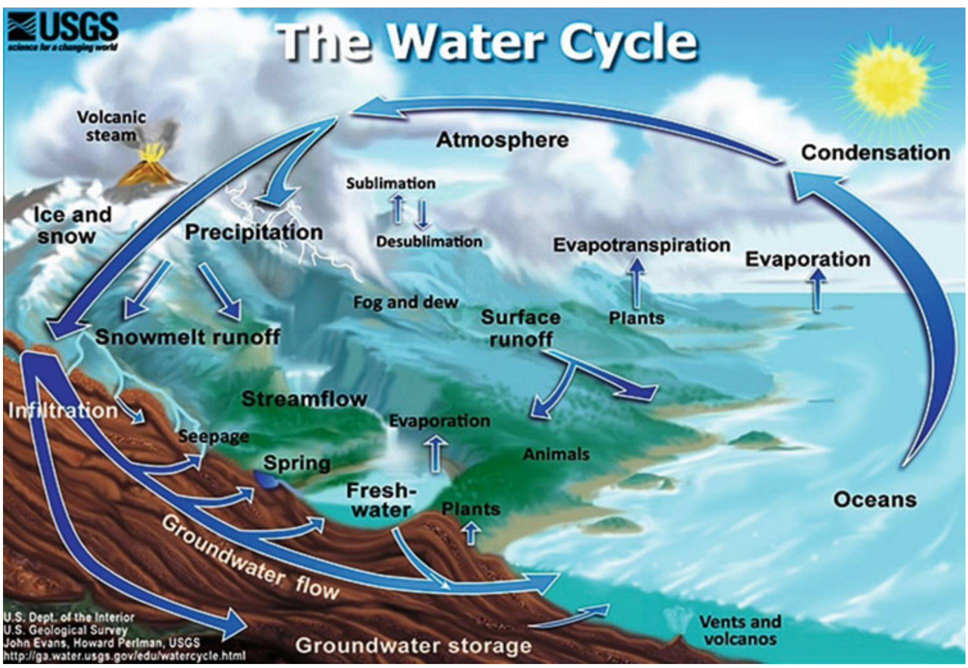
Traditionally the term "mineral water" is referred to specific groundwaters used for curative purposes which are characterized by higher than average concentrations of dissolved solids or gases or higher temperature. The formal definition of mineral water has been proposed for the first time by L. Grünhut in 1911 during the International Balneological Congress in Nauheim, Germany. It was established that this term referred only to groundwater which contains at least 1 g/kg of dissolved solids [4]. In fact, the great majority of spas and baths based on curative waters at that time utilized waters with mineralization higher than 1 g/kg [4, 5]. During the same Congress, the lower limits of concentrations of pharmacologically active compounds (e.g., F^-, Br^-, I^-, HBO_2, CO_2, Fe, H_2SiO_3, etc.) in water were agreed upon. Groundwater which contains less than 1 g/kg of dissolved solids but has concentrations of pharmacologically active compounds above the lower limit can be counted among curative mineral waters as the so-called specific waters. These criteria adapted in Nauheim are widely used today in many countries, especially those with well-developed spa treatment and balneological industries. Nevertheless, there are many reservations and critical remarks with regard to such definition of mineral waters stemming from medical, hydrogeological, or food premises.

In hydrogeology, in the area of groundwater research, exploration, and exploitation, groundwater is considered as a part of the geological environment remaining in constant interaction

Mineral and Thermal Waters, Fig. 1 Water constitutes nearly three-fourths of the Earth's surface



Mineral and Thermal Waters, Fig. 2 Distribution of Earth's water (Data after [1]; graph adapted from [2])



Mineral and Thermal Waters, Fig. 3 Hydrological cycle – continuous movement of water masses (adapted from [3])

with rocks. This water–rock interaction forms groundwater chemical composition and is responsible for concentration of specific ions and gas phases, as well as the content of total dissolved solids. From a hydrogeological point of view, groundwater is mineralized because it always contains some amount of dissolved solids. However, the term "mineral water" is exclusively used for groundwater in which total dissolved solids are higher than that of waters commonly used for drinking or household purposes. There are many classifications of groundwater which take into account the concentration of total dissolved solids (TDS) to distinguish between freshwaters and mineralized (i.e., mineral) waters; some of them are more logical and more convenient than others and have been used in hydrogeology for a long period of time (Table 1).

As can be seen from Table 1, the TDS of 1 g/kg is the most widely accepted boundary between fresh and mineral water.

Groundwater with elevated concentrations of dissolved solids also may reveal elevated temperatures in the place of their occurrence – and in such case, they also become thermal water if the temperature is above 20 °C at the outflow.

The beginnings of bottling mineral waters for the consumer market date back to the seventeenth century with rapid growth in the nineteenth century. The development of bottling water industry and bottling water market and possibilities of selling packaged water on the global market led to the necessity of implementing specific legal regulations for marketed mineral waters together with their own definitions. According to such regulations, bottled mineral waters have to comply with internationally recognized standards and recommendations relating to food (i.e., bottled water is treated as a food), food production, processing, health safety, chemical composition, packaging, labeling, transportation, etc. Marketed bottled waters are defined as natural mineral waters, springwaters, drinking waters, sparkling waters, etc. and may contain less than 1 g/L of total dissolved solids. Mineral water is treated separately from any other type of water, and it is not regulated by standards for drinking water from municipal sources or other bottled waters. Mineral water is not considered to be processed, distilled, or demineralized water; it is rather groundwater from a spring or a well that contains minerals naturally present in the water as it flows from the underground geological source. About 97% of all bottled water sold in Europe is either natural mineral water or springwater. Both types of water comply with the standards of the European Union Directive 2009/54/EC. Each brand of natural mineral and springwater has its own distinctive taste, a unique set of properties, and specific mineral composition which is derived from the geological conditions present in the area where the water is abstracted [9].

Various aspects of mineral water in the bottling industry are discussed in more detail in section "Mineral Water in Bottling Industry: Legislation and Consumption."

Thermal waters, also called **geothermal waters**, are defined as groundwaters with a temperature of at least 20 °C at the outflow. The isotherm of 20 °C lies at depths of about 1,500–2,000 m in the Earth's crust in regions of permafrost, rising to about 100 m or higher in subtropical regions; at the boundary with the tropics, it emerges onto the Earth's surface [10]. Thermal waters are extracted primarily for their heat content and only secondarily for their mineral content [17] High crustal temperatures and good aquifer permeability are the most important factors giving rise to thermal water systems from which heat can be obtained. The heat is extracted from liquid water or steam, which may be brought to the surface at natural emergences,

Mineral and Thermal Waters, Table 1 Exemplary classification of groundwater based on the total dissolved solids (TDS). Classification proposed by Davis [6] is more widely used and corresponds with majority of hydrogeological applications

Classification of Davis [6] TDS (ppm)		Classification of Krieger [7] TDS (ppm)		Classification of Pazdro and Kozerski [8] TDS (g/L)	
Freshwaters	0–1,000	Freshwaters	0–1,000	Ultra-freshwaters	<0.1
Brackish waters	1,000–10,000	Slightly saline	1,000–3,000	Freshwaters	0.1–0.5
Saline waters	10,000–100,000	Moderately saline	3,000–10,000	Acratopegae	0.5–1.0
Brines	>100,000	Very saline	10,000–35,000	Mineral waters	1.0–35.0
		Brines	>35,000	Brines	>35.0

such as fumaroles, solfataras, mofettes, geysers, and springs, or via wells and boreholes, which may be several thousand meters deep in sedimentary basins [17].

Most of high-enthalpy and many low-enthalpy geothermal systems are artesian, at least in the initial state. In artesian basins, waters with temperatures of 70–100 °C or more emerge from fissures that extend to depths of around 2,000–3,000 m. In mountainous regions, such as the Alps, the Caucasus, the Tien-Shan, and the Pamirs, thermal waters reach the surface in numerous hot springs, with temperatures of up to 50–90 °C [10].

In the regions of present-day volcanism, they are manifested in the form of geysers and steam vents, which emit steam-water mixtures and vapors upon emergence onto the surface; in this case, waters with temperatures of 150–250 °C emerge from fissures that extend to depths of around 500–1,000 m. Examples of these types of thermal waters are Pauzhetka on the Kamchatka Peninsula, Big Geysers in the USA, Wairakei in New Zealand, the soffione at Larderello in Italy, and geysers in Iceland.

Modern utilization of thermal waters began with people's awareness that the heat produced with thermal waters can be treated as geothermal energy resources which can be captured and harnessed for space heating or cooling, gardening, greenhouses, fish farming, etc. The fourteenth century was the beginning of utilization of thermal waters for space heating; a district heating network was applied for the first time in the French village of Chaudes-Aigues Cantal. Nowadays, space heating with geothermal energy is one of the most common and widespread direct uses of geothermal resources. The production of electricity requires high-temperature thermal water resources or dry steam with temperatures above 150 °C. Italian Prince Piero Ginori Conti was the first to test geothermal steam energy conversion to electric power between 1904 and 1905 at a geothermal field in Larderello, Italy (thermal water and steam temperatures in the range 180–260 °C); the first commercial power plant of 250 kWe was commissioned in Larderello in 1913. Nowadays, the use of thermal waters as geothermal heat carriers saves the equivalent of more than two million tonnes of oil per year [11].

Mineral and Thermal Water as a Curative Agent: From Past to Present

From the beginning of time, water has been an essential resource for the survival of man. It is not surprising that evidence of the earliest civilizations has been found along the banks of the rivers: the Tigris and Euphrates in Mesopotamia, the Nile in Egypt, the Indus in India, and the Huang-He (Yellow River) in China [9]. Besides freshwater for drinking and household purposes, thermal and mineral waters have been used by people for bathing, healing, therapy, or religious rites for thousands of years. Balneology – the practice of using natural mineral and thermal springwater for the treatment and cure of disease – also has a long history.

Archeological findings show that mineral and thermal waters have been used for bathing since the Bronze Age (about 12,000–3,000 years BC). There are many findings and examples showing that people preferred to settle in the vicinity of hot springs, e.g., in Japan – as early as 11,000 years BC. Location of human settlements in the vicinity of hot springs took place in many areas in the world and in many periods of human history; the beginnings of such cities like Budapest, Sofia, and Reykjavik are directly connected with the occurrence of thermal springs [12–14]. Hot springs have been used in connection with religious rites in ancient Egypt and by the Jews of the Middle East [15]. From the earliest times, humanity has associated certain mineral and thermal water springs with divine powers of healing [16]. In the first millennium BC, the Greeks believed in divinities associated with thermal and mineral waters and their curative properties. On the island of Kos, the healing powers of the local spring were attributed to Asclepios, the god of medicine and son of Apollo, and a magnificent sanctuary was built for him at the end of the fourth century BC [17]. Greek physician Hippocrates of Kos, called "the father of medicine," applied thermal water for medical treatment. Ancient Greeks developed the school of physiotherapy and balneology based upon geothermal waters and herbs. They are also the authors of the popular maxim "health from water." Romans took over this motto, which is currently known as "SPA" – *salus per aquam* in Latin. The word "spa" also traces its origin to a town "Spa" near Liege in southern Belgium near the German border. The mineral springs of Spa, mentioned already by Pliny the Elder, were rediscovered in 1326 when an ironmaster used the iron-bearing springwater to cure his ailments. He founded a health resort at the spring called Espa (meaning fountain in the Walloon language). Espa became so popular that the word known in English as "spa" became the common designation for similar health resorts around the world [17].

The Romans were attracted by the curative properties of thermal springs and often settled in their vicinity. In the Roman Empire, hot baths were a rite, an art of relaxation, and an important element of social and political life. Baths played an important role in the formation of urban societies as well as in the development of trade connections and economic relationships. Ancient Romans disseminated the idea of hot baths in many parts of the empire. Many spas famous today have a long history dating back to Roman times, for example, Baden-Baden in Germany was originally known as Aquae Aureliae, Wiesbaden in Germany known as Aquae Mattiacorum, Bath in England known as Aquae Sulis, and in France Aix-les-Bains known as Aquae Allobrogum and Aix-en-Provence known as Aquae Sextiae [17–19]. Successors of Roman baths were Turks who restored or implemented this tradition in various parts of their Ottoman Empire.

In Europe, after the fall of the Roman Empire, the majority of the baths they established were neglected. However, many early churches were built on sites of ancient healing. In the thirteenth and fourteenth centuries, mineral and thermal water baths started to be rediscovered, and newer ones developed across Europe; in the eighteenth century, many baths were rebuilt, and many new "watering places" or "spas" were established. Spas then became fashionable centers of resort for the upper classes and nobility [17]. The thermal springs at Buda were known in thirteenth-century Hungary due to the Knights of the Order of St John, who started to use them for balneal treatment. Interest in the therapeutic use of thermal and mineral waters in Bohemia increased markedly in the fourteenth century, culminating in 1348 when the spa of Karlovy Vary (Var meaning "boiling") was established by order of Charles IV of Luxembourg. The town, formerly known by its German name of Carlsbad, and its treatment facilities developed rapidly on the basis of 12 naturally carbonated thermal springs [17, 20]. In present-day Romania, the gradual development of spa towns in the vicinity of thermal and mineral water sources occurred from the thirteenth to nineteenth centuries. In medieval Transylvania, the warm baths southeast of Oradea were first described in 1221.

Geothermal waters were also used in Chinese medicine, which developed principles of medical treatment: so-called "cold" diseases (e.g., rheumatism) should be healed with "warm" cures (commonly geothermal waters), whereas the diseases connected with high temperature require "cold" cures and thus could not be treated with warm water [21].

Specific traditions of usage of geothermal waters have also been developed throughout the centuries in Japan. They developed the so-called onsen – special center – unknown from other countries in which physical and mental regeneration was practiced along with medical treatment at the site of geothermal springs [22].

In the USA, the use of natural springs, especially geothermal ones, has gone through three stages of development: (i) use by Indians as sacred places, (ii) development by the early European settlers to emulate the spas of Europe, and (iii) as a place of relaxation and fitness [20]. The Indians of the Americas considered hot springs as sacred places and believed in the healing powers of the heat and mineral waters. Montezuma, the great Aztec leader, spent time at a spa, Aqua Hedionda, to recuperate from his strenuous duties; it was later developed into a fashionable spa by the Spaniards. Every major hot spring in the USA has some record of use by the Indians, some for over 10,000 years. These springs were also known as neutral ground, to which warriors could travel and rest unmolested by other tribes. Here they would recuperate from battle [17, 20, 23]. European settlers in the USA and Canada in seventeenth and eighteenth century found and used these natural hot springs and later, realizing their commercial value, developed many into spas after the tradition in Europe.

The popularity of spas as fashionable resorts declined rapidly in Europe and the USA in the early part of the twentieth century. However, their use for therapeutic treatment, often as part of the national health services, has continued through today in many countries, especially in Central and Eastern Europe, Russia, Germany, France, Austria, Hungary, Romania, the Czech Republic, Slovakia, and also in China [15, 24]. The USA, unlike Europe, did not have the government, trade unions, social security, and national health insurance program to support developments as spas or healing centers. Interest in spas languished, and many of the majestic resorts went into decline and closed at this time [20].

The use of medically supervised spas, especially in Europe and Japan, has long been accepted. They are used for both treatment and preventive therapy. At the end of the twentieth century, a significant resurgence in the use of hot spring and thermal bathing as a method of cure and prevention was observed. Famous spa towns came back to life as modernization led the hot springs toward becoming commercial ventures. Two such examples are the well-known thermal bathing center of Karlovy Vary in the Czech Republic and the Piestany spa in Slovakia. The Karlovy Vary spa became one of the leading thermal spring providers in the early part of the century. In 1911 AD, it recorded a guest count of 70,935 guests, adding to its fame and attracting many European elites of the day. Natural hot springs and their benefits were also rediscovered and developed in many countries famous today for their hot springs, including European countries like Hungary, France, Slovakia, Czech Republic, Germany, Romania, Italy, Poland, Russia, Portugal, Spain, Turkey, and many others like Iceland, Japan, China, Taiwan, and Argentina.

In the former Czechoslovakia, there are 52 mineral water health spas and more than 1,900 mineral springs, to which every year about 220,000 citizens are granted free spa treatment for 3 weeks, paid by the national health insurance program [14, 15, 17].

Poland has more than 42 prosperous health resorts based on mineral and thermal waters and curative muds which are well established in the system of public health care. Over 70% of health resorts and towns with curative waters are located in southern Poland in the Sudetes and Carpathian regions, together with Carpathian Depression.

As the world's largest country by area, Russia possesses a diverse landscape and with it comes copious natural resources, ranging from thermal water in the highlands to abundant muds, salts, herbs, and gases suitable for holistic treatments. Russia has hundreds of mineral springs, many of which have been developed into fully operational recreation and treatment facilities. Banya, the famous sauna culture in Russia dating back 1,000 years, is also ripe for development. According to Spa and Wellness International Council (SWIC, Russia), there are 2,000 traditional medical-focused health resorts or sanatoriums in the country which are often located on mineral springs. There are also more than 1,500 spas; however, only about one-third correspond to international standards [25]. Today, there are approximately 210 spas in the USA with 4.5 million persons attending a spa in 1997 [15]. In Japan there are over 1,500 spas that are used by over 100 million visitors every year.

Spa and health resorts which belong to the system of public health care and public balneotherapy in European countries usually are based on the mineral and thermal waters which are called curative waters or specific waters. Curative water is usually mineral or thermal water, or groundwater, with no chemical and microbiological contamination and with natural diversity of physical and chemical properties, containing specific concentrations of pharmacologically active compounds above the lower limit agreed in the special regulations and legislation in particular countries. For example in Poland, curative water must comply with the following requirement according to Act of 9 June 2011 Geological and Mining Law:

– Total solid dissolved mineral content at least 1,000 mg/dm^3
– Ferrous ion content – at least 10 mg/dm^3 (ferruginous waters)

– Fluoride ion content – at least 2 mg/dm^3 (fluoride waters)
– Iodine ion content – at least 1 mg/dm^3 (iodide waters)
– Bivalent sulfur ion content – at least 1 mg/dm^3 (sulfide waters)
– Meta-silicic acid content – at least 70 mg/dm^3 (silicic waters)
– Radon content – at least 74 Bq (radon waters)
– Carbon dioxide content – at least 250 mg/dm^3 (250–1,000 mg/dm^3 carbonic acid waters, >1,000 mg/dm^3 carbonated water)

The regulatory concentration of pharmacologically active compounds is derived from medically observed curative and healing effects; such values may be different in other countries.

Nowadays however, some of the most common reasons for bathing in mineral thermal waters are to revitalize the skin, calm the nerves, detoxify the body, and refresh the oxygen levels. Many also turn to thermal mineral waters for health issues such as arthritis, fibromyalgia, skin conditions, depression, respiratory illnesses including asthma, and locomotor and circulatory diseases. Improving health and appearance and escaping stress to refresh and revitalize the body and the mind are the main reasons why people go to spas and why spas are becoming an increasingly important part of life today.

Thermal Water as Renewable Energy Source

Geothermal Energy, Geothermal Systems, and Geothermal Resources

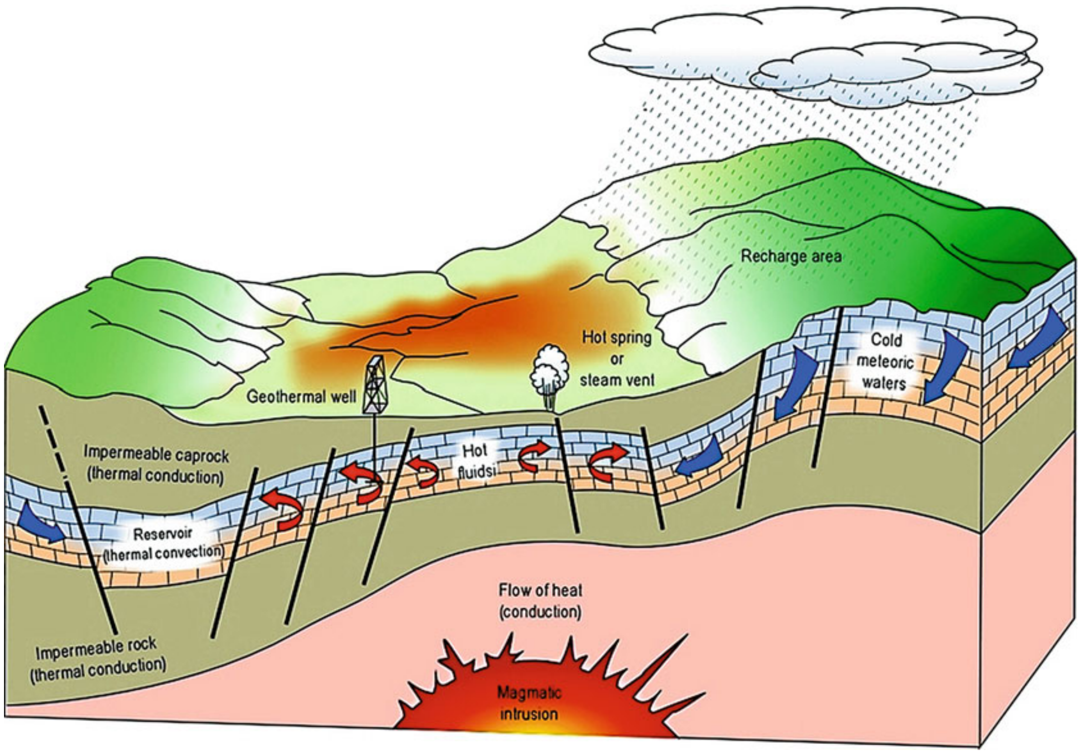
Geothermal energy is heat energy generated and stored in the Earth. According to current concepts, there are four main sources of geothermal energy [26–28]: (i) initial heat of the Earth related to impact of particles at the accretion of protoplanetary nebula and its enthalpy (i.e., primordial energy of planetary accretion), (ii) change in potential gravitational energy of the planet due to density differentiation, (iii) change of kinetic energy of the Earth's gyration accompanied by the transformation of mechanical energy into thermal ones by means of tidal friction, and (iv) release of nuclear energy at decay of radioactive elements. The intensity of energy release by each of these sources varies in time. Common energy release in the Earth is estimated at 5.3–7.2 $\times$ 10^{13} W [28]. The term "geothermal energy" should be understood not only as the "heat generated and stored within the Earth" but also as an indication of that part of the Earth's heat that can be extracted and used by man as a resource [29].

A **geothermal system** consists of three elements [29] (Fig. 4):

– Heat source
– Reservoir (i.e., aquifer)
– Fluid – i.e., thermal water (or steam) as a heat carrier

The geothermal system comprises the entire hydrogeologically connected area with the recharge zone, reservoir aquifer, and outflow zone (Fig. 4). A heat source enables constant conductive transfer of heat energy to the surrounding rocks and pore (or fissure) groundwaters increasing their temperature and the geothermal gradient of the area. The geothermal aquifer, which plays a role of **geothermal reservoir**, is the volume of heated and permeable rocks which can be exploited by extraction of thermal water (or steam). The geothermal reservoir is usually overlain by a cover of impermeable rocks and connected to the recharge zone, where the meteoric waters can partly replace the fluids that escape from the reservoir through springs or are extracted by boreholes. Fluid (water) convection heated from the heat source at the base of the circulation system is a driving force in the geothermal reservoir; heated fluid of lower density tends to rise upward leaving place for the colder fluid of higher density coming from the margins of the system [29]. Depending on various hydrogeological and geological settings and thermal conditions, geothermal systems may vary considerably. According to [17] the following geothermal systems may be distinguished:

(a) According to the manner of heat transport:
   – Conductive systems
   – Convective systems



Mineral and Thermal Waters, Fig. 4 Schematic illustration of classical geothermal system (Adapted from [13])

(b) According to the mobility of the components:
  – Systems with mobile components (water, gases, and even magma)
  – Systems without mobile components (i.e., hot dry rock)
(c) According to the nature of the components:
  – Hot dry rock systems
  – Magmatic systems
  – Thermal water systems
(d) According to the state of the components:
  – Systems of solid components
  – Systems of magmatic melts with and without gases and vapor
  – High-enthalpy systems (generally dominated by steam)
  – Low-enthalpy systems (liquid water dominated)
(e) According to the hydraulic closure of the components:
  – Closed systems – where the thermal fluid is confined by strata of low permeability, recharge of water is slow, and over long distances, the discharge of water or steam is intermittent
  – Open systems – where the water flows from an area of recharge to one of discharge and is not confined by a low permeability layer

Geothermal energy resources are defined as the amount of thermal energy (heat) accumulated in the Earth's crust down to a given depth, referred to as the particular area for which the calculations are made and for the mean annual temperature at that point on the Earth's surface [30, 31].

Geothermal energy resources can be of two types [28]:

(a) Hydrothermal resource – when the geothermal energy is accumulated in hot groundwater, i.e., the heat carrier is thermal water or steam extracted with the wells.
(b) Petrothermal resource – when the geothermal energy is accumulated in hot rocks (including

magmatic melts); the energy carrier is a media (usually water) injected through wells into the hot rock formations (the so-called hot dry rocks – HDR) and/or salt diapirs.

The geothermal gradient, which is the difference in temperature between the core of the planet and its surface, drives a continuous conduction of thermal energy in the form of heat from the core to the surface. The output of geothermal energy to the atmosphere and hydrosphere occurs also with advective mechanism as a result of volcanic and hydrothermal activity. The **conductive heat flow** (i.e., heat flow density – flow of energy per unit of area per unit of time) is characterized by considerable regional variations. Polyak and Smirnov [32, 33] showed that abyssal heat flow is naturally related to the age of continental tectonic units and tectonomagnetic activity: the densities of the average heat flow in the Precambrian, Paleozoic, and Mesozoic units are about 45, 55, and 70 mW/m^2, respectively, and reach 90 mW/m^2 in areas of Cenozoic volcanism; i.e., heat flow density increases as the age of tectonomagnetic activity decreases [28]. In active zones of the continent – ocean transition – the heat flow density values vary abruptly from 20 to 40 mW/m^2 in trenches up to 80–400 mW/m^2 in volcanic arcs and back-arc basins; the highest heat flow occurs in the zones of mid-ocean ridges, where the total heat flux reaches 400–800 mW/m^2. The global mean value of conductive heat flow density is estimated at 85 mW/m^2 (e.g., [34, 35]).

The heat flow density distribution predetermines the variations of the depth temperatures, which determines the geothermal gradient, geothermal regions, and groundwater temperatures. The **geothermal gradient** expresses the increase in temperature with depth in the Earth's crust. The average geothermal gradient is about 2.5–3.0 °C/100 m, but ranges vary widely in different parts of the Earth's crust; in geothermal areas the geothermal gradient is much higher than the average value.

Geothermal energy resources are extracted from geothermal systems. They are usually (and traditionally) classified among **renewable resources of heat energy**. However, the ability to renew the energy resources depends on the rate of energy recharge. In the exploitation of natural geothermal systems, energy recharge takes place by advection of thermal waters on the same time scale as production from the resource. In such cases, the geothermal resources are renewable. Reaching and maintaining this equilibrium in the geothermal system indicates also the sustainable use of the energy resources. On the other hand, there are geothermal resources (e.g., hot, dry rocks, and some of the hot water aquifers in sedimentary basins) where energy recharge is only by thermal conduction. Due to the slow rate of thermal conduction processes, the rate of the heat energy consumption is greater than the rate of energy recharge; such geothermal systems should be considered as finite energy resources (e.g., [29]).

Classification of Geothermal Resource

Currently, there is no standard, uniform classification of geothermal energy resources in the world literature (e.g., [29, 31, 36]). The main factor determining the potential use of particular geothermal resources is the reservoir temperature, i.e., the enthalpy (temperature) of the heat carrier.

In relation to the reservoir temperature, the geothermal resources are usually divided into three groups (Table 2): (i) low-enthalpy resources, (ii) intermediate-enthalpy resources, and (iii) high-enthalpy resources. As can be seen in Table 2, there is no full agreement in relation to temperature ranges for particular group of resources.

Such simple classification qualitatively shows the energetic value of various thermal water resources. It is useful only to qualify the subsequent potential application of thermal water to direct use or to electricity production. Taking into account the state of actual technological development in the area of geothermal energy utilization, the water at temperatures greater than around 100 °C can be used to produce electricity (e.g., binary systems), although the efficiency of electricity generation (and therefore, the economic viability) will increase with higher temperatures. It is widely accepted that the high-enthalpy resources correspond to reservoir temperatures higher than 150 °C and can be applied to conventional electric power generation. Thermal water with temperature below 150 °C is recommended

Mineral and Thermal Waters, Table 2 Geothermal resources classification in relation to enthalpy of the heat carrier

Geothermal resources	Temperature (°C)				
	Muffler and Cataldi [30]	Hochstein [37]	Benderitter and Cormy [38]	Nicholson [39]	Axelsson and Gunnlaugsson [40]
Low-enthalpy resources	<90	<125	<100	≤150	≤190
Intermediate-enthalpy resources	90–150	125–225	100–200	–	–
High-enthalpy resources	>150	>225	>200	>150	>190

to be utilized in a direct way, e.g., district heating, greenhouses, aquaculture, etc. In relation to this classification, distinction is often made between water-dominated systems and vapor-dominated (i.e., dry steam) systems. In water-dominated systems, liquid water is the pressure-controlling fluid phase. These geothermal systems with temperature range <125–>225 °C are the most widely distributed in the world and depending on temperature may produce hot water, water and steam mixture, wet steam, and sometimes dry steam [29]. In vapor-dominated systems, liquid water and vapor coexist in the system, with vapor as the continuous pressure-controlling phase. Geothermal systems of this type are high-temperature systems, usually produce dry steam or superheated steam, and are rather rare in the world, e.g., Geysers in California and Larderello in Italy [29].

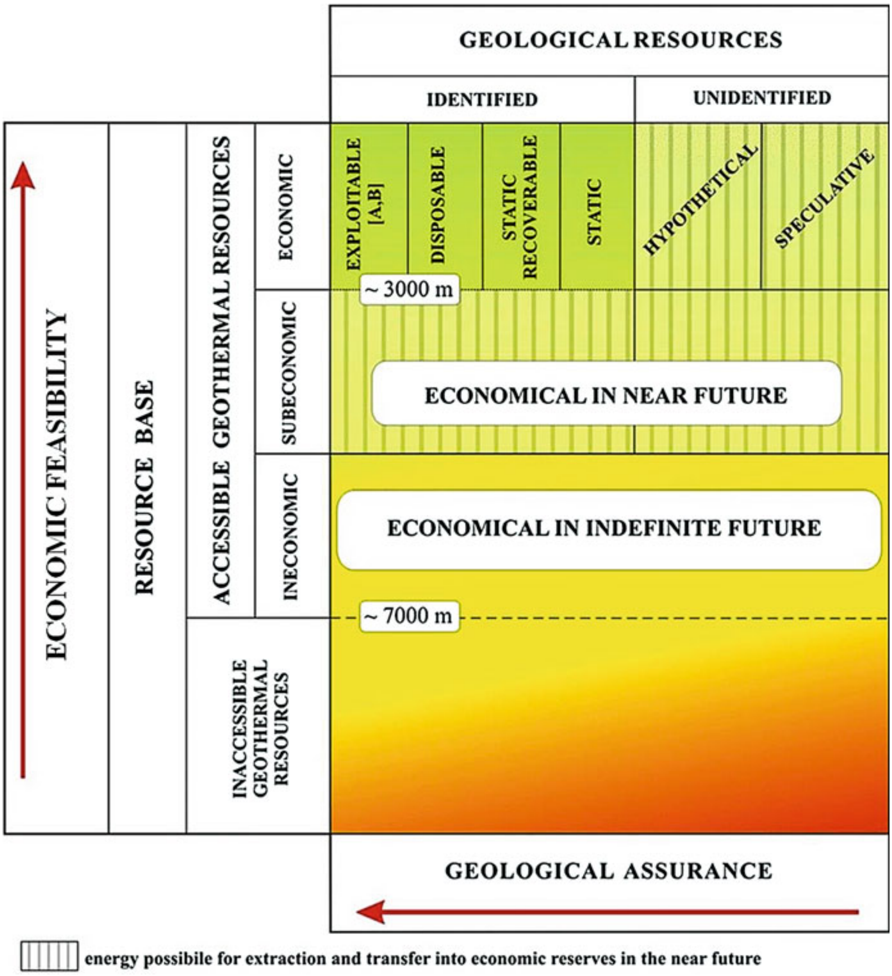
Another much more advanced and more "quantitative" example of classification of geothermal resources worth presenting here takes into account the geological recognition of the geothermal resources together with ecological, technological, and economic aspects of their development, exploitation, and utilization. The classification shown here was developed by Wojciech Górecki from the University of Science and Technology in Kraków, Poland, in the 1990s and later improved via its application in many geothermal projects including successful evaluation and classification of geothermal resources of Poland. Results and details of this application can be found in Górecki et al. [31, 36, 41–44] and Hajto [45].

The classification (Fig. 5) is based on the well-known format of the McKelvey diagram. The McKelvey diagram utilizes the resources and reserves to assess the amount of any minerals on the Earth in the light of two distinct parameters: degree of certainty and the profitability (technically: these parameters are used as axes of the diagram [31, 46]).

The presented classification combines the concepts of geothermal resources assessment and classification used in European countries [37, 47–51, 66]. Moreover, the classification takes into account the latest requirements of classification of geothermal resources proposed by the United Nations Framework Classification for Fossil Energy and Mineral Reserves and Resources 2009 [52] and Hajto [45].

The horizontal axis of the diagram (Fig. 4) represents the accuracy of geological recognition (i.e., degree of certainty), whereas the vertical one displays the depth of occurrence and the economic efficiency of geothermal energy use (i.e., profitability). The methodology takes also into account the level of geological confidence (hydrogeological and thermal parameters) of the study area, which enables estimation of resources according to the classification as accessible, static, static-recoverable, disposable, and exploitable resources and reserves. These particular types of the geothermal resources are defined as follows [31, 36]:

(a) *The accessible geothermal energy resources* – the amount of thermal energy accumulated in the Earth's crust down to 3,000 m depth or to the top of the crystalline basement, referred to the mean annual Earth's surface temperature expressed in [J].



Mineral and Thermal Waters, Fig. 5 The McKelvey diagram presenting classification of geothermal resources (After [31, 45])

(b) *The static resources of geothermal waters and energy* – the amounts of free (gravitational) geothermal water hosted in pores, fractures, or caverns of given hydrogeothermal horizon, expressed in [m^3] or [km^3], and recalculated into the energy units [J]. These resources are calculated when the recognition of continuous groundwater reservoirs or horizons is possible in a given area. Based on the determined properties of the aquifer (i.e., lithology, thickness, porosity, and permeability), the identification of producing reservoirs and horizons is possible.

(c) *The static, recoverable geothermal waters and energy resources* – constitute only a part of the static resources diminished by the recovery index R_o, expressed in [m^3] or [km^3], and recalculated into the energy units [J].

(d) *The disposable geothermal waters and energy resources* – the amount of free (gravitational) geothermal water within the horizon or other calculation unit, which can be developed under given conditions but without detailed localization as well as technical and economic specification of an intake and expressed in [m^3/day], [m^3/year], [J/year], or [TOE/year] (TOE – tonnes of oil equivalent).

(e) *The exploitable geothermal waters and energy resources* – the amount of free (gravitational) geothermal water, which can be produced at given geological and environmental settings with intakes of optimum technical and economic parameters, expressed in [m^3/h] and [m^3/day] at relevant drawdown, and recalculated into [J/year] or [TOE/year].

Details concerning the calculation of particular resources can be found in Górecki et al. [31, 36, 41].

The classification according to the McKelvey diagram was modified in order to better adjust the terminology and definitions to the Polish assessment concepts and prevailing low-enthalpy resources. The volume method of determination of the "heat in place" of the geothermal resources was employed taking into account assumptions suggested by Muffler and Cataldi [30], Gringarten and Sauty [53], and Lindal [54] and with strong concerns for both the economic and ecological aspects of geothermal energy assessment. The accessible and the static geothermal energy resources have exclusively cognitive meanings, whereas the disposable, and particularly the exploitable, resources are of practical importance.

Utilization of Thermal Waters

The resources of heat energy accumulated in thermal waters (including steam) can be used in many different ways. To illustrate the possible utilization of thermal waters with respect to their temperature, the classical Lindal [54] diagram can be used as shown in Fig. 6.

As can be seen from the diagram, thermal waters at temperatures around 20 °C at the outflow are rarely used, and if so, only in specific conditions or in heat pump applications. The usefulness of thermal waters to electricity production begins at temperatures around 100 °C: application of binary systems is required. At temperatures around 140–150 °C and higher, thermal water resources can be used to generate electric power in conventional ways. Two facts can be deduced from the Lindal diagram [55]: (i) the temperature of thermal water resources may limit their possible use and (ii) the cascading utilization of thermal resources may increase the feasibility and profitability of geothermal projects.

The utilization of heat energy accumulated in thermal waters and steam falls into two traditional categories:

1. Electric power generation
2. Direct use – where the space (district) heating is dominant

Direct Use of Thermal Waters: Heating and Cooling

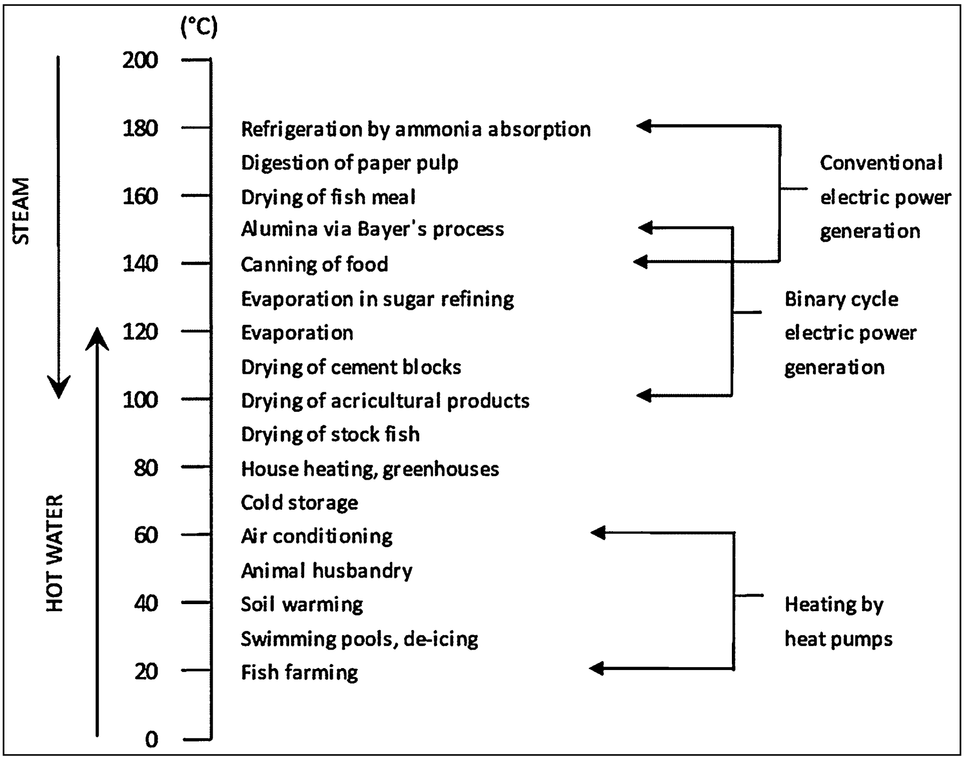
Direct or non-electric utilization of geothermal energy refers to the immediate (i.e., direct) use of heat energy rather than its conversion to some other forms, for example, electrical energy [56]. Direct applications utilize low- to intermediate-enthalpy geothermal resources which correspond with temperature range between 50 and 150 °C. Thermal waters are the principal heat carrier at the low-enthalpy geothermal resources, and their reservoirs can be exploited by conventional water well drilling equipment.

The low-enthalpy systems are much more widespread than the high-enthalpy ones, and they are more likely to be located near potential users. In the USA, for example, of the 1,350 known geothermal systems, only 5% have temperatures above 150 °C, and 85% are below 90 °C [56, 57]. In fact, almost every country in the world has some low-enthalpy geothermal resources.

Direct utilization of thermal waters is one of the oldest, the most versatile, and also the most common forms of usage of geothermal energy [15, 29]. Direct utilization consists of various forms of heating and cooling. The major areas of direct utilization of thermal waters [56] are:

(a) Swimming, bathing, and balneology (therapeutic use)
(b) Space heating and cooling
(c) Agriculture applications (mainly greenhouse heating and some animal husbandry)
(d) Aquaculture applications (mainly fish pond and raceway heating)
(e) Industrial processes
(f) Heat pumps (for heating and cooling)

Typical component of a direct-use heating system is presented in Fig. 7.



Mineral and Thermal Waters, Fig. 6 Lindal diagram depicting the possible use of thermal waters and steam in relation to temperature (After [54])

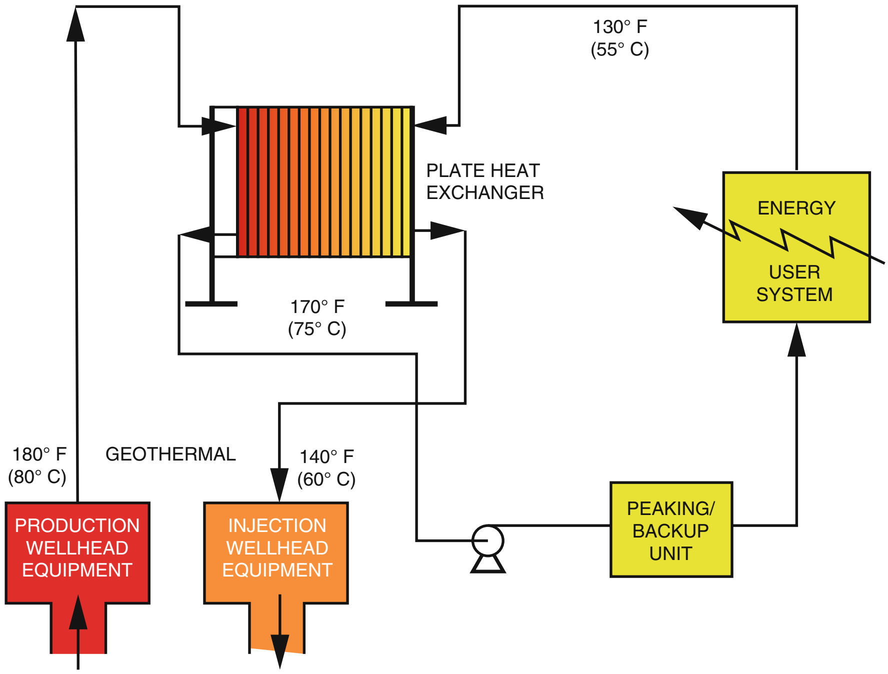
The comprehensive summary of the current world's direct utilization of geothermal heat and thermal waters may be found in Lund and Boyd [59] and Lund et al. [60].

Worldwide total installed geothermal power of direct use of geothermal energy is around 70,329 MWt, and energy use is about 587,786 TJ/year (163,287 GWh/year) distributed among 82 countries [59]. Direct use of geothermal energy in particular countries in 2015 is summarized in Table 3.

The leading countries in terms of installed geothermal power (i.e., installed capacity, MWt) are the USA, China, Sweden, Turkey, and Germany, accounting for 68.5% of the world geothermal power; in terms of annual energy use (TJ/year), the leading countries are China, the USA, Sweden, Turkey, and Iceland, accounting for 63.6% of the world usage. The largest percent increase in geothermal installed capacity (MWt) over the past 5 years was in Thailand, Egypt, India, South Korea, and Mongolia. In terms of annual energy use (TJ/year), the largest percent of increases over the past 5 years was in Thailand, Egypt, the Philippines, Albania, and Belarus. Most of these increases were due to geothermal heat pump installations and/or better reporting on bathing and swimming use [60].

The distribution of direct utilization of geothermal energy among the various forms of use worldwide in 2015 is presented in Figs. 8 and 9.

District heating (including heat pumps) represents nearly 88% of the installed geothermal power (MWt) and nearly 89% of the annual geothermal energy use (TJ/year). The world's largest consumption of direct-use geothermal heat is connected with application of geothermal heat pumps; annual geothermal energy use accounts for 55.3% and installed geothermal power for 70.95%. Although most of the installations occur in North America, Europe, and China, the number of countries with installations increased from 26 in



Mineral and Thermal Waters, Fig. 7 Typical components of a direct-use heating system (Adapted from [58])

2000 to 33 in 2005 to 43 in 2010 and to 48 in 2015 [15].

The equivalent number of installed 12 kW units (typical of the USA and Western Europe homes) is approximately 4.16 million. In contrast to Europe, in the USA most units are sized for peak cooling load and are oversized for heating; in Europe, most units are sized for the heating load [59].

Space heating has increased 44% in installed geothermal power and in annual energy use over 2010. The installed geothermal power in 2015 reached 7,556 MWt and the annual energy use 88,222 TJ/year. The leaders in district heating in terms of annual geothermal energy use are China, Iceland, Turkey, France, and Germany, whereas Turkey, the USA, Italy, Slovakia, and Russia are the major users in the individual space heating sector.

Bathing and swimming are the second largest area of direct usage of geothermal heat. Almost every country has spas and resorts that have swimming pools heated with geothermal water, including balneology – the treatment of diseases with thermal water. However, it is a common practice that the water is allowed to flow continuously, regardless of use (such as at night when the pool is closed). As a result, the actual usage and capacity figures may be high.

Actual installed geothermal power in this sector is about 9,140 MWt, and the annual energy use is 119,381 TJ/year [59].

Average **capacity factors** (i.e., load factors) determined for each country (Table 3) vary from 0.09 to 0.99. The lower values usually indicate countries in which geothermal heat pump usage predominates; capacity factor for the category "geothermal heat pumps" is low and estimated to be 0.21 [59]. The higher values of capacity factors are for countries with high industrial use (e.g., New Zealand) or continuous operation of pools for swimming (e.g., Algeria, Caribbean Islands, Madagascar, and Mexico). The worldwide capacity factor dropped from 0.40 in 2000 to the current 0.265 in 2015. It is assumed to be a result of the increase in geothermal heat pump usage [60]. Capacity factors for the various categories of use remain approximately constant when compared to 2010, except for industrial uses which dropped from 0.70 to 0.54 [59]. The capacity factor is calculated as follows:

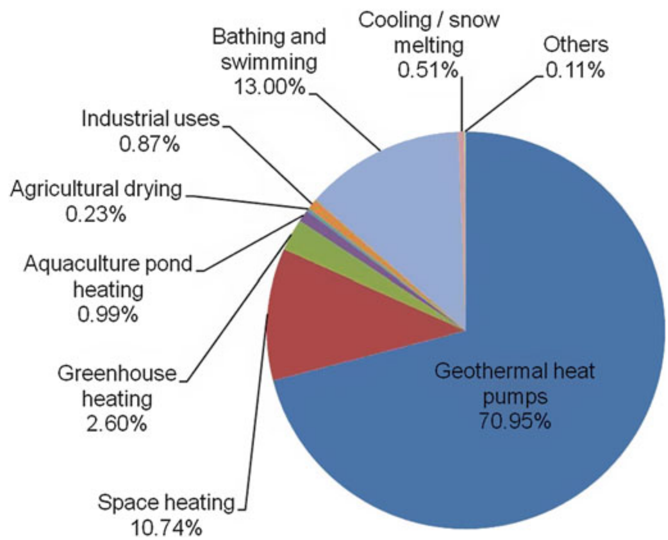
Mineral and Thermal Waters, Table 3 Direct use of geothermal energy worldwide in 2015 after Lund and Boyd [59]. *CF* capacity factor for each country; for additional explanation, see the text

Country	Installed geothermal power (MWt)	Energy use (TJ/year)	Energy use (GWh/year)	CF	Country	Installed geothermal power (MWt)	Energy use (TJ/year)	Energy use (GWh/year)	CF
Albania	16.23	107.59	29.89	0.21	Jordan	153.30	1,540.00	427.81	0.32
Algeria	54.64	1,699.65	472.25	0.99	Kenya	22.40	182.62	50.73	0.26
Argentina	163.60	1,000.03	277.81	0.19	Korea (South)	835.80	2,682.65	745.24	0.10
Armenia	1.50	22.50	6.25	0.48	Latvia	1.63	31.81	8.84	0.62
Australia	16.09	194.36	53.99	0.38	Lithuania	94.60	712.90	198.04	0.24
Austria	903.40	6,538.00	1,816.26	0.23	Macedonia	48.68	601.11	166.99	0.39
Belarus	4.73	113.53	31.54	0.76	Madagascar	2.81	75.59	21.00	0.85
Belgium	206.08	864.40	24.01	0.13	Mexico	155.82	4,171.00	1,158.70	0.85
Bosnia and Herzegovina	23.92	252.33	70.10	0.33	Mongolia	20.16	340.46	94.58	0.54
Brazil	360.10	6,622.40	1,839.70	0.58	Morocco	5.00	50.00	13.89	0.32
Bulgaria	93.11	1,224.42	340.14	0.42	Nepal	3.32	81.11	22.53	0.78
Canada	1,466.78	11,615.00	3,226.65	0.25	The Netherlands	790.00	6,426.00	1,785.14	0.26
Caribbean Islands	0.10	2.78	0.77	0.85	New Zealand	487.45	8,621.00	2,394.91	0.56
Chile	19.91	186.12	51.70	0.30	Norway	1,300.00	8,260.00	2,294.63	0.20
China	17,870.00	174,352.00	48,434.99	0.31	Pakistan	0.54	2.46	0.68	0.14
Columbia	18.00	289.88	80.50	0.51	Papua New Guinea	0.10	1.00	0.28	0.32
Costa Rica	1.00	21.00	5.83	0.67	Peru	3.00	61.00	16.95	0.64
Croatia	79.94	684.49	190.15	0.27	Philippines	3.30	39.58	11.00	0.38
Czech Republic	304.50	1,790.00	497.26	0.19	Poland	488.84	2,742.60	761.89	0.18
Denmark	353.00	3,755.00	1,043.14	0.34	Portugal	35.20	478.20	132.84	0.43
Ecuador	5.16	102.40	28.45	0.63	Romania	245.13	1,905.32	529.30	0.25
Egypt	6.80	88.00	24.45	0.41	Russia	308.20	6,143.50	1,706.66	0.63
El Salvador	3.36	56.00	15.56	0.53	Saudi Arabia	44.00	152.89	42.47	0.11

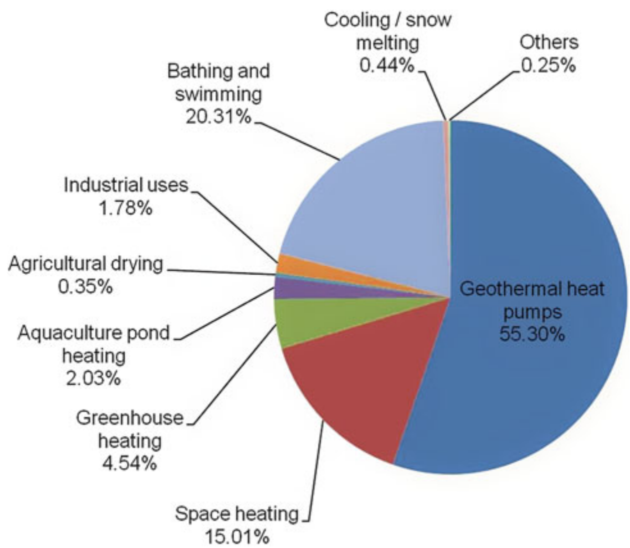
(*continued*)

Mineral and Thermal Waters, Table 3 (continued)

Country	Installed geothermal power (MWt)	Energy use (TJ/year)	Energy use (GWh/year)	CF	Country	Installed geothermal power (MWt)	Energy use (TJ/year)	Energy use (GWh/year)	CF
Estonia	63.00	356.00	98.90	0.18	Serbia	115.64	1,802.48	500.73	0.49
Ethiopia	2.20	41.60	11.56	0.60	Slovak Republic	149.40	2,469.60	686.05	0.52
Finland	1,560.00	18,000.00	5,000.40	0.37	Slovenia	152.75	1,137.23	315.93	0.24
France	2,346.90	15,867.00	4,407.85	0.21	South Africa	2.30	37.00	10.28	0.51
Georgia	73.42	695.16	193.12	0.30	Spain	64.13	344.85	95.80	0.17
Germany	2,848.60	19,531.30	5,425.80	0.22	Sweden	5,600.00	51,920.00	14,423.38	0.29
Greece	221.88	1,326.45	368.49	0.19	Switzerland	1,733.08	11,836.80	3,288.26	0.22
Greenland	1.00	21.00	5.83	0.67	Tajikistan	2.93	55.40	15.39	0.60
Guatemala	2.31	56.46	15.68	0.78	Thailand	128.51	1,181.20	328.14	0.29
Honduras	1.93	45.00	12.50	0.74	Tunisia	43.80	364.00	101.12	0.26
Hungary	905.58	10,268.06	2,852.47	0.36	Turkey	2,886.30	45,126.00	12,536.00	0.50
Iceland	2,040.00	26,717.00	7,422	0.42	Ukraine	10.90	118.80	33.00	0.35
India	986.00	4,302.00	1,195.10	0.14	United Kingdom	283.76	1,906.50	529.63	0.21
Indonesia	2.30	42.60	11.83	0.59	United States	17,415.91	75,862.20	21,074.52	0.14
Iran	81.50	1,103.12	306.45	0.43	Venezuela	0.70	14.00	3.89	0.63
Ireland	265.54	1,240.54	344.62	0.15	Vietnam	31.20	92.33	25.65	0.09
Israel	82.40	2,193.00	609.22	0.84	Yemen	1.00	15.00	4.17	0.48
Italy	1,014.00	8,682.00	2,411.90	0.27	**Worldwide total**	**70,328.98**	**587,786.43**	**163,287.07**	**0.27**
Japan	2,186.17	26,130.08	7,258.94	0.38					



Mineral and Thermal Waters, Fig. 8 Worldwide percentage distribution of total geothermal power installed (MWt, including heat pumps) among various forms of direct use in 2015 (Adapted from [60]). For additional explanation, see the text



Mineral and Thermal Waters, Fig. 9 Worldwide percentage distribution of total annual geothermal energy used (JT/year, including heat pumps) among various forms of direct use in 2015 (Adapted from [60]). For additional explanation, see the text

$$[(\text{annual energy use in TJ/year})/ (\text{installed capacity in MWt})] \times 0.03171.$$

This number reflects the equivalent percentage of equivalent full load operating hours per year (i.e., CF = 0.70 is 70% equivalent to 6,132 full load hours per year.

Electric Power Generation

Electric power generation is a crucial utilization of geothermal resources in general. The first application of geothermal energy to production of electric power was in Italy, with experimental work by Prince Ginori Conti between 1904 and 1905; the first commercial power plant (250 kWe) was established in 1913 at Larderello, Italy. After that, similar experimental plants were installed in the Geysers, California, in 1932; at Wairakei, New Zealand, in 1958; and at Pathe, Mexico, in 1959 [58].

Technological development allows efficient use of high-enthalpy geothermal resources to produce electrical energy. Depending on the characteristics of geothermal systems, electricity can be generated with application of (i) conventional steam turbines (Fig. 10) or (ii) binary plants (Fig. 11). Conventional steam turbines require fluids at temperatures of at least 150 °C and are available with either atmospheric (back pressure) or condensing exhausts [29]. Atmospheric exhaust turbines are simpler and cheaper. The steam, direct from dry steam wells or, after separation, from wet wells, is passed through a turbine and exhausted to the atmosphere.

In the case of low-enthalpy resources, generally below 150 °C, the binary plants (i.e., organic Rankine cycle – ORC plants) are used to generate electric power (Fig. 10). The binary plants utilize a secondary working fluid, usually an organic fluid (typically *n*-pentane), that has a low boiling point and high vapor pressure at low temperatures when compared to steam. The secondary fluid is operated through a conventional organic Rankine cycle (ORC): the geothermal fluid yields heat to the secondary fluid through heat exchangers, in which this fluid is heated and vaporizes; the vapor produced drives a normal axial flow turbine and is then cooled and condensed, and the cycle begins again [58, 61].

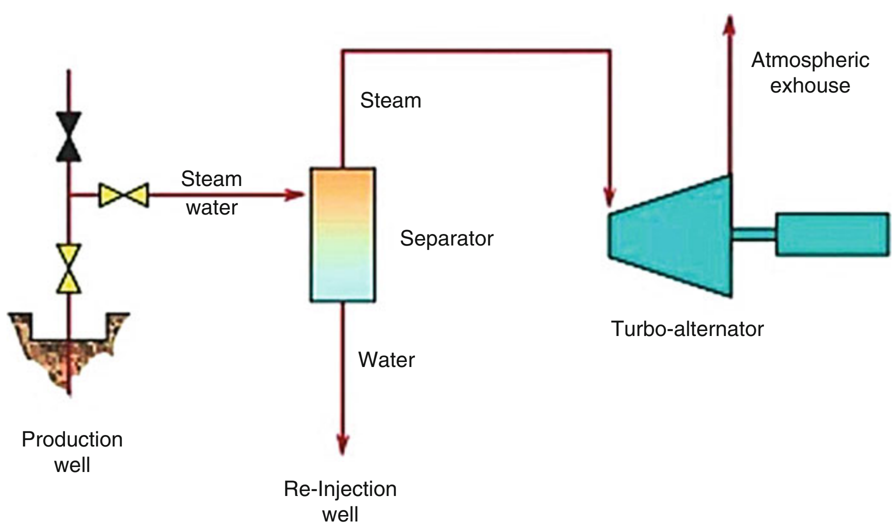
Another binary system, the Kalina cycle, was developed in the 1990s. A water–ammonia mixture as working fluid is used in the Kalina cycle plants. The working fluid is expanded, in superheated conditions, through the high-pressure turbine and then reheated before entering the low-pressure turbine [29]. After the second expansion, the saturated vapor moves through a recuperative boiler before being condensed in a water-cooled condenser. The Kalina cycle is more efficient than existing geothermal ORC binary power plants but is of more complex design.

According to the latest report by Lund et al. [60], electric power is produced from geothermal energy in 26 countries (Tables 4, 5 and 6).

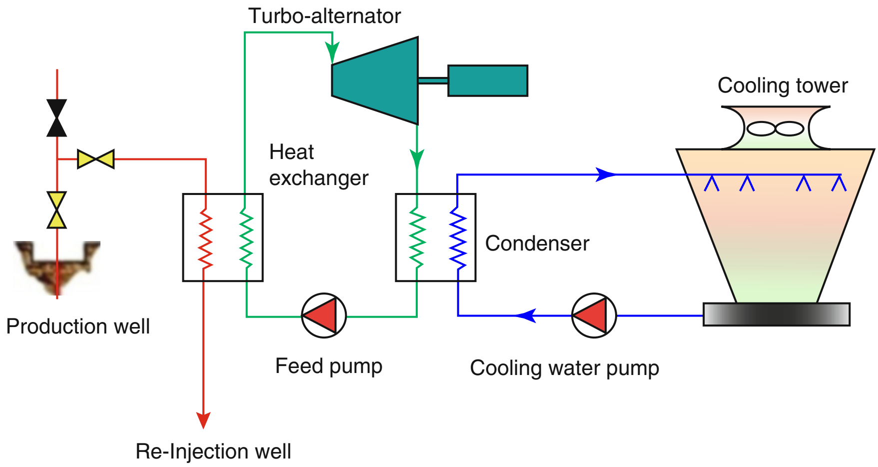
As can be seen from Tables 4, 5 and 6, the top five countries for capacity and produce energy are the USA, the Philippines, Indonesia, Mexico, and New Zealand. For absolute value increase in the last 5 years, Kenya, the USA, and Turkey are the leaders; especially in relation to Turkey, after a period of stagnation, they have revealed significant growth in geothermal electricity production. Taking into account the percentage increase of installed electrical capacity, the leaders are Turkey (336%), Germany (280%), and Kenya (194%). Turkey's growth is mainly due to large installations, whereas Germany's is due to installations of many smaller units [60].

Mineral Water in Bottling Industry: Legislation and Consumptions

Bottled water can be simply defined as the water intended for drinking packed into glass or plastic bottles.



Mineral and Thermal Waters, Fig. 10 Exemplary scheme of conventional steam turbine (Adapted from [61])



Mineral and Thermal Waters, Fig. 11 Exemplary scheme of the geothermal binary power plant (Adapted from [61])

Mineral and Thermal Waters, Table 4 Installed electrical capacity and electrical energy produced from geothermal resources for 2010 and 2015 in particular countries (After Lund et al. [58])

Country	Installed in 2010 (MWe)	Energy in 2010 (GWh)	Installed in 2015 (MWe)	Energy in 2015 (GWh)
Australia	0.1	0.5	1.1	0.5
Austria	1.4	3.8	1.2	2.2
China	24	150	27	150
Costa Rica	166	1,131	207	1,511
El Salvador	204	1,422	204	1,422
Ethiopia	7.3	10	7.3	10
France	16	95	16	115
Germany	6.6	50	27	35
Guatemala	52	289	52	237
Iceland	575	4,597	665	5,245
Indonesia	1,197	9,600	1,340	9,600
Italy	843	5,520	916	5,660
Japan	536	3,064	519	2,687
Kenya	202	1,430	594	2,848
Mexico	958	7,047	1,017	6,071
New Zealand	762	4,055	1,005	7,000
Nicaragua	88	310	159	492
Papua New Guinea	56	450	50	432
Philippines	1,904	10,311	1,870	9,646
Portugal	29	175	29	196
Romania			0.1	0.4
Russia	82	441	82	441
Taiwan			0.1	
Thailand	0.3	2	0.3	1.2
Turkey	91	490	397	3,127
USA	3,098	16,603	3,450	16,600
Total	**10,897**	**67,246**	**12,635**	**73,549**

Mineral and Thermal Waters, Table 5 Worldwide installed electrical capacity and electrical energy produced from geothermal resources for 2010 and 2015 and forecast for 2020 (After Lund et al. [58])

Continent	Installed in 2010 (MWe)	Energy in 2010 (GWh)	Installed in 2015 (MWe)	Energy in 2015 (GWh)	Forecast for 2020 (MWe)
Europe	1,643	11,371	2,133	14,821	3,385
Africa	209	1,440	601	2,858	1,601
America	4,565	26,803	5,085	26,353	8,305
Asia	3,661	23,127	3,756	22,084	6,712
Oceania	818	4,506	1,056	7,433	1,440
Total	**10,897**	**67,246**	**12,635**	**73,549**	**21,443**

Note: Europe includes Guadeloupe (France), Kamchatka (Russia), and Turkey

Mineral and Thermal Waters, Table 6 Plant category per country: installed electrical capacity MWe (After Lund et al. [58])

Country	Back Pressure	Binary	Double flash	Dry steam	Hybrid	Single flash	Triple flash	Total
Australia		1						1
Austria		1						1
China		3	24			1		28
Costa Rica	5	63				140		208
El Salvador		9	35			160		204
Ethiopia		7						7
France		2	5			10		16
Germany		27						27
Guatemala		52						52
Iceland		10	90			564		665
Indonesia		8		460		873		1,340
Italy		1		795		120		916
Japan		7	135	24		355		520
Kenya	48	4				543		594
Mexico	75	3	475			466		1,019
New Zealand	44	265	356			209	132	1,005
Nicaragua	10	8				142		160
Papua New Guinea						50		50
Philippines		219	365			1,286		1,870
Portugal		29						29
Romania		0						0
Russia						82		82
Taiwan		0						0
Thailand			0					0
Turkey		198	178			20		397
USA		873	881	1,584	2	60		3,450
Total	**181**	**1,790**	**2,544**	**2,863**	**2**	**5,079**		**12,640**

The beginnings of large-scale water bottling are connected exclusively with mineral and thermal waters used for healing and medicinal purposes from the sources of their occurrences. The necessity for bottling the water was directly connected with the successful development of the thermal and mineral water resorts and the conscious recognition of curative and therapeutic benefits of such waters: people sought to take the curative waters from their sources (e.g., mineral water spas, thermal baths, etc.). They visited the sources and resorts and took back waters to their homes, to continue benefiting from the waters' healing properties. According to the European Federation of Bottled Waters (EFBW), the bottling and commercialization of mineral waters first began in Europe in the middle of sixteenth century, but large-scale water bottling and transport began in the late nineteenth century (EFBW 2017). The first time water was bottled for the consumer market water was in 1622 at the Holy Well, Malvern, UK [62]. As early as 1767, the waters of Jackson's Spa in Boston, Massachusetts, USA, were bottled and sold commercially. The waters of a mineral spring near Albany, New York, were bottled about 1800; the Saratoga Springs water was bottled in 1820; in Europe the Apollinaris water of Germany was bottled in 1892; and the San Pellegrino waters of Italy were bottled in 1899 ([9], EFBW 2017). These bottled waters were sold as medicinal remedies in pharmacies until the twentieth century.

Today, bottled water is readily available as a convenient and healthy beverage in a wide range of formats and packaging materials.

Different types of water are marketed, each strictly defined by regulations. The bottled waters sold nowadays may come from natural sources (e.g., mineral water, artesian, spring, groundwater, sparkling) or straight from a tap – drinking water from municipal source or community water systems.

All bottled waters in the market intended for drinking are treated as a food and must comply with special requirements and quality standards established in bottled water legal regulations on the national specifics for each country, as well as an international level.

Codex Alimentarius: FAO and WHO Standards

On the international level, the **Codex Alimentarius** is such a collection of internationally recognized standards, codes of practice, guidelines, definitions, and recommendations relating to food, food production and processing, and food safety, including standards for bottled drinking waters and natural mineral waters. The Codex Alimentarius is developed and maintained by the Codex Alimentarius Commission established in November 1961 by the Food and Agriculture Organization of the United Nations (FAO). The World Health Organization (WHO) joined the Commission in June 1962. The first session of the Commission was held in Rome in October 1963. The high importance of the Codex Alimentarius stems from the fact that it is recognized by the World Trade Organization (WTO) as an international reference point for the resolution of disputes concerning food safety and consumer protection. Codex Standard 108-1981 (adapted 1981, amendment 2001, 2011, revised 1997, 2008) entitled *Standard for natural mineral waters* applies to all packaged natural mineral waters offered for sale as food.

Natural mineral water is defined as follows:

- It is characterized by its content of certain mineral salts and their relative proportions and the presence of trace elements or other constituents.
- It is obtained directly from natural or drilled sources from underground water-bearing strata for which all possible precautions should be taken within the protected perimeters to avoid any pollution of, or external influence on, the chemical and physical qualities of natural mineral water.
- It shows constancy of composition and stability of discharge and temperature.
- It is collected under conditions which guarantee the original microbiological purity and chemical composition of essential components.
- It is packaged close to the point of emergence of the source with particular hygienic precautions.
- It is not subjected to any treatment other than those permitted by this standard; permitted treatments include separation from unstable constituents, such as compounds containing iron, manganese, sulfur, or arsenic (by decantation and /or filtration).

Additionally the Codex Standard 108-1981 defines different types of natural mineral waters in relation to CO_2 content. The following types of natural mineral water are distinguished:

(a) **Naturally carbonated natural mineral water** – which has the same content of CO_2 spontaneously and visibly given off under normal conditions of temperature and pressure in the source
(b) **Non-carbonated natural mineral water** – which, by nature, does not contain free CO_2 in excess of the amount necessary to keep the hydrogen carbonate salts present in the water dissolved
(c) **Decarbonated natural mineral water** – which has less CO_2 content than that at emergence and does not visibly and spontaneously give off CO_2 under normal conditions of temperature and pressure
(d) **Natural mineral water fortified with CO_2 from the source** – which has more CO_2 content than that at emergence
(e) **Carbonated natural mineral water** – which has been made effervescent by the addition of CO_2 from another origin

Codex Standard 227-2001 entitled *General standard for bottled/packaged drinking waters* applies to waters for drinking purposes other than natural mineral waters. This standard additionally defines two very important types of bottled waters, namely:

(a) **Waters defined by origin** – groundwaters or surface waters that share the following characteristics:
  – They originate from specific environmental resources without passing through a community water system.
  – Precautions have been taken within the vulnerability perimeters to avoid any pollution of, or external influence on, the chemical, microbiological, and physical qualities of water at origin.
  – Collecting conditions which guarantee the original microbiological purity and essential elements of their chemical makeup at origin.
  – From the microbiological standpoint, are constantly fit for human consumption at their source and are kept in that state with particular hygienic precautions until and while packaging.
  – Are not subject to any modification or treatment other than those permitted (i.e., generally the quality of such water must comply with the *Guidelines for drinking water quality* published by the World Health Organization).
(b) **Prepared waters** – waters that do not comply with all the provisions set for waters defined by origin, may originate from any type of water supply (including municipal water), and can be subjected to any treatment that modifies the original water in order to comply with chemical, microbiological, and radiological safety requirements for prepackaged water.

EU Regulations

In the European Union, specific legislation applies to the three different categories of bottled water: natural mineral water, springwater, and bottled drinking water (Table 8).

Directive 2009/54/EC (2009) on the exploitation and marketing of natural waters defines directly natural mineral water and springwater. The bottled water can be called **natural mineral water** when:

– It is originating from an underground water table or deposit (i.e., groundwater) and emerging from a spring tapped at one or more natural or bore exits.
– It is exploited under certain conditions specified in Annex II of this Directive, and its exploitation shall be subject to permission from the responsible authority of the country where the water is extracted.
– Its water bottled at the source.
– In its state at source, may not be the subject of any treatment other than (i) the separation of its unstable elements such as iron and sulfur compounds, manganese, and arsenic through decantation, filtration, or treatment with ozone-enriched air, in so far as this treatment does not alter the composition of the water as regards the essential constituents which give it its properties, and (ii) the total or partial elimination of free carbon dioxide by exclusively physical methods.

– In its state at source, may not be the subject of any addition other than the introduction or the reintroduction of carbon dioxide under certain conditions specified in Annex I of the directive.
– At the source and during its marketing, a natural mineral water shall be bacteriologically pure and free from (a) parasites and pathogenic microorganisms, (b) *Escherichia coli* and other coliforms and fecal streptococci in any 250 mL sample examined, (c) sporulated sulfite-reducing anaerobes in any 50 mL sample examined; and (d) *Pseudomonas aeruginosa* in any 250 mL sample examined.
– Any disinfection treatment by whatever means or the addition of bacteriostatic elements shall be prohibited.

The term **springwater** is reserved for a water which is intended for human consumption in its natural state and bottled at the source, which:

– Satisfies the conditions of exploitation laid down in Annex II
– Has not undergone any treatment other than those referred to natural mineral waters
– Satisfies the microbiological requirements similar as natural mineral waters
– Shall comply with the provisions of Directive 98/83/EC on the quality of water intended for human consumption

According to Directive 98/83/EC, a "water intended for human consumption" means:

– All water either in its original state or after treatment, intended for drinking, cooking, food preparation, or other domestic purposes, regardless of its origin and whether it is supplied from a distribution network, from a tanker, or *in bottles* or containers
– All water used in any food production undertaking for the manufacture, processing, preservation, or marketing of products or substances intended for human consumption unless the competent national authorities are satisfied that the quality of the water cannot affect the wholesomeness of the foodstuff in its finished form

Bottled drinking waters, also known as table waters, may originate from various sources, including groundwater, surface water, and municipal supply. Bottled drinking water must comply with national and EU drinking water regulations (Directive 98/83/EC), which are different from the rules governing natural mineral water and springwaters. Bottled drinking water is commonly treated and disinfected for taste. Purification by chemical and physical treatment, such as chlorination and reverse osmosis, is common practice. Carbon dioxide may be added to create a sparkling water. Minerals may be restored to this water (EFBW 2017).

Taking into account the EU regulations in relation to bottled waters, natural mineral water can be clearly distinguished from ordinary drinking water (i) by its nature, which is characterized by its naturally stable mineral content, trace elements, or other constituents, and, where appropriate, by certain effects and (ii) by its original purity. The main differences between springwater and natural mineral water are the following (Table 7):

(a) Stability of chemical composition is not a requirement for springwaters, and mineral composition need not to be stated on the label.
(b) Springwater must only meet standards for water intended for human consumption (Directive 98/83/EC).
(c) There is no formal requirement for a recognition process for springwaters as there is for natural mineral water, but quality monitoring and protection of the source must be maintained.

Legal Aspects in the USA

In the USA, two primary federal laws protect the public from contaminants in drinking water: the Safe Drinking Water Act (SDWA) and the Federal Food, Drug, and Cosmetic Act (FFDCA). The SDWA of 1974 gave the Environmental Protection Agency (EPA) federal jurisdiction to regulate quality standards in drinking water delivered by public water supplies. The Food and Drug Administration (FDA), an agency of the US Department, Health and Human Services, is responsible for

Mineral and Thermal Waters, Table 7 Differences of standards for the three main types of bottled waters in EU legislation (After [65])

	Natural mineral water	Springwater	Bottled drinking water
Water origin	Protected underground source and surrounding environment	Protected underground source	Various: groundwaters as well as surface water
Minerals origin	Naturally present in water from the source	Naturally present in water from the source	Can be removed or modified for taste purposes
Disinfection	Prohibited; microbiologically safe at source	Prohibited; microbiologically safe at source	Permitted; usually applied
Chemical treatment	Prohibited	Prohibited	Allowed; usually applied to comply with legislations
Uniqueness	Unique source and mineral balance	Unique source	No
Stability	Mineral balance always naturally stabile	Mineral balance may vary naturally	No restrictions
Bottling and transport	Bottled at source	Bottled at source	May be transported to a separate facility
Safety	Suitable for human consumption at source	Suitable for human consumption at source	Water usually filtered and/or chemically treated and disinfected before being safe to drink

Mineral and Thermal Waters, Table 8 Compilation of selected European Union directives and regulations connected with bottled water industry (Recommended by EFBW)

Subject of regulation	Regulatory act
Natural mineral water	Directive 2009/54/EC on the exploitation and marketing of natural mineral waters
Springwater	Directive 2009/54/EC on the exploitation and marketing of natural mineral waters Directive 98/83/EC on the quality of water intended for human consumption
Bottled drinking water (table water)	Directive 98/83/EC relating to the quality of water intended for human consumption
Food safety and food hygiene	Regulation 178/2002/EC laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety Regulation 852/2004/EC on the hygiene of foodstuffs Regulation 882/2004/EC on official controls performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules
Packaging and labeling	Regulation No 282/2008/EC on recycled plastic materials and articles intended to come into contact with foodstuffs Directive 13/2000 of the European Parliament and the Council relating to the Labelling, Presentation, and Advertising of Foodstuffs

regulations of quality standards of food and drinks. The major role of the FDA in the regulation of bottled water stems from the classification of bottled water as “food” under the FFDCA ([9], IBWA 2017).

The FDA has issued comprehensive bottled water standards of identity (Code of Federal Regulation – 21 C.F.R. Part 165.110), which provide uniform requirements and definitions for the following types of bottled water: artesian water, groundwater, mineral water, purified water, sparkling water, and springwater. FDA has also established bottled water standards of quality for more than 90 substances (21 C.F.R. Part 165.110b).

The FDA’s quality standards for bottled water are compatible with EPA standards for tap water.

Each time EPA establishes a standard for a contaminant, FDA either adopts it for bottled water or finds that the standard is not necessary for bottled water [9].

FDA regulations pertaining to bottle water are contained in Codes of Federal Regulations 21 C.F.R. Part 103 and 21 C.F.R. Part 129. They provide for microbiological, physical, chemical, and radiological quality standards of bottled water.

According to 21 C.F.R. Part 129.3, **bottled drinking water** means all water which is sealed in bottles, packages, or other containers and offered for sale for human consumption, **including bottled mineral water**.

According to 21 C.F.R. Pat 165.110, **mineral water** is defined as water containing not less than 250 parts per million (ppm) total dissolved solids (TDS), coming from a source tapped at one or more bore holes or springs, and originating from a geologically and physically protected underground water source. Mineral water shall be distinguished from other types of water by its constant level and relative proportions of minerals and trace elements at the point of emergence from the source, due account being taken of the cycles of natural fluctuations. No minerals may be added to this water.

Other types of bottled water defined by regulations of 21 C.F.R. Pat 165.110 are as follows:

(a) **Artesian water** (artesian well water) is the water from a well that taps a confined aquifer (a water-bearing underground layer of rock or sand) in which the water level stands at some height above the top of the aquifer.
(b) **Groundwater** is the water from a subsurface saturated zone that is under a pressure equal to or greater than atmospheric pressure. Groundwater must not be under the direct influence of surface water as defined in 40 C.F.R. Part 141.2.
(c) **Purified water** or demineralized water is the water that has been produced by distillation, deionization, reverse osmosis, or other suitable processes and that meets the definition of "purified water" in the United States Pharmacopeia, 23d Revision, 1 January 1995 and 1 C.F.R. Part 51. Alternatively, the water may be called "deionized water" if the water has been processed by deionization, "distilled water" if it is produced by distillation, "reverse osmosis water" if the water has been processed by reverse osmosis, and "X drinking water" with the X being filled in with one of the defined terms describing the water in this paragraph (e.g., "purified drinking water" or "deionized drinking water").
(d) **Sparkling bottled water** is the water that, after treatment and possible replacement with carbon dioxide, contains the same amount of carbon dioxide that it had as it emerged from the source. Sparkling bottled waters may be labeled as "sparkling drinking water," "sparkling mineral water," "sparkling springwater," etc.
(e) **Springwater** is the water derived from an underground formation from which water flows naturally to the surface of the Earth. Springwater shall be collected only at the spring or through a bore hole tapping the underground formation feeding the spring. There shall be a natural force causing the water to flow to the surface through a natural orifice. The location of the spring shall be identified. Springwater collected with the use of an external force shall be from the same underground stratum as the spring, as shown by a measurable hydraulic connection using a hydrogeologically valid method between the bore hole and the natural spring, and shall have all the physical properties, before treatment, and shall be of the same composition and quality, as the water that flows naturally to the surface of the Earth. If springwater is collected with the use of an external force, water must continue to flow naturally to the surface of the Earth through the spring's natural orifice. Plants shall demonstrate, on request, to appropriate regulatory officials, using a hydrogeologically valid method that an appropriate hydraulic connection exists between the natural orifice of the spring and the bore hole.
(f) **Well water** is the water from a hole bored, drilled, or otherwise constructed in the ground, which taps the water aquifer.

The FDA regulation to 21 C.F.R. Pat 165.110 allows water from municipal system to be bottled. When bottled water comes from a community water system, as defined in 40 CFR Part 141.2, except when it has been treated to meet the definitions of "purified water or sterile water," and is labeled as such, the label shall state "from a community water system" or, alternatively, "from a municipal source."

As can be seen from an overview of the water bottled standards promoted by the UN FAO and WHO, and regulations implemented by directives of the EU in the member states, or by regulations of the FDA in the USA, bottled water, including mineral water, has unique characteristics that justify its regulation separate and apart from tap drinking water.

Generally, standards and regulations concerning mineral waters or other bottled waters do not take into account the amount of dissolved solids and minerals in the water; only the FDA standards in the USA require the amount of TDS to be not less than 250 ppm for mineral water. According to the EU and FAO regulations, bottled mineral water can be classified as the specific type of water in relation to content of TDS, even for waters with very low concentration of dissolved minerals.

Market and Consumption Overview

Water is used in many different ways, not only for drinking but also in farming, manufacturing, and the production of energy. The consumption of water, including bottled water, varies from one country to another depending on many climatic, societal, and cultural factors.

Global consumption of bottled water is growing in almost every major geographical region of the world. The bottled water market initially emerged as a large, mainstream commercial beverage category in Western Europe, where consumption of it has long been part of many residents' routines [63]. Nowadays, the bottled water industry stands as a truly global beverage market with consumption estimated at 92 billion gallons in 2016. Several Asian markets achieved strong growth to become major bottled water markets during the 2000s. In fact, Asia itself became the largest regional market in 2011 (Fig. 10, Table 9).

Mineral and Thermal Waters, Table 9 Bottled water consumption in ten leading countries. Data from 2011 and 2016 for comparison. *CAGR* compound annual growth rate [63]

Rank	Country	2011	2016	2011/ 2016 CAGR (%)
		Millions of gallons		
1	China	12,117.6	22,146.9	12.8
2	United States	9,107.2	12,781.9	7.0
3	Mexico	7,227.2	8,514.3	3.3
4	Indonesia	4,728.7	7,156.4	8.6
5	Brazil	4,503.8	5,507.4	4.1
6	India	3,045.1	5,193.9	11.3
7	Thailand	3,120.8	3,841.4	4.2
8	Germany	2,956.1	3,134.1	1.2
9	Italy	2,831.1	2,909.3	0.5
10	France	2,249.8	2,389.7	1.2
	All others	15,927.2	18,538.0	3.1
	World total	67,814.6	92,113.3	6.3

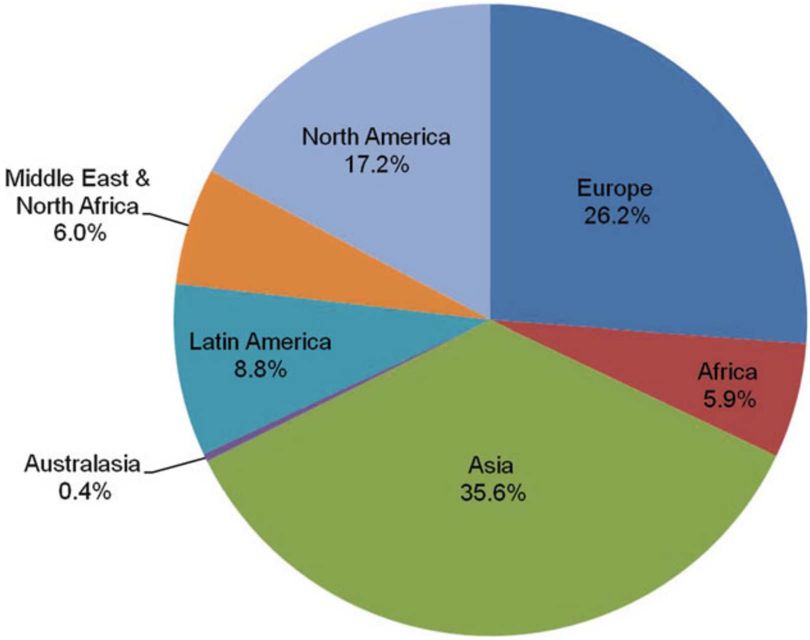
China became a leader in consumption of bottled water in 2013. By 2016, China accounted for close to one quarter of global bottled water volume (Table 9, Fig. 12).

More than a quarter of the global bottled water production is located in Europe, where consumers purchase and drink more than 13.7 billion gallons of bottled water every year.

The sector contributes significantly to the European economy by providing direct employment to 54,000 people and investing particularly in less prosperous regions such as rural areas [64]. Leading European countries represent firmly established bottled water markets, and their growth tends to be smaller than those where bottled water has much less tradition of use.

Within the packaged water market, the European consumer has a preference for naturally sourced waters which have not been disinfected nor chemically processed and which are associated with a specific place of origin. In Europe, 97% of all bottled water produced is either natural mineral water (82%) or springwater (15%). This is

Mineral and Thermal Waters, Fig. 12 Global bottled water consumption (Plot adapted from [64])



unique compared to other continents where the bulk of bottled water sold is processed water (e.g., the USA, Asia).

Consumption per capita by individual region or country diverges significantly from total consumption in particular countries (Table 10).

As can be seen from Table 10, several Western European countries have per capita consumption levels far above 25 gallons, and the Mexican market had an average intake of about 67 gallons in 2016 [63]. On the other hand, it is well known that a considerable part of the developing world, where the majority of the world's population resides, finds its per capita consumption figures still in the single-digit range. One optimistic fact is that bottled water's international growth signals demand for it in diverse markets. Consumers have demonstrated a thirst for it in highly developed markets, in less developed ones, and in economies in transition.

Water is one of very few vital needs for human beings. It is assumed that a healthy sedentary adult living in a temperate climate should drink at least 1.5 L of water per day. This level of water intake balances water loss and helps keep the body properly hydrated. According to the International Bottled Water Association (IBWA), consumers

Mineral and Thermal Waters, Table 10 Per capita consumption of bottled water in leading countries in 2011 and 2016 (After [63], IBWA 2017)

Rank	Country	2011 (gal)	2016 (gal)
1	Mexico	60.5	67.2
2	Thailand	46.9	56.9
3	Italy	46.6	47.5
4	United States	29.2	39.3
5	Germany	35.9	38.0
6	France	35.4	36.6
7	Belgium-Luxembourg	34.6	35.3
8	Saudi Arabia	26.7	31.3
9	Spain	31.5	31.0
10	United Arab Emirates	24.6	30.7
11	Hungary	28.7	29.7
12	China, Hong Kong	24.5	28.8
13	Indonesia	19.4	27.7
14	Argentina	26.6	27.6
15	South Korea	18.7	27.1
16	Lebanon	31.2	27.1
17	Brazil	22.9	26.8
18	Austria	24.7	24.8
19	Switzerland	25.0	24.8
20	Poland	20.3	24.7
	Global average	**9.7**	**12.4**

choose bottled water for several reasons: taste, quality, and convenience. They are not buying bottled water because of elaborate marketing campaigns; they are choosing bottled water instead of less healthy packaged beverages (IBWA 2017). Markets have positioned bottled waters as very good alternative for carbonated soft drinks or fruit juice drinks. In the developed world, consumers have come to see bottled water not only as a way of achieving hydration but also as functional beverage. In developing countries, bottled water serves as a partial solution to the problem of access to water of good quality.

Bibliography

1. Shiklomanov I (1993) World fresh water resources. In: Gleick PH (ed) Water in crisis: a guide to the world's fresh water resources. Oxford University Press, New York, pp 14–24
2. Perlman H (2016) Educational materials. https://water.usgs.gov/edu/watercycle.html
3. Evans J, Perlman H (2016) USGS educational materials. https://water.usgs.gov/edu/watercycle.html
4. Dowgiałło J, Karski A, Potocki I (1969) Geologia surowców balneologicznych. Wydawnictwa Geologiczne, Warszawa, p 296. [in Polish only]
5. Grünhut L, Hintz E (1907) Einteilung der mineralwässer. Deutsches Bäderbuch, Leipzig
6. Davis SN (1964) The chemistry of saline waters by R.A. Krieger – discussion. Groundwater 2(1):51–51
7. Krieger RA (1963) The chemistry of saline waters. Groundwater 1(4):7–12
8. Pazdro Z, Kozerski B (1990) Hydrogeology. Wydawnictwa Geologiczne, Warszawa, p 624
9. LaMoreaux PE, Tanner JT (2001) Springs and bottled waters of the world. Ancient history, source, occurrence, quality and use. Springer, Berlin, p 315
10. Prochorov A (ed) (1969–1978) Great Soviet encyclopedia, III edn. Soviet Encyclopedia Publishing House, Moscow. http://bse-lib.com
11. Harrison R, Mortimer ND, Smarason OB (1990) Geothermal heating. A handbook of engineering economics. Pergmon Press, Brussels–Luxembourg, p 558
12. Cataldi R (1999) The year zero in geothermics. In: Cataldi R, Hodgson S, Lund JW (eds) Stories from a heated Earth. Our geothermal heritage. Geothermal Resources Council and International Geothermal Association, Sacramento, pp 7–17
13. Fridleifsson IB (2000) Geothermal energy for the benefit of the people worldwide. In: Proceedings of the world geothermal congress 2000, Kyushu-Tohoku, Japan
14. Kępińska B (2006) Geothermal waters in the history of civilization. In: Górecki W, Hajto M (eds) Atlas of geothermal resources of Mesozoic formations in the Polish Lowlands. AGH, Goldruk, Kraków, p 482
15. Lund, JW (2005) Balneological use of thermal waters, short courses of the world geothermal congress. Course on integration use of geothermal waters, Antalya, 2005
16. Ross K (2001) Health tourism: an overview HSMAI Marketing review. Online document. https://www.hospitalitynet.org/news/4010521.html
17. Albu M, Banks D, Nash H (1997) Mineral and thermal groundwater resources. Springer, Dordrecht, p 431
18. Alderson F (1973) The inland resorts and spas of Britain. David & Charles, Newton Abbott, Devon
19. Rockel I (1986) Taking the waters – early spas in New Zealand. Government Printing Office Publishing, Wellington, p 195
20. Lund JW (1993) Spas and balneology in the United States. GHC Q Bull 14:1–3
21. Wang Ji-Yang (1995) Historical aspects of geothermal use in China. Proceedings of the world geothermal congress, Florence, 18–31 May 1995
22. Sekioka M (1995) Geothermal energy in history. The case of Japan. Proceedings of the world geothermal congress, Florence, 18–31 May 1995
23. Swanner GM (1988) Saratoga – queen of spas. North Country Books, Inc., Utica, p 304
24. Jianli N, Jinkai L, Chengzhi W, Shuqiang Q (1993) The control and protection of groundwater in fractured gneiss during coal mining, in hydrogeology of hard rocks. In: Banks SB, Banks D (eds) Proceedings XXIVth congress of the international association of hydrogeologists, 28 June–2 July 1993, Oslo, pp 259–264
25. Phillips J (2004) Russia. Spa Bus Mag (4). www.spabusiness.com
26. Kropotkin PN, Polyak BG (1973) The energy balance of the Earth. In: Zemnaya kora seismoopasnyh zon (The Earth's crust of seismohazardous zones). Nauka, Moscow, pp 7–24
27. Polyak BG (1988) Teplomassopotok iz mantii v glavnyh strukturah ziemnoj kory (Heat-mass flow from the mantle in main structures of the Earth's crust). Nauka, Moscow, p 192. [in Russian only]
28. Kononov VI (2002) Geothermal resources of Russia and their utilization. Lithol Miner Resour 37(2):97–106
29. Dickson MH, Fanelli M (2003) Geothermal energy: utilization and technology. UNESCO and Taylor & Francis Group, Paris, p 66
30. Muffler LJP, Cataldi R (1978) Methods for regional assessment of geothermal resources. Geothermics 7:53–89
31. Górecki W, Hajto M et al (eds) (2006) Atlas of geothermal resources of Paleozoic formations in the Polish Lowlands. AGH, Goldruk, Kraków, p 240
32. Polyak BG, Smirnov YB (1966) Heat flow on continents. Dokl Akad Nauk SSSR 168(1):170–172
33. Polyak BG, Smirnov YB (1968) Relationship between terrestrial heat flow and tectonic structure of continents. Geotektonika 4:3–19
34. Smirnov YB, Kononov VI (1991) Geothermal studies and superdeep drilling. Soviet Geol 8:28–37

35. Kononov VI, Yudahin FN, Svalova WB (1993) Geotermia seismicheskih i aseismicheskih zon. Nauka, Moscow, p 400. [in Russian only]
36. Górecki W, Hajto M et al (eds) (2006) Atlas of geothermal resources of Mesozoic formations in the Polish Lowlands. AGH, Goldruk, Kraków, p 484
37. Hochstein MP (1990) Classification and assessment of geothermal resources. In: Dickson MH, Fanelli M (eds) Small geothermal resources: a guide to development and utilization, UNITAR, New York, pp 31–57
38. Benderitter Y, Cormy G (1990) Possible approach to geothermal research and relative costs. In: Dickson MH, Fanelli M (eds) Small geothermal resources: a guide to development and utilization. UNITAR, New York, pp 59–69
39. Nicholson K (1993) Geothermal fluids, vol XVIII. Springer, Berlin, p 264
40. Axelsson G, Gunnlaugsson E (2000) Background: geothermal utilization, management and monitoring. In: Long exploitation, IGA, WGC 2000 short courses, Kokonoc, Kyushu
41. Górecki W et al (1993) Metodyka oceny zasobów energii wód geotermalnych w Polsce. Ekspertyza 12/93 MOŚZNiL, Arch. ZSE AGH, Kraków [expertise, in Polish only]
42. Górecki W et al (1994) Określenie odnawialnych zasobów energii geotermalnej na Niżu Polskim. Spraw. z wykonania projektu badawczego KBN nr 901279101. Arch. ZSE AGH, Kraków [scientific project report; in Polish only]
43. Górecki W et al (1995) Atlas zasobów energii geotermalnej na Niżu Polskim. ZSE AGH, Towarzystwo Geosynoptyków "GEOS," Kraków [in Polish only]
44. Górecki W, Sowiżdżał A et al (eds) (2012) Geothermal atlas of Carpathian Foredeep. AGH UST, Kraków, p 418
45. Hajto M (2016) A brief glossary of Polish and the UNFC-2009 classifications and nomenclature of geothermal resources assessment. Proceedings of the European geothermal congress, Strasbourg, 19–24 September 2016, pp 1–6
46. McKelvey VE (1974) Mineral resource estimates and public policy. Am Sci 60:32–40
47. Gosk E (1982) Geothermal resources assessment. In: Čermak V, Haenel R (eds) Geothermics and geothermal energy. E. Schweizerbart'sche Verlagsbuchhandlug, Stuttgard, p 299
48. Haenel R (1982) Geothermal resource and reserve assessment. Report NLfB, Archive No 95 100, Hannover
49. Haenel R, Staroste E (1988) Atlas of geothermal resources in the European community. Austria and Switzerland. Verlag Th Schäfer, Hannover, 110 plates, p 74
50. Haenel R, Hurter S (eds) (2002) Atlas of geothermal resources in Europe. Commission of the European Communities, Taf. Brüssel, Luxemburg, p 74
51. Sorey ML, Nathenson M, Smith C (1983) Methods for assessing low temperature geothermal resources. U.S. Geological Survey, Circural 892
52. ECE 2009 United Nations framework classification for fossil energy and mineral reserves and resources 2009. Economic commission for Europe, ECE series no. 39, p 20
53. Gringarten AC, Sauty JP (1975) A theoretical study of heat extraction from aquifer with uniform regional flow. J Geophys Res 80(35):4956–4962
54. Lindal B (1973) Industrial and other applications of geothermal energy. In: HCH A (ed) Geothermal energy. UNESCO, Paris, pp 135–148
55. Gudmundsson JS (1988) The elements of direct uses. Geothermics 17:119–136
56. Lund JW (2004) Geothermal direct-heat utilization. In: Kępińska B, Popovski K (eds) Proceedings of the international geothermal days Poland 2004, Kraków and Skopje, September 2004, pp 19–33
57. Muffler LJP (ed) (1979) Assessment of geothermal resources of the United States – 1978. USGS Circular 790, Arlington, p 163
58. Lund JW, Bjelm L, Bloomquist G, Mortensen AK (2008) Characteristics, development and utilization of geothermal resources – a Nordic perspective. Episodes 31(1):140–147
59. Lund JW, Boyd TL (2015) Direct utilization of geothermal energy 2015. Worldwide review. Proceedings of the world geothermal congress 2015, Melbourne, 19–25 April
60. Lund JW, Bertani R, Boyd T (2015) Worldwide geothermal energy utilization 2015. GRC Trans 39:79–92
61. Fanelli M, Dickson MH (2004) What is geothermal energy? Information provided in the official website of the International Geothermal Association (IGA). www.geothermalenergy.org
62. MHDC (2008) Great Malvern conservation area. Appraisal and conservation strategy. Malvern Hills District Council, Planning Services, April 2008, p 83
63. Rodwan JG Jr (2017) Bottled water 2016: no. 1 and growing. U.S. and international developments and statistics. Bottled water reporter, IBWA, May/June, pp 12–21
64. Deffis JP, Fosselard P (2016) Natural mineral and spring waters. The natural choice for hydration. EFBW industry report. www.efbw.org, p 13
65. EFBW (2017) EU legislations on bottled waters. European Federation of Bottled Water. www.efbw.org
66. Nathenson M, Muffler LJP (1975) Geothermal resources in hydro thermal convection systems and conduction dominated areas. In: White DE, Williams DL (eds) Assessment of geothermal resources of the United States – 1975. U.S. Geological Survey Circular 726, p 155
67. EFBW (2017) History of bottled water. European Federation of Bottled Water. www.efbw.org
68. IBWA (2017) Regulations of bottled water. International Bottled Water Association. www.bottledwater.org
69. IBWA (2017) Bottled water report. International Bottled Water Association. www.bottledwater.org

Water in Loess

Peiyue Li and Hui Qian
School of Environmental Science and Engineering, Chang'an University, Xi'an, Shaanxi, China
Key Laboratory of Subsurface Hydrology and Ecological Effects in Arid Region of the Ministry of Education, Chang'an University, Xi'an, Shaanxi, China

Article Outline

Glossary

Available water capacity Available water capacity is the soil water content within a range between field capacity and wilting point. It is the soil water content that can be retained in the soil and is available for plants.

Confined groundwater Confined groundwater is the counterpart of phreatic groundwater or unconfined groundwater. It refers to the groundwater stored in confined aquifers, which are overlain by a confining layer made up of low permeability materials such as clay.

Ecological civilization Ecological civilization is a term that describes a new stage of the development of human civilization. It represents a new stage of civilization following industrial civilization and is the final goal of an environmental reform within a given society. Ecological civilization can be described as the sum of all material and ideological achievements obtained by employing strategies for a harmonious development of the human society and the environment. It involves a synthesis of economic, educational, political, agricultural, and other societal reforms toward sustainability.

Field capacity Field capacity is the amount of soil moisture or water content retained in the soil after the drainage of excess water by gravity.

Loess Plateau The Loess Plateau is one of the four highlands of China and one of the birthplaces of the ancient Chinese civilization. It was and still is an important center of the Silk Road and is the most concentrated and largest area on earth in terms of loess accumulation, covering 630,000 km^2. It spreads across the seven Chinese provinces Qinghai, Gansu, Ningxia, Inner Mongolia, Shaanxi, Shanxi, and Henan and is mainly composed by the Shanxi Plateau, the Shaanxi-Gansu-Shanxi Plateau, the Longzhong (mid-Gansu) Plateau, the Ordos Plateau, and the Hetao Plain.

Loess Loess refers to the predominantly silt-sized sediment formed by the accumulation of wind-blown dust under dry climatic conditions. It is usually homogeneous, porous, slightly coherent, and non-stratified. The loess grains are angular and composed of crystals of quartz, feldspar, mica, and other minerals.

Phreatic groundwater Phreatic groundwater refers to the groundwater stored in phreatic aquifers. Phreatic aquifers are also referred to as unconfined aquifers whose upper boundaries are provided by the phreatic surface where the pore water pressure is under atmospheric conditions.

Water resources vulnerability Water resources vulnerability is the sensitivity and capability of a water resource system to adapt to the changes of water system structures, the decrease of water quantity, and the deterioration of water

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_968

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_968-1

quality in the context of climate change and human activities as well as the consequent changes in water supply, water demand, and water management and the occurrence of water-related hazards. It involves the sensitivity of water resources to internal and external changes and the adaptability to these changes.

Wilting point Wilting point is defined as the minimal soil moisture that the plant requires to avoid wilting. If moisture decreases to this point or below, a plant will wilt and can no longer recover its turgidity.

Definition of the Subject

Loess is a fine-grained windblown (aeolian) sediment which is homogeneous, porous, friable, pale yellow or buff, slightly coherent, typically non-stratified, and often calcareous [1, 2]. It is widely distributed in China, Argentina, Europe, the United States, and Middle Asia and mainly occurs in arid and semiarid regions with severe water resource shortages and fragile ecological environments. Mineral resources such as oil, gas, and coal, however, are rich in these areas.

The Loess Plateau of China is the most typical loess distribution area in the world. However, as a result of increasing human activities such as urbanization, industrialization, and energy exploitation, water demands are high, resulting in soil and water pollution [3]. The Loess Plateau is currently facing serious challenges in water resources development and environmental protection. It is therefore crucial to understand the characteristics of loess and the status of water resources in the Loess Plateau area.

This entry reviews the concept/origin of loess and its universal distribution, introduces its physicochemical and geotechnical characteristics, summarizes the problems associated with water resources development in the Loess Plateau of China, and discusses the importance of surface water, groundwater, and soil water in this region.

Introduction

Loess is a fine-grained windblown (aeolian) sediment which is widely distributed. Loss deposits contain buried evidence of Paleolithic occupations [4]. Along the Rhine Valley in Germany, where the deposit was first recognized, local residents named the widely distributed soil “Löss,” and this German word is the origin of the English term loess [2].

As early as 3000 years ago, there were some brief descriptions about loess in the ancient Chinese book “Yu Gong.” However, modern scientific research on loess began only after the publication of the book *Principles of Geology* by the English geologist Charles Lyell in the 1830s [5, 6]. Charles Lyell proposed that “the present is the key to the past,” suggesting the study of the past based on present characteristics of geological formations. Today, this proposal is still canonized by many researchers. Sun [5], however, thought that such an approach should be expressed as “the past is the key to the future,” because predicting the future from the past records is more meaningful to modern citizens. Nowadays, with an increasing awareness of ecological issues, people are much more concerned about the development of the geological and natural environment in the context of economic development, population growth, and environmental problems. However, the research conducted by Charles Lyell on loess marks the beginning of modern loess research, and the history of modern loess research in China can be divided into the following three phases [6–8]:

Phase I: From the 1840s to the 1940s

During this period, loess research has moved away from the fields of soil science and geography toward geology. At the same time, research on loess gradually changed from considering general problems to loess genesis, stratigraphy, and other specific studies. The research approach in this period was simple and relied mainly on fossils. International loess research was dominant in this period, and many internationally recognized geologists conducted their studies in China. For example, Pumpelly [9], Richthofen [10], and Willis et al. [11] carried out investigations on loess and published their results in books or journals. Representative publications in this period include the five-volume German book by Richthofen *China:*

Ergebnisse eigener Reisen und darauf gegründeter Studien and the book *Research in China* by Willis. These publications show that research in this period was focused on the genesis of loess.

Phase II: From the 1950s to the 1970s

A large number of studies on the chemical, mineral, and grain composition as well as on the physical and mechanical properties of loess were carried out in this period. Due to massive geotechnical engineering projects in the loess areas of China, Chinese scholars became the leading loess research community, especially in the field of loess collapsibility and construction techniques on loess basement. Some Chinese monographs on loess were also published in this period, including *Composition and Structure of the Loess*, *Loess in the Middle Reaches of the Yellow River*, *Chinese Loess Accumulation*, *Research on the Basic Nature of the Loess*, and *Chinese Loess and Loess Rock*. In particular, Zhu [12] proved that the dark layer imbedded in the loess layers is paleosol, which is beneficial for the stratigraphic division of the loess strata and the recovery of the paleoclimate.

Phase III: From the 1980s to the Present

In this period, the latest technical methods such as paleomagnetism, isotopic chronology, and ^{14}C dating became available, providing the possibility to determine loess age and for stratigraphic division of the loess strata [13]. In addition, environmental geochemistry, paleoclimatology, and microstructure analysis were also applied in loess research and enriched the scope of loess studies. Numerous scientific contributions were published in this period, such as *Loess and the Environment*, *Soil and Agriculture in the Loess Plateau*, and *Loessology*. Research efforts in this period have significantly enhanced our understanding of the formation and development history of the Chinese loess strata and the paleoclimate. Studies on the collapsibility of loess, the mechanical constitutive of loess, the dynamic characteristics of loess, and the engineering of geological and geo-environmental issues in loess areas have also achieved great progress [6].

In the past two decades, many ecological and environmental studies have been carried out on the Loess Plateau. Loess water erosion constitutes a great threat for environmental safety and local residents in loess areas. Wang et al. [14] investigated the impacts of water erosion on a gas pipeline in the loess area through aerial photo interpretation. They suggested an approach of controlling loess water erosion using soil solidified material that could improve the mechanical properties and anti-erosion ability of loess. Research also showed that land use has significant effects on soil moisture and soil water storage [15]. Climate change significantly affects the water budget balance in loess areas [16, 17], and the total annual imbalance of the land-surface water budget could reach 20.6% across the Loess Plateau of China [17]. Such studies provide a significant contribution to the ecological preservation and the environmental restoration of loess areas.

However, due to the current rapid population growth and fast urbanization in loess areas of China, many land creation projects have been implemented on the Loess Plateau [18]. These projects may produce potentially negative effects on the local ecology and water resources, although they provide more land for urbanization and agriculture. Such projects should therefore be carried out with caution to sustain the results of ecological civilization construction, which took several decades to achieve.

Loess and Its Distribution

Loess is a wind-transported sediment formed mainly during the Quaternary and sometimes recorded in older geologic intervals [2, 5, 19]. It is mainly composed of fine sand and clay, with grain sizes ranging from 0.005 to 0.05 mm. The loess is loose and porous with well-developed vertical joints, which enables vertical water flow rather than horizontal flow. In addition, loess contains a variety of soluble substances, which facilitate the formation of eroded valleys and cause subsidence and collapse. The porosity of loess sediments generally ranges from 42% to 55% [5].

The mineral compositions of loess include three categories: clastic minerals, clay minerals, and authigenic minerals. Clastic and clay minerals are dominant, while authigenic minerals are minimal in loess. The clastic minerals mainly include quartz, feldspar, and mica, which account for 80% of the clastic minerals, followed by pyroxene, amphibole, chlorite, and magnetite. In addition, loess is rich in carbonates such as calcite. The clay minerals in loess are mainly illite, montmorillonite, kaolinite, goethite, and aqueous hematite. The chemical composition of loess is dominated by SiO_2, Al_2O_3, and CaO, followed by Fe_2O_3, MgO, K_2O, Na_2O, and FeO. The contents of TiO_2 and MnO in loess, however, are small. It should be noted that the mineral contents and the chemical composition of loess of different regions may vary with depth, age, and forming environment of the loess (Fig. 1).

Loess is extremely unstable and may collapse when wetted, seismically shaken, or disturbed by human activities [18]. The most significant feature of the loess is its great collapsibility. Worldwide, about half of the loess sustains collapsibility, and collapsible loess in China accounts for 60% of the total loess [20]. Collapsibility refers to the effect or ability of soil, under deadweight or additional load, to produce a sharp sinking and cause ground deformation in the case of constant pressure under the influence of moisture content changes [20]. The collapsibility of loess depends on multiple factors such as particle composition, porosity, depth, changes in atmospheric precipitation and temperature, and anthropogenic effects. The regional variation of loess collapsibility in China is controlled by climate and particle composition [20]. In northwestern China, the climate is dry and cool and becomes wet and hot in southeastern

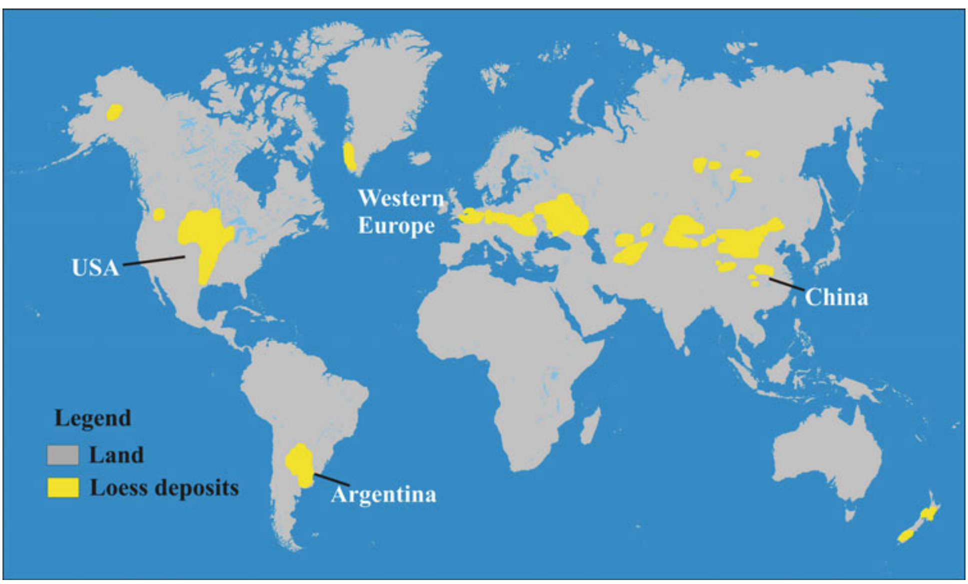
Water in Loess, Fig. 1 Photos of loess geomorphology and profiles from the Loess Plateau of China

China, which corresponds to the regional loess collapsibility variation trend in China (high loess collapsibility in northwestern China to no collapsibility in southeastern China). Loess particle size also shows a descending trend from northwestern to southeastern China; the increase of clay minerals in loess decreases soil porosity, thereby reducing loess collapsibility.

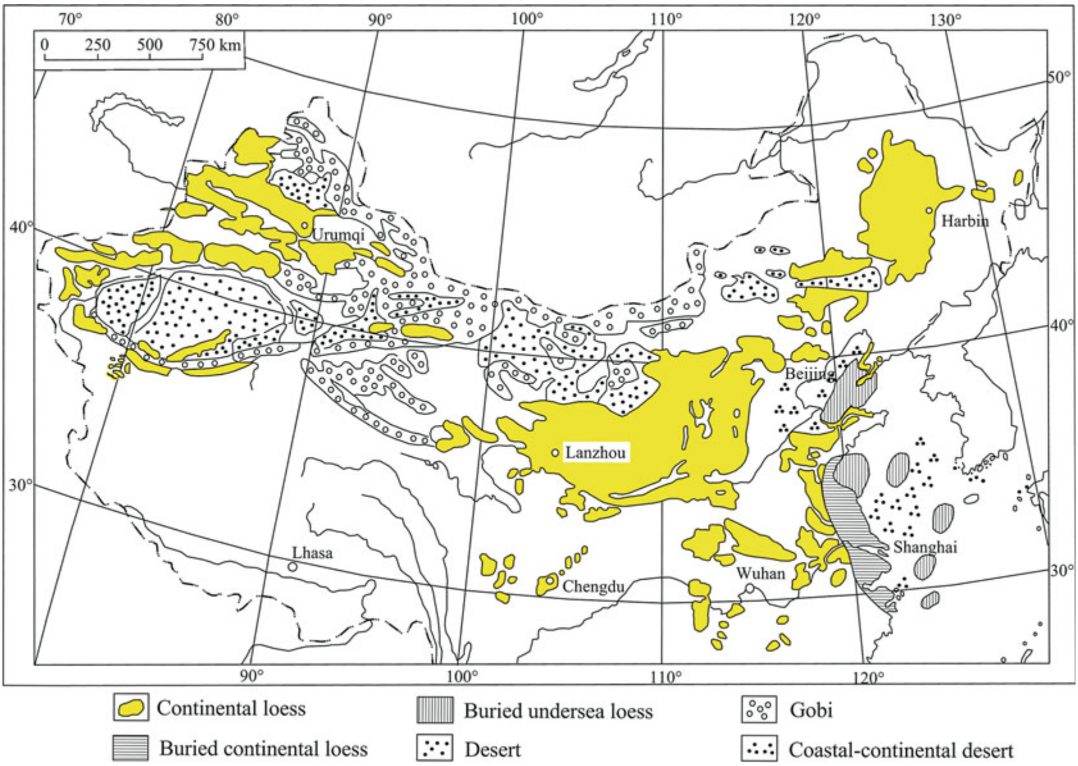
Loess is widely distributed over the world, accounting for 10% of the world land area [19]. It is mainly distributed in the mid-latitude arid and semiarid areas of the northern hemisphere. In the southern hemisphere, loess is mainly distributed in some countries of South America and in New Zealand (Fig. 2). Particularly in Europe and North America, the northern boundary of the area covered by loess is roughly connected to the Pleistocene glaciers, and the loess areas are mainly distributed in the United States, Canada, Germany, France, Belgium, the Netherlands, Central and Eastern Europe, and the Ukraine. In Asia and South America, the loess areas are adjacent to the desert and the Gobi and are mainly found in China, Iran, the Central Asian part of the Russia, and in Argentina [5]. China has the largest and thickest loess cover of the world (Fig. 3). Here, loess covers an estimated area of 630,000 km^2 and accounts for approximately 7% of the total territorial area [6]. The Loess Plateau, the main loess distribution area in China, is well known for its highest loess thickness and most complete loess stratigraphy and has become the center of world loess research.

Water Resources Development on the Loess Plateau

Water is the key element for economic development and ecological conservation in the Loess Plateau of China. Drought and water shortage are the most serious eco-environmental problems on the Loess Plateau and severely constrain ecological civilization and economic development [23]. On the Loess Plateau, river water and groundwater are the most widely used water resources and are applied in a variety of uses. According to the literature [23], river water accounts for 77.3% of the total water supply in the Loess Plateau, while groundwater represents 22.7%. River water, which is mainly derived from the Yellow River and its tributaries, is predominantly used for agricultural irrigation and industrial production, while the domestic water supply



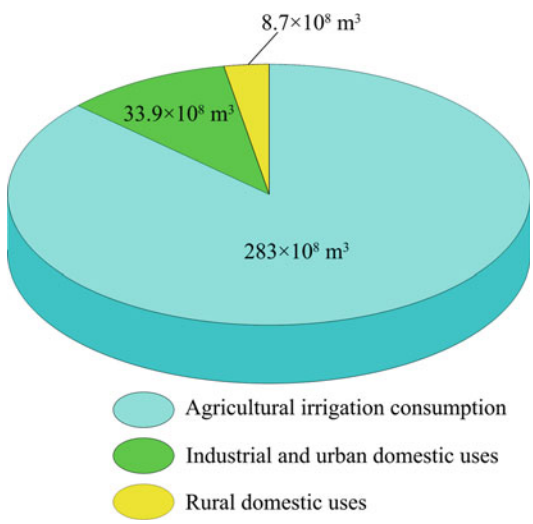
Water in Loess, Fig. 2 Global loess distribution (Modified after Pye [21] and Smalley et al. [22])



Water in Loess, Fig. 3 Loess distribution in China (Modified after Wu [6])

relies mainly on groundwater. As shown in Fig. 4, agriculture is the largest sector in terms of water use, and each year, around 87% of the total water supplies are used for irrigation, while industrial and urban domestic uses account for 10% and rural domestic use accounts for 3% of the total water supply [23]. These numbers indicate that agriculture has the largest water-saving potential. For the development of sustainable water resources on the Loess Plateau, it is crucial to adopt water-saving irrigation techniques and to implement vegetation structure adjustments.

According to the national allocation of the Yellow River water [23], the total amount of available surface water resources for the Loess Plateau is only 224.2×10^8 m^3, of which about 80% are used for irrigation, leaving only 44.84×10^8 m^3 for other uses such as ecological maintenance. In terms of groundwater, only water resources stored underneath the alluvial plains on the Loess Plateau are available for development. There are several large irrigation districts on the



Water in Loess, Fig. 4 The use of water resources on the Loess Plateau

Loess Plateau, and extensive irrigation has induced groundwater level rise, causing secondary soil salinization [24, 25]. There are also several irrigation districts experiencing groundwater level declines due to intense groundwater abstraction for multiple uses, causing damages to wells and pumps. Considering the severe situation of water resources, some Chinese scholars have proposed integrated water resources management for maintaining the sustainability of water resources [26]. It should be noted that groundwater is far more important on the Loess Plateau than other water resources, and thus, the scientific and sustainable development of groundwater is critical for the local economy and society. Soil water is also important for the sustainable development of the Loess Plateau. In the Loess Plateau area, the loess deposit is thick and porous, favoring the storage and occurrence of soil water. It has been estimated that the total soil water amount stored in a 2 m loess layer over the Loess Plateau can reach up to 1785.5×10^8 m^3 [23]. However, due to intensive evaporation in this area, the loess layers are usually in the state of water deficit.

In addition to drought and water shortage, another serious water-related problem associated with the development of economy and society on the Loess Plateau is water pollution. The Yellow River is the largest surface water source on the Loess Plateau, and its water quality has been significantly degraded because of human activities and natural environmental changes. The main stream and the tributaries of the Yellow River have been subject to contamination of different degrees. In addition, the groundwater on the Loess Plateau is undergoing severe quality deterioration. Various researches have shown that nitrate, fluoride, sulfate, and trace metals, due to agriculture/industry and natural release from soil, are major groundwater pollutants in the Loess Plateau area [27–30]. Issues related to water pollution cause dramatic economic losses each year and pose potential risks to residents and the environment [31, 32]. A study by Wu and Sun [32] has shown that health risks of local residents are mainly a result of ingestion of contaminants, with children being more vulnerable than adults. Ma et al. [33] reported a rapid decline in both surface water and groundwater quality during the past 20 years in the Malian River Basin of the Longdong Loess Plateau; petroleum contamination caused by the oil industry in the Longdong oil field is the largest source of pollution and affects surface water and groundwater quality. In addition, some studies have shown that agriculture on the Loess Plateau is the main contributor to nonpoint source pollution [34, 35]. Tourism in the Loess Plateau area, which is the economic basis of a number of cities, should also be adequately regulated. The fast development of tourism in many parts of the Loess Plateau has caused significant water pollution and environmental degradation [36].

For the sustainable development and protection of water resources in the Loess Plateau, it is required to allocate all water resources conjunctively, including surface water, groundwater, and soil water. To achieve this goal, it is mandatory to obtain a full understanding of the quality and quantity state of these water resources.

Surface Water on the Loess Plateau

The uneven distribution of rainfall causes the deficiency of annual surface water resources, hampering vegetation restoration and ecological conservation on the Loess Plateau [37]. The Yellow River and its tributaries are the main surface water resources on the Loess Plateau. In this area, surface water is characterized by three factors: First, it is scarce, and such scarcity is not in harmony with the local population and land resources. The average annual runoff in this area is only 71.1 mm per year, which is less than one third of the average runoff in China. The per capita water resource in this area is only 546 m^3, which is only 30% of the national per capita water resource, and the average water resource per hectare is only 2625 m^3, less than 10% of the national water resource per hectare [23]. Second, surface water in this area is spatiotemporally uneven. Local surface water on the Loess Plateau is mainly recharged by precipitation. Precipitation, however, decreases gradually from south to north and from east to west, making the northwestern part of the area more arid than the southeastern part. Heavy rains are relatively more common in north Shaanxi, west Shanxi, and Inner Mongolia on the Loess Plateau and less common in Western

Gansu and Ningxia. In addition, as the Loess Plateau is characterized by a continental monsoon climate, the temporal rainfall distribution in the area is highly uneven, and the amount of precipitation in July to September (summer season) occupies 50–70% of the annual total rainfall [38]. The Yellow River is the largest river flowing through the Loess Plateau and provides a large volume of surface water for irrigation and industrial uses. However, the spatiotemporal distribution of the Yellow River is also uneven. The Yellow River drainage can be divided into four zones: water abundance zone, transition zone, water shortage zone, and drought zone. The Loess Plateau mostly belongs to the transition and the water shortage zones, and the area of the Loess Plateau accounts for approximately 62% of the Yellow River drainage, but the volume of river water accounts for only 30% of the total Yellow River runoff [23], indicating an unevenness of spatial surface water distribution. Third, the surface water in the Loess Plateau contains large volumes of sediment. It has been estimated that the annual sediment load of the Yellow River reached 16×10^8 tons each year in the 1990s. In recent years, a variety of engineering and biological measures for water and soil conservation have been implemented on the Loess Plateau; as a result, the annual sediment load of the Yellow River is declining and soil erosion is alleviated in some areas [39]. However, due to the rapid economic development and urbanization, some local governments now neglect soil and water conservation, and many cities on the Loess Plateau focus on urban expansion, resulting in severe damage to vegetation planted 20 years ago [18, 40, 41]. Such an approach is likely to impede water and soil conservation achievements obtained by local governments. Shi and Shao [39] and Sun et al. [40] state that soil and water loss in the Loess Plateau region are caused by natural erosion and accelerated erosion. Cultivation, uncontrolled development, overgrazing, mining, road construction, and urbanization are important anthropogenic factors accelerating erosion. To maintain the soil and water conservation achievements on the Loess Plateau, it is important and compulsory to strictly control human activities in this region.

The annual surface water resource in the Loess Plateau was 321.5×10^8 m^3 from 1950 to 1989 [23]. However, the amount of surface water varies from year to year and is influenced by recharge source, land cover/land use, and drainage size. The river runoffs measured at the main stations of the Loess Plateau from 1919 to 1979 are shown in Table 1 [23]. Annual runoff varies from year to year and from station to station. For the Sanmenxia station located in the downstream of the Loess Plateau, the average annual runoff is 504×10^8 m^3, with a maximum annual runoff of 823×10^8 m^3 and a minimum annual runoff of 242×10^8 m^3. The minimum annual runoff is less than half of the average annual runoff, while the maximum annual runoff is 3.5 times that of the minimum annual runoff. For the Lanzhou station in the upstream, the average annual runoff is 326×10^8 m^3, which is 64.7% of that at the Sanmenxia station. The maximum and minimum annual runoffs are also lower than those at the downstream stations, except for Hekou station, which indicates that local surface water joins the main Yellow River. The ratios between the maximum and minimum annual runoff at all four stations are larger than 3.0, suggesting significant interannual variation of the Yellow River runoff.

The overall water quality of the surface water on the Loess Plateau is not a reason for optimism, although water quality protection measures have been launched and some local improvements have been witnessed. Water quality is classified into five grades according to the Chinese water quality standards. Grades 1 and 2 represent excellent and good quality water which can be used for various purposes. Water falling into grade 3 is of fair quality and acceptable for domestic uses, while grade 4 water is of poor quality but can be used for irrigation. Grade 5 represents very-poor-quality water that cannot be used for any purpose [42, 43].

The river water quality assessment results for the Yellow River drainage from 1998 to 2007 are shown in Table 2 [44, 45]. The assessment was performed according to the surface water quality standard released by the Bureau of Quality and Technical Supervision of China [43]. As shown in Table 2, more than half of the rivers were characterized by very poor water quality (grade 5), and almost 80% of the river water was of poor and very poor water quality (grades 4 and 5) and therefore unsuitable for domestic uses in the first several years of the twenty-first century, indicating serious water pollution because of rapid economic development without adequate consideration of ecological issues. Particularly, in 2002, only less than 20% of the river water was acceptable for domestic purposes, and excellent and good quality water accounted for only about 5% of the total river length. This is, in fact, a poor condition for residents living in this area. In 2004, the river length with excellent and good quality exceeded 10% for the first time, and since then, the river water quality has gradually improved. In addition, since 2005, the total river length monitored and assessed has been accounting for over 10,000 km and since 2012, for more than 20,000 km, indicating a development in monitoring techniques and increasing investment in water quality protection. Most importantly, the river length with acceptable water quality for domestic uses (grades 1 to 3) has been over 50% of the total river length monitored and assessed since 2012, which is regarded as a great success in river water pollution remediation.

The river water quality trend from 2002 to 2015 on the Loess Plateau is shown in Fig. 5. The percentage of very-poor-quality river water has been generally decreasing since 2002, while that of grades 1–2 water has been increasing, clearly demonstrating an improvement of river water quality. The percentages of fair-quality water and poor-quality water are also slightly decreasing. In particular, the water quality improvement since 2012 is significant probably because of the national strategy on ecological civilization construction.

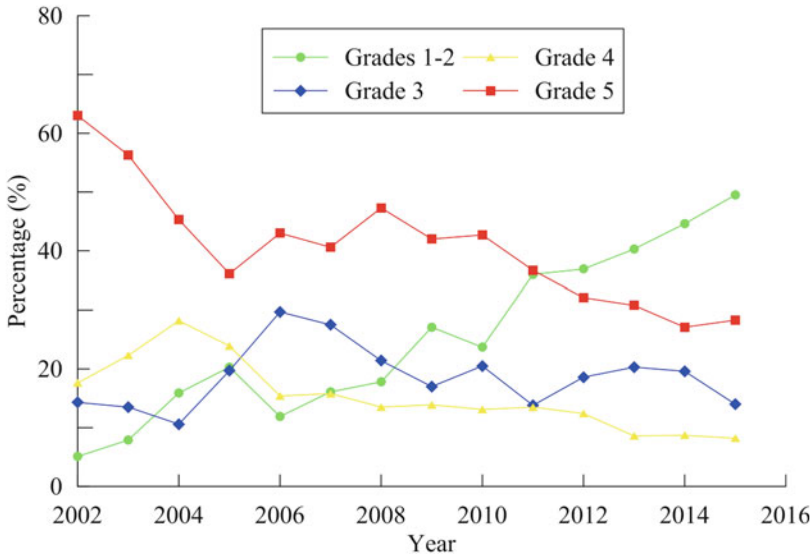
The significant water quality improvement achieved in the past 10 years is encouraging. However, ecological civilization construction still needs to be further developed. As shown in Table 4, in 2015, over 30% of the rivers were still contaminated, and recovering these rivers requires long-term efforts and large amounts of investments as well as endeavor from all parts including governments, scientists, and the public.

Water in Loess, Table 1 Annual surface water runoff measured in main stations of the Loess Plateau from 1919 to 1979 ($\times 10^8$ m^3)

Station	Average annual runoff	Maximum annual runoff	Minimum annual runoff	Ratio (maximum/minimum)
Lanzhou	326	517	163	3.2
Hekou town	319	532	156	3.4
Longmen	390	632	197	3.2
Sanmenxia	504	823	242	3.4

Water in Loess, Table 2 River water quality assessment results for the Yellow River drainage from 2002 to 2015

Year of assessment	Total length of the rivers assessed (km)	Percentage of different water quality grades (%)			
		Grades 1 and 2	Grade 3	Grade 4	Grade 5
2002	7497.0	5.1	14.3	17.6	63.0
2003	7497.0	7.9	13.5	22.3	56.3
2004	7497.0	15.9	10.6	28.2	45.3
2005	13228.4	20.3	19.7	23.9	36.1
2006	12510.8	11.9	29.7	15.4	43.0
2007	13492.7	16.1	27.5	15.8	40.6
2008	13847.7	17.8	21.4	13.5	47.3
2009	14039.3	27.1	17.0	13.9	42.0
2010	14295.4	23.7	20.5	13.1	42.7
2011	19734.2	36.0	13.8	13.5	36.7
2012	20545.3	36.9	18.6	12.4	32.1
2013	21538.7	40.3	20.3	8.6	30.8
2014	20209.7	44.6	19.6	8.7	27.1
2015	21655.0	49.5	14.0	8.2	28.3



Water in Loess, Fig. 5 River water quality trend from 2002 to 2015 on the Loess Plateau

Water in Loess, Table 3 Natural groundwater resources of different types on the Loess Plateau [46]

Aquifer	Area (10^4 km^3)	Groundwater amount (10^8 m^3 per year)	Percentage (%)	Groundwater modulus (10^4 m^3/km^2 per year)
Pore water in loose rocks	16.59	195.64	58.67	11.79
Fissure water in crystalline rocks	10.43	48.63	14.58	4.66
Karst water in carbonate rocks	4.51	44.98	13.49	9.97
Pore-fissure water in clastic rocks	9.94	25.59	7.67	2.57
Pore-fissure water in loess and underlying rocks	20.00	18.61	5.58	0.93
Total	61.47	333.45	100.00	5.42

Groundwater on the Loess Plateau

As surface water on the Loess Plateau is scarce and mostly of poor quality, groundwater has long been the most important source of water in many parts of the arid and semiarid Loess Plateau [26, 42]. The occurrence, flow, and spatial distribution of groundwater on the Loess Plateau are controlled by geological and hydrogeological settings. Climate and geomorphological and hydrological conditions are also vital factors influencing its occurrence, distribution, and transformation. Groundwater on the Loess Plateau can be divided into four main types according to aquifer properties: pore water in loose rocks, karst water in carbonate rocks, pore-fissure water in clastic rocks, and fissure water in crystalline rocks [23], with pore water in loose rocks being the most abundant type. The estimated amount of total natural groundwater resources on the Loess Plateau is 333.45 × 10^8 m^3 (Table 3) [46]. Precipitation is the main source of groundwater recharge for groundwater in the region, and in some areas, it is the only recharge source. However, soil conservation measures on the Loess Plateau can alter the hydrologic cycle and change the water fluxes through the land surface to groundwater [47]. Influenced by climate zoning, groundwater is abundant in the south and scarce in the north of the Loess Plateau. The loess pores and fissures are the main sites for the storage and transportation of loess groundwater. Among these pores and fissures, small but uniform pores and microfractures represent the main water storage spaces in the loess layers, while large, nonuniform, but well-connected fissures and voids are the dominant migration channels for groundwater in loess layers [48].

There are several large basins (alluvial plains) on the Loess Plateau, such as the Yinchuan Basin in Ningxia, the Guanzhong Basin in Shaanxi, the Hetao Plain in Inner Mongolia, and the Fenhe Plain in Shanxi. These basins are densely populated, and groundwater is the most important water source for domestic water supply. The groundwater in these basins is mainly pore water, and the aquifers are usually thick and characterized by multilayer structures. The groundwater modulus in these basins is high, indicating abundant groundwater resources. For example, the groundwater modulus in the Yinchuan Basin ranges from 43.3×10^4 m^3/km^2 per year to 55.7×10^4 m^3/km^2 per year, being higher than the average groundwater modulus of pore water in loose rocks (Table 3). Typically, precipitation is an important recharge source for groundwater underneath the basins. However, irrigation infiltration and percolation from irrigation channels are major recharge sources for groundwater in some basins. Irrigation infiltration and percolation from irrigation channels account for over 70% of the total recharge for groundwater in both the Yinchuan Basin and the Weining Plain [49, 50], and in the Guanzhong Basin, irrigation is also one of the major recharge sources for groundwater [29, 31]. According to Li et al. [23], the total natural groundwater resources of the Loess Plateau amount up to 177×10^8 m^3 per year.

Groundwater in the loess tablelands and hilly-gully regions shows different characteristics from that in the basins. In these areas, the rocks underlying loess develop large numbers of fissures, facilitating water seepage. As such, the loess and underlying rocks usually form a unified aquifer characterized by a dual structure with pore water in the upper loess and fissure water in the underlying rocks. Phreatic groundwater underneath the tablelands usually flows from the center of the tablelands to the edges and outcrops as springs in the foot of the loess slope. Much of the phreatic groundwater with great abundance in the tablelands has been exploited for multiple uses. In the hilly-gully regions of the Loess Plateau, the terrain is strongly cut by gullies, enabling groundwater discharge. Most of the fissure water outcrops and is discharged as springs in the slopes, forming seasonal surface water in the valleys. The groundwater modulus in the arid hilly-gully regions of the Loess Plateau is usually smaller than 1.0×10^4 m^3/km^2 per year, while that in the semiarid and semi-humid hilly-gully regions of the Loess Plateau ranges from 1.0×10^4 to 5.0×10^4 m^3/km^2 per year, indicating that groundwater recharge in the loess tablelands and the hilly-gully regions is limited. According to the literature [23], the amount of total natural groundwater resources in the loess tablelands and the hilly-gully regions of the Loess Plateau is only 63×10^8 m^3 per year.

Groundwater in the mountainous parts of the Loess Plateau is mainly fissure water and karst water. Fissure water is widely distributed on the Ordos Plateau, in the Guide and Qinshui Basins, and in the water division zones of the Taihang, Lvliang, and Liupan mountains. Karst water is mainly found in the Taihang, Lvliang, and Beishan mountains [48]. Vertical infiltration of precipitation is the most important recharge source for groundwater in the mountainous regions [49]. Groundwater in the mountainous regions is more abundant than in the loess tablelands and the hilly-gully regions, and the groundwater modulus in the mountainous regions ranges from 5.0×10^4 to 15.0×10^4 m^3/km^2 per year [23]. The amount of total natural groundwater resources in the mountainous regions of the Loess Plateau is approximately 94×10^8 m^3 per year.

Compared to surface water, groundwater can be protected by the upper unsaturated zone and therefore is usually less contaminated and of better quality than surface water. In general, groundwater in the Loess Plateau is mainly freshwater with a salinity below 1000 mg/L. Brackish water with a salinity higher than 1000 mg/L is mainly distributed in northern Yinchuan, Western Gansu, and the Xihaigu region of Ningxia (including Xiji, Haiyuan, and Guyuan). The freshwater resource accounts for about 86% of the total natural groundwater resources, while brackish water accounts for 10% and saline water with a salinity higher than 3000 mg/L for about 4% [46, 48]. Natural groundwater quality anomalies and anthropogenic groundwater pollution are two main issues occupying local decision-makers and scientists. In many parts of the Loess Plateau, the natural background contents of harmful elements such as fluoride, arsenic, Mn, and Cr^{6+} are abnormally high. For example, abnormally high concentrations of fluoride have been reported in the Yinchuan Plain [30, 50, 51], southern Ningxia [52, 53], the Guanzhong Basin [29], and the Datong Basin [54, 55]. High arsenic levels in groundwater are also common in the northern Yinchuan Plain [51], the Datong Basin [56], and the Hetao Plain [57]. These groundwater quality anomalies are most likely the results of specific geological and hydrogeological conditions and water-rock interactions [58].

Anthropogenic groundwater contamination has also been widely reported in many parts of the Loess Plateau and has gradually become the most serious risk for human health. Common anthropogenic factors affecting groundwater quality include agricultural activities, industrial effluents, coal and oil mining activities, tourism, and wastes from livestock. Agriculture is an important factor influencing shallow groundwater quality and is mainly responsible for groundwater nitrate pollution in many agricultural production regions of the Loess Plateau. Guo et al. [59] studied the impacts of fertilization practices on environmental risks of nitrate in semiarid farmlands of the Chinese Loess Plateau by analyzing soil profiles. Wu and Sun [31] reported shallow groundwater pollution in an agricultural and industrial region of the Guanzhong Basin and concluded that agricultural and industrial activities are important factors affecting nitrate pollution of shallow groundwater. As the Loess Plateau is rich in coal and oil resources, the exploitation of these resources has entailed serious groundwater pollution. Pan et al. [60] and Ma et al. [33] have reported petroleum pollution in both loess and groundwater underneath the loess. They concluded that petroleum contamination caused by the oil industry is the largest source of pollution, leading to the deterioration of the water quality.

Considering the significance of groundwater pollution in the loess areas, groundwater quality research to alleviate water pollution in northwest China is crucial. Li [61] and Li et al. [62] raised concerns about groundwater quality research in Western China and proposed to enhance groundwater monitoring networks, increase research investments from central and local governments, improve basic environmental education, and strengthen the collaboration among different organizations and parties.

Soil Water on the Loess Plateau

In addition to surface water and groundwater, soil water is another important and valuable water resource for agriculture on the Loess Plateau, especially in semiarid and semi-humid areas. It is a bond between surface water and groundwater conversion [63]. Because the unsaturated zone on the Loess Plateau is thick, only a small portion of precipitation can flow into rivers or infiltrate into aquifers, recharging groundwater. Most of the precipitation on the Loess Plateau is stored in the unsaturated zone as soil water. Soil water is an important component of water resources and a direct source and key element for vegetation growth and, therefore, highly important for ecological civilization construction. As such, research on soil water resources has become a hot topic in hydrogeological and ecological communities [64, 65].

The thick loess cover on the Loess Plateau has created a unique condition for the conversion of rainwater resources into soil water. Loess soil is thick and porous, and its effective porosity can reach 25–30% [23]. The soil water capacity for a 200 cm soil layer can reach 551.1–847.4 mm, that is, the theoretical soil water amount can reach 55.1×10^8 m^3–84.7×10^8 m^3 for each 10,000 km^2 of loess land. However, the time of soil water retention is relatively short due to gravity and strong evaporation, and therefore, not all soil water can be used by vegetation. When the water content is larger than a threshold, soil water will infiltrate into groundwater. This threshold is called field capacity or field moisture capacity and represents the highest amount of water content that can be retained in the soil under field conditions. The water content under which plants cannot liberate the remaining moisture from the soil particles is called wilting point. Water available for vegetation growth is defined as available water capacity, the soil water within the range between field capacity and wilting point [66]. On the Loess Plateau, the total amount of the available soil water resources is 1785.54×10^8 m^3 [23]. It should be noted that the soil water storage capacity considerably varies under different land cover/ land use types [67], and soil water content also changes with soil depth [68].

Overall, the total amount of regional water resources in the loess areas is controlled by precipitation. Surface water and groundwater are both derived from precipitation and readily available for multiple uses. Soil water is also derived from precipitation but cannot be used directly by humans. Due to the thick loess layer on the Loess Plateau, where surface water and groundwater are either limited or contaminated, soil water resources provide a supply for the majority of rain-fed agriculture and forestry. Such resources must therefore be protected as surface water and groundwater and should be considered in regional water resources allocation.

Future Directions

The Loess Plateau is a significant part of the Belt and Road Initiative, which acts as a bridge between Asia and Europe [3]. The water resources problems along the road have always been the concern of international scholars [69, 70], as the Belt and Road Initiative will increase the demand for water resources and the intensified human activities negatively impact the environment. However, based on the Belt and Road Initiative, many national and international organizations and institutes are involved in various scientific research projects to minimize the negative effects. Today, the Loess Plateau, as the most important Chinese part of the Belt and Road Initiative, draws much more attention from scholars than ever.

Water is the key to the success of the initiative and of the ecological civilization construction. A number of significant results have been achieved regarding water resources allocation, water conservation, and water quality protection on the Loess Plateau. However, the water-related issues in the loess areas of China still need to be evaluated in detail. The following issues or research fields are important and should be seriously considered:

- Spatiotemporal evolution and drought history of the Loess Plateau. This topic is of great value to track the water environment when loess was deposited and the cause of drought. Such research will require a long record of historical climate data and geochemical data.
- Impacts of climate change on water resources and water resources vulnerability on the Loess Plateau. Based on the analysis on the impacts of climate change on the water demand and water supply, this research topic should help to establish a new balance between water demand and water supply and assess water resources vulnerability.
- Relationship between human activities and drought on the Loess Plateau. This research field focuses on the impacts of land use/land cover changes on hydrological processes at different scales. It links human activities and climate change, which is quite complex and therefore requires international collaboration.
- Impacts of human activities on the permeability of loess. Large-scale human activities will completely change the physical properties of loess, affecting loess permeability and soil water dynamic. Research on this topic, via experiments and numerical modeling, seeks answers to reduce these negative impacts.
- Groundwater research on the Loess Plateau. The thick soil layer and the diverse and complex geological structures and geomorphology on the Loess Plateau impede groundwater research. It is therefore necessary to find ways to accelerate groundwater research on the Loess Plateau by proposing key groundwater research topics and introducing international collaborations and advanced technologies.

Bibliography

Primary Literature

1. Pan D, Jia H, Yuan Y (2015) A GIS-based ecological safety assessment of Wushen banner, China. Hum Ecol Risk Assess Int J 21(2):297–306. https://doi.org/10.1080/10807039.2014.913441

2. Kovács J, Varga G (2013) Loess. In: Bobrowsky PT (ed) Encyclopedia of natural hazards. Springer, Dordrecht, pp 637–638
3. Li P, Qian H, Howard KWF, Wu J (2015) Building a new and sustainable silk road economic belt. Environ Earth Sci 74(10):7267–7270. https://doi.org/10.1007/s12665-015-4739-2
4. Davis RS, Holliday VT (2017) Loessic Paleolithic, Tajikistan. In: Gilbert AS (ed) Encyclopedia of Geoarchaeology. Springer, Dordrecht, pp 492–494
5. Sun J (2005) Loessology (Vol. 1). Hong Kong Archaeological Society, Hong Kong. (in Chinese)
6. Wu J (2015) Evolution and regulation of groundwater environment affected by land creation projects in loess area: a case study from the construction of Yan'an new district. Ph.D. dissertation of Chang'an University, Xi'an.(in Chinese)
7. Liu D (1965) Chinese loess accumulation. Science Press, Beijing. (in Chinese)
8. Zhang Z, Zhang Z, Wang Y (1989) Chinese loess. Geologic Publishing House, Beijing. (in Chinese)
9. Pumpelly R (1866) Geological researches in China, Mongolia and Japan during the years 1862 to 1865. Smithsonian Contributions to Knowledge, Washington
10. Richthofen F (1877–1885) China: Ergebnisse eigener Reisen und darauf gegründeter Studien, 5 vols, Reimer, Berlin
11. Willis B, Blackwelder E, Sargent RH (1907) Research in China, volume 1, part 1: descriptive topography and geology. Carnegie Institution of Washington, Washington DC
12. Zhu X (1958) Discussion on the red layer in loess layer. J Quat Sci 1(1):74–82. (in Chinese)
13. Frechen M (2011) Loess in Europe. J Quat Sci 60(1):3–5. https://doi.org/10.3285/eg.60.1.00
14. Wang J, Xiang W, Zuo X (2010) Situation and prevention of loess water erosion problem along the west-to-east gas pipeline in China. J Earth Sci 21(6):968–973. https://doi.org/10.1007/s12583-010-0150-9
15. Jiao F, Wen Z-M, An S-S (2013) Soil water storage capacity under chronosequence of revegetation in Yanhe watershed on the Loess Plateau, China. SpringerPlus 2(Suppl 1):S15. https://doi.org/10.1186/2193-1801-2-S1-S15
16. Li Z, Liu WZ, Zhang XC, Zheng FL (2010) Assessing and regulating the impacts of climate change on water resources in the Heihe watershed on the Loess Plateau of China. Sci China Earth Sci 53(5):710–720. https://doi.org/10.1007/s11430-009-0186-9
17. Zhang Q, Wang S, Xi W, Li H (2011) Experimental study of the imbalance of water budget over the Loess Plateau of China. Acta Meteor Sin 25(6):765–773. https://doi.org/10.1007/s13351-011-0607-5
18. Li P, Qian H, Wu J (2014) Accelerate research on land creation. Nature 510(7503):29–31. https://doi.org/10.1038/510029a
19. Zárate MA (2017) Eolian settings: loess. In: Gilbert AS (ed) Encyclopedia of Geoarchaeology. Springer, Dordrecht, pp 233–238
20. Sun J, Wang L, Men Y, Dong Z, Yi X, Wang Q, Zhao X, Li Z (2013) Loessology (Vol. 2). Xi'an Cartographic Publishing House, Xi'an. (in Chinese)
21. Pye K (1984) Loess. Prog Phys Geogr 8:176–217. https://doi.org/10.1177/030913338400800202
22. Smalley I, Marković SB, Svirčev Z (2011) Loess is [almost totally formed by] the accumulation of dust. Quat Int 240:4–11. https://doi.org/10.1016/j.quaint.2010.07.011
23. Li R, Yang W, Li B (2008) Research and future prospects for the Loess Plateau of China. Science Press, Beijing. (in Chinese)
24. Li P, Wu J, Qian H (2016) Regulation of secondary soil salinization in semi-arid regions: a simulation research in the Nanshantaizi area along the silk road, northwest China. Environ Earth Sci 75(8):698. https://doi.org/10.1007/s12665-016-5381-3
25. Wu J, Li P, Qian H, Fang Y (2014) Assessment of soil salinization based on a low-cost method and its influencing factors in a semi-arid agricultural area, northwest China. Environ Earth Sci 71(8):3465–3475. https://doi.org/10.1007/s12665-013-2736-x
26. Li P (2012) Theory and practice of water science. Science Press, Beijing. (in Chinese)
27. Xing M, Liu W (2016) Using dual isotopes to identify sources and transformations of nitrogen in water catchments with different land uses, Loess Plateau of China. Environ Sci Pollut Res Int 23:388–401. https://doi.org/10.1007/s11356-015-5268-y
28. Li P, Wu J, Qian H, Zhou W (2016) Distribution, enrichment and sources of trace metals in the topsoil in the vicinity of a steel wire plant along the silk road economic belt, northwest China. Environ Earth Sci 75(10):909. https://doi.org/10.1007/s12665-016-5719-x
29. Li P, Qian H, Wu J, Chen J, Zhang Y, Zhang H (2014) Occurrence and hydrogeochemistry of fluoride in shallow alluvial aquifer of Weihe River, China. Environ Earth Sci 71(7):3133–3145. https://doi.org/10.1007/s12665-013-2691-6
30. Wu J, Li P, Qian H (2015) Hydrochemical characterization of drinking groundwater with special reference to fluoride in an arid area of China and the control of aquifer leakage on its concentrations. Environ Earth Sci 73(12):8575–8588. https://doi.org/10.1007/s12665-015-4018-2
31. Wu J, Sun Z (2016) Evaluation of shallow groundwater contamination and associated human health risk in an alluvial plain impacted by agricultural and industrial activities, mid-west China. Expo Health 8(3):311–329. https://doi.org/10.1007/s12403-015-0170-x
32. Li P, Wu J, Qian H (2014) Origin and assessment of groundwater pollution and associated health risk: a case study in an industrial park, northwest China. Environ Geochem Health 36(4):693–712. https://doi.org/10.1007/s10653-013-9590-3
33. Ma J, Pan F, He J, Chen L, Fu S, Jia B (2012) Petroleum pollution and evolution of water quality in the Malian River basin of the Longdong Loess Plateau,

northwestern China. Environ Earth Sci 66(7):1769–1782. https://doi.org/10.1007/s12665-011-1399-8

34. Wu L, Liu X, Ma X (2016) Spatio-temporal variation of erosion-type non-point source pollution in a small watershed of hilly and gully region, Chinese Loess Plateau. Environ Sci Pollut Res 23(11):10957–10967. https://doi.org/10.1007/s11356-016-6312-2
35. Wu L, Li P, Ma X (2016) Estimating nonpoint source pollution load using four modified export coefficient models in a large easily eroded watershed of the loess hilly–gully region, China. Environ Earth Sci 75:1056. https://doi.org/10.1007/s12665-016-5857-1
36. Wu J, Xue C, Tian R, Wang S (2017) Lake water quality assessment: a case study of Shahu Lake in the semiarid loess area of northwest China. Environ Earth Sci 76:232. https://doi.org/10.1007/s12665-017-6516-x
37. Huang J, Wen J, Osamu H, Hiroshi Y (2014) Runoff and water budget of the Liudaogou catchment at the wind–water erosion crisscross region on the Loess Plateau of China. Environ Earth Sci 72:3623–3633. https://doi.org/10.1007/s12665-014-3273-y
38. Cai Q (2001) Soil erosion and management on the Loess Plateau. J Geogr Sci 11(1):53–70
39. Shi H, Shao M (2000) Soil and water loss from the Loess Plateau in China. J Arid Environ 45:9–20. https://doi.org/10.1006/jare.1999.0618
40. Sun H, Gan Z, Yan J (2001) The impacts of urbanization on soil erosion in the Loess Plateau region. J Geogr Sci 11(3):282–290
41. Zhao J (2005) The five major changes in the evolution of the Loess Plateau. J Geogr Sci 15(4):475–483. https://doi.org/10.1360/gs050411
42. Li P, Wu J, Qian H (2014) Hydrogeochemistry and quality assessment of shallow groundwater in the southern part of the Yellow River alluvial plain (Zhongwei section), China. Earth Sci Res J 18(1): 27–38. https://doi.org/10.15446/esrj.v18n1.34048
43. Bureau of Quality and Technical Supervision of China (2002) Environmental quality standards for surface water, (GB 3838–2002). China Environmental Science Press, Beijing. (in Chinese)
44. Li P, Gao X (2010) Water resources status and change rule in Yellow River basin. Water Conserv Sci Technol Eco 16(1):45–47. (in Chinese)
45. Yellow River Conservancy Commission of the Ministry of Water Resources of China (2009–2016) Yellow River Water Resources Bulletin for the year of 2008–2015. http://www.yellowriver.gov.cn/other/hhgb/. Accessed 21 May 21 2017
46. Teng Z, Zhang Y, Hu W, Li L (2000) Water resource and water quality evaluation of underground water in Loess Plateau. J Northwest Univ (Nat Sci Ed) 30(1):60–64. (in Chinese)
47. Gates JB, Scanlon BR, Mu X, Zhang L (2011) Impacts of soil conservation on groundwater recharge in the semi-arid Loess Plateau, China. Hydrogeol J 19:865–875. https://doi.org/10.1007/s10040-011-0716-3
48. Wang Z, Wang Z, Wang G (1990) Study on rational development and utilization of groundwater resources in Loessial plateau area. Academy Press, Beijing. (in Chinese)
49. Li P, Zhang Y, Yang N, Jing L, Yu P (2016) Major ion chemistry and quality assessment of groundwater in and around a mountainous tourist town of China. Expo Health 8(2):239–252. https://doi.org/10.1007/s12403-016-0198-6
50. Qian H, Li P (2011) Hydrochemical characteristics of groundwater in Yinchuan plain and their control factors. Asian J Chem 23(7):2927–2938
51. Chen J, Qian H, Wu H, Gao Y, Li X (2017) Assessment of arsenic and fluoride pollution in groundwater in Dawukou area, Northwest China, and the associated health risk for inhabitants. Environ Earth Sci 76:314. https://doi.org/10.1007/s12665-017-6629-2
52. Wu J, Li P, Qian H (2012) Study on the hydrogeochemistry and non-carcinogenic health risk induced by fluoride in Pengyang County, China. Int J Environ Sci 2(3):1127–1134. https://doi.org/10.6088/ijes.00202030001
53. Wei C, Guo H, Zhang D, Wu Y, Han S, An Y, Zhang F (2016) Occurrence and hydrogeochemical characteristics of high-fluoride groundwater in Xiji County, southern part of Ningxia Province, China. Environ Geochem Health 38(1):275–290. https://doi.org/10.1007/s10653-015-9716-x
54. Li L, Wang Y, Wu Y, Li J (2013) Major geochemical controls on fluoride enrichment in groundwater: a case study at Datong Basin, northern China. J Earth Sci 24(6):976–986. https://doi.org/10.1007/s12583-013-0385-3
55. Guo Q, Wang Y, Ma T, Ma R (2007) Geochemical processes controlling the elevated fluoride concentrations in groundwaters of the Taiyuan Basin, northern China. J Geochem Explor 93(1):1–12. https://doi.org/10.1016/j.gexplo.2006.07.001
56. Yu Q, Wang Y, Ma R, Su C, Wu Y, Li J (2014) Monitoring and modeling the effects of groundwater flow on arsenic transport in Datong Basin. J Earth Sci 25(2):386–396. https://doi.org/10.1007/s12583-014-0421-y
57. Guo H, Zhang Y, Jia Y, Zhao K, Kim K (2013) Spatial and temporal evolutions of groundwater arsenic approximately along the flow path in the Hetao basin, Inner Mongolia. Chin Sci Bull 58:3070–3079. https://doi.org/10.1007/s11434-013-5773-7
58. Ali S, Thakur SK, Sarkar A, Shekhar S (2016) Worldwide contamination of water by fluoride. Environ Chem Lett 14(3):291–315. https://doi.org/10.1007/s10311-016-0563-5
59. Guo S, Wu J, Dang T, Liu W, Li Y, Wei W, Keith Syers J (2010) Impacts of fertilizer practices on environmental risk of nitrate in semiarid farmlands in the Loess Plateau of China. Plant Soil 330:1–13. https://doi.org/10.1007/s11104-009-0204-x
60. Pan F, Ma J, Wang Y, Zhang Y, Chen L, Edmunds WM (2013) Simulation of the migration and transformation of petroleum pollutants in the soils of the Loess Plateau: a case study in the Maling oil field of northwestern China.

Environ Monit Assess 185:8023–8034. https://doi.org/10.1007/s10661-013-3152-0
61. Li P (2016) Groundwater quality in western China: challenges and paths forward for groundwater quality research in western China. Expo Health 8(3):305–310. https://doi.org/10.1007/s12403-016-0210-1
62. Li P, Tian R, Xue C, Wu J (2017) Progress, opportunities and key fields for groundwater quality research under the impacts of human activities in China with a special focus on western China. Environ Sci Pollut Res Int 24(15):13224–13234. https://doi.org/10.1007/s11356-017-8753-7
63. Zhang W, Zhang P, Li J, Ma Z (2012) Evaluation of soil water resource and study of ecological recovery in Loess Plateau. Yellow River 34(10):100–102. https://doi.org/10.3969/j.issn.1000-1379.2012.10.030. (in Chinese)
64. Liu X, He B, Yi X, Zhang L, Han F (2016) The soil water dynamics and hydraulic processes of crops with plastic film mulching in terraced dryland fields on the Loess Plateau. Environ Earth Sci 75:809. https://doi.org/10.1007/s12665-016-5670-x
65. Wang Z, Liu B, Zhang Y (2009) Soil moisture of different vegetation types on the Loess Plateau. J Geogr Sci 19:707–718. https://doi.org/10.1007/s11442-009-0707-7
66. Yang W (2001) Soil water resources and afforestation in Loess Plateau. J Nat Resour 16(5):433–438. (in Chinese)
67. Jiao F, Wen Z-M, An S-S (2013) Soil water storage capacity under chronosequence of revegetation in Yanhe watershed on the Loess Plateau, China. SpringerPlus 2(Suppl 1):S15. https://doi.org/10.1186/2193-1801-2-S1-S15
68. Wang ZQ, Liu BY, Liu G, Zhang YX (2009) Soil water depletion depth by planted vegetation on the Loess Plateau. Science in China series D: earth. Sciences 52(6):835–842. https://doi.org/10.1007/s11430-009-0087-y
69. Howard KWF, Howard KK (2016) The new "silk road Economic Belt" as a threat to the sustainable management of Central Asia's transboundary water resources. Environ Earth Sci 75:976. https://doi.org/10.1007/s12665-016-5752-9
70. Li P, Qian H, Zhou W (2017) Finding harmony between the environment and humanity: an introduction to the thematic issue of the silk road. Environ Earth Sci 76:105. https://doi.org/10.1007/s12665-017-6428-9

Books and Reviews

An Z (2014) Late Cenozoic climate change in Asia: loess, monsoon and monsoon-arid environment evolution. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-7817-7. ISBN:978-94-007-7816-0

Bakels CC (2009) The western European Loess Belt: agrarian history, 5300 BC - AD 1000. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-9840-6. ISBN:978-1-4020-9839-0

Derbyshire E, Dijkstra T, Smalley IJ (1995) Genesis and properties of collapsible soils. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-011-0097-7. ISBN:978-94-010-4047-1

Li T, Wang C, Li P (2012) Loess deposit and loess landslides on the Chinese Loess Plateau. In: Wang F, Miyajima M, Li T, Shan W, Fathani TF (eds) Progress of geo-disaster mitigation Technology in Asia. Springer, Berlin/Heidelberg, pp 235–261

Tsunekawa A, Liu G, Yamanaka N, Du S (2014) Restoration and development of the degraded Loess Plateau, China. Springer Japan, Tokyo. https://doi.org/10.1007/978-4-431-54481-4. ISBN:978-4-431-54480-7

Wang X, Jiao F, Li X, An S (2017) The Loess Plateau. In: Zhang L, Schwärzel K (eds) Multifunctional land-use Systems for Managing the nexus of environmental resources. Springer, Heidelberg, pp 11–27

Yang X, Liu T, Yuan B (2009) The Loess Plateau of China: Aeolian sedimentation and fluvial erosion, both with superlative rates. In: Migon P (ed) Geomorphological landscapes of the world. Springer Netherlands, Dordrecht, pp 275–282

Part IV

Environmental Remediation and Sustainability

Desertification and Impact on Sustainability of Human Systems

David Mouat, Scott Thomas and Judith Lancaster
Division of Earth and Ecosystem Sciences, Desert Research Institute, Reno, NV, USA

Article Outline

Glossary

Adaptation Changes made by organisms, including humans, to enable them to be more suitable for different conditions or situations.

Desertification The gradual degradation of habitable land which affects soils, flora, and fauna and reduces productivity and an ecosystem's ability to adapt. It is caused by various factors, including natural dynamics, climatic condition, and human activities. Direct causes of desertification include wind and water erosion, deforestation, overgrazing, land conversion to crop land, irrigation leading to salinization, and other physical stressors leading to soil loss, soil compaction, loss of vegetative cover, loss of biodiversity, and degradation in ecosystem productivity. Indirect causes include climate variation, poverty, political instability, lack of education, or a combination of factors.

Drought A period of reduced precipitation resulting in prolonged shortages in the water supply. It can have a substantial impact on the ecosystem, agriculture, and economy of the affected region.

Drylands Include arid, semiarid, and dry subhumid areas and have a ratio of mean annual precipitation to mean annual potential evapotranspiration ranging between 0.05 and 0.65.

Ecosystem services The benefits people obtain from ecosystems. These may be divided into provisioning (food and water), regulating (controlling floods and diseases), cultural (recreational, spiritual), and supporting services such as nutrient cycling.

Land-based natural capital Includes the properties of the soil and geomorphological, biotic, and hydrological features that interact with each other and with climate to determine the quantity and nature of ecosystem services provided by the land.

Land degradation A temporary or permanent lowering of the land-based natural capital and/or economic productive capacity of land, including changes to soil and vegetation.

Land degradation neutrality A state whereby the amount and quality of land resources necessary to support ecosystem functions and services and enhance food security remain stable or increase within specified temporal and spatial scales and ecosystems.

Resilience The ability of a system to absorb disturbance and retain essentially the same function, structure, and feedbacks.

Sustainability (as it relates to desertification) Is the conservation of land-based natural capital and ecosystem services so that people are able to meet the needs of the present without compromising the ability of future generations to meet their own needs.

Vulnerability A condition of being at risk. Being susceptible to harm or damage.

Definition of the Subject

Desertification is not a new phenomenon, but it appears to be increasing in severity and extent.

Judith Lancaster has retired.

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_268

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_268-3

The Millennium Ecosystem Assessment estimated that 10–20% of drylands were degraded in 2005 [1]. More recent estimates are that up to 25% of all land worldwide is currently highly degraded, 36% is slightly or moderately degraded, and 10% is improving [2].

The term "desertification" originates from the 1940s [3], although it was not widely recognized until the West African sub-Sahelian droughts of the 1960s and 1970s. As Dietz and coauthors [4] discuss, the West African area has recovered from those droughts, not only because natural fluctuations in climate resulted in more rainfall but, and most importantly, because of human adaptation to the desertification "situation." In fact, desertification is not caused by a lack of rain – but by numerous interrelated contributing factors from both natural and social systems, which are acknowledged in this definition:

> Land degradation in arid, semi-arid and dry sub-humid areas resulting mainly from negative human impacts combined with difficult climatic and environmental conditions. [5]

Introduction

Desertification occurs on every continent and affects more than two billion people who live in these areas [1]. It is caused by various factors, including natural dynamics, climatic condition, and human activities. Direct causes of desertification vary widely and include wind and water erosion, deforestation, overgrazing, land conversion to crop land, irrigation leading to salinization, and other physical stressors leading to soil loss, soil compaction, loss of vegetative cover, loss of biodiversity, and a degradation in ecosystem productivity. Indirect causes include climate variation, poverty, political instability, lack of education, or a combination of factors [1–7].

Desertification is estimated to affect about 1.9 billion hectares worldwide [6, 7]. These are inherently fragile ecosystems often close to the tipping point between continued production of ecosystem services [1] and a spiral of change leading to barren landscapes where people can no longer survive. Dryland regions tend to receive minimal attention from local or national agencies, and they are typically poorly provided for in terms of education, health, and infrastructure services. Dryland peoples are thus marginalized [8], going unnoticed unless war, natural disasters, or famine draw attention to them. Desertification and poverty are strongly linked [1], and this situation is exacerbated by the political instability of many dryland regions.

Faced with intensifying land degradation at the global scale, the United Nations Convention to Combat Desertification (UNCCD) was adopted in 1994, emphasizing sustainable development at the community level [9]. Regional and national assessments have been carried out (e.g., [10, 11]) and research addresses many issues. The UNCCD is committed to "bottom-up" action, recognition of traditional knowledge, and the importance of the roles of women in finding solutions, yet disagreement among scientists, policy makers, and communities still exist in many dryland areas hampering efforts to address this environmental, economic, and social problem.

Desertification and Human Systems

Changes in Production Systems

Desertification leads to diminished sustainability in a most basic manner: degraded lands suffer diminished capacity to meet the needs of present or future generations. Effects vary widely; the extent, severity, and impacts of desertification vary in both space and time, driven by pressures people put on dryland ecosystems combined with the intensity of aridity [1]. Therefore, dryland natural capital and the provision of goods and services – such as water, wood for fuel and building, and fodder for grazing – also vary.

People have developed a range of coping mechanisms in response to the natural fluctuations of ecosystems, which include nomadism, shifting cultivation, and surplus accumulation. As pointed out by Reynolds and Stafford Smith [12], dryland peoples are not the "problem" in desertification, nor are they "victims;" they are one part of an integrated system, and their responses to environmental change vary depending upon the severity, duration, and scale of the change. However, growth in both population and poverty may render previously effective coping mechanisms inadequate and result in increased vulnerability to hunger, disease, and political pressures. With two billion people involved, these issues are serious and are likely to become more so as the implications of climate change are factored in.

Long-term trends to more intense and longer droughts and overall drying have been observed in the Sahel, the Mediterranean, Southern Africa, North and South America, Australia, and areas throughout Asia over the period 1900–2005 [13] and 2001–2010 [14]. Climate change models for Africa indicate that temperatures are likely to increase from 0.2 to 0.5° C per decade [15], which will contribute to increasing rainfall variability by more than one standard deviation from normal in many areas [14] and will, in turn, significantly decrease perennial surface runoff. In the Western Cape of South Africa, for example, up to half the present perennial water supply is likely to be lost – even based on a relatively optimistic climate model used by de Wit and Stankiewicz [15]. The Intergovernmental Panel on Climate Change (IPCC) [14] found that decreases in runoff totaled about 17% during the 1990s, which indicates that the trend has already started. Furthermore, the IPCC states that model projections of changes in average surface temperatures show marked warming between 1986–2005 and 2081–2100 and changes in precipitation show increased precipitation between the two periods [14].

The IPCC [in [13, 14]] reports that reduced rainfall when coupled with temperature increases may not only lead to diminished surface water availability and less recharge but may also affect plants – to the point where current crop varieties may suffer reduced yields or not produce at all. IPCC Working Group 1 suggests that reduced rainfall in some areas juxtaposed with increased flooding in others may make rainfed agriculture, which is practiced in many dryland areas, a precarious undertaking [13, 14]. Rangeland, or natural pastures, upon which pastoralists depend for stock grazing are also likely to be affected with reductions in forage quantity and extent, and water points will become less reliable and productive.

These changes to natural resource viability will have, or are having, impacts beyond food shortages for dryland peoples. Competition for resources may weaken reciprocal arrangements which are inherent components of coping mechanisms – increasing vulnerability, the possibility of conflict and migration, and the failure of social institutions such as tenure and inheritance systems, markets, and subsidies [12].

Response Strategies

Vital for developing response strategies is the recognition of the complex interconnectedness of humans and their environment – called a social-ecological system [16, 17]. Response strategies and social resilience to environmental change have been shown to be partially dependent upon the range of response options available ("response diversity") – often at the individual household level [18]. In their analysis of building social resilience in arid ecosystems, Vogel and Smith [18] suggest that households in the most marginal of environments tend to possess less resilience – they lack options, wherewithal for making changes, backup resources, and competencies to adapt.

In Namibia, unpredictable precipitation and concomitant productivity can result in cycles of abundance and paucity of natural resources and dryland agriculture crops. In the good years everyone "banks" for the future, whether it is literally in monetary terms (as is the case for the wealthier individuals) or in terms of famine food stored for lean times (by poorer families and communities). The stored food option is adequate until the natural cycles are changed by a prolonged drought or climate variation, in which case options run out and hunger prevails.

Croppers and herders in Senegal have different levels of social resilience and decision-making options yet coexist and have developed some intergroup adaptive strategies [19]. In describing their Social Resilience Model, Bradley and Grainger [19] discuss both social and environmental resilience, concluding that desertification in their study region was less severe than might be expected due to human behavior [cf. 4]. By implication, adaptive strategies in the region are successful and maintain livelihoods despite climatic fluctuations, negative impacts due to misguided polices, and the inherent variability of ecosystem resilience in time and space.

The role of women in generating solutions is increasingly being recognized. In Senegal, the National Action Plan initiated as part of the country's ratification of the UNCCD established a national forum on the involvement of women. In Kenya, 30–50% of the participants in the formulation of the national action plan were women. Northeast Brazil is particularly active in promoting women's role in sustainable development [20]. Addressing gender inequalities in land tenure, inheritance, and decision making and recognizing the importance of women for dryland communities and ecosystems is a positive trend, which is gathering momentum in many countries [21].

Adaptation Systems and Improving Livelihoods

Most adaptation systems are of grass roots origin and local implementation, but this does not preclude their inclusion or performance within more formal organizational structures. National, Regional, and Subregional Action Programmes are key to implementing the UNCCD, and although such programs might appear to be "imposed" or "top-down," the UNCCD has recognized the importance of community level involvement, local knowledge, the role of women, and the value of synergistic activities between other UN programs [22]. The programs implemented by the UNCCD and many in-country NGOs operate at the local level and aim to improve livelihoods within the context of desertification, rather than (or sometimes as well as) instigating remediation.

In several areas of northwestern China, desertification has been reduced in recent years due to the application of new policies designed to change human activity and protect the environment. The "Grain for Green" and "Grazing Prohibition" policies are improving the environment, but other challenges remain – such as how to improve farmer's income and develop the economy. These questions were driving forces behind Zhou and Mouat's hypothetical study [23], which identified five uncertainties facing the Minqin region:

- Will there be sufficient water for agricultural and domestic use?
- What effect will climate change have?
- Is soil salinization likely to increase?
- What changes to the economy will occur, possibly as a result of government policy?
- Are land use patterns going to change?

Based upon the present land use status, one of the potential alternative futures for the region developed by Zhou and Mouat [23] focuses upon the development of high-tech industry in the city of Wuwei. As a result of this, the city expands somewhat into the agricultural area, but the rural population remains on the land while taking advantage of new opportunities in the city. A policy to import high water use commodities initiated by the administration in Wuwei, combined with mandated water conservation, results in lowered water use for that area – and agriculture downstream, around Minqin, improves. Although based upon a hypothetical study, such diversification is an example of successful social resilience, which, in conjunction with policy decisions, enhances environmental resilience and minimizes desertification [cf. 19].

While China and other countries have developed both policies and programs to "green" their deserts, these efforts have had mixed results and questionable sustainable outcomes. Planting trees in arid regions having less than 100 mm of annual precipitation will almost certainly be unsustainable, while other efforts that rely on appropriate agricultural/grazing strategies can have more sustainable outcomes [24].

Management strategies – for land, water, and human resources – are critical for human existence in all ecosystems, but particularly those such as drylands where system resilience is fragile. Traditional policies and organizations can be effective, for example, longstanding policies for managing wells in southern Ethiopia [25]. However, Emiru [25] reports that newly constructed water points are often administered poorly due to ineffective, newly initiated management committees.

An example of a relatively new management strategy that appears effective is the development of community-based conservation efforts called conservancies (e.g., [26]). Many countries in

Africa are adopting the conservancy approach, which is usually focused on promoting conservation of natural resources and wildlife, plus sustainable utilization and sharing of resources. These organizations are a diversification from the single-family approach to management, and recalling the arguments put forward by Bradley and Grainger [19], it is those who are able to maintain options and diversify who are most successful in living with desertification. Research in the field of resilience theory [16, 17, 27] suggests that resilience can be enhanced through pursuit of connectivity, collaborative capacity, response diversity, avoidance of false subsidies, enhancement of situational awareness, development of leadership capacity, increasing of reserves, and other strategies, all of which are available in varying degrees to those living in drylands.

In western Senegal there are several strategies which are adopted by the pastoralist and agricultural groups in times of reduced productivity [19]. What is particularly interesting is how the groups vary in their perceptions and adaptations. The pastoralists move between their grazing, watering, and trade sites on a seasonal and long distance basis, both as an anticipatory and a response strategy. During these migrations, social networks and use rights are maintained on a regular basis. Both agriculturalists and pastoralists accumulate surpluses. Agriculturalists expand the area under cultivation; pastoralists build up herds with the idea that at least some of the animals would survive a future drought.

Both groups have some diversity in livelihoods including collection of forest products, trade, and spiritualist activities in addition to cropping and/or raising livestock. However, Bradley and Grainger [19] report that the pastoralists diversified more easily and had a confident approach to this adaptive strategy. In response to changing environmental conditions, both groups maintained their customary production mode as long as possible, but, under pressure, pastoralists would increase their mobility, diversify stock type, or enter the gum arabic trade. However, the response by agriculturalists is typically for some of the group to migrate to urban areas to look for work. It would appear that the pastoralists adapt readily, are innovative in their solutions, and possess more autonomy, possibly as a result of the mobility and therefore flexibility inherent in their "normal" lifestyle.

Recognizing the multiple benefits of halting and reversing land degradation, the concept of "Zero Net Land Degradation" has been proposed [22]. The UNCCD defines land degradation neutrality (LDN) as "a state whereby the amount and quality of land resources necessary to support ecosystem functions and services and enhance food security remain stable or increase within specified temporal and spatial scales and ecosystems." The goal is to maintain or enhance the stocks of natural capital associated with land resources and the ecosystem services that flow from them. The objectives of LDN are to:

- Maintain or improve the sustainable delivery of ecosystem services
- Maintain or improve productivity, in order to enhance food security
- Increase resilience of the land and populations dependent on the land
- Seek synergies with other social, economic, and environmental objectives
- Reinforce responsible and inclusive governance of land

LDN was conceived to encourage a dual-pronged approach of measures to avoid or reduce land degradation combined with measures to rehabilitate already degraded land so as to achieve no net loss of healthy and productive land.

Discussion

Human adaptations which improve resilience – such as diversification and development of reciprocal networks (i.e., response diversity, connectivity, and collaborative capacity) – are likely to be more successful, for human and ecosystems, than abandoning the land and migrating to cities. The latter strategy increases the potential for desertification, adds to the problem of urban poor, and presents a security risk for many regions.

Populations living in drylands face challenges from climate change, decreased productivity and biological diversity, poverty, and changing social institutions, but Reynolds et al. [8] suggest a framework which can be used by managers and policy makers to address the complexity of interlinked systems and recognize what is important to change and where research can help. Based on the major characteristics of dryland social-ecological systems, they discuss how maintenance of local environmental knowledge is key to functional coadaptation in drylands – the implication is that accelerating integration of science and local environmental knowledge at local and regional levels could yield better outcomes.

Global climate change models operate at scales that do not permit a local level of analysis (although downscaling is improving), and given the great spatial variability of dryland ecosystems, there will be some areas which are less affected than others – even within the same ecosystem and climate regime [14]. This internal variability will impose additional social pressures upon dryland peoples, adding a further element of uncertainty to society, sustainability, and human and environmental security. It is imperative that desertification is thought of as a process, a problem, and a phenomenon that affects people, not just in those places directly involved but globally as well.

Future Directions

In this entry, the process of desertification, especially as a social-ecological system phenomenon integrating biophysical and human uses, has been described. As Reynolds and Stafford Smith have stated, people are neither the problem nor the victims; they are a part of an integrated system. Causes and effects have been discussed as have international efforts to "combat," mitigate, and prevent desertification (UNCCD). Finally, response strategies and adaptation systems are discussed.

What will be the complexion of efforts to deal with desertification as the twenty-first century progresses? Increasing populations and concomitant competition for resources including water and arable land will require strategies to ameliorate a likely worsening situation. These include understanding how to use water and other resources more efficiently; developing salt-, drought-, and pest-resistant crops, developing better systems for providing market information, lessons learned regarding best practices for drylands agriculture, developing local and green energy resources, and focusing on methods for enhancing resilience. As pointed out, at the local level people faced with increasing land degradation and decreasing livelihoods adapt, revolt, migrate, or die. International, national, and non-governmental organizations are increasingly striving to communicate success stories across communities. These same organizations, at the very least, are drawing awareness to the problems that people living in drylands face. They are also raising awareness of the importance of developing bottom-up approaches, involving women, and of incorporating traditional knowledge into solutions. Developing infrastructure will assist. While early warning systems for forecasting drought have been around for decades, similar methods for forecasting degradation are still only in a nascent state, although the UNCCD is striving to develop effective initiatives.

Increasing our understanding of the causes and processes of desertification is important, but so is the importance of figuring out how to use what is already known. It is important to ask questions about how drylands might change in the future including plausible alternatives. Assessing and analyzing the alternatives will allow one to pick pathways of development that will allow for positive futures.

In our view, it is important to remain focused on the sustainability of desert ecosystems, and it is essential to focus on the impacts of climate variation on human use systems in the context of sustainability. While much attention is focused on ameliorating and mitigating the effects of desertification, it is important to prevent drylands from becoming desertified in the first place.

Bibliography

1. MEA (Millennium Ecosystem Assessment) (2005) Ecosystems and human well-being: desertification synthesis. World Resources Institute, Washington, DC, p 25

2. FAO (2011) The state of the world's land and water resources for food and agriculture (SOLAW) – managing systems at risk. Food and Agriculture Organization of the United Nations/Earthscan, Rome/London
3. Aubreville A (1949) Climats, forest, et desertification de l'Afrique Tropicale. Societe de Editions Geographiques, Maritime et Coloniales, Paris, p 255
4. Dietz AJ, Ruben R, Verhagen A (eds) (2004) The impact of climate change on drylands: with a focus on West Africa, Environment and policy series, vol 39. Springer, New York, p 468
5. UNCED (United Nations Conference on Environment and Development) (1992) Earth summit agenda 21: programme of action for sustainable development. United Nations Department of Public Information, New York
6. UNCCD (2004) Preserving our common ground. UNCCD: 10 years on. Secretariat of the United Nations Convention to Combat Desertification. Available at: http://www.preventionweb.net/files/678_7539.pdf. Accessed Jul 2017
7. Low PS (ed) 2013 Economic and social impacts of desertification, land degradation and drought. White Paper I. UNCCD 2nd scientific conference, prepared with the contributions of an international group of scientists. Available at: http://2sc.unccd.int/fileadmin/unccd/upload/documents/WhitePapers/White_Paper_1.pdf
8. Reynolds JF, Stafford Smith DM, Lambin EF, Turner BL II, Mortimore MN, Batterbury SPJ, Downing TE, Dowlatabadi H, Fernandez RJ, Herrick JE, Huber-Sannwald E, Jiang H, Leemans R, Lynam T, Maestre FT, Ayarza M, Walker B (2007) Global desertification: building a science for dryland development. Science 316:847–851, and www.sciencemag.org. 11 May 2007
9. United Nations Convention to combat desertification in those countries experiencing serious drought and/or desertification, particularly in Africa 1994. https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=XXVII-10&chapter=27&clang=_en. Accessed 8 July 2017
10. European Commission Joint Research Center and UNEP (2015) World atlas of desertification, introductory brochure, p 16 (Atlas forthcoming)
11. Nwokocha CO (2017) An appraisal of the strategies implored by government to combating drought and desertification in the North-East geo-political zone, 2004-2014. Rev Pub Adm Manage 5:205. https://doi.org/10.4172/2315-7844.1000205
12. Reynolds JF, Stafford Smith DM (2002) Do humans cause deserts?, pp 1–21. In: Reynolds JF, Stafford Smith DM (eds) Global desertification: do humans cause deserts? Dahlem University Press, Berlin, p 437
13. CCDC (Commission on Climate Change and Development) (2008) Climate change and drylands. http://www.ccdcommission.org
14. IPCC (2014) Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change [Core Writing Team, Pachauri RK, Meyer LA (eds)]. IPCC, Geneva, p 151
15. de Wit M, J S (2006) Changes in surface water supply across Africa with predicted climate change. Science 311:1917–1921
16. Walker B, Salt D (2006) Resilience thinking: sustaining ecosystems and people in a changing world. Island Press, Washington, DC, pp 1–151
17. Walker B, Holling CS, Carpenter SR, Kinzig A (2004) Resilience, adaptability and transformability in social-ecological systems. Ecol Soc 9(2): Article 5. Available online: http://www.ecologyandsociety.org/vol9/iss2/art5/ Accessed July 2017
18. Vogel CH, Smith J (2002) Building social resilience in arid ecosystems, pp 149–165. In: Reynolds JF, Stafford Smith DM (eds) Global desertification: do humans cause deserts? Dahlem University Press, Berlin, p 437
19. Bradley D, Grainger A (2004) Social resilience as a controlling influence on desertification in Senegal. Land Degrad Dev 15:451–470
20. UNEP (2004) Women and desertification: a dynamic relationship. Chapter 4, in UNEP, 2004, Women and the environment, pp 49–59. http://www.unep.org/pdf/women/chapterfour.pdf
21. Hart SL (2010) Capitalism at the crossroads: next generation business strategies for a post-crisis world, 3rd edn. Pearson, Upper Saddle River, p 175
22. Orr BJ, Cowie AL, Castillo Sanchez VM, Chasek P, Crossman ND, Erlewein A, Louwagie G, Maron M, Metternicht GI, Minelli S, Tengberg AE, Walter S, Welton S (2017) Scientific conceptual framework for land degradation neutrality. A Report of the science-policy interface. United Nations Convention to Combat Desertification (UNCCD), Bonn
23. Zhou L, Mouat D (2007) Whose decisions impact land use change: the people or the government? Thoughts from Northwest China. Proceedings of the American Association for the Advancement of Science, San Francisco
24. Hannaway D, Mouat D Brewer L (2017) Greening the desert: strategies and tools for sustainability. Proceedings of the 6th Kubuqi International Desert Forum (KIDF), Ordos
25. Emiru N (2010) Role of traditional institutions in water resource governance in the Borana lowlands, southern Ethiopia. Haramata 55:13–15
26. Who is CANAM? http://www.canam.iway.na/home.htm. Accessed July 2017
27. Kerner DA, Thomas JS (2014) Resilience attributes of social-ecological systems: framing metrics for management. Resources 3:672–702. https://doi.org/10.3390/resources3040672

Geochemical Modeling in Environmental and Geological Studies

Chen Zhu
Department of Geological Sciences, Indiana University, Bloomington, IN, USA

Article Outline

Glossary

Activation energy The energy that must be overcome in order for a chemical reaction to occur.

Mass transport The net movement of mass from one location to another due to hydrological processes such as advection, dynamic dispersion, chemical reactions, and microbial activities.

Rate law A rate law is a statement about how the rate of a reaction depends on the concentrations of the participating species.

Solubility product Equilibrium constants for various kinds of reactions with a solid phase on one side and its constituent ions on the other.

Species A chemical entity distinguishable from other entities by molecular formula and structure, e.g., CO_2 and O_2 in a gas, and HCO_3^-, $H_2CO_3{}^{o}{}_{(aq)}$, $CO_3{}^{2-}$, $NaHCO_3{}^{o}{}_{(aq)}$.

Definition of the Subject and Its Importance

Geochemical modeling uses a set of mathematical expressions thought to represent chemical and transport processes in a particular geological system. The predictions of the model are partially observable or experimentally verifiable. Geochemical modeling has found applications in studies of chemical reactions in geological and environmental systems because of its utilities for synthesis of data, testing scenarios, and predicting long-term consequences of chemical reactions.

Introduction

Geochemical modeling is a powerful and indispensable tool for research and investigations of environmental sustainability science and technology. It allows quantitative evaluation of complex processes that often have feedback loops and it can predict the extent and consequences of geochemical reactions in the order of thousands to tens of thousands of years, beyond the range of laboratory experiments.

An excellent example of the utility of geochemical modeling in environmental sustainability science and engineering and its unique characteristics is illustrated by geological carbon sequestration – the injection of carbon dioxide (CO_2) into deep geological formations as a climate mitigation tool. Upon injection of CO_2 into a geological formation, CO_2 is dissolved into the native brine. The carbonated brine becomes acidic and corrosive, aggressively reacting with host rocks (e.g., a sandstone or carbonate rocks). Some primary (native) minerals are dissolved and secondary minerals precipitated. These dissolution-precipitation reactions can drastically change the porosity and permeability of the host rocks, and thereby impact the injectivity and storage safety in the long-term.

In the above example, there are many processes that are coupled. The flow of the separate supercritical CO_2 phase transports the CO_2 in the aquifer, determined by the viscosity, density, and the relative permeability of CO_2 in contrast to the brine. CO_2 dissolves into the brine when CO_2 makes contact with it. A brine with dissolved

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_202

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC, 2012
https://doi.org/10.1007/978-1-4419-0851-3_202

CO_2, probably in the form of HCO_3^-, has higher density than the rest. Density-driven vertical convection can occur, and convection brings about chemical gradients that result in more reactions.

The above example demonstrates that the system is complex and that the chemical reactions are coupled with each other and also coupled to transport processes, such as advection and dispersion. It is difficult to quantitatively evaluate these reactions without a computer model. There may be some ideas about what reactions will occur and how fast they will occur, but by developing a model, ideas can be formulated explicitly and thoughts can be tested quantitatively. Often, ideas are restricted to a particular aspect of chemical reactions, gleaned from specific laboratory experiments or field observations. Whether such ideas are valid when they are examined with measurable or observable consequences of the overall chemical systems is not known. In other words, whether the ideas of a subsystem hold up in the overall scheme of things must be tested. The system of concern often not only involves the coupling of hundreds of reactions, but also to processes like diffusion, advection, and dispersion, biological activities, thermal conduction, mechanical stress, and deformation.

The geological carbon sequestration example also illustrates the necessity of the prediction of chemical reactions for all practical purposes. Although federal or national regulations for safe underground CO_2 storage still need to emerge, it is reasonable to assume that national and local regulations will demand risk assessments of wellbore integrity, well injectivity, and long-term performance in the order of thousands of years. The geology of each injection site differs, and geological heterogeneities at a given site are a fact of life. It is not possible to conduct laboratory experiments either for each possible geochemical system or for durations of more than a few months. Performance assessments necessary at all stages of CO_2 storage operations (site assessment/selection, design, installation, operations and monitoring, and closure/post-closure) have to be partly based on geochemical modeling predictions.

Thus, it is clear that geochemical modeling is an indispensable tool in geochemical research and engineering. This entry introduces the basic concepts of geochemical modeling, provides information for accessing modeling codes and further readings, and shows applications in the field of environmental sustainability sciences and technology.

What Is Geochemical Modeling?

To modify Zhu and Anderson's [1] definition a little bit, a *geochemical model* is "an abstract object, described by a set of mathematical expressions thought to represent chemical and transport processes in a particular system. The predictions of the model are partially observable or experimentally verifiable." A geochemical model typically includes a *geochemical reaction network,* which means the finite array of reactions in a geochemical system and transport processes for a reactive transport system.

Mathematically, for a geochemical system that has n species, the following ordinary differential equations completely define the geochemical reaction network [2],

$$\frac{dm_i}{dt} = \sum_j \upsilon_{i,j} r_{i,j}, i \in n, \tag{1}$$

where m_i denotes the concentrations of ith species, t the time, υ_j the stoichiometric coefficient for ith species in the jth reaction, and $r_{i,j}$ the production or consumption rate of the ith species in the jth reaction. For a reactive transport system, the geochemical reaction network is defined by the transport equations,

$$\frac{\partial m_i}{\partial t} + L(C_i) = \sum_j v_j r_i \tag{2}$$

where L is the advection, dispersion, diffusion operator [3].

For historic development of geochemical modeling, the readers are referred to Zhu and Anderson [1] and Nordstrom [4]. To develop or apply a geochemical model, the modeler needs three parts: (1) a computer code that solves Eqs. 1 and/or 2; (2) a thermodynamic and kinetics database; (3) an input file that supplies the chemical analysis and the design or conceptual model of the geochemical model. Zhu and Anderson [1] called it "the part of three."

The Part of Three

Numerous computer programs for geochemical modeling have been developed. Without exception, the Newton-Raphson iteration method is used to solve the highly nonlinear Eq. 1. These computer codes include EQ3/6 [5], SOLMINEQ.88 [6], PHREEQC [7], and MINTEQA2 [8]. The codes that are sponsored by government agencies are essentially free of charge and are widely used. With ever more increasing computing power, tremendous advancements have been seen in code developments in the last decade. Zhu [9] pointed out that computer code developments are ahead of underlying science.

Most widely distributed computer codes come with a database of equilibrium constants for chemical reactions. For example, a database called DATA 0.DAT was initially developed for EQ3/6 by Tom Wolery [5] and Jim Johnson [10, 11]. This database was later adopted for the program PHREEQC as LLNL.DAT, GWB as THERMO.DAT, and the database in ToughReact. Similarly, thermodynamic databases have been developed for MINTEQA2 as MINTEQA2.DAT for PHREEQC as PHREEQC.DAT.

The compilations of equilibrium constants for chemical reactions draw from the available databases of standard state properties for minerals, such as those from Helgeson et al. [12], Wagman et al. [13], Berman [14, 15], Holland and Powell [16], Nordstrom et al. [17], and Robie and Hemingway [18]. More specialized databases are available for uranium [19]. Because these internally consistent databases only contain a limited number of minerals, while applications of geochemical modeling to a variety of geological and environmental topics require a wider range of minerals and solids, additional minerals are added to these equilibrium constant databases MINTEQA2 and PHREEQC.

For aqueous species, an internally consistent database developed by Harold Helgeson and collaborators, which includes a large number of aqueous species, has been widely used in the field of geochemistry [20–23]. In this database, the temperature and pressure dependences of thermodynamic properties for aqueous species were predicted using the parameters of the revised Helgeson–Kirkham–Flowers (HKF) equations of state for aqueous species [20, 24, 25]. Activity coefficients for the charged aqueous species were calculated from the extended Debye–Hückel equation or B-dot equation fitted to mean salt NaCl activity coefficients [26]. The computer program SUPCRT92 can be used to generate equilibrium constants at elevated temperatures and pressures [11]. Equilibrium constants have been calculated using the standard state properties from this database and compiled into databases that accompany the program EQ3/6 which was then adopted to other programs, e.g., LLNL.DAT in PHREEQC and THERMO.DAT in GWB©.

It has been known for quite some time that the internal consistency of thermodynamic data for modeling calculations is important. Nordstrom and Munoz [27] elaborated on the topic of internal consistency and readers are urged to consult their writing on the topic. Note that the standard state thermodynamic properties in the mineral databases [12, 14–17] are *internally consistent*, but some added minerals may not necessarily be so. The aqueous species collected in the Helgeson–Sverjensky–Shock compilations are internally consistent, but others collected may not be.

In the deep parts of the subsurface and in the shallow parts where evaporate sediment beds are located, groundwater becomes briny and can have concentrations of dissolved solids up to 300,000 mg/L. In dealing with concentrated solutions, Pitzer's ion interaction approach using virial specific interaction equations is generally preferred over the ion association theory when calculating ionic activities. The Pitzer's model, commonly the Harvie–Moller–Weare (HMW) formulation of it [28], has been incorporated into geochemical modeling codes EQ3/6, PHRQPITZ [29], and TOUGHREACT [30]. Although progress has been made in compiling the Pitzer interaction parameters [31], the lack of Pitzer's activity coefficient parameters at elevated temperatures and for minor or trace elements remains a barrier to accurate calculation of solubility and saturation indices for highly saline fluids.

The input file is where the modeler defines the composition of the chemical system, includes or excludes certain types of chemical reactions, and assigns the boundary and initial conditions. In other words, this is where the modeler translates the conceptual model into recognizable formats by computer programs. The modeler also develops the conceptual model in this way [1].

From the discussion of each of the "part of three," it should be clear that a computer program should always be distinguished from a geochemical model. For example, PHREEQC is a geochemical modeling computer program or code; it is not a geochemical model. It is always true that the modeler, not the computer code, produces a geochemical model. A geochemical modeling code is a utility, e.g., an oven. The oven does not bake the cake. The chef bakes the cake.

Speciation–Solubility Modeling

Speciation modeling calculates the distribution of aqueous species and mineral saturation indices according to the solutions of the mass and charge balance and mass action equations. From the activities of the aqueous species, one can calculate Saturation Indices (SI),

$$SI = \log\left(\frac{IAP}{K}\right) \tag{3}$$

where K stands for the equilibrium constant of the dissolution reaction and IAP stands for the Ion Activity Product. When $SI = 0$, the mineral is at equilibrium with the aqueous solution. When $SI < 0$, the aqueous solution is undersaturated with respect to the mineral of concern and the mineral will dissolve. When $SI > 0$, the aqueous solution is supersaturated with respect to the mineral and the mineral will precipitate.

Speciation-solubility modeling has become a routine exercise since Garrels and Thompson [32] first calculated the aqueous speciation in seawater and saturation states with respect to mineral solubility. The principles are well known, and the numerical modeling techniques and their tweaking are mature. There are hundreds of computer codes available for this kind of calculation.

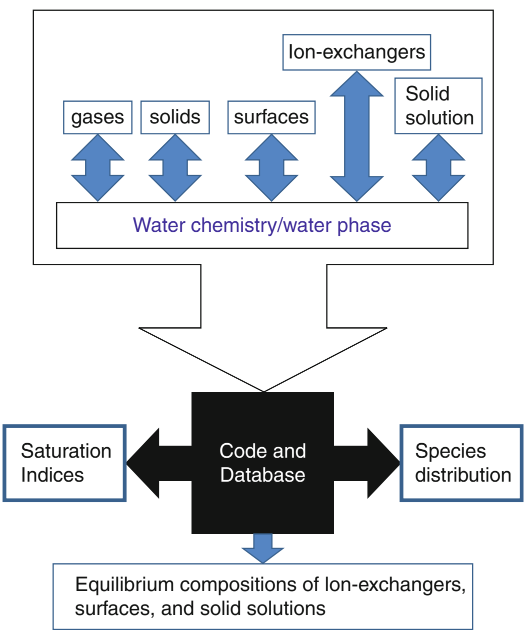
The results of speciation and solubility calculations are used for a number of issues in environmental sustainability science and technology. The toxicity of many chemicals is not only related to the total concentrations, but also to the species. Liu et al. [33] measured the total concentrations of antimony (Sb) with valence of V. They then used speciation calculations and found the dominant aqueous Sb species was $Sb(OH)_6^-$. In general, it is assumed that aqueous species in the solution are in mutual equilibrium (homogeneous equilibria). The exception to this rule is redox species, which are well known for not being at equilibrium in surficial water bodies [34, 35].

The calculated saturation indices give directions of chemical reactions. As elaborated in Zhu and Anderson [1] and in numerous books on thermodynamics, saturation indices show the direction of the chemical reactions, as dictated by the second law of thermodynamics, but tell nothing about the rate of reactions. For example, it is well known in geochemistry that natural waters are often supersaturated with respect to crystalline quartz, but the rate of quartz precipitation is too slow for these waters to reach equilibrium with quartz even after thousands of years at low temperatures. However, reaction direction is the first thing that one must know before one can extract reaction rates (one needs to know which direction the reactions will go before one can estimate how fast or slow they will do it).

One could extend the scope of traditional speciation– solubility modeling to include the equilibrium partitioning of a chemical between the aqueous solution and mineral surfaces (Fig. 1). Many computer codes (e.g., MINTEQA2 and PHREEQC) now allow the calculations of surface adsorption according to the surface complexation theory [35–37]. One can find the details of the surface complexation theories in the textbooks cited above. In terms of modeling, when the total concentration of an ion (e.g., Pb^{2+}) in an aqueous solution is given, the codes can calculate how much Pb^{2+} is partitioned onto the mineral surface(s), how much Pb^{2+} remains in the aqueous solution, what the relative percentage of surface bound Pb^{2+} among different surface species is, and what the dominant aqueous Pb^{2+} species are.

Similarly, one can calculate the equilibrium partitioning of chemicals between an aqueous solution and ion-exchangers. Unlike the surface adsorption and the surface complexation theory that describes it, modeling of ion-exchange reactions lacks both theoretical footing and internally consistent databases [1]. Readers are encouraged to read more in Appelo and Postma [38]. Equally, one can also calculate the equilibrium distribution between an aqueous solution and solid solution phases [39, 40], and gas phases (Fig. 1).

Speciation–solubility modeling provides a "snapshot" of a dynamic system, and the basic building block for more advanced process modeling. The calculated activities of the various ionic and molecular species give the *IAP* for the saturation state evaluation. This type of calculation represents the majority of applications of geochemical modeling to the field of environmental sustainability science and technology. However, despite the maturity of the modeling techniques, a number of challenges exist which make the evaluation of the reaction directions a nontrivial task (cf. [9]).



Geochemical Modeling in Environmental and Geological Studies, Fig. 1 Schematic configuration of speciation–solubility models

Reaction Path Modeling

Once the results of aqueous speciation and solubility calculations are obtained (i.e., calculated species distribution and saturation indices) for a given temperature, pressure, and instance of time, processes can be modeled. The simplest next step is *reaction path modeling*, tracing the evolution of aqueous solution composition and speciation and mineral paragenesis through time or the reaction progress as a result of irreversible reactions (e.g., feldspar dissolution) or processes (e.g., titration, mixing, or increase or decrease of temperature or pressure). The modeling is accomplished by applying the principle of mass balance, thermodynamics that govern the equilibrium between species, and kinetics that govern the rate of mass transfer among phases. The concept and mathematical foundation of reaction path modeling was introduced to geochemistry by Harold Helgeson [41]. Numerous articles and books have described this approach [2, 5, 42, 43]. Computer codes EQ3/6, PHREEQC, MINTEQA2, SOLMINEQ.88 and GWB© can all perform these kinds of calculations.

In the past a few years, advancement has been seen in computing powers and development of computer programs, which make it possible to perform reaction path modeling involving complicated reaction networks. This includes incorporation of various forms of rate laws and inclusion of an almost unlimited number of reactions into a single model. Conceptual developments in geochemical reaction networks now allow one to explore the intricacies of feedback mechanisms for complex inorganic geochemical systems [44].

One area that has been rapidly advanced in recent years is the linking of microbial activities with a network of redox and non-redox reactions and the exploration of the complex feedback loops in biogeo-chemistry. For example, Istok et al. [45] carried out a *biogeochemical reaction path modeling* for simulating an in situ field experiment which investigated the bioreduction of uranium near Oak Ridge National Laboratory in Tennessee, USA. In the field experiment, ethanol was injected into the aquifer to stimulate microbial activities at the field site and the reduction of nitrate, U(VI), Fe(III), Mn(IV), and sulfate were observed to proceed concomitantly.

Historically, microbial mediated reactions were modeled as semiempirical kinetics, decoupled into a whole suit of inorganic geochemical reactions as a consequence of biostimulation or bioremediation [45, 46]. The Monod-type kinetic expressions are used to describe rates of substrate utilization and biomass production. Istok et al. [45] developed a new approach, dubbed as "thermodynamically based." In their approach, the actual microbial community is represented by a synthetic microbial community consisting of a collection of microbial groups, each with a unique growth equation that couples a specific pair of energy yielding redox reactions. Simulations monitored temporal changes of microbial biomass, community composition, aqueous speciation, and oxidation states of multivalent chemicals, as well as the dissolution and precipitation of minerals.

Istok et al.'s [45] modeling results are shown in Fig. 2. Simulations predicted that acetate addition will result in the growth of only 8 of the 25 microbial groups during the experiment. The increasing biomass and changing community composition occurred as increasing amounts of acetate were reacted. Early on in the reaction path, predicted biomass increase was dominated by predicted growth of manganese reducers, iron reducers, and denitrifiers, with much smaller predicted growth of other groups, reflecting the relatively larger initial amounts of Mn(IV)- and Fe(III)-bearing minerals and nitrate compared to oxygen and sulfate. As additional acetate was reacted, growth of other groups was predicted to become energetically favorable, especially sulfate reducers. Predicted patterns of the growth of the various groups resulted in predicted changes in community composition. The results are generally consistent with clone libraries developed from groundwater samples.

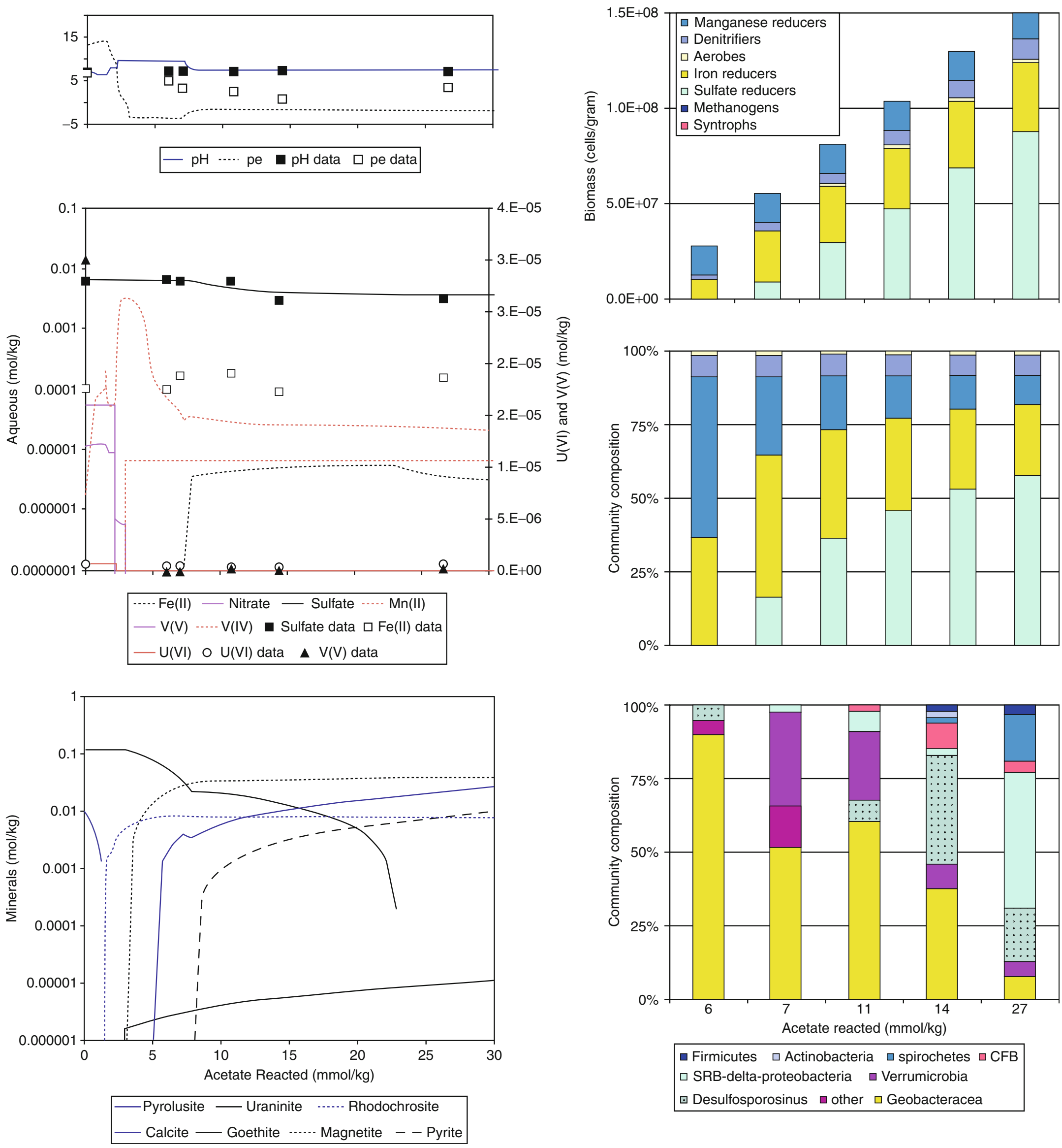
While some computer software is fully capable of performing such complex computations, the requirements of modeling parameters for the biogeochemical reactions are daunting [48]. Numerous nonunique interpretations of field data may be possible. However, geochemical modeling can help to integrate processes and reactions in a quantitative manner and provide a more comprehensive understanding of biogeochem-ical processes than simply providing apparent zeroth-order and first-order rates.

Coupled Reactive Mass Transport Modeling

Geochemical models included in this category all solve the advection-dispersion-reaction (ADR) mass transport equations, typically using the sequential iteration approach. The reaction term is fully coupled to chemical equilibrium and kinetics (and microbial activities). In other words, at each time step and at each node or in each grid cell, reaction path calculations described above are performed. Many transport codes on the market are termed "reactive transport model," but only the partitioning coefficient or K_D approach is used. In the K_D approach, all chemical reactions pertinent to a chemical are described by a single parameter [49]. In this entry, only multiple component mass transport models with the nonlinear mass balance equations for speciation and mass transfer coupled to the ADR equation are called coupled reactive mass transport (CRMT) models. For the fundamentals of the subject, readers are referred to Yeh and Tripathi [50].

Some codes in this category are coupled to fluid flow and also have the capabilities of simulating multiphase flow (air, water, and CO_2), density-dependent flow, the feedback of dissolution – precipitation reactions and changes of permeability and flow patterns – and thermal and mechanical stress. Applications of CRMT include fate and transport of metal and radionuclides in groundwater systems, geologic carbon sequestration, sediment diagenesis, geothermal energy exploration and production, sea water intrusion, and formation of ore deposits.

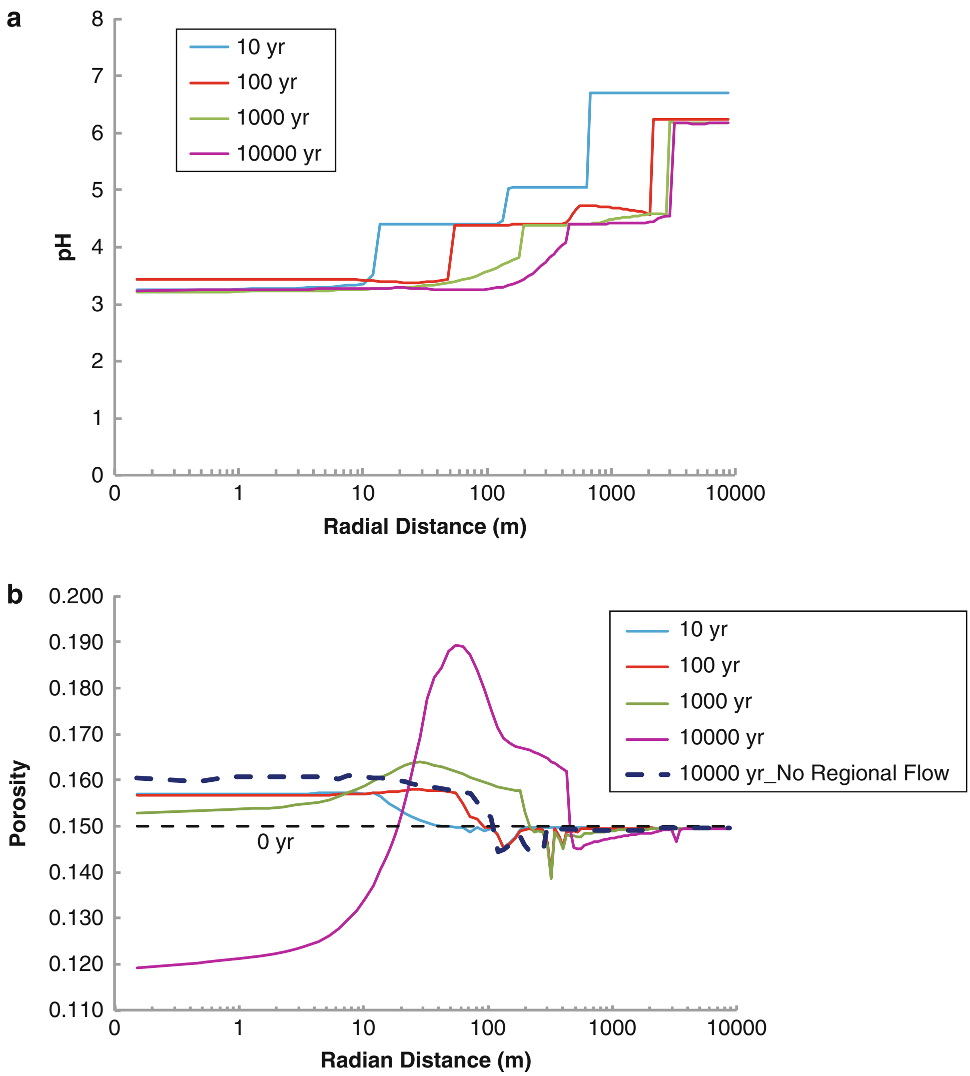
As an example, Liu et al. [51] used multiphase reactive flow and transport modeling to simulate large-scale CO_2 injection (a million tons per year for 100 years) into Mt. Simon sandstone, a major candidate saline reservoir in the Midwest of USA. The long-term fate of CO_2 was simulated by extending the modeling period to 10,000 years (the predictive utility of geochemical modeling). The results indicate that most of the injected CO_2 remains



Geochemical Modeling in Environmental and Geological Studies, Fig. 2 Comparison of flush simulations with geochemical and clone library data from field natural gradient experiment of Anderson et al. [47]. Figure 3 of Istok et al. [45]

within a radius of 3,300 m lateral distribution. Four major trapping mechanisms and their spatial and temporal variations are evaluated in the simulations: hydrodynamic, solubility, residual, and mineral trapping. A strongly acidified zone (pH 3–5) forms in the areas affected by the injected CO_2(0–3,300 m), and consequently causes extensive mineral precipitation and dissolution (Fig. 3).

Coupled reactive mass transport models are great tools to further understand geochemical reaction networks. Typically, CRMT modeling generates a great amount of numerical experimental data. Significant amounts of time and energy are necessary to dissect and distill the information on what reactions have happened and how reactions are coupled.



Geochemical Modeling in Environmental and Geological Studies, Fig. 3 (**a**) Simulated pH variations as a function of radial distance at year 10, 100, 1,000, and 10,000 with regional flow rate of 0.3 m/year. The initial pore fluids had a pH of 6.9. (**b**) Simulated porosity variations as a function of radial distance at year 10, 100, 1,000, and 10,000 with a regional flow rate of 0.3 m/year, compared at year 10,000 with no regional flow

Future Directions

There is no doubt that geochemical modeling plays an increasingly important role in environmental sustain-ability sciences and technology, as environmental decisions (carbon dioxide storage, remediation) are made partly based on model predictions. Computer code development has given flexibility to most applications. However, modelers need to have a grasp of the fundamental understanding of the underlying sciences and the limitations of geochemical models.

Zhu [9] gave a list of research needs to model geochemical reactions:

1. Internally consistent standard state thermodynamic properties for minerals, particularly minerals with complex and variable chemical compositions and structures like smectite. More accurate thermodynamic properties for common

minerals like feldspars would help to resolve the controversy on the Al-bearing minerals.
2. More experimental data and resolution of ambiguities surrounding the speciation of aqueous elements like that for Al species.
3. Improvements in sampling and filtration of natural and laboratory samples for better saturation state assessments.
4. Solid solution models for feldspar and clay minerals.
5. More measurements of rate – free energy of reaction relations at a variety of temperature and pH conditions, leading to accurate theoretical or empirical correlations.
6. Rates and rate laws for precipitation reactions, and improved understanding of nucleation process.
7. Pitzer activity coefficient parameters for trace elements and for all elements at elevated temperatures.
8. More rigorous treatment of experimental data with statistical analysis.
9. Assessment of error propagations.

Acknowledgments The writing of this entry was also made possible with continued financial support from the US National Science Foundation (EAR0423971, EAR0509775, EAR 0809903) and the US Department of Energy (DE-FG26-04NT42125, DE-FE0004381). Any opinions, findings, and conclusions or recommendations expressed in this material, however, are those of the authors and do not necessarily reflect the views of the US Government or any agency thereof.

Bibliography

Primary Literature

1. Zhu C, Anderson GM (2002) Environmental applications of geochemical modeling. Cambridge University Press, London, p 304
2. Helgeson HC et al (1970) Calculation of mass transfer in geo-chemical processes involving aqueous solutions. Geochim Cosmochim Acta 34:569–592
3. Fang YL, Yeh GT, Burgos WD (2003) A general paradigm to model reaction-based biogeochemical processes in batch systems. Water Resour Res 39(4):1083
4. Nordstrom DK (2007) Modeling low-temperature geochemical processes. In: Drever JI (ed) Surface and ground water, weathering and soils, treatise on geochemistry. Elsevier, New York, pp 1–38, online update
5. Wolery TJ (1992) EQ 3/6, A software package for geochemical modeling of aqueous systems: package overview and installation guide (version 7.0). URCL-MA-110662-PT-I, University of California/Lawrence Livermore Laboratory, Livermore, p 41
6. Kharaka YK et al (1988) SOLMINEQ.88: a computer program for geochemical modeling of water-rock interactions. Water-resources investigations report 88–4227, US Geological Survey
7. Parkhurst DL, Appello AAJ (1999) User's guide to PHREEQC (version 2)-a computer program for speciation, batch-reaction, one dimensional transport, and inverse geochemical modeling. Water-resource investigation report, US Geological Survey, p 312
8. Allison JD, Brown DS, Novo-Gradac KJ (1991) MINTEQA2/PRODEFA2, a geochemical assessment model for environmental systems, version 3.0 user's manual
9. Zhu C (2009) Geochemical modeling of reaction paths and geochemical reaction networks. In: Oelkers EH, Schott J (eds) Thermodynamics and kinetics of water-rock interaction. Mineralogical Society of America, Washington, pp 533–569
10. Johnson JW, Lundeen SR (1994) GEMBOCHS thermodynamic data files for use with the EQ 3/6 software package. Lawrence Livermore National Laboratory, p 99
11. Johnson JW, Oelkers EH, Helgeson HC (1992) SUPCRT92 – a software package for calculating the standard molal ther-modynamic properties of minerals, gases, aqueous species, and reactions from 1-bar to 5000-bar and 0°C to 1000°C. Comput Geosci 18(7):899–947
12. Helgeson HC et al (1978) Summary and critique of the ther-modynamic properties of rock forming minerals. Am J Sci 278A:569–592
13. Wagman DD et al (1982) The NBS tables of chemical thermo-dynamic properties – selected values for inorganic and C-1 and C-2 organic-substances in SI units. J Phys Chem Ref Data 11(Supplement 2):392
14. Berman RG (1988) Internally-consistent thermodynamic data for minerals in the system Na2O-K_2O-CaO-MgO-FeO-Fe_2O3-Al_2O3-SiO2-TiO2-H_2O-CO2. J Petrol 29(2):445–522
15. Berman RG (1990) Mixing properties of Ca-Mg-Fe-Mn garnets. Am Mineral 75:328–344
16. Holland TJB, Powell R (1998) An internally consistent thermodynamic data set for phases of petrological interest. J Metamorph Geol 16:309–343
17. Nordstrom DKetal (1990) Revised chemical equilibrium data for major water-mineral reactions and their limitations. In: Melchior DC, Bassett RL (eds) Chemical modeling of aqueous systems II. American Chemical Society, Washington, pp 398–413
18. Robie RA, Hemingway BS (1995) Thermodynamic properties of minerals and related substances at 298.15 K and 1 bar (10^5pascals) pressure and at higher temperatures. US Geological Survey Bulletin 2131, p 456
19. Grenthe I et al (1992) The chemical thermodynamics of uranium. Elsevier, New York
20. Helgeson HC, Kirkham DH, Flowers GC (1981) Theoretical prediction of the thermodynamic behavior of aqueous electrolytes at high pressures and temperatures. IV. Calculation of activity coefficients, osmotic coefficients, and apparent molal and standard and relative partial molal properties to 600°C and 5 kb. Am J Sci 281:1249–1516

21. Shock EL, Helgeson HC (1988) Calculation of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: correlation algorithms for ionic species and equation of state predictions to 5 kb and 1000°C. Geochim Cosmochim Acta 52:2009–2036
22. Shock EL, Helgeson HC, Sverjensky DA (1989) Calculations of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: standard partial molal properties of inorganic neutral species. Geochim Cosmochim Acta 53:2157–2183
23. Sverjensky DA, Shock EL, Helgeson HC (1997) Prediction of the thermodynamic properties of aqueous metal complexes to 1000°C and 5 kb. Geochim Cosmochim Acta 61(7):1359–1412
24. Shock EL et al (1992) Calculation of thermodynamic and transport properties of aqueous species at high pressures and temperatures. Effective electrostatic radii, dissociation constants and standard partial molal properties to 1000°C and 5 kb. J Chem Soc London, Faraday Trans 88: 803–826
25. Tanger JC, Helgeson HC (1988) Calculations of the thermodynamic and transport properties of aqueous species at high pressures and temperatures: revised equation of state for the standard partial molal properties of ions and electrolytes. Am J Sci 288:19–98
26. Oelkers EH, Helgeson HC (1990) Triple-ion anions and polynuclear complexing in supercritical electrolyte-solutions. Geochim Cosmochim Acta 54(3):727–738
27. Nordstrom DK, Munoz JL (1994) Geochemical thermodynamics, 2nd edn. Blackwell, Oxford
28. Harvie CE, Moller N, Weare JH (1984) The predication of mineral solubilities in natural waters: the Na-K-Mg-Ca-H-Cl-SO_4-OH-HCO_3-CO_3-CO_2-H_2O system to high ionic strength at 25°C. Geochim Cosmochim Acta 48(4):723–751
29. Plummer LN et al (1988) A computer program incorporating Pitzer's equations for calculation of geochemical reactions in brines. Water resources investigations report 88–4153, US Geological Survey, p 310
30. Xu T et al (2004) TOUGHREACT user's guide: a simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media (V1.2). Paper LBNL-55460. Lawrence Berkeley National Laboratory
31. Wolery T et al (2004) Pitzer database development: description of the Pitzer geochemical thermodynamic database data 0.ypf. Appendix I in In-Drift precipitates/salts model (P. Mariner) report ANL-EBS-MD-000045 REV 02. Bechtel SAIC Company, Las Vegas
32. Garrels RM, Thompson ME (1962) A chemical model for sea water at 25°C and one atmospheric pressure. Am J Sci 260:57–66
33. Liu FY et al (2010) Antimony speciation and contamination of waters in Xikuangshan Sb mining and smelting area, China. Environ Geochem Health. https://doi.org/10.1007/s10653-010-9284-z
34. Lindberg RD, Runnells DD (1984) Groundwater redox reactions – an analysis of equilibrium state applied to Eh measurements and geochemical modeling. Science 225(4665):925–927
35. Stumm W, Morgan JJ (1996) Aquatic chemistry, chemical equilibria and rates in natural waters. Wiley, New York, p 1022
36. Stumm W (1992) Chemistry of solid-water interfaces: processes at the mineral-water and particle-water interface in natural systems, 1st edn. Wiley, New York
37. Dzombak DD, Morel FMM (1990) Surface complex modeling: hydrous ferric oxide. Wiley, New York, 393
38. Appelo CAJ, Postma D (2005) Geochemistry, groundwater and pollution. A. A. Balkema, Leiden
39. Zhu C (2004) Coprecipitation in the barite isostructural family: 1. Binary mixing properties. Geochim Cosmochim Acta, 68(16):3327–3337
40. Zhu C (2004) Coprecipitation in the barite isostructural family: 2. Binary mixing properties. Geochim Cosmochim Acta, 68(16):3339–3349
41. Helgeson HC (1968) Evaluation of irreversible reactions in geochemical processes involving minerals and aqueous solutions-1. Thermodynamic relations. Geochim Cosmochim Acta 32:853–877
42. Helgeson HC (1979) Mass transfer among minerals and hydro-thermal solutions. In: Barnes HL (ed) Geochemistry of hydrother-mal ore deposits. Wiley, New York, pp 568–610
43. Anderson GM, Crerar DA (1993) Thermodynamics in geochemistry: the equilibrium model. Oxford University Press, New York, 588
44. Zhu C et al (2010) Coupled alkali feldspar dissolution and secondary mineral precipitation in batch systems: 4. Numerical modeling of kinetic reaction paths. Geochim Cosmochim Acta 74(14):3963–3983
45. Istok JD et al (2010) A thermodynamically-based model for predicting microbial growth and community composition coupled to system geochemistry: application to uranium bioreduction. J Contam Hydrol 112(1–4):1–14
46. Liu C et al (2001) Kinetic analysis of the bacterial reduction of goethite. Environ Sci Technol 35(12):2482–2490
47. Anderson TT, Vrionis HA, Ortiz-Bernard I, Resch CT, Long PE, Dayvault R, Karp K, Marutzky S, Metzler DR, Peacock A, White DC, Lowe M, Lovley DR (2003) Stimulating the in situ activity Geobacter species to remove uranium from the groundwater of a uranium-contaminated aquifer. Appl Environ Microb 69:5884–5891
48. Roden EE (2008) Microbiological controls on geochemical kinetics 1: fundamentals and case study on microbial Fe(III) oxide reduction. In: Brantley SL, Kubicki J, White AF (eds) Kinetics of water-rock interaction. Springer, New York, pp 335–415
49. Zhu C (2003) A case against Kd-based transport model: natural attenuation at a mill tailings site. Comput Geosci 29:351–359
50. Yeh GT, Tripathi VS (1989) A critical evaluation of recent development of hydrogeochemical transport models of reactive multi-components. Water Resour Res 25(1):93–108
51. Liu FY et al (2010) Coupled reactive transport modeling of CO_2 sequestration in the Mt. Simon sandstone formation, Midwest U.S.A. Int J Greenh Gas Con 5:294–307

Geologic Carbon Sequestration: Sustainability and Environmental Risk

Curtis M. Oldenburg
Energy Geosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Article Outline

Glossary

Carbon dioxide capture and storage (CCS) The capture and compression of CO_2 from fossil fuel power plants and other industrial point sources followed by its transport to wells for injection into deep geologic formations for permanent storage.

Carbon dioxide capture, utilization, and storage (CCUS) The capture and compression of CO_2 from fossil fuel power plants and other industrial point sources followed by its beneficial utilization, most commonly for injection into mature oil fields for enhanced oil recovery, during which process the CO_2 is eventually permanently stored.

Consequence An impact arising from the occurrence of an event or process. For example, the consequence of high CO_2 concentrations in the atmosphere is global warming.

Geologic carbon sequestration (GCS) = Geologic CO_2 storage (GCS) The last step of CCS in which CO_2 is injected through wells into deep subsurface formations for permanent storage.

Hazard A potential impact or consequence of an event or process. For example, CO_2 emissions were first recognized as a hazard to the climate by John Tyndall in the mid-nineteenth century.

Likelihood The probability or degree of potential for an event or process to occur. For example, the likelihood of large CO_2 emissions continuing is very high given population growth and worldwide increases in standard of living.

Risk The product of likelihood and consequence of an event or process. For example, the risk of climate change is very high because both the likelihood of continued CO_2 emissions and the consequences of elevated atmospheric CO_2 concentrations are high.

Storage reserve (capacity) Effective pore-space volume available for CO_2 storage given existing or projected economic, technological, legal, environmental, and regulatory factors.

Storage resource (capacity) Physical pore-space volume available for CO_2 storage independent of economics, extraction technology, or regulations.

Definition of Subject and Its Importance

Carbon dioxide (CO_2) capture and storage (CCS) is a combination of technologies that reduces the risk of climate change by directly reducing the net CO_2 emissions arising from the use of fossil fuels as the main global primary energy source [1]. In CCS, as commonly envisioned, CO_2 will be captured and compressed from flue gases at point sources such as coal-fired power plants, transported by pipeline, and injected into deep geologic formations on- or offshore for permanent storage (i.e., geologic sequestration)

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_200

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC (outside the USA) 2018, https://doi.org/10.1007/978-1-4939-2493-6_200-3

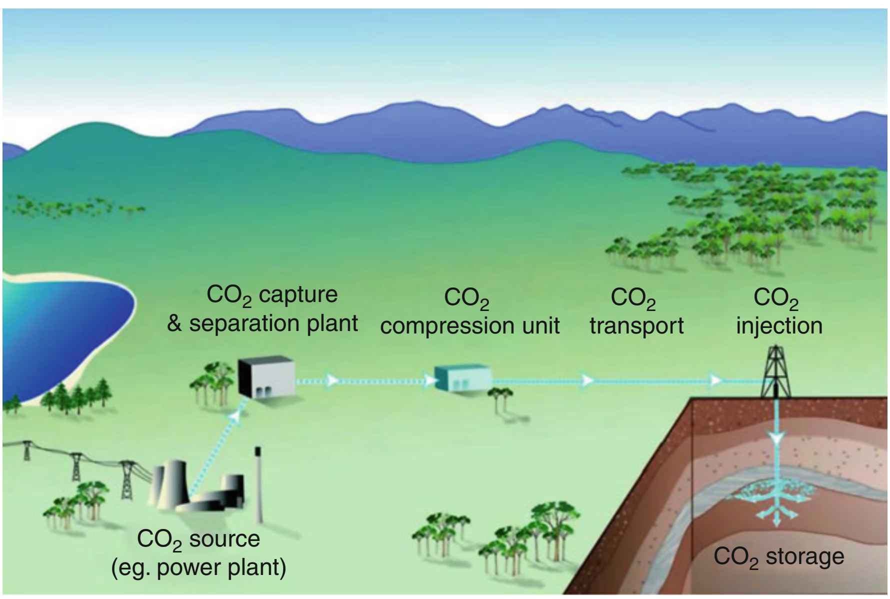
(Fig. 1). Note that in this context, *permanent* is loosely defined as several millennia, the idea being that if the majority of the injected CO_2 is prevented from entering the atmosphere over this time frame or longer, the global warming impacts of the CO_2 are avoided. Carbon dioxide (CO_2) capture, utilization, and storage (CCUS) adds a beneficial utilization of captured CO_2 that assures non-emission to the atmosphere. Because the subject presented here is geologic carbon sequestration (GCS), we consider only the large-scale subsurface uses of CO_2 as relevant utilization options, and we consider only the most common of these, CO_2 for enhanced oil recovery [2], which carries with it most of the same containment concerns as CCS, making separate discussion of CCUS here unnecessary.

Processes for capturing of CO_2 vary depending upon whether hydrocarbons are burned with air or pure oxygen. Capture from the former involves either the use of liquid sorbents, solid membranes, or other engineered materials that can extract CO_2 from a mixture of gases resulting from power generation or other industrial processes in which CO_2 is often a minor component at relatively low pressure. Capture after combustion with pure oxygen more simply involves separating the CO_2 from water vapor by condensation. However this comes at the cost of building and operating a facility to separate pure oxygen from air.

As of 2017, there are only two large (>800,000 tonnes/year) capture operations related to power generation in existence, but there are another dozen large capture facilities operating on other industrial plants. Half of these are related to natural gas processing (see next paragraph). Another two are related to hydrogen and one to biofuel production facilities.

CO_2 occurs in many natural gas (methane, CH_4) fields at concentrations above those required for delivery to customers. Natural gas processing to remove CO_2 has been carried out for decades, and at several gas fields, captured CO_2 is reinjected for geologic sequestration [3, 4].

In power plant capture, extraction of CO_2 can be done after combustion, so-called post-combustion capture, or during pre-combustion steps, which has the advantage of higher pressures and higher CO_2 concentrations than capture after combustion with air [1]. Higher pressures and



Geologic Carbon Sequestration: Sustainability and Environmental Risk, Fig. 1 Schematic of an onshore carbon dioxide capture and storage (CCS) process (CO2CRC, http://www.co2crc.com.au/aboutccs/)

higher concentrations lower the energy needed to effect capture, e.g., using membranes or adsorption approaches. Capture after combustion with pure oxygen has the benefit of relatively pure CO_2 exhaust with water vapor, which is readily separated by condensation.

There are also direct-air capture approaches for CCS that make use of solid sorbents to capture CO_2 from ambient air rather than specifically at point sources [e.g., 5, 6]. Direct-air capture has the advantage that it addresses emissions from all sources, including mobile CO_2 sources such as automobiles and trucks, but also the disadvantages of much lower CO_2 concentration and pressure, which results in a much higher cost per tonne captured.

Biofuels can be produced by fermentation of the biomass from crops that have taken carbon out of the atmosphere during their growth. Carbon dioxide capture during fermentation is attractive because fermentation produces nearly pure CO_2 making capture straightforward, and subsidies for using biofuels are often correlated with CO_2 emission reduction via use of plant matter as a feedstock for biofuel production. Additional reductions in carbon intensity beyond uptake of carbon during growth can be realized by capturing CO_2 from fermentation and injecting it for geologic sequestration. This type of process is called BECCS (BioEnergy with CCS), and the only large plant of this kind recently began operations in the USA [7, 8].

Regardless of how CO_2 capture is accomplished, the process must be capable of providing a stream of CO_2 for compression and subsequent transport to sequestration sites. Although direct injection of CO_2 into the deep oceans has received a large amount of attention [e.g., 9], and numerous processes to accelerate uptake of atmospheric CO_2 by the oceans have been discussed [e.g., 10, 11, 12], concerns about permanence and impact to marine ecosystems are larger for ocean sequestration than for geologic sequestration [e.g., 13]. This leaves geologic carbon sequestration (GCS) as the main approach under consideration for isolating the vast quantities of CO_2 from the atmosphere needed to avert catastrophic climate change. It is important to note that GCS can be carried out both onshore and offshore. The most famous and long-lived GCS project (~20 years) is the Sleipner project located off the coast of Norway [2]. There is a growing interest in the USA to investigate offshore options [14].

The extra expense involved in capturing, transporting, and injecting CO_2 in the CCS or CCUS process can be expressed in terms of an energy penalty, i.e., the amount of energy that must be expended above business-as-usual fossil fuel energy use. Estimates of the energy penalty for CCS vary over a wide range depending on combustion process, age of facility, distance to geologic storage site, etc., but are likely around 40% [1, 15]. Whether stated in terms of dollars or energy penalty, the largest expense in CCS and CCUS is capture (which also includes compression) currently projected to account for more than 60% of the cost of CCS [16].

Introduction

Fossil fuels are abundant, inexpensive to produce, and easily converted to usable energy by combustion as demonstrated by humanity's dependence on fossil fuels for over 80% of its primary energy supply [17]. This reliance on fossil fuels comes with the cost of carbon dioxide (CO_2) emissions that exceed the rate at which CO_2 can be absorbed by terrestrial and oceanic systems worldwide resulting in increases in atmospheric CO_2 concentration as recorded by direct measurements over more than five decades [18]. Carbon dioxide is the main greenhouse gas linked to global warming and associated climate change, the impacts of which are currently being observed around the world, and projections of which include alarming consequences such as water and food shortages, sea level rise, and social disruptions associated with resource scarcity [19]. The current situation of a world that derives the bulk of its energy from fossil fuel in a manner that directly causes climate change equates to an energy-climate crisis.

Although governments around the world have only recently begun to consider policies to avoid the direst projections of climate change and its impacts, sustainable approaches to addressing the crisis are available. The common thread of feasible strategies to the energy-climate crisis is the simultaneous use of multiple approaches based on available technologies [e.g., 20]. Efficiency improvements (e.g., in building energy use), increased use of natural gas relative to coal, and increased development of renewables such as solar, wind, and geothermal, along with nuclear energy, are all available options that will reduce net CO_2 emissions. While improvements in efficiency can be made rapidly and will pay for themselves, the slower pace of change and greater monetary costs associated with increased use of renewables and nuclear energy suggests an additional approach is needed to help bridge the time period between the present and the future when low-carbon energy is considered cheap enough to replace fossil fuels. Carbon dioxide capture and storage (CCS) is one such bridging technology [1].

CCS has been the focus of an increasing amount of research over the last 20–25 years and is the subject of several books [e.g., 21] and a comprehensive IPCC report that thoroughly covered the subject [1]. This IPCC report was effectively updated in key areas in 2015 by a special issue in the *International Journal of Greenhouse Gas Control* [22]. The Global CCS Institute (GCCSI) reports that CCS is currently being carried out in several countries around the world in conjunction with natural gas extraction, enhanced oil recovery, fertilizer production, hydrogen production, iron and steel production, and biofuel production [23]. We note that although a large amount of CO_2 has been used for enhanced oil recovery since the mid-1970s in the USA, most of it has been natural CO_2 produced from CO_2 domes (e.g., Bravo Dome (New Mexico), Jackson Dome (Mississippi), and Sheep Mountain and McElmo Domes (Colorado)) rather than anthropogenic CO_2. One prominent exception is the Weyburn CO_2-enhanced oil production project in Saskatchewan, Canada, which uses anthropogenic CO_2 [24]. Despite progress, widespread deployment of CCS or CCUS remains the subject of research and planning rather than action on the gigatonne-CO_2-injected-per-year scale needed to mitigate emissions from the perspective of climate change. The reasons for delay in deploying CCS more widely are concerns about cost [25], regulatory and legal uncertainty [26], and potential environmental impacts [27], including induced seismicity (earthquakes caused by fluid injection) [28].

This entry discusses the long-term (decadal) sustainability and environmental hazards associated with the geologic CO_2 storage (GCS) component of large-scale CCS/CCUS [e.g., 29]. Discussion here does not focus on capture and transport of CO_2 because these will occur above ground and are similar to existing engineering, chemical processing, and pipeline transport activities and are therefore easier to evaluate with respect to risk assessment and feasibility. The focus of this entry is on the more uncertain part of CCS/CCUS, namely, geologic storage. The primary concern for sustainability and viability of GCS is whether there is sufficient capacity in sedimentary basins worldwide to contain the large of amounts of CO_2 needed to address climate change. But there is also a link between sustainability and environmental impacts. Specifically, if GCS is found to cause unacceptable impacts that are considered worse than its climate change mitigation benefits, the approach will not be widely adopted. Hence, GCS has elements of sustainability insofar as capacity of the subsurface for CO_2 is concerned and also in terms of whether the associated environmental risks are acceptable or not to the public.

Geologic CO_2 Storage (GCS): How Does It Work?

In order to understand the main environmental hazards and sustainability issues associated with GCS, the basic principles of CCS must be understood. First, CO_2 gas compresses into a relatively high-density form at the pressures and temperatures encountered below a depth of approximately 1 km in the Earth. In this dense form, called its *supercritical* form because it is neither strictly liquid nor strictly gas, a larger amount of CO_2 can be stored per unit volume than if CO_2 is stored as a gas at shallower depths. The density of CO_2 at depths greater than 1 km below the ground surface onshore ranges from around 200 to 800 kg/m^3 depending on the increases in pressure and temperature with depth (termed the pressure and geothermal gradients, respectively).

The depths targeted for GCS are typically in the range of 1–4 km, with the maximum depth dictated by the economics of deep wells and typically decreasing permeability of deep GCS target formations. The density of CO_2 at a particular location is nearly constant across these depths as the effects on CO_2 density of increasing temperature approximately compensate for increasing pressure in typical sedimentary basins [27]. Although CO_2 is very dense at depth relative to its gaseous form at the ground surface and can therefore be volumetrically sequestered efficiently in the deep subsurface, it will always be buoyant relative to the native fluids (saline groundwater or brine) in the onshore subsurface and tend to rise up through them if a flow path is available. In offshore environments, there are areas in ocean basins where low seawater temperatures cause CO_2 density in the sediments below the seafloor to be higher than surrounding formation water thereby removing the upward buoyancy force that is present in onshore GCS systems [30].

Second, global tectonics have created sedimentary basins on all of the continents in which sediment deposition over geologic time scales has produced thick sequences of sedimentary rock capable of storing CO_2 [31]. There is a vast amount of pore space in these sedimentary rocks arising from the imperfect packing of individual rock grains and incomplete cement filling of the space (pores) between the grains. Significant space can also sometimes arise from pervasive fracturing of the rock. In addition, sedimentary rocks commonly exist in alternating sequences of sandstones (relatively coarse-grained, with high permeability) and shales and mudstones (fine-grained, with low permeability) making a configuration in which some sedimentary layers are permeable and others are relatively impermeable. The fine-grained and low-permeability formations are the cap rocks that provide the upper seal for the high-porosity and permeability reservoirs into which CO_2 can be injected in the process of GCS.

Four different primary trapping mechanisms are recognized to occur in the deep subsurface to permanently sequester CO_2 [29]. These include:

1. Structural and stratigraphic trapping: This occurs when buoyant CO_2 flows up and becomes trapped against fine-grained and very low-permeability overlying cap rock, often in dome-shaped structures in the case of structural trapping. This is the same mechanism that traps oil and natural gas.
2. Residual gas trapping: The process in which CO_2 bubbles are left trapped in the pores of the rock as CO_2 and water flows through the reservoir (e.g., by buoyancy forces) and water in-fills the pores previously occupied by CO_2. This is the same process that occurs in oil reservoirs as water replaces oil and prevents full recovery motivating various enhanced oil recovery approaches. It is akin to why a sponge is not dry after it is wrung out.
3. Solubility trapping: The process in which CO_2 dissolves into the saline water or brine in the reservoir rock. This same process of CO_2 dissolution occurs to create both natural and manufactured carbonated beverages.
4. Mineral trapping, which occurs as CO_2 dissolved in the native water, reacts with minerals and other dissolved constituents to form new carbonate minerals. This is analogous to the precipitation of travertine that forms in some hot (and cold) spring waters.

CO_2 injected into the deep subsurface will tend to be trapped by all four of these mechanisms in proportions that vary over time. For example, mineral trapping depends on dissolution [e.g., 32] and precipitation of mineral phases that can take on the order of 100–1000 s of years [33, 34]. Considered together, the fractions of trapping by residual gas and solubility and mineral precipitation processes tend to increase over time, while the fraction of CO_2 trapped by structural and stratigraphic trapping decreases [29]. As sequestered CO_2 progresses over time through the sequence of structural and stratigraphic, residual gas, solubility, and mineral trapping mechanisms, CO_2 storage is considered to become more permanent [29].

One process that has similarities to GCS is underground (natural) gas storage (UGS), carried out at over 450 sites in the USA [35]. In this process, methane (CH_4) produced from natural gas reservoirs in one location is reinjected into depleted natural gas and oil reservoirs, aquifer storage reservoirs, or solution-mined salt caverns for temporary storage until market demand (e.g., a cold or hot spell) exceeds supply at which time extra gas is withdrawn from the storage reservoir. In the USA, the amount of natural gas stored at any one time is much smaller than the amount of CO_2 that is produced from fossil fuel power plants per year (approximately six times less CH_4 by volume (7.5 Tcf $= 1.4 \times 10^8$ tonnes [36]) is stored overall than there is CO_2 produced at fossil fuel power plants (47 Tcf $= 2.4 \times 10^9$ tonnes) per year). Furthermore, the natural gas storage industry uses the same reservoir for decades of injection and withdrawal cycles, whereas the GCS industry would need to continuously develop new reservoirs. So while the processes are very similar and much can be learned from the natural gas storage industry, the scale of the GCS industry will need to be much larger [e.g., 37, 38] in order for it to contribute substantially to averting climate change.

Opportunity and Capacity

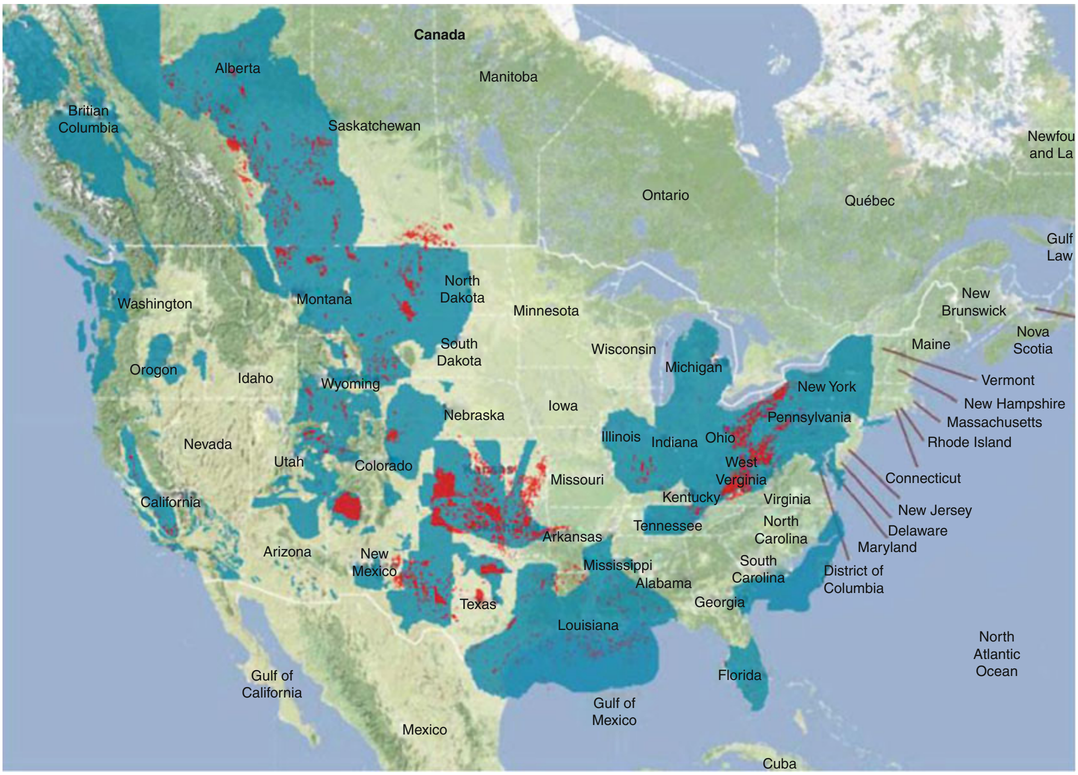
As mentioned above, sedimentary basins in the USA and around the world are the primary targets for large-scale GCS [31, 1]. Shown in Fig. 2 are sedimentary basins (blue) in the USA and Canada with hydrocarbon-producing regions shown in red. As shown, there are large areas of the USA and Canada that are potential sinks for CO_2. Most of the opportunity is in sedimentary basins on the continent, but offshore opportunities are also being pursued [e.g., 39, 14]. Economics and regulatory and environmental considerations will govern the extent to which offshore options are viable. Current efforts in North America are mostly aimed at onshore GCS opportunities, while in Europe primarily offshore opportunities are pursued.

Sustainability and feasibility of GCS are largely dependent on capacity. Evaluations have shown there is more than enough capacity to store point-source CO_2 emissions for hundreds of years or more [e.g., 1]. However, large capacity is a necessary but not sufficient condition for GCS feasibility. First, large capacity does not equate to adequate injectivity, i.e., there may be large porosity in some formations that have low permeability or are highly compartmentalized. This would require more wells to inject CO_2 than the economics of a project could support. Second, capacity may not be available in close proximity to large CO_2 sources necessitating long pipeline transport distances and associated extra costs [41]. Some of this transport cost can be accommodated under reasonable projections of CO_2 storage economics, but at the extreme, it could make CCS unviable due to economics or societal acceptance of greater pipeline length.

This discussion points out that there are two different types of capacity, namely, resource and reserve capacity [e.g., 42]. Most evaluations to date have focused on resource capacity, i.e., the total amount of pore space available regardless of where it is located or what it takes to access it. Although different methods to estimate capacity have led to wide variations in capacity estimates over various regions [e.g., 43], there is no doubt that there is an enormous amount of resource capacity available. In short, resource capacity does not at present appear to limit the long-term (decadal) sustainability of GCS. On the other hand, resource capacity is not the only measure of feasibility.

As described in the glossary, reserve capacity is the more practical measure of capacity, because it includes economic, technological, and regulatory restrictions and limitations on capacity. By this definition, reserve capacity is a fraction of resource capacity, and reserve capacity can change over time as economics, technology, or regulations change.

The remainder of this entry discusses potential environmental impacts associated with GCS, such as the possibility of groundwater contamination and induced seismicity. Environmental risks and costs may be unacceptable to the public, leading



Geologic Carbon Sequestration: Sustainability and Environmental Risk, Fig. 2 Sedimentary basins (blue) in the USA and Canada considered good targets for potential geologic storage of CO_2, with oil- and gas-producing regions shown in red [40]

to regulation that along with economics (which depends in part on policy drivers such as a tax on greenhouse gas pollution or tax credits for GCS) determines the reserve capacity.

Potential Impacts

The injection of large quantities of CO_2 into the deep subsurface through wells imposes a large perturbation to the local natural system in terms of changing the composition and pressure of the native fluids [44]. Specifically, CO_2 will partially dissolve into the native saline groundwater or brine while also pushing these native fluids outward away from the well. It will also produce a pressure wave that moves rapidly outward through the native fluids beyond where injected CO_2 is in contact with those fluids.

The deep fluid injection process is well known and practiced widely for injection of various fluids today [45, 35, 26], and the reverse, production of fluids through wells, such as oil, gas, and groundwater, are similarly practiced widely under regulatory frameworks aimed at protecting against adverse consequences. Nevertheless, the novelty of GCS associated with the large volume of CO_2 that needs to be injected motivates discussion of what can go wrong and what environmental impacts are possible. This discussion will serve to evaluate which impacts are the most likely and which have the greatest consequences.

Broadly, environmental impacts of CCS can be broken down into those occurring at depth with no discharge of CO_2 into the atmosphere (i.e., the CO_2 storage objective is achieved even as other consequences occur) and those that involve CO_2 discharging into the atmosphere. Presented in

Table 1 are potential impacts of GCS broken down into these two broad categories.

While the impacts arising from CO_2 leaking upward into the vadose zone, root zone, surface water, and out of the ground may be very serious, such occurrences all require a conduit or flow pathway from the deep injection zone to the near-surface environment, such as an improperly abandoned well or transmissive fault. Any GCS project that had moderate to high potential for the leakage scenarios in the upper part of Table 1, as defined by society through the development of regulations, would presumably not be undertaken assuming effective risk management, insurance, and regulatory processes are in place. Furthermore, theoretical studies aimed at finding ways that CO_2 could be catastrophically released from CO_2 storage sites leading to the most serious impacts at the ground surface have found self-limiting fluid interference behaviors rather than runaway behaviors [60]. Finally, the impacts described in the upper part of Table 1 are associated with failures of GCS in that CO_2 will enter the atmosphere negating the sequestration objective. For projects receiving value through greenhouse gas pollution abatement markets, this provides an additional incentive for project operators to avoid leakage as it would cause a direct reduction in revenue. Assuming an adequate monitoring program is in place, these leakage events would be relatively obvious, and appropriate changes in operations and remedial actions could be carried out.

In the exceptional case of the occurrence of an uncontrolled CO_2 leak from a well into the atmosphere, the main consequence of concern is asphyxiation of workers or bystanders. Documented CO_2 well blowouts associated with oil production indicate the asphyxiation hazard is low for blowouts occurring in open environments [e.g., 61, 62]. Modeling studies of open-air scenarios have also found that the area of asphyxiation hazard around a blowout is small because turbulent mixing and dispersion acts to rapidly decrease concentrations [63].

In contrast to a well blowout or the scenarios in the upper part of Table 1, it may be much more difficult unless frequent seismic monitoring is

Geologic Carbon Sequestration: Sustainability and Environmental Risk, Table 1 Shallow (top part of the table) and deep (bottom part of the table, shaded) processes and potential impacts of GCS [27]

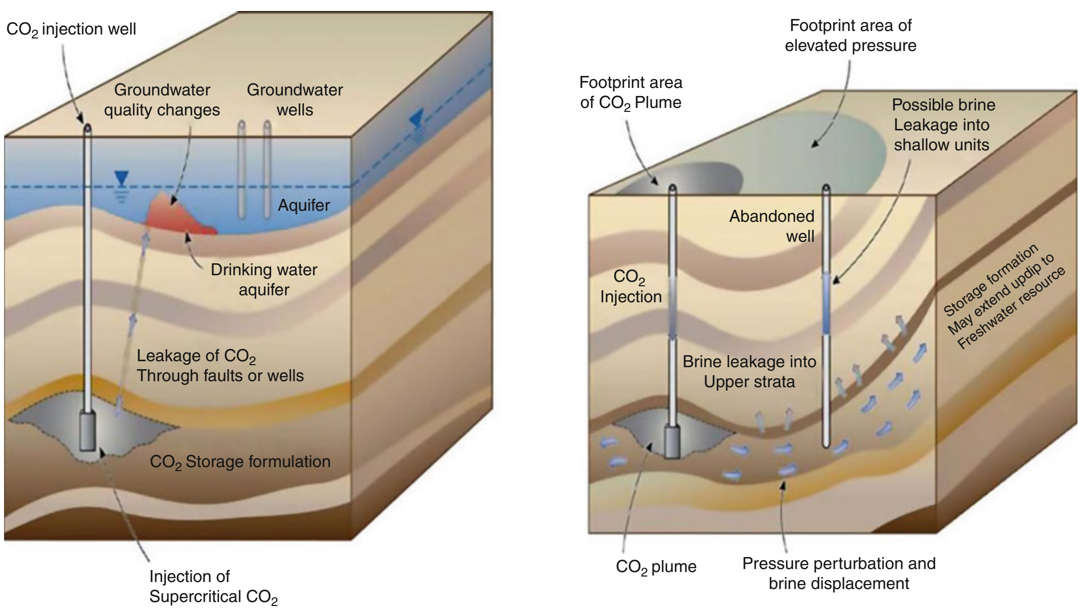
Category	Scenario	Significance	References
CO_2 enters the atmosphere	Root zone impacts	Profound, visible impact on plants, trees, crops	[46, 47]
	Migration into vadose zone	May include root zone and entry into buildings	[48]
	Bubbling through surface water	Alters water quality (e.g., lowering pH)	[49]
	Accumulation in topographic lows	Very hazardous due to possibility of asphyxiation	[50, 48]
	Seepage into basements and homes	Very hazardous due to possibility of asphyxiation	[51]
	Ground plumes	Very hazardous due to possibility of asphyxiation	[50, 52, 53]
CO_2 may or may not enter atmosphere	Intrusion of CO_2 into potable water	Lowers pH, dissolves minerals potentially releasing heavy metals	[54, 55]
	Intrusion of CO_2 into hydrocarbon, mineral, or geothermal resources	Lowers value of natural gas or mineral resources such as potash	
	Displacement of saline groundwater or brine into potable water by regional pressurization	Saline water intrusion into potable water degrades water quality	[55, 56, 57]
	Induced seismicity	CO_2 injection pressure may cause felt earthquakes	[58, 59]

carried out to detect the onset and development of the scenarios listed in the lower part of the table in order to take early action to limit impacts. Although the scenarios listed in the lower part of Table 1 do not involve CO_2 entering the atmosphere and thus do not involve outright failure of GCS, the intrusion of CO_2 or saline groundwater or brine into groundwater resources and injection-induced seismicity are considered the hazards associated with GCS that are sufficiently likely enough to warrant risk assessment and related regulatory measures in order to minimize the likelihood of their occurrence and their consequences. These two categories of risk, described in more detail below, must be assessed [e.g., 64, 65] and managed as part of widespread, long-term, and sustainable GCS deployment.

Potential Impacts to Potable Groundwater

CO_2 that leaks upward out of the storage region through wells [e.g., 66, 67], or faults and fractures [68], can potentially enter potable groundwater resources as shown schematically in Fig. 3a. Degradation of groundwater quality is possible through indirect contamination. As CO_2 dissolves into groundwater, it partitions into species comprising dissolved inorganic carbon (DIC) as CO_2 (aq), HCO_3^-, and CO_3^{2-}, resulting in a decrease in the solution pH. At the same time, alkalinity is controlled by HCO_3^- and CO_3^{2-}, which can increase upon CO_2 dissolution. Control over the geochemical changes in the water is provided by the composition and mineralogy of the mineral grains, coatings, and cements present in the rock. For example, a carbonate mineral such as calcite ($CaCO_3$) in the rock will dissolve by the reaction $CO_2 + H_2O + CaCO_3 = Ca^{2+} + 2HCO_3^-$, resulting in the doubling of dissolved inorganic carbon (DIC) (i.e., one mole of CO_2 reacts to produce two moles of HCO_3^-) and a release of Ca^{2+} to solution. Similar reactions are possible involving alteration of biotite, plagioclase and alkali feldspar, and other common minerals in sedimentary rocks [e.g., 34].

CO_2 leakage into groundwater aquifers will also give rise to impacts on microbiological communities [69]. Although cell density declines by three to six orders of magnitude from the ground surface to 4 km depth, microbes at the depths of



Geologic Carbon Sequestration: Sustainability and Environmental Risk, Fig. 3 Potential groundwater impact scenarios. Left-hand figure from [55] and right-hand figure from [57] potable groundwater can be affected if CO_2 intrudes into this region. The alteration of minerals such as feldspars by acidic groundwater can release iron which can stimulate Fe^{3+}-reducing communities and result in methanogenesis. Microbial processes can affect geochemistry and vice versa.

Assuming the reaction kinetics allow it, geochemical reactions can further alter pH, DIC, isotopic composition, and trace element concentrations in solution. For example, trace elements in the minerals, in coatings, or in ion exchange sites in clays (including heavy metals such as lead) may be released into groundwater as biogeochemical conditions change with associated degradation of groundwater quality [70, 54]. Observations of such effects have been made during CO_2 injection experiments at field sites [e.g., 71, 72] and in the laboratory [73]. Recent work has further assessed the potential for such reactions by examining actual groundwater compositions and aquifer mineralogy from across the USA and found that increases in the concentration of As and Pb could be a concern if widespread CO_2 leakage into groundwater resources were to occur [55]. Buffering reactions may serve to moderate pH decline and may serve to diminish groundwater degradation as observed in a natural analog study in New Mexico [74]. In summary, it is recognized that impacts of CO_2 leakage on potable groundwater may be significant and costly if they occur, and therefore careful GCS site selection, operation, and monitoring [e.g., 75] are essential to reduce groundwater contamination risk.

Another hazard to groundwater resources is the potential intrusion of displaced saline groundwater or brine or CO_2-charged water into potable groundwater as shown in Fig. 3b. In addition to the above biogeochemical impacts arising from the CO_2 itself, there is the first-order degradation arising from the presence of dissolved solids (e.g., NaCl, $CaCl_2$, KCl) in the saline groundwater or brine along with whatever trace elements it may contain. Protected groundwater in the USA is defined on the basis of total dissolved solid (TDS) content equal to 10,000 mg/L or less. Injection into deep aquifers is regulated in the USA by the Environmental Protection Agency (EPA) under the Underground Injection Control (UIC) program to protect this groundwater from degradation [e.g., 76, 45]. The hazard arising from GCS is that deep saline water or brine pressurized by CO_2 injection may migrate upward into protected groundwater aquifers, thereby increasing TDS and degrading the resource.

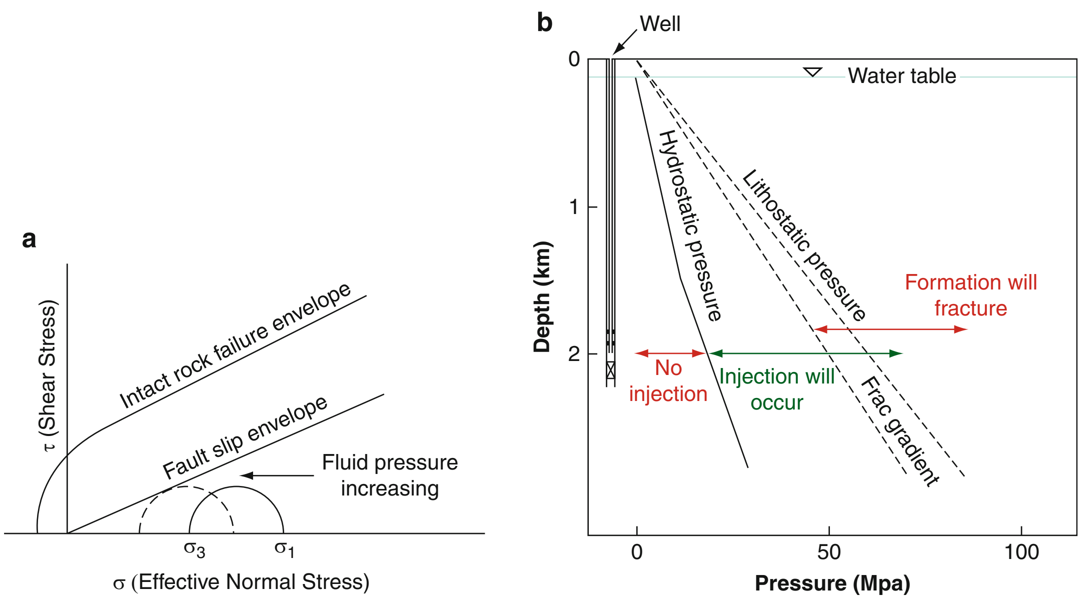
The main reason that saline groundwater or brine intrusion arising from GCS is such a concern is that pressure increases associated with CO_2 injection can occur at great distances (~10–100 km) from the injection site [77, 56, 57, 78]. So while characterization of a given site may have demonstrated that CO_2 will be contained within a well-defined CO_2 storage region, there will generally be a large region of pressure increase in the formation that may not have been characterized to the same degree because of the large distance from the injection site. Because of this, it is possible that the cap rock may not be continuous over these large distances or may not have the same integrity as the region targeted for CO_2 storage. Nevertheless, in order for upward saline groundwater or brine intrusion to occur, there must be a driving force in addition to a conduit or pathway (e.g., improperly abandoned well or fault or fracture zone). Although the pressure increase is high near the CO_2 injection wells, it falls off rapidly away from the wells. In addition, brines with high TDS require larger overpressures to be driven upward into potable groundwater through wells or other conduit (e.g., conductive fault) due to their high density and resistance to flow [79]. Furthermore, once in the potable aquifer, the higher density of the brine will tend to limit the extent of its mixing with potable groundwater [80].

Induced Seismicity

The phenomenon of induced seismicity due to fluid injection has been recognized for approximately 50 years starting with the well-known example of fluid waste disposal by injection at the Rocky Mountain Arsenal in Colorado [81, 82, 83]. Induced seismicity due to injection is understood from experience in the fields of injection for deep disposal of liquid waste, and injection for geothermal energy extraction [84, 58, 59], and for GCS [28]. One cause of induced seismicity is the reduction in effective stress that accompanies an increase in pore pressure. The potential for this seismicity is determined by the Mohr-Coulomb criterion which quantifies the amount of normal stress reduction provided by fluid pressure that is needed before shear failure occurs (i.e., reactivation of existing faults or slippage along fractures). The Mohr-Coulomb criterion is given by the relation $\tau = C + \mu\,(\sigma_n - p)$ where τ is the shear strength of the rock, C is the Coulomb criterion (undrained, unconfined compressive strength), μ is the coefficient of internal friction, σ_n is the normal stress, and p is the fluid pressure [e.g., 85]. When the right-hand side (normal stress) is smaller than the left-hand side (shear stress), the rock is likely to slip along fracture planes of optimal orientation, which can release seismic energy (i.e., causes an earthquake). The Mohr-Coulomb equation shows that injection pressure reduces the effective normal stress in the rock, hence the tendency for injection to cause slippage along existing faults and fractures as shown in Fig. 4a.

A simple graphical representation of pressures as a function of depth helps elucidate the processes active near an injection well. Shown in Fig. 4b are the variations with depth of hydrostatic pressure, so-called fracture pressure (or commonly frac pressure) and lithostatic pressure. As shown, fluid pressure at an injection well must be larger than hydrostatic pressure in order for injection to occur. However, if the pressure exceeds the frac pressure at a given depth, the injection process will tend to fracture the formation. By the Mohr-Coulomb criterion, seismicity can be induced at injection pressures below the frac pressure as effective stress decreases and existing faults are reactivated. Either the generation of new fractures or slippage along existing faults and fractures can be manifested as induced seismicity.

It is important to note that while the word *earthquake* evokes fear and a certain image of destruction in most people's minds, the term encompasses a wide range of magnitudes, from microseismic earthquakes that cannot be felt by humans to great earthquakes that damage structures and imperil life. Earthquakes tend to follow a logarithmic frequency distribution such that very



Geologic Carbon Sequestration: Sustainability and Environmental Risk, Fig. 4 (**a**) Mohr circle representation of fault slip (induced seismicity) as fluid pressure increases and (**b**) pressure-depth depiction showing hydrostatic, frac, and lithostatic pressure gradients small earthquakes are orders of magnitude more frequent than very large earthquakes [86]. Experience from water injection into geothermal systems shows that the majority of induced seismicity is microseismicity, with felt earthquakes much rarer and moderate to large earthquakes rarer still [58]. Despite the fact that large earthquakes are not expected to be induced by CO_2 injection in carefully chosen sites [87], the hazard of induced seismicity is currently the most difficult to assess with regard to GCS. For instance, in addition to induced seismicity caused by increasing fluid pressure at a fault, there is currently an ongoing discussion in the research literature about the ability of stress changes transmitted through rock resulting from fluid pressure changes at the well to cause induced seismicity ("poroelastic triggering"). This mechanism appears capable of inducing events at greater distances than those induced by fluid pressurization at a fault, but consensus regarding the existence or absence of this phenomenon among researchers has not yet occurred.

Aside from the hazard of ground acceleration at the surface, induced seismicity also creates the possibility that a cap rock seal could fracture or a fault could become permeable giving rise to a leakage pathway for CO_2 [e.g., 88, 85] and such a fault could be too small to have been recognized during site characterization [89]. Fracturing affecting cap rock is a well-recognized failure mode, and injection regulations are aimed at preventing them from happening. However, induced seismicity of critically stressed rocks on preexisting faults is possible even when the frac pressure is not exceeded [90]. The extent to which the risk of induced seismicity, objectively considered to be a small risk, outweighs the benefits of reducing climate change that CCS affords is one of the questions that must be addressed by the public and decision-makers to determine the extent to which CCS is employed as a mitigation.

Future Directions

As the discussion above suggests, one approach that can aid in addressing the energy-climate crisis is CCS/CCUS. There are significant costs to CCS/CCUS, primarily associated with capture, and the process also brings with it recognized environmental risks, the most uncertain of which are associated with the geologic storage component of the process. The main risks in GCS are threat to potable groundwater and induced seismicity, two areas of active research. Despite the need for greater understanding of these hazards, mitigation measures are available today. For example, if contamination of groundwater were to occur, the water could be treated, or alternate sources could be found if treatment is found impractical [75]. As for induced seismicity, the hazard can be reduced by reducing injection pressure (e.g., through use of more wells over a larger area for a given CO_2 source), by carrying out pressure management through saline groundwater or brine extraction, by careful site selection that avoids heavily faulted areas, and by establishing and enforcing building codes.

The path forward for demonstrating and deploying CCS/CCUS at the gigatonne scale as a sustainable part of the portfolio of energy production and use changes that are needed to mitigate the energy-climate crisis consists of six broad steps. First, testing and demonstration capture projects [e.g., 91, 92] from anthropogenic sources could expand rapidly and by many factors so that the cost of capture can be reduced. Second, large GCS projects could be undertaken in different regions and geologic settings to determine if the demonstration GCS projects that preceded them scale up as anticipated. These multiple projects would show if capture can be economic and if GCS performs as envisioned, in which case additional CCS deployments can be added over time. Third, research on alternative capture and combustion approaches that enable more efficient capture could be accelerated [e.g., 93]. Fourth, a large program of site characterization and capacity studies [e.g., 94] could be undertaken so that the large basin-scale sites are understood and operational plans can be put in place quickly at the time when large-scale capture facilities come on line and anthropogenic CO_2 streams become available for sequestration. Fifth, research on injection, trapping, migration, long-term fate, leakage impacts, mitigation, monitoring, and modeling could be continued so that GCS can be optimized and related technologies can be commercialized and deployed in a cost-effective manner. Sixth, governments at all levels could promulgate regulatory and economic policies that answer the current questions and uncertainty faced by businesses who foresee the broad outlines of a carbon-constrained future but do not yet have the clear ground rules provided by government that are necessary for making the large capital investments required for CCS/CCUS.

The decision to take on the costs and risks of CCS, with the accompanying promise of contributing to reductions in the extent of climate change, should be made based on an objective comparison against the climate and environmental risks of carrying on business as usual with fossil fuel use and unabated CO_2 emissions. The public and decision-makers should keep in mind that the environmental risks of CCS are local to the basin where GCS is carried out, whereas the projected impacts of climate change are global-to-regional in scale and are expected to have profound consequences for the social, physical, and natural systems on Earth. Support for CCS technology will come in the form of policy decisions about carbon pricing, injection regulations, and legal frameworks that encourage commercial applications of CCS/CCUS. The decision about whether to adopt these policies will ultimately fall on the public or its representatives in democratic societies and ruling parties and people in authoritarian societies. The risks to the Earth's environment and social systems of doing nothing about the energy-climate crisis must be communicated effectively to the public and the decision-makers so that they can make informed decisions about acceptable risks and costs of climate change and of the various options available for reducing those risks and costs.

Bibliography

Primary Literature

1. IPCC special report on carbon dioxide capture and storage (2005) Metz B, Davidson O, de Coninck HC, Loos M, Meyer LA (eds). Prepared by Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge
2. Blunt M, Fayers FJ, Orr FM (1993) Carbon dioxide in enhanced oil recovery. Energy Convers Manag 34(9):1197–1204
3. Torp TA, Gale J (2004) Demonstrating storage of CO_2 in geological reservoirs: the Sleipner and SACS projects. Energy 29(9–10):1361–1369
4. Eiken O, Ringrose P, Hermanrud C, Nazarian B, Torp TA, Høier L (2011) Lessons learned from 14 years of CCS operations: Sleipner, in Salah and Snøhvit. Energy Procedia 4:5541–5548
5. Keith DW, Ha-Duong M, Stolaroff JK (2006) Climate strategy with CO_2 capture from the air. Clim Chang 74(1–3):17–45
6. Lackner K (2010) Washing carbon out of the air. Sci Am 302:66–71. https://doi.org/10.1038/scientificamerican0610-66.
7. Finley RJ (2014) An overview of the Illinois Basin–Decatur project. Greenhouse Gases Sci Technol 4(5):571–579
8. Mooney C (2017) The quest to capture and store carbon – and slow climate change – just reached a new milestone, Washington Post, April 10, 2017. https://www.washingtonpost.com/news/energy-environment/wp/2017/04/10/the-quest-to-capture-and-store-carbon-and-slow-climate-change-just-reached-a-new-milestone/?utm_term=.83ff9466bf29. Accessed 11 Sept 2017
9. Brewer PG, Friederich G, Peltzer ET, Orr FM Jr (1999) Direct experiments on the ocean disposal of fossil fuel CO_2. Science 284(5416):943–945
10. Caldeira K, Rau GH (2000) Accelerating carbonate dissolution to sequester carbon dioxide in the ocean: geochemical implications. Geophys Res Lett 27(2):225–228
11. Chisholm SW, Falkowski PG, Cullen JJ (2001) Discrediting ocean fertilization. Science 294(5541):309–310
12. Buesseler KO, Boyd PW (2003) Will ocean fertilization work? Science 300(5616):67–68
13. Shaffer G (2010) Long-term effectiveness and consequences of carbon dioxide sequestration. Nat Geosci 3:464–467
14. Meckel TA, Trevino R, Carr D, Nicholson A, Wallace K (2013) Offshore CCS in the northern Gulf of Mexico and the significance of regional structural compartmentalization. Energy Procedia 37:4526–4532
15. House KZ, Harvey CF, Aziz MJ, Schrag DP (2009) The energy penalty of post-combustion CO_2 capture & storage and its implications for retrofitting the U.S. installed base. Energy Environ Sci. https://doi.org/10.1039/b811608c
16. McKinsey and Co (2008) Carbon capture and storage: assessing the economics. McKinsey & Company, p 49
17. IEA, Key World Energy Statistics (2009) International Energy Agency (IEA), Paris. http://www.iea.org/textbase/nppdf/free/2009/key_stats_2009.pdf
18. Keeling RF, Piper SC, Bollenbacher AF, Walker JS (2009) Atmospheric CO_2 records from sites in the SIO air sampling network. In trends: a compendium of data on global change. Carbon dioxide information analysis center, Oak Ridge National Laboratory,

U.S. Department of Energy, Oak Ridge. doi:https://doi.org/10.3334/CDIAC/atg.035
19. IPCC, Climate Change (2007) Synthesis report. Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change (2007) [Core Writing Team, Pachauri RK, Reisinger A (eds)]. IPCC, Geneva, p 104
20. Pacala S, Socolow R (2004) Stabilization wedges: solving the climate problem for the next 50 years with current technologies. Science 305(5686):968–972
21. Smit B, Reimer JA, Oldenburg CM, Bourg IC (2014) Introduction to carbon capture and sequestration, vol 1. World Scientific. Imperial College Press, London
22. Gale J, Abanades JC, Bachu S, Jenkins C (2015) Special issue commemorating the 10th year anniversary of the publication of the intergovernmental panel on climate change special report on CO_2 capture and storage. Int J Greenhouse Gas Control 40:1–458
23. GCCSI (Global CCS Institute) (2017) http://www.globalccsinstitute.com/sites/www.globalccsinstitute.com/files/content/page/122973/files/status-ccs-project-database-current-08-09-2017.xls. Accessed 11 Sept 2017
24. White DJ, Burrowes G, Davis T, Hajnal Z, Hirsche K, Hutcheon I, Majer E, Rostron B, Whittaker S (2004) Greenhouse gas sequestration in abandoned oil reservoirs: the International Energy Agency Weyburn pilot project. GSA Today 14(7):4–11
25. Rubin ES, Chen C, Rao AB (2007) Cost and performance of fossil fuel power plants with CO_2 capture and storage. Energy Policy 35(9):4444–4454
26. Wilson EJ, Gerard D (2007) Risk assessment and management for geologic sequestration of carbon dioxide. In: Wilson EJ, Gerard D (eds) Carbon capture and sequestration integrating technology, monitoring, and regulation. Blackwell Publishing, Ames, pp 101–125
27. Oldenburg CM (2007) Migration mechanisms and potential impacts of CO_2 leakage and seepage. In: Wilson EJ, Gerard D (eds) Carbon capture and sequestration integrating technology, monitoring, and regulation. Blackwell Publishing, Ames, pp 127–146. LBNL-58872
28. White JA, Foxall W (2016) Assessing induced seismicity risk at CO_2 storage projects: recent progress and remaining challenges. Int J Greenhouse Gas Control 49:413–424
29. IPCC special report on CO_2 capture and storage, Chapter 5. Benson SM, Cook PJ (eds) (2005) Cambridge University Press, Cambridge
30. House KZ, Schrag DP, Harvey CF, Lackner KS (2006) Permanent carbon dioxide storage in deep-sea sediments. Proc Natl Acad Sci 103(33):12291–12295
31. Bachu S (2003) Screening and ranking of sedimentary basins for sequestration of CO_2 in geological media in response to climate change. Environ Geol 44(3): 277–289
32. Spycher N, Ennis-King J, Pruess K (2003) CO_2-H_2O mixtures in the geological sequestration of CO_2. Assessment and calculation of mutual solubilities from 12 to 100 C and up to 600 bar. Geochemica et Cosmochimica Acta 67:3015–3031
33. Gunter WD, Bachu S, Benson S (2004) The role of hydrogeological and geochemical trapping in sedimentary basins for secure geological storage of carbon dioxide. Geol Soc Lond Spec Publ 233:129–145
34. Xu T, Apps JA, Pruess K (2005) Mineral sequestration of CO_2 in a sandstone-shale system. Chem Geol 217(1–4):295–318
35. Benson SM, Hepple R, Apps J, Tsang CF, Lippmann M (2002) Lessons learned from natural and industrial analogues for storage of carbon dioxide in deep geological formations, E.O. Lawrence Berkeley National Laboratory Report LBNL-51170
36. Katz DL, Tek MR (1981) Overview on underground storage of natural gas. J Pet Technol 33(6):943–951
37. USGS (2010) http://energy.er.usgs.gov/health_environment/co2_sequestration/co2_illustrations.html. Accessed 6 Oct 2010
38. Haszeldine RS (2009) Carbon capture and storage: how green can black be? Science 325(5948): 1647–1652
39. Shrag DP (2009) Storage of carbon dioxide in offshore sediments. Science 325:1658–1659
40. NATCARB (2008) U.S. Department of Energy, Carbon Sequestration Atlas of the United States and Canada, Office of Fossil Energy, National Energy Technology Laboratory. http://geoportal.kgs.ku.edu/natcarb/atlas08/gsinks.cfm. Accessed 6 Oct 2010
41. McCoy ST, Rubin ES (2008) An engineering-economic model of pipeline transport of CO_2 with application to carbon capture and storage. Int J Greenhouse Gas Control 2(2):219–229
42. Brennan ST, Burruss RC, Merrill MD, Freeman PA, Ruppert LF (2010) A probabilistic assessment methodology for the evaluation of geologic carbon dioxide storage: U.S. Geological Survey Open-File Report 2010–1127, 31 p. http://pubs.usgs.gov/of/2010/1127. Accessed 6 Oct 2010
43. Bradshaw J, Bachu S, Bonijoly D, Burruss R, Holloway S, Christensen NP, Mathiassen OM (2007) CO_2 storage capacity estimation: issues and development of standards. Int J Greenhouse Gas Control 1:62–68
44. Birkholzer JT, Oldenburg CM, Zhou Q (2015) CO_2 migration and pressure evolution in deep saline aquifers. Int J Greenhouse Gas Control 40:203–220
45. US EPA (United States Environmental Protection Agency), Technical Program Overview, Underground injection control regulations, Office of Water 4606, EPA 816-R-02-025, revised July 2001. http://www.epa.gov/safewater/uic/index.html. Accessed 10 Oct 2010
46. Qi J, Marshall JD, Matson KG (1994) High soil carbon dioxide concentrations inhibit root respiration of Douglas Fir. New Phytol 128:435–441

47. Farrar CD, Sorey ML, Evans WC, Howie JF, Kerr BD, Kennedy BM, King C-Y, Southon JR (1995) Forest-killing diffuse CO_2 emission at Mammoth Mountain as a sign of magmatic unrest. Nature 376:675–677
48. Oldenburg CM, Unger AJA (2003) On leakage and seepage from geologic carbon sequestration sites: unsaturated zone attenuation. Vadose Zone J 2(3):287–296
49. Oldenburg CM, Lewicki JL (2005) Leakage and seepage of CO_2 from geologic carbon sequestration sites: CO_2 migration into surface water. Lawrence Berkeley National Laboratory Report LBNL-57768
50. Giggenbach WF, Sano Y, Schmincke HU (1991) CO_2-rich gases from Lakes Nyos and Monoun, Cameroon; Laacher See, Germany, Dieng, Indonesia, and Mt. Gambier, Australia- variations on a common theme. J Volcanol Geotherm Res 45:311–323
51. Robinson AL, Sextro RG, Riley WJ (1997) Soil-gas entry into houses driven by atmospheric pressure fluctuations–the influence of soil properties. Atmos Environ 31(10):1487–1495
52. Hanna SR, Steinberg KW (2001) Overview of Petroleum Environmental Research Forum (PERF) dense gas dispersion modeling project. Atmos Environ 35: 2223–2229
53. Britter RE (1989) Atmospheric dispersion of dense gases. Ann Rev Fluid Mech 21:317–344
54. Wang S, Jaffe PR (2005) Dissolution of a mineral phase in potable aquifers due to CO_2 releases from deep formations; effect of dissolution kinetics. Energy Convers Manag 45:2833–2848
55. Apps JA, Zheng L, Zhang Y, Xu T, Birkholzer JT (2010) Evaluation of potential changes in groundwater quality in response to CO_2 leakage from deep geological storage. Transp Porous Media 82:215–246
56. Birkholzer JT, Zhou Q, Tsang C-F (2009) Large-scale impact of CO_2 storage in deep saline aquifers: a sensitivity study on the pressure response in stratified systems. Int J Greenhouse Gas Control 3(2):181–194
57. Birkholzer JT, Zhou Q (2009) Basin-scale hydrogeologic impacts of CO_2 storage: capacity and regulatory implications. Int J Greenhouse Gas Control 3(6):745–756
58. Majer EL, Baria R, Stark M, Oates S, Bommer J, Smith B, Asanuma H (2007) Induced seismicity associated with enhanced geothermal systems. Geothermics 36(3):185–222
59. Majer E, Baria R, Stark M (2008) Protocol for induced seismicity associated with enhanced geothermal systems. Report produced in Task D Annex I (9 April 2008), International Energy Agency-Geothermal Implementing, Agreement (incorporating comments by: C Bromley, W Cumming, A Jelacic and L Rybach) (http://www.iea-gia.org/publications.asp)
60. Pruess K (2005) Numerical studies of fluid leakage from a geologic disposal reservoir for CO_2 show self-limiting feedback between fluid flow and heat transfer. Geophys Res Lett 32(14):L14404
61. Skinner L (2003) CO_2 blowouts: an emerging problem. World Oil 224(1):38–42
62. Gouveia FJ, Johnson M, Leif RN, Friedmann SJ (2005) Aerometric measurement and modeling of the mass of CO2 emissions from Crystal Geyser, Utah. NETL 4th annual carbon capture and sequestration conference, Alexandria, 2–5 May
63. Aines RD, Leach MJ, Weisgraber TH, Simpson MD, Friedman SJ, Bruton CJ (2008) Quantifying the potential exposure hazard due to energetic releases from a failed sequestration well. In: Proceedings of the ninth international conference on greenhouse gas control technologies GHGT-9, Washington DC, 16–20 Nov 2008
64. Pawar RJ, Bromhal GS, Chu S, Dilmore RM, Oldenburg CM, Stauffer PH, Zhang Y, Guthrie GD (2016) The National Risk Assessment Partnership's integrated assessment model for carbon storage: a tool to support decision making amidst uncertainty. Int J Greenhouse Gas Control 52:175–189
65. Keating E, Bacon D, Carroll S, Mansoor K, Sun Y, Zheng L, Harp D, Dai Z (2016) Applicability of aquifer impact models to support decisions at CO_2 sequestration sites. Int J Greenhouse Gas Control 52: 319–330
66. Gasda SE, Bachu S, Celia MA (2004) Spatial characterization of the location of potentially leaky wells penetrating a deep saline aquifer in a mature sedimentary basin. Environ Geol 46:707–720
67. Scherer GW, Celia MA, Prevost J-H, Bachu S, Bruant R, Duguid A, Fuller R, Gasda SE, Radonjic M, Vichit-Vadakan W (2005) Leakage of CO_2 through abandoned wells: role of corrosion of cement. In: Thomas DC, Benson SM (eds) Carbon dioxide capture for storage in deep geologic formations, vol 2. Elsevier, Kidlington, Oxford, pp 827–848
68. Shipton ZK, Evans JP, Kirschner D, Kolesar PT (2004) AP Williams and J Heath. Analysis of CO_2 leakage through 'low-permeability' faults from natural reservoirs. Geol Soc Lond Spec Publ 233:43–58
69. Onstott TC (2005) Impact of CO_2 injections on deep subsurface microbial ecosystems and potential ramifications for the surface biosphere. In: Thomas DC, Benson SM (eds) Carbon dioxide capture for storage in deep geologic formations, vol 2. Elsevier, Kidlington, Oxford, pp 1217–1249
70. Schuett H, Wigand M, Spangenberg E (2005) Geophysical and geochemical effects of supercritical CO2 on sandstones. In: Thomas DC, Benson SM (eds) Carbon dioxide capture for storage in deep geologic formations, vol 2. Elsevier, Kidlington, Oxford, pp 767–786
71. Kharaka Y, Cole DR, Hovorka SS, Gunther WD, Knauss KG, Freifield BM (2006) Gas-water-rock interactions in Frio formation following CO_2 injection: implications for the storage of greenhouse gases in sedimentary basins. Geology 34:577–580
72. Kharaka YK, Thordsen JJ, Kakouros E, Ambats G, Herkelrath WN, Birkholzer JT, Apps JA, Spycher NF, Zheng L, Trautz RC, Rauch HW, Gullickson K (2010) Changes in the chemistry of shallow groundwater related to the 2008 injection of CO_2 at the ZERT

field site, Bozeman, Montana. Environ Earth Sci 60:273–284
73. Carroll S (2009) Trace metal release from Frio sandstone reacted with CO_2 and 1.5 N NaCl Brine at 60°C. In: Proceedings of the 8th annual conference on carbon capture and sequestration, Pittsburgh, May 2009
74. Keating EH, Fessenden J, Kanjorski N, Koning DJ, Pawar R (2010) The impact of CO_2 on shallow groundwater chemistry: observations at a natural analog site and implications for carbon sequestration. Environ Earth Sci 60(3):521–536
75. Price PN, Oldenburg CM (2009) The consequences of failure should be considered in siting geologic carbon sequestration projects. Int J Greenhouse Gas Control 3(5):658–663
76. Wilson EJ, Johnson TL, Keith DW (2003) Regulating the ultimate sink: managing the risks of geologic CO_2 storage. Environ Sci Technol 37(16):3476–3483
77. Nicot J-P (2008) Evaluation of large-scale CO_2 storage on fresh-water sections of aquifers: an example from the Texas Gulf Coast Basin. Int J Greenhouse Gas Control 2(4):582–593
78. Zhou Q, Birkholzer JT, Mehnert E, Lin Y-F, Zhang K (2009) Modeling basin- and plume-scale processes of CO_2 storage for full-scale deployment. Ground Water 48(4):494–514
79. Nicot J-P, Oldenburg CM, Bryant SL, Hovorka SD (2009) Pressure perturbations from geologic carbon sequestration: area-of-review boundaries and borehole leakage driving forces. Energy Procedia 1(1):47–54
80. Oldenburg CM, Rinaldi AP (2011) Buoyancy effects on upward brine displacement caused by CO_2 injection. Transp Porous Media 87(2):525–540
81. Hollister JC, Weimer RJ (eds) (1968) Geophysical and geological studies of the relationships between the Denver earthquakes and the Rocky Mountain Arsenal well, Q. Colorado School of Mines 63, Golden, 251 pp
82. Hoover DB, Dietrich JA (1969) Seismic activity during the 1968 test pumping at the Rocky Mountain Arsenal disposal well, circular 613. U.S. Geological Survey, Washington, DC
83. Herrmann RB, Park S-K, Wang C-Y (1981) The Denver earthquakes of 1967–1968. Bull Seismol Soc Am 71(3):731–745
84. Cypser DA, Davis SD (1998) Induced seismicity and the potential for liability under U.S. law. Tectonophysics 289(1–3):239–255
85. Rutqvist J, Birkholzer J, Cappa F, Tsang C-F (2007) Estimating maximum sustainable injection pressure during geological sequestration of CO_2 using coupled fluid flow and geomechanical fault-slip analysis. Energy Convers Manag 48(6):1798–1807
86. Gutenberg B, Richter CF (1944) Frequency of earthquakes in California. Bull Seismol Soc Am 17:185–188
87. Sminchak J, Gupta N, Byrer C, Bergman P (2003) Aspects of induced seismic activity and deep-well sequestration of carbon dioxide. Environ Geosci 10(2):81–89
88. Wiprut D, Zoback M (2000) Fault reactivation and fluid flow along a previously dormant normal fault in the northern North Sea. Geology 28(7):595–598
89. Mazzoldi A, Rinaldi AP, Borgia A, Rutqvist J (2012) Induced seismicity within geological carbon sequestration projects: maximum earthquake magnitude and leakage potential from undetected faults. Int J Greenhouse Gas Control 10:434–442
90. Zoback MD, Gorelick SM (2012) Earthquake triggering and large-scale geologic storage of carbon dioxide. Proc Natl Acad Sci 109(26):10164–10168
91. Hovorka SD, Benson SM, Doughty C, Freifeld BM, Sakurai S, Daley TM, Kharaka YK, Holtz MH, Trautz RC, Nance HS, Myer LR, Knauss KG (2006) Measuring permanence of CO_2 storage in saline formations: the Frio experiment. Environ Geosci 13(2):105–121
92. Litynski J, Plasynski S, McIlvried HG, Mahoney C, Srivastava RD (2008) The United States Department of Energy's regional carbon sequestration partnerships program validation phase. Environ Int 34(1):127–138
93. Allam RJ, Fetvedt JE, Forrest BA, Freed DA (2014) The oxy-fuel, supercritical CO2 Allam cycle: new cycle developments to produce even lower-cost electricity from fossil fuels without atmospheric emissions. In: ASME Turbo Expo 2014: turbine technical conference and exposition (pp. V03BT36A016-V03BT36A016), American Society of Mechanical Engineers
94. Friedmann SJ, Dooley JJ, Held H, Edenhof O (2006) The low cost of geological assessment for underground CO2 storage: policy and economic implications. Energy Convers Manag 47(13–14):1894–1901

Books and Reviews

Baines SJ, Worden RH (eds) (2004) Geologic storage of carbon dioxide. Geol Soc Lond Spec Publ 233:107–247

Eide LI (2009) Carbon dioxide capture for storage in deep geological formations, vol 3. CPL Press/BP, Newbury/Berkshire

Thomas DC, Benson SM (eds) (2007) Carbon dioxide capture for storage in deep geologic formations-results from the CO_2 capture project, vol 2. Elsevier, Kidlington

Wilson EJ, Gerard D (eds) (2007) Carbon capture and sequestration integrating technology, monitoring, and regulation. Blackwell Publishing, Ames

Marine Life Associated with Offshore Drilling, Pipelines, and Platforms

Martin Hovland
Centre for Geobiology, University of Bergen, Bergen, Norway
Statoil ASA, Stavanger, Norway

Article Outline

Glossary

Cold seep A location on the seafloor where natural fluids (gas and liquids) seep upward from the substratum, into the overlying water column.

Cold-water coral reef A mounded natural structure on the seafloor consisting of live animals and dead remains and sediments. The mound is partly constructed by colonizing corals that are not dependent on sunlight (i.e., ahermatypic corals) such as the most common species: *Lophelia pertusa.*

Fish sighting The underwater visual detection (recording) of fish (here, larger than 0.5 m in length) using submersible vehicles with lights and cameras, such as ROVs.

Iceberg ploughmark Up to 100 m wide and many kilometer long furrows in the seafloor, produced by the action of drifting grounded icebergs. Off Mid-and Northern-Norway and several other places such (relict) furrows remain from the last glaciation.

OHI The Offshore Hydrocarbon Industry (OHI) searches for natural accumulations (reservoirs) of oil and gas (hydrocarbons) and develops the means to extract and distribute (transport) them.

Platform An artificial structure designed to drill for hydrocarbons and/or produce (extract) and distribute hydrocarbons offshore in water depths up to 3 km. A platform can either be floating, semi-submersible, or fixed to the seafloor.

ROV Remotely Operated Vehicle (ROV) is a remotely controlled underwater vehicle of variable size (from <1 m long, up to about 3 m in length). The vehicle is normally fitted with propellers (thrusters), lights, cameras, manipulator arms, and other sensors and devices depending on its operational task.

Subsea template A structure normally constructed of steel tubing, designed for a variety of purposes within the OHI. A normal subsea production template has up to four wellheads and has typical dimensions of 20 m × 20 m × 10 m.

Trunk pipeline A pipeline designed to transport large quantities of natural gas or oil over long distances (up to 1000 km). Normally, they have diameters between 20″ and 44″ (inner diameter of the steel pipe). Before laid on the seafloor, they are coated with varying thicknesses of concrete coating for added weight, to prevent them from becoming buoyant.

Umbilical A specially designed flexible, multipurpose cable used for powering underwater equipment (including ROVs and subsea templates) and also used for sending and receiving control and sensor signals. Umbilicals can contain combinations of electrical cables and optical fibers.

Definition of the Subject

The offshore hydrocarbon, "oil," industry (OHI) searches for oil and natural gas by drilling exploration wells as deep as 10 km below the seafloor.

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_478

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC, 2012
https://doi.org/10.1007/978-1-4419-0851-3_478

When a commercial oil or gas field has been documented by such drilling, the exploitation of the resource will start by developing the field and the construction of production units and transportation infrastructure. Until only 15 years ago, this meant the construction of large, concrete-based or steel "jacket" production platforms. Because of intense research and technological development, many of the new offshore hydrocarbon fields are developed with smaller remotely controlled subsea steel structures placed directly on the seafloor, often without any infrastructure visible above the water. These new fields are produced remotely over distances of up to 150 km, with fiber optical cables, satellite communication, umbilicals, and pipelines.

In contrast to the other main (traditional) offshore industry, for example, the fishery industry, the OHI has employed strict environmental rules and regulations, which are efficiently practiced in most countries. These ensure little harm to sensitive marine organisms during normal field development and production. In addition to obeying the imposed rules and regulations, the OHI is, by tradition, constantly developing new and more cost-effective and environmentally friendly technology and infrastructure.

With knowledge and experience from 30 years of underwater detailed mapping and visual observations, mainly from the North Sea, spanning from predrilling seafloor surveys to annual surveys of pipelines and platforms, it is found that the marine life (the visual mega-fauna, at least) apparently benefits from the OHI-related installations on the seafloor. The reason being improved shelter conditions for large fish and also for spawning fish, and also an increased amount of energy (nutrients and seston) available in the water mass near these human-made structures, some of which act as artificial reefs [1]. The future needs for improved management of the marine biological resources, including the valuable deepwater corals and natural fish stocks, can be done by increased awareness of underwater life in general, via live video footage released to the public by, for example, the OHI. Furthermore, it also calls for academic scientific research into how best visual documentation of the seafloor can be used for an improved understanding of the complex underwater ecology and biodiversity change.

Introduction

In an increasingly energy hungry society, the quest for finding and exploiting underground oil and gas (hydrocarbon) resources is being continuously improved. Whereas the world's total onshore hydrocarbon production is gradually decreasing, the OHI is currently increasing its production volume. According to Maurer [2], ecological systems are complex and combine both idiosyncratic and unpredictable outcomes with strong constraints on system structure that makes them paradoxically both deterministic and unpredictable at the same time. Because of this, there has been no universal theory to guide research on ecological phenomena.

Over the last 50 years, the OHI has, unfortunately, inflicted several enormous oil-spills on the marine and coastal environments. There have at least been five such episodes that should never have occurred: The blowout and spill in the Santa Barbara basin, off Los Angeles (January, 1969), the Ekofisk Bravo blowout in the Norwegian-sector of the North Sea (April, 1977), the Ixtoc blowout, Mexican-sector of the Gulf of Mexico (GoM) (June, 1979), the Piper Alpha disaster, UK-sector of the North Sea (July, 1988), and lastly, but not least, the Deepwater Horizon blowout and disaster, US-sector of the GoM (May, 2010). Apart from these unfortunate, generally short-lived (less than 2 years), environmental inflictions, the OHI at large appears to be environmentally friendly, as will be discussed herein. This notion has been documented by extensive seafloor mapping and annual visual inspections of platforms, pipelines, and other infrastructure. Thus, rather than representing a threat to marine life in general, the OHI is, at least in the North Sea, a benefit to marine life in general. This is not only because, by its design, it protects numerous fish against industry fishing and trawling, but also because the large artificial underwater steel and concrete constructions represent geometrically complex structures in an otherwise mostly structureless seafloor environment. Furthermore, the industry is continuously improving its methods for underwater mapping, inspection, and monitoring of the environment.

This assessment of marine life associated with normal offshore drilling, pipelines, and platforms stems from over 30 years of unique visual observation by manned submarines (1977–1981) and ROVs (remotely operated vehicles), (1979–2010). It is based on the active participation and responsibility for conducting detailed mapping surveys of the seafloor, visual documentation, coupled with remotely sensed (geophysical data). The current experience covers large expanses of virgin seafloor, stretching from the Shtokman field at 73.6 °N, in the eastern Barents Sea, south to 51 °N, off Dunkerque, France. A total of 522 fixed production-related structures (platforms and subsea templates) have been installed on the Norwegian Continental Shelf, and over 7,000 km of trunk pipelines have been constructed in these regions during this time-span. Thus, there is a unique variety of first-hand specific knowledge that can be shared from the numerous site surveys of platform locations, pre-lay visual surveys (conducted before the laying of the long, trunk pipelines) of the seafloor, to annual inspections of the constructed pipelines.

However, the main difficulty is how to describe and disseminate this unique visual OHI-related underwater experience and information in a way that can be used by marine scientists in a quantitative manner. This task is envisioned to resemble that faced by the pioneering land-explorers after their long treks across previously unknown parts of the globe, during the "age of discovery," a couple of centuries ago. The narration, therefore, will be fragmentary, as most of what is observed on the seafloor is new, and as most of the water and seafloor bordering onto the visually observed space is virtually unknown, despite it occurring in some of the world's most fished and scientifically studied oceanic regions (the North Atlantic Ocean).

The Offshore Hydrocarbon Industry (OHI) and the "Second Surface"

A Brief History of the OHI

The onshore hydrocarbon industry started moving out into shallow waters sometime in the early 1930s, offshore Venezuela (Lake Maracaibo), offshore the states of California and Louisiana, USA, and in the Caspian Sea, offshore Baku, Azerbaijan. The first installations were simple steel and wood constructions built in knee-deep waters. However, their size and complexity was gradually increased with increasing water depth, up to several tens of meters. Simple steel jacket drilling towers were constructed and there were bridges and roads built on piled steel and wood foundations, often in a hap-hazard manner. After sometime, there were many accidents and mishaps, before improvements were made and special standards were invoked. The one single event that hit the OHI and aroused the world's environmental conscience was the big blowout oil spill in Santa Barbara, offshore Los Angeles, California on January 29, 1969. This also had immense consequences for the stricter regulations imposed on offshore drilling and the exploitation of offshore oil and gas. Even though no people were killed, this event made such a graphic impression on the population of southern California that in the following spring, "Earth Day" was born. Many consider the publicity surrounding the oil spill a major impetus to the environmental movement.

In Europe, the OHI started with the development of the UK southern gas fields off the east coast of England, in the mid-1960s. Here, the platforms and pipelines met the tough environment of the North Atlantic. New rules and regulations, British North Sea Standards were imposed. In 1967, the OHI moved even further north, in the North Sea, to Norwegian waters. The Norwegian Petroleum Directory (NPD) and The Norwegian state oil company, Statoil, were born, some years later. Although the giant oil field Ekofisk was developed with similar standards as in the UK southern gas fields and the Forties and Piper fields of the mid-UK North Sea, the Norwegian fields still further north, such as Statfjord, Gullfaks, and Troll, had to withstand even tougher environmental conditions. These fields, located at water depths between 130 and 320 m were therefore developed with giant concrete "gravity base" platforms, as the underwater technology evolution was not ready for moving delicate equipment like pumps, electronics, and gauges under water. When the Troll A concrete platform was towed out from Stavanger, and placed on the seafloor over the giant Troll field on May 17, 1995, it became the largest human-made structure ever to be moved. It measures a total of 472 m in height, from the top of the drilling tower to the bottom of the concrete skirts that penetrate 30 m into the soft clays at the location where it is still producing oil and gas, off Bergen, Norway. The platform houses about 200 workers who stay on board for 14 days at a time, rotating in and out by helicopter.

The rotation occurs all year round, even during the darkest and stormiest winter months (December through February). This platform produces about 18% of the total gas consumption of Germany and is, therefore, of immense value both for the owners (Statoil, Shell, and the Norwegian government) and for the consumers in Germany and surrounding countries. The gas is transported through 36″ (36 in.) and 40″ concrete-coated steel pipelines, welded together on board huge offshore pipe-laying vessels and placed carefully onto the seafloor. Such huge "trunk"-pipelines criss-cross from the Norwegian fields to processing plants on mainland Norway, and are rerouted from there, through other gas export pipelines to England, Germany, Belgium, and France. The largest pipe, the Langeled pipeline, was constructed between 2004 and 2006. It is a 40″ and 44″ diameter pipeline of 1,200 km length, which originates from the onshore processing plant at Nyhavna, south of Trondheim. From there, it runs south to the Sleipner field in the middle of the North Sea, and continues to Immingham on the east coast of the UK.

During the last 15 years, subsea technology has developed fast and most modern fields are constructed solely with remote-controlled subsea structures. The Snehvit field in the Barents Sea is, for example, produced through three subsea templates (steel structures) with several production wells in each template. The field lies 135 km from shore and is remotely operated from the onshore production plant at Melkøya near Hammerfest, the world's northernmost city. At present, Norway is the third world's largest exporter of crude oil, and it runs 522 offshore fixed production-related structures of which 365 are subsea and fully submerged. In year 2009, Norway also exported a total volume of 96.6 billion Standard cubic meters (Bn Sm^3) of natural gas to Europe through the trunk pipelines.

The Second Surface of Earth

The seafloor is the "Second Surface" of Earth, indicating that it is hidden in many ways. It covers an area which is about three times larger than the visible land surface. Most of this surface is still unknown – because water is a "black body" substance when it comes to the electromagnetic spectrum. It absorbs most of the visual light that encounters it. Therefore, in contrast to sound waves, the light rays have very low transit ranges through water. The photic zone of the ocean, into which solar rays can penetrate, are reckoned to be down to a maximum of just over 100 m in the clearest waters, that is, water with little seston and other particles. This is not very deep, considering the mean depth of the ocean being about 3,500 m. Because of this lack of efficient visual access, the Second Surface is only beginning to be explored in detail. This surface ranges in depth, from 0 m at the coastlines, to about 300 m on the continental shelves, then down to 5,000 m (5 km) on the great abyssal plains, to over 11 km in the deepest trenches. Because most of Earth's surface is covered by water, the Second Surface represents a very significant and essentially important entity. So far, remotely sensed (acoustic) surveys only cover about 10% of it (i.e., indirectly, with data that needs interpretation by geophysicists) and only less than 1% visually (i.e., directly with cameras). This means there exists more visual documentation of the surfaces of both Moon and Mars, than of the immensely more important Second Surface of Earth. However, from sediment sampling, fishing (trawling), scientific scraping (dredge sampling), and drilling, in all oceans, over time, it is currently known that the Second Surface mostly consists of mud (clay), sand, rock, and in some areas metals and salts. But, because being flooded by water, it is both pressurized, and buoyed at the same time, and behaves accordingly, which is often totally different to the well-known land surface. Furthermore, on average, the Second Surface has a much thinner crust than the onshore continental crust and is more likely to be exposed to high heat flow from the Earth's interior. Along the tectonic plate boundaries (mid-ocean spreading zones and subduction zones), the high heat flow induces underground convection fluid currents and the venting of warm fluids in hydrothermal vent systems [3].

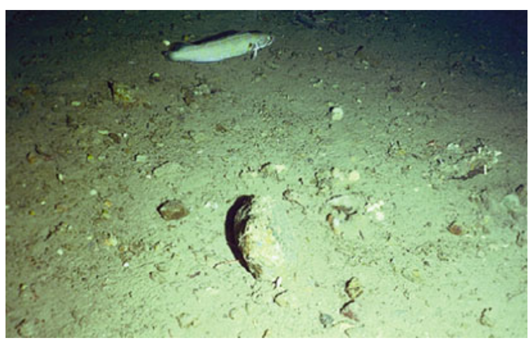
Although academic research institutions and consortia, such as the Integrated Ocean Drilling Program (IODP), perform many types of investigations at water depths to about 6 km, the mining industry is working at depths to about 5 km, and the fishing fleet is gradually trawling to depths greater than 2 km, the OHI is currently working to water depths of about 3 km. However, about 90% of its activity still occurs at water depths between 100 m and 1,500 m. These are the water depths, therefore, which will be addressed herein. From the extensive mapping-, construction-, and visual inspection-work performed by OHI at these depths, several places in the world, there is some general and also specific knowledge about processes and marine life that can be disseminated, including some new discoveries.

Whereas the fishing industry tends to operate in a “blind” mode when it comes to the seafloor, the OHI naturally operates with more caution, partly in order of preventing damage to sensitive and costly equipment and structures, and partly because of law enforcement (at least in US and European, including Norwegian waters). Thus, no drill-site is drilled without a proper predrilling assessment of the seafloor, whereby any significant physical obstructions and known sensitive organisms, including chemosynthetic fauna and coral reefs are documented beforehand. The problem with the bottom trawling of the fishery industry is the insensitivity to what is down there, that is, “indiscriminate obliteration” on the seafloor. As long as such bottom trawling is legal practice, all sessile organisms in the world are actually threatened by trawl-board disturbance, at least those living at water depths shallower than 2,000 m water depth.

The General Background Seafloor Life

In order to set the OHI-related marine fauna observations into perspective, the general background seafloor has to be described. In the depth interval 100–1,500 m, the general background seafloor (about 90% of the total area within this depth range) is drab and appears relatively “uninteresting” (Fig. 1), just like the enormous sand fields of a desert on land. However, in most areas, the drabness is spotted with small hubs of life, like oases in the same desert. There are vast areas with level, muddy bottoms. Several studies have shown, however, that, for the deep-sea biota, there is a distinct decrease in population differentiation and species diversity with depth [4, 5]. Any large erratic boulder, rock outcrop, or wreck is colonized by invertebrates, which seem to attract also other marine life, including fish (Fig. 2).

Piepenburg and co-workers [6] used classical marine biological methods to study patterns and determinants of the distribution and structure of benthic faunal assemblages in the northern North Atlantic. Using a suite of sampling methods including corers, trawls, and seabed imaging (benthic spot photography), they managed to “adequately probe various benthic community fractions, such as foraminifers, poriferans, macrobenthic endofauna, peracarid crustaceans, and megabenthic epifauna.” The general patterns they found were, not unexpectedly, a depth zonation, and also a significant decline in biomass and abundance by as much as two and three orders of magnitude. These were the most conspicuous general patterns detected. However, in terms of species richness, no common trend for water



Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 1 The general seafloor off Mid-Norway looks like this. In the background is a Tusk (*Brosme brosme*). The stones (cobbles and gravel) have been washed out of the underlying clay-dominated till. Some of the larger stones can be seen to be colonized by invertebrates of different colors

Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 2 A large erratic boulder located inside a pockmark crater at 280 m water depth off the Island of Fugløy off Northern-Norway, north of the Polar Circle. The organisms seen colonizing the boulder are mainly filterfeeders, sea anemones, and serpulid tube worms

depth or latitude was perceivable. They especially studied the East Greenland continental shelf margin between 68 °N and 81 °N at water depths between 40 and 3,700 m. Here they found relatively productive hydrographic zones being the marginal ice zones, polynyas, and anti-cyclonic gyres. They interpret this as being evidence for the importance of water column processes for subsequent food availability being the major determinants for the benthic assemblages and the significance of pelago-benthic coupling in the study in general [6]. This is not surprising, food availability being the most necessary ingredient for life in general. When it came to the distribution of megafaunal species, such as echino-derms, it was found that community patterns on a 10 km scale and the dispersion of organisms on a 100 m scale were best explained by the seafloor properties. This means that macrofauna is dependent on structure and type of the seafloor sediments and topography. Furthermore, they found no evidence for a direct pelago-benthic coupling, irrespective of water depths. These contrasting findings emphasize that the relative importance of potential community determinants can change with both spatial scale and life traits, for example, body size, mobility and feeding ecology, of organisms considered [6]. Thus, the stage is set for a narration of discoveries made by the OHI and the associated biological and physical research performed for this industry by academic researchers and institutions during the period 1980–1998. After 1998, the academic institutions have picked up many of the research leads pioneered by OHI-activity within marine geophysics and biology.

Unique Processes and Biotypes Initially Studied due to OHI-Activity

Since the late 1960s, new processes and features have been discovered on the Second Surface. Some of them are completely unique to the underwater world at water depths of 100–1,500 m. Perhaps one of the most surprising revelations is that this type of mapping and areal seafloor documentation became instrumental for the discovery of several previously unknown natural conditions of the seafloor. Thus, at least three unique discoveries were made as a result of such surveys: (1) A biological seepage relationship, (2) The discovery of pockmark craters and their potential significance for marine life, and (3) The discovery of myriads of large, cold-water coral reefs (also called deepwater coral reefs). Although the results were not possible without cooperation with academic institutions, especially in the UK, USA, Germany, France, and Norway, the pioneering discoveries were instigated due to OHI-activities, mainly in the Gulf of Mexico and the North Sea. The discoveries are mentioned here, as they sometimes are relevant to the marine life observed on the seafloor. These processes and features are:

1. Venting of reduced organic fluids (fluid flow, or "cold seeps")
2. Crater formation, by fluid flow
3. Bioherms, including the cold-water coral reefs

One of the recent wide-scope books on benthic life in the North Atlantic [7], actually fails to mention the existence of prolific deepwater coral reefs occurring there. In the book, which is aptly titled: "The Northern North Atlantic – A Changing Environment" neither deepwater coral reefs, *Lophelia pertusa*, nor "*Lophelia*-reefs" are found in the index, or at all mentioned in the text. Why is it that thousands of large coral reefs, some known to science for at least 200 years and from the early 1990s published by OHI-related scientists, manage to avoid mention (recognition) in such apparently authoritative scientific literature? Couldit bethat only "Classical marine biological" results are recognized? The publisher claims that: "the Greenland-Iceland-Norwegian Seas can now be considered one of the best studied subbasins of the world's oceans" [7]. But, even so, the information published in this book is important, as it provides the necessary background knowledge about life in general on the seafloor "desert," outside the coral reefs.

Venting of Reduced Organic Fluids Generally, the ocean floor is covered in thick sediments that deposit by gravitation, with particles sinking through the water column and accumulating in thick layers on the Second Surface. The fluids, including petroleum gas and liquids (hydrocarbons) trapped underneath such sediments are lighter than the solids and, therefore, move upward to surface at discrete locations due to buoyancy. This process is also called "migration" and where the flow penetrates the Second Surface from below, it is called marine fluid flow [8]. The discrete locations where the fluids occur at the surface are called "cold seep" locations. Depending on the geological setting, the distance between each cold seep location on the seafloor varies considerably, from kilometers or miles, to only several meters. However, cold seeps are important for life within, on, and above the Second Surface because they represent transport pathways for dissolved chemical constituents and sustain unique oasis-type ecosystems at the seafloor [9]. Fluids expelled through seeps contain re-mineralized nutrients (silica, phosphate, ammonia, and alkalinity) and hydrogen sulfide, as well as dissolved and free methane from microbial degradation of sedimentary organic matter. Because methane gas molecules (CH_4) have the highest relative hydrogen content (four hydrogen atoms to one carbon atom) of any organic compound, it represents a valuable energy source to certain primary producers: archaea and bacteria, that is, the methanotrophs and the methane oxidizers. Apart from near-cold seep locations, seawater has generally very low concentrations of methane and other light hydrocarbons, such as ethane (C_2H_6), propane (C_3H_8), butane (C_4H_{10}), and pentane (C_5H_{12}). Perhaps the single most important reaction associated with cold seeps is the anoxic oxidation of methane (AOM) by archaea and sulfate reducing bacteria (SRB), with secondary reactions involving the precipitation of carbonate ($CaCO_3$), in the form of inorganic aragonite and calcite [9]. The OHI has long been interested in these seafloor processes, for various reasons, not least because they contain tell-tale indications of where deep-seated hydrocarbons (reservoirs) may be found.

During a predrilling geophysical site survey, in 1977, at the Tommeliten field in the central North Sea at 78 m water depth, side scan sonar data showed numerous bubble streams emanating from the seafloor, immediately above a buried salt dome [10]. Subsequently, this location was investigated with ROV by Statoil in 1983. The gases leaking naturally through the seafloor were documented to be the reduced organic light hydrocarbon gases (methane to pentane) which also continually charged the upper, porous sediments. An intimate relationship was found between organisms, such as anthozoans and the visual bubbling of gas through the sediments [11]. This visual inspection and sampling of the naturally leaking gas produced several interesting results: (a) documenting small "reefs" or "bioherms" consisting of many kinds of filter-feeders, scavengers, and predators occurring adjacent to the seeps; (b) small depressions (so-called eyed-pockmarks), in the seafloor, with high-density macrofaunal communities in their centers, and (c) white patches of bacterial mats also occurring over relatively large seafloor areas, where the sediments were charged with gas seeping up from deeper layers.

Later studies of this Tommeliten site also showed two other important aspects, relevant to marine life: (1) that the bacterial mats were easily torn apart and carried up into the water column by slight disturbances of the near-bottom water by the ROV and (2) that the seafloor had been partly cemented by methane derived authigenic carbonate rock [8, 11] both within the eyed-pockmarks and elsewhere on the otherwise flat sea-floor, near the bacterial mats [12]. Academic research at this active seepage site and also at one similar site near the Gullfaks field, further north, identified a microbial community dominated by sulfur-oxidizing and sulfate-reducing bacteria (SRB) as well as methanotrophic bacteria and archaea. Stable carbon isotope values of specific, microbial fatty acids and alcohols from both the Tommeliten and Gullfaks sites were found to be highly depleted in the heavy isotope ^{13}C, indicating that the microbial community readily incorporated seeping methane or its metabolites [13].

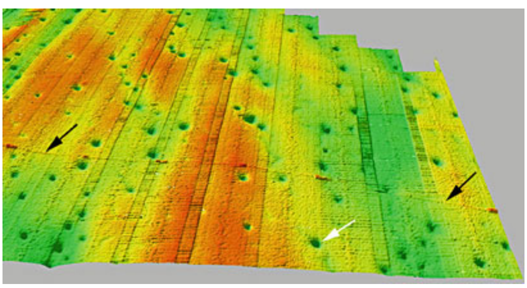
At Tommeliten and Gullfaks there is, therefore, no doubt that the dense bioaccumulations on the seafloor, including, the bioherms, are a direct result of seeping gas (energy and nutrients) from deep below the sea-floor. This may also be one of the reasons why there is plenty fish and heavy trawling activity at the Gullfaks location [14]. Although most marine ecologists, environmental scientists, and biologists are used to assess traditional phytoplankton concentrations dependent only on top-down linkages in the food chain, the modern seep studies find more and more bottom-up links [13, 15].

Mud volcanoes are locations where fluids and solids (water, mud, gas, and petroleum) well up through the Earth's surface, driven by overpressures in the subsurface. They both occur on the land surface, as in Azerbaijan [16] and many places on the Second Surface [17, 18]. From studies and surveys at the underwater Håkon Mosby Mud Volcano (HMMV) located at the boundary between the Barents Sea and the Norwegian-Greenland Sea, a relatively prolific and complex ecology has been found. Although only a minor portion of this ecosystem relies directly upon chemosynthetic energy, this portion is probably very important for the sustenance of the system. A simplified food chain for HMMV was published by Vogt et al. [19]: The primary producers are suspected to be methanotrophic bacteria and anoxic methane oxidizers (archae), besides other "conventional" microorganisms relying on added heat and continuous sediment disturbance by the turbid flow of mud from the mud volcanic vent. The secondary consumers were made up of benthic suspension feeders and deposit feeders, such as asteroids, holothurians, etc. In addition, there was a chemosynthetically based food web, relying on symbionts, such as the pogonophoran tube worms, that host endogenic chemosynthetic bacteria [10, 19]. The primary consumers consist of filter-feeders and other predatory invertebrates and vertebrates, such as sponges, crinoids, pycnogonids, basket-stars, and fish, mainly consisting of Eelpout and Skate (Raja spp). A maximum density of one Eelpout for every square meter was documented above portions of the mud volcano [19]. This example also demonstrates the very important influence of substances originating from below the seafloor in modifying and "fertilizing" the immediate seafloor environment, to result in enhanced productivity.

Seabed Pockmarks: Craters Formed by Focused Fluid Flow On the land surface it is often very difficult to detect surface degassing (seepage) events, apart from those associated with water and mud flows, such as at mud volcanoes [20]. On the sediment-covered Second Surface and in some lakes, however, the situation is different. In the late 1960s, numerous craters were found off Nova Scotia. They were called "pockmarks" by their discoverers, Lew King and Brian MacLean of the Bedford Institute of Oceanography (BIO), Canada [21]. Today, it is known that pockmarks occur in certain portions of the seafloor, the world over, and even in some lakes [8, 10]. Also a very close relationship between the pockmarks and local increase in visible seston (including plankton) was found [11]. Although pockmarks ranging in size from the small, "unit-pockmarks" (<5 m diameter, 1 m depth), to normal-pockmarks, complex-pockmarks, and giant-pockmarks (up to 500 m in diameter and 30 m depth), have been found, very little is known about their formation and sustaining mechanisms [8, 22]. Although they are known to be formed as a consequence of buried gas reservoirs and fluid flow, there are only very few pockmarks known to be producing continuous visible bubble streams [23].

After having noticed the marine life occurring inside some of the pockmark explored with ROVs in the North Sea [10], it was recognized that the "classic" (conventional) science of marine geology would probably never have discovered pockmark craters (Fig. 3). The reason being, that this science relied on acquiring numerous spot samples of the seafloor using corers and grab samplers, and determining the nature of the sea-floor mainly by measuring the grain-size of the spot sampled sediment grains. Thus, if only one seafloor sample was acquired at 1 km spacing in the Norwegian Trough of the North Sea, where there is an average of 15 pockmarks per square km, it is likely that only one sample out of 100 would happen to sample inside a pockmark crater. This sample would probably turn out to be anomalous (i.e., containing unexpected biota and sediments). Thus, the chances of this sample (one-in-a-hundred) being discarded or recorded as an "anomaly" (or only of "curiosity value") are great. At least it would not turn out in the statistics of the survey, and the chances of it being taken seriously are therefore meager.

From modern seafloor remote sensing with multi-beam (swathe) echo-sounders to long-range side scan sonars, it is known that the seafloor can be covered in a high density of pockmark craters. The density ranges from zero to more than 20 pockmarks per square km of mapped seafloor (Fig. 3). Because recent studies conclude that



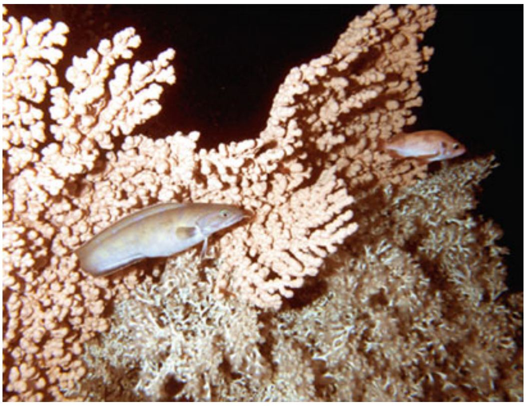
Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 3 In the 300 m deep Norwegian Trough area of the northern North Sea, there are thousands of normal-pockmarks. They are shown here on a perspective image of a seafloor relief map, with artificial vertical enhancement (x5). The size of the largest pockmark craters is 100 and 8 m depth. The *white arrow* points to a normal-pockmark of about 70 m diameter. The *black arrows* point at a trunk pipeline of 30″ diameter. Lines running north south represent noise in the digital data set, due to inaccuracies in sound velocity and tidal correction

pockmarks form due to focused fluid flow [22, 23], the seafloor can generally be divided into "hydraulically active" and "passive" areas. Thus, the active seafloor will have seep manifestations, like pock-mark craters and the passive areas will be devoid of craters.

There are two important corollaries to this "hydraulic theory": (a) It is valid for all volumes of soil, which have a porosity system partly filled by liquid and partly filled by gas. It is, therefore, also valid for all ocean depths, lakes and swamps, as the driving gas-type (methane, carbon dioxide, hydrogen sulfide, or hydrogen) is immaterial. (b) Because seeping fluids through the seafloor can be regarded as enhanced energy input to the marine fauna (primary producers especially), the enhanced hydraulic activity manifest by seeps and high density of pockmarks indicates the likelihood of enhanced marine biological productivity.

Cold-Water Coral Reefs of the North Atlantic It is known that the presence of solitary corals, sea pens, sea lilies, and sponges on the deep-sea floor offers rare, firm substrates for sessile organisms in an otherwise generally featureless environment. The relative importance of such biotic habitats for the local biodiversity may, therefore, be greater for the deep-sea than for shallower regions [24]. Those modern reef structures that seem to defy all normal reasoning with respect to location and environment are the deepwater coral and cold-water coral reefs (also named "ahermatypic colonial scleractinians"). In the North Atlantic and Gulf of Mexico, the most common types are those built by the *Lophelia* sp. stony coral, found on both sides of the Atlantic Ocean, including the Reykjanes Ridge, south of Iceland (Fig. 4). Because they are also found as far north as the Polar Circle, in the Barents Sea and off Mid- and Northern-Norway, they must somehow be independent of seasonal sunlight variations. Some of these reefs, found off Ireland and on parts of the Norwegian Continental Shelf had been known to fishermen and biologists for over 200 years [25–30].

Even so, the visual documentation by OHI-surveys during the 1980s and mid-1990s of thousands of large, deepwater coral reefs was completely unexpected to most marine biologists [31–36].



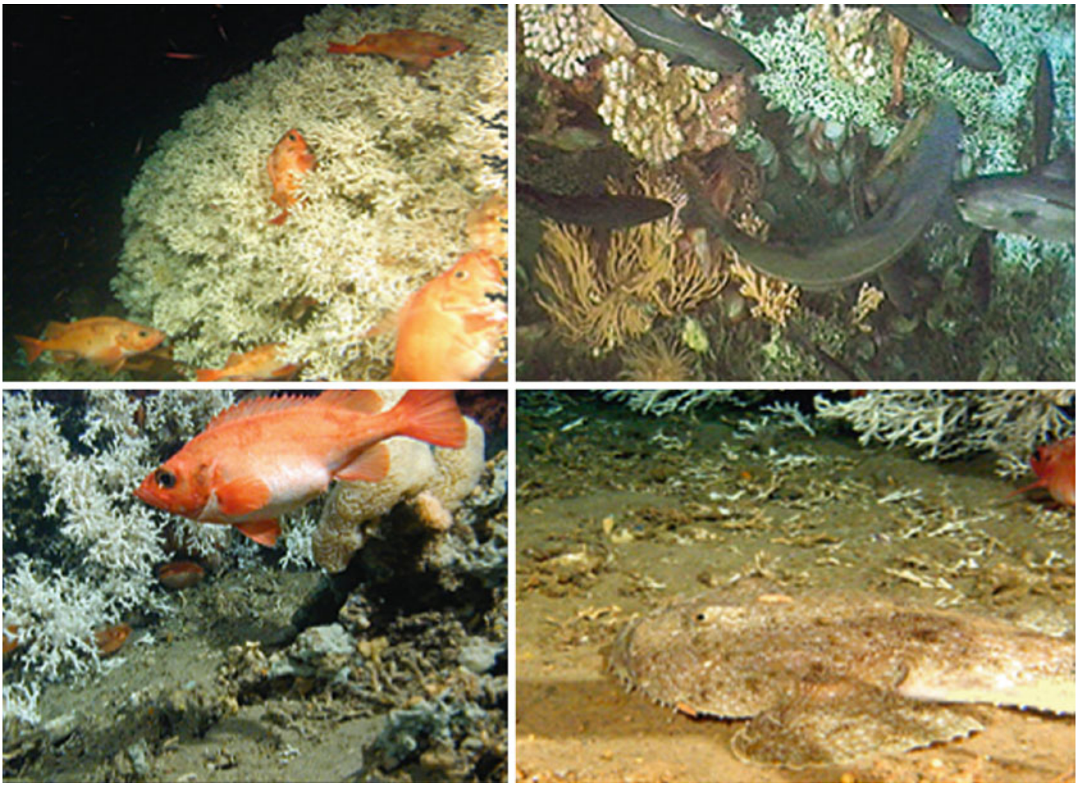
Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 4 Photograph from a typical Norwegian deep-water coral reef showing two types of corals: the soft-branched octocoral *Paragorgia arborea* (closest) and the reef-forming ahermatypic stony coral *Lophelia pertusa*, (*white and pink*). The two most common reef-related fish are also shown here, the Tusk (*Brosme* sp) and the Redfish (*Sebastes* sp) [47]. The height of this coral reef is about 3 m above the surrounding, undisturbed, even seafloor

Their modern rediscovery with broad-scale geophysical mapping and detailed video footage and photographs became a major eye-opener – a revelation that caused an improved reassessment of the complex and dynamic biological production of the North Atlantic Ocean. This modern documentation actually started in June, 1982, with an OHI-related detailed investigation for a potential pipeline route from the Askeladden field in the Barents Sea to Norway. During which, a 15 m high and 50 m wide *Lophelia*-reef was found and visually documented by Statoil [31, 33]. Further pipeline route surveys off Mid-Norway between 1985 and 1990 documented hundreds of similar reefs, until then unknown to science. Until June, 1991, when Statoil invited researchers from five Scandinavian academic and Norwegian authorities to view their unique data set, consisting of detailed geophysical and photographic/video results, these biotic structures had mainly been treated as curiosities and oddities on the seafloor [37]. But, even after this seminar, it was not the Norwegian scientists who first managed to mobilize an academic detailed visual study of the reefs, but the Germans. Freiwald et al. [34] had previously been studying carbonate secreting algae along the northern Norwegian coast. Having learnt from Statoil where the reefs occurred, they targeted them during their next planned cruise, already in 1992 [34]. The main surge of academic interest in the reefs of the North Atlantic came immediately after the publication of large carbonate reefs off Southwest Ireland. This was also partly based on OHI-related exploration geophysical data [33, 35, 38]. Subsequently, the first International conference on cold-water corals was staged in Canada, 2001 [39].

One of the main and controversial questions still remaining to be answered with respect to these impressive biological structures is why they occur in deep and cold water, even north of the Polar Circle, where there is hardly any photosynthesis occurring during the winter months. Even though there is ample evidence suggesting that they rely on extra nutrients and energy originating from below ground, that is, the "hydraulic theory" [11, 31, 35, 36, 40, 41], modern marine biological research has still not found sufficient evidence to support this theory. The main theory of the modern marine ecologists is, according to Buhl-Mortensen et al. [24], that they exist in high latitudes and deep waters because this is where the right water masses occur: "…it is not the deep water par se, but the distribution of intermediate and deep water masses that controls the bathymetric distribution of these corals. Corals typically create habitats reaching from decimetres to meters above the surrounding seabed and occur on mixed bottoms in areas with relatively high currents." [24], and the argumentation continues: "Colonial scleractinians need hard substrate for settlement. This substrate can be a shell or a pebble, and as soon as one colony is present it provides new hard substrate for subsequent colonisation" [24]. So, the question remains as to why only less than 0.1 per mille (‰) of the total area in the depth zones where they occur is covered by cold-water coral reefs? Why are there, for example, no more of them in the Norwegian and New Zealand fjords where the distribution of intermediate and deep water masses is right, and where there is ample suitable hard substrate (rock bottom) with high current speeds? However, some recent microbial studies seem to point the way further. Yakimov et al. [42], for example, recently found metabolically active microbial communities associated with deepwater corals in the Mediterranean. Also recent OHI-related research of the microbial food chain surrounding the coral reefs off Mid-Norway has provided some interesting new findings. The contrast between coral-associated and free-living bacteria may suggest that few free-living bacteria are directly ingested by the coral and that instead, corals feed on non-bacterial plankton. Small (100–200 mm) zooplankton has been suggested important in the diet of corals [43]. In addition, the tissue-associated bacterial communities potentially provide a direct translocation of nutrients through metabolism of particulate and dissolved organic matter in the seawater. One *Lophelia pertusa* associate was studied in more detail and named "Candidatus Mycoplasma corallicola" [44]. This bacterium was abundant in *L. pertusa* from both sides of the Atlantic Ocean and is considered an organotrophic commensalist [44]. Given the importance of chemo-synthesis in deepwater ecosystem development and functioning, cold-water coral reef communities may be linked to a diversity of chemoautotrophic microorganisms that synthesize organic compounds from inorganic compounds by extracting energy from reduced substances and by the fixation of dissolved CO_2. Just a tiny fraction of microorganisms associated with deep-water coral reefs have yet been identified, and even less assigned to a function. Although no nutritional symbiosis based on chemosynthesis [45] are known to have been documented on deepwater coral reefs, primary producers affiliated with chemoautotrophs (utilizing H_2S, NO_2^-) and methanotrophs (utilizing CH_4) have been found associated with the reef animals and their ambient environment [41, 42, 46]. Thus, also light hydrocarbons can probably stimulate the growth and the high biodiversity found on the *Lophelia* reefs associated with some Norwegian hydrocarbon fields [47]. Only further detailed studies of the reefs will be able to answer these important questions.

The cold-water coral reefs and carbonate mounds represent exceedingly valuable habitats for numerous species, besides the corals themselves. Also for fish, they represent shelter and nursing homes for juveniles, as one of the few comparative studies of on-reef versus off-reef fish-counts for deepwater coral reefs documents. Fish species' richness and abundance was found to be greater on the reef than over the surrounding seabed, as 92% of species, and 80% of individual fish were associated with the reef. The results indicate that the reefs have a very important functional role in deepwater ecosystems as fish habitats [48]. In particular, visual monitoring of coral reefs at the Morvin and Kristin fields, off Mid-Norway, has shown that the redfish (*Sebastes* sp.) probably spawns at the *Lophelia* colonies (Fig. 5). During the month of May, numerous fish with wide bellies, obviously ready to spawn (this species spawns live sprats), congregate very close to the *Lophelia* colonies (Fig. 5). The juveniles can immediately after spawning find full protection against predator fish within the complex structure of the live and dead *Lophelia* skeleton meshwork. Furthermore, the Norwegian coral reefs have long been known to fishermen as "Uerbakker" [26], meaning "Redfish slopes," and Furevik et al. [49] were the first to report scientific evidence that long-line catches of redfish, ling, and tusk can be significantly greater on the reefs than in off-reef areas. In addition, Husebø et al. [50] set long-lines in coral habitats and found significantly more fish than elsewhere and also that the fish were generally larger than those caught in the non-coral areas. The importance of these habitats and their internal ecological dynamics has been discussed in more detail by Buhl-Mortensen et al. [24].

According to authoritative assessments of possible threats to the cold-water coral reefs, the main threats according to Roberts et al. [51] are: (1) Bottom trawling, (2) Hydrocarbon drilling and seabed mining, and (3) Ocean acidification (i.e., global climate change [52, 53]). In addition, for each individual coral reef, there is always the possibility of dramatic environmental changes by natural causes, such as nearby underwater avalanches (burial), and in the case of the hydraulic theory being viable, that the seepage or venting is naturally depleted or exhausted, and becomes



Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 5 Fish sightings on *Lophelia*-reefs at Norwegian hydrocarbon fields Kristin and Morvin. *Upper and lower left: Sebastes* sp. (redfish) congregating for spawning on a 1.5 m tall *Lophelia* colony. Upper right: Seithe feeding near a *Lophelia*-reef at Kristin. The fish often swims around the ROV when operating off Mid-Norway. *Lower right*: A monk fish resting on the seafloor next to a *Lophelia* colony at Morvin. Notice its white protrusions under its head that resemble the *Lophelia* branches. This fish is sometimes found to be resting on top of *Lophelia* colonies, where it is well camouflaged

"turned off." However, the OHI seems to come out of such assessments far better than the fishery industry: "Compared with widespread evidence for physical damage to reef structures from bottom trawling, there is little evidence that hydrocarbon exploitation substantially threatens cold-water coral ecosystems. *L. pertusa* colonizes North Sea oil platforms and seems to have a self-seeding population, despite proximity to drilling discharge. Greatest concern is over the potential for drill cuttings to smother reef fauna, but such effects would be highly localized when compared with the extent of seabed affected by bottom trawling" [51]. This latter scenario was carefully tested by a 3-month-long drilling and monitoring campaign of four production wells at the Morvin field off Mid-Norway, where numerous coral reefs exist on-site. The results of this campaign are reviewed in a later chapter.

To sum up the threats, based on OHI-related observations and modern publications, it is found that the coral reefs thrive, despite them being located close to OHI-structures and to sporadic drilling activity. They even colonize parts of the OHI-platforms, as seen in Fig. 5, from the Statfjord field [47, 54]. The main threats to the deepwater coral reefs are therefore mostly mechanical, as they are delicate structures and cannot sustain the mechanical indiscriminate stress imposed by portions of the fishing industry and to a much lesser extent, the OHI. This has been amply documented off Norway and off New Zealand [55–57]. Furthermore, based on abundant visual and geophysical data mainly acquired from the OHI, it was possible for Norwegian authorities to be first to officially protect and conserve a large portion of cold-water coral reefs. The first area to be protected against bottom trawling was the Sula Reef in the late 1990s and the Haltenpipe reefs off Mid-Norway with a 970 sq. km large protected area [36]. Today, awareness of the ecological significance of deepwater corals is growing rapidly, as it is known that colonial corals provide important habitats and could play a critical role in the life history of many marine species, including fish of commercial interest [58, 59]. This awareness has lead to a general call for the establishment of marine protected areas (MPAs) and especially to protect the most important cold-water coral habitats [60].

Marine Life Affected by the OHI

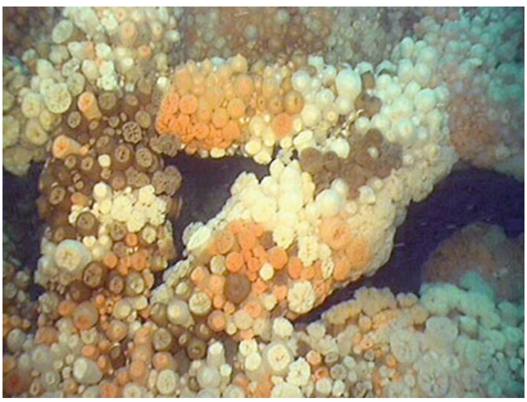
Methods of Observation

Before an offshore hydrocarbon field is developed, rules and regulations say that biological baseline studies shall be carried out. Over the last two decades, the focus of such studies has shifted from being purely based on detecting pollution effects from OHI-activity to also include assessments of biodiversity and ecology [61]. A standard outline of a station pattern for sampling seafloor biology is either a grid, covering a large area, or samples obtained along lines forming a cross, with the longest axis downstream with respect to the prevailing current. Several replicate samples are collected at each station. Because this type of sampling was found to be relatively "blind" to any local features on the seafloor, such as pockmarks and cold-water coral reefs, it has now been complemented with an initial seafloor mapping campaign. The macrobenthos fauna of interest in baseline sampling surveys comprises the following main taxonomic groups: Polychaeta, Crustacea, Mollusca, Echinodermata, and Varia (remaining groups). Only animals larger than 1 mm (macrobenthos) are included in such analyses. Subsequent to the development and production at offshore hydrocarbon fields, the sampling of sediments and macrobenthos is repeated perhaps every other year.

The study of benthic communities can provide an indication of pollution from offshore activities. Similarly, the geochemical analysis of sampled sediments can provide levels of pollutants and their distribution around the offshore installations. Any changes in species composition and densities of individuals can also provide tell-tale signals about damaging pollution. Experience from many years of such work has proved that the benthic fauna near OHI-installations can be affected by a number of factors, including the discharges of drilling fluids, cuttings, and others, including the accidental release of oil and also physical disturbances. One of the current challenges of planning future environmental/baseline surveys in the OHI is to improve the cross-departmental communication, as the detailed seafloor mapping is performed by others than those responsible for environmental surveys [61].

During initial mapping and inspection along potential pipeline routes at the Morvin field, off Mid-Norway, numerous coral reefs were found and visually inspected. Here, it was also decided to perform a detailed geochemical investigation prior to the production drilling activity. One of the largest reefs, the MRR ("Morvin Reference Reef") was systematically investigated and sampled (water, sediment, and organisms). This reef is located inside a large (130 m × 80 m × 10 m) pockmark depression (Figs. 6 and 7). The reef occupies about one third of the pockmark, growing from the maximum depth (at 370 m), up along the northern side, to the rim of the depression (at 360 m water depth). The MRR is about 80 m long, 25 m wide, and spans the elevation interval: 360–370 m. *Lophelia pertusa* is the main reef-building coral [47]. The geochemical analyses of sediments at Morvin proved that they contained varying concentrations of light hydrocarbons (methane-butane). Because the sediments in which these hydrocarbons were sampled are located within the oxygenated upper portion of the seafloor (i.e., only 40 cm below surface), any hydrocarbons remaining there over long time would have been reduced (oxidized) relatively rapidly. Therefore, there must be a natural seepage of light hydrocarbons from below, via molecular and fluid migration. This type of fluid flow is called "micro-seepage" of light hydrocarbons [8, 62].

Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 6 Special features on the seafloor attract and protect large fish. Here is an aircraft wreck from World War II off Mid-Norway, which is now utilized as shelter for large Lings. At least five larger than 1 m length can be seen here. The wreck was detected with side scan sonar and inspected by ROV as part of a site survey for exploration drilling



Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 7 A subsea steel template structure in the central North Sea, Danish Sector. The steel structure has been completely colonized by sea anemones. This is also a perfect hiding place for fish of all sizes

However, because no more than eight samples were acquired at Morvin, it is not possible to state any significant statistical variation over a regionally significant area.

The highest interstitial hydrocarbon concentrations were found in the upper sediments at location S8, which is inside a unit-pockmark, located just to the north of MRR (Fig. 4). Combined with modern results from molecular biological methods, there is here factual support for the notion that a nutrient-rich, “fertile” substratum represents one of the keys to understanding the location of this deep-water coral reef. Previous studies of *Lophelia* O-C stable isotopes [63, 64] show that there are relatively large variations in the dC -value (ranging from — 2‰ to — 9‰ PDB). For corals, these negative values have so far been interpreted as being caused by ambient pH-variations, among other factors. However, similar variations in stable isotopes in bivalves at seep sites in the Gulf of Mexico were interpreted as being caused by hydrocarbon uptake in microorganisms, subsequently ingested by the bivalves. This suggests that future studies of the coral reefs should also include systematic stable isotope analyses, combined with more geochemical sediment analyses.

The Impact of Exploratory Drilling

Scientific Drilling When the Deep Sea Drilling Project (DSDP) initiated its research with the drilling vessel “Challenger,” in 1974, it found indications of oil at nearly 3 km water depth on its second drilling hole on the first exploration leg (DSDP, Leg 1, Hole 2). When the first cores came onto the deck, there was an unmistakable smell of crude, and small droplets of oil could clearly be seen in the carbonate rock sampled. The objective of this hole was to determine the nature of the so-called Challenger Knoll on the abyssal plain of the mid-Gulf of Mexico. By drilling 144 m into this feature, the scientists determined it was a carbonate capped salt dome. However, because the DSDP was a pure scientific drilling project, it should not be looking for oil. A panel of experts was immediately established to make sure that the coming drilling legs should not target hydrocarbon-prone areas. This is how the “Pollution Prevention and Safety Panel” of the DSDP and later the Ocean Drilling Program (ODP) was borne. Today, this panel of experts of the IODP (Integrated ODP) has been re-named to EPSP (Environmental Protection and Safety Panel).

After exploration drilling for oil had come underway, on the Norwegian Continental Shelf (NCS), during the late 1960s, strict rules and regulations were imposed on the industry. For example, no drilling was allowed on the NCS unless a special "predrilling site survey" was performed. This includes a full seafloor coverage of side scan sonar acoustic images, which is able to detect all types of features on the seafloor larger than about 3 m in extent, such as wrecks, munition, rock outcrops, and many other features, including man-made features on the seafloor.

OHI-Related Drilling On the Norwegian Continental Shelf, it has been normal practice to discharge the sediments produced from drilling ("cuttings") and drill-mud from drilling the first 600–800 m of the well (the "top-hole") directly to the water column, where they are dispersed by currents. The largest particles (about 90% of the cuttings) will then be deposited in a slight mound within about 200 m of the well location and will form a slight mound on the seafloor. Not only do deepwater coral reefs challenge modern science with numerous questions on how they can exist in the middle of a seemingly barren shelf environment [47] – they also present a challenge on how to drill safely without damaging them. At the Morvin field, mentioned earlier, this was a major challenge to field development.

Production Drilling at Morvin, a Special Case The Morvin-field is more densely populated by significant deepwater coral reefs than any other Norwegian offshore field [47]. Whereas pipeline routes and the final location of steel templates could easily be adjusted to avoid the coral reef occurrences, one of the major challenges was how to drill the production wells without harm to the reefs. Because there are significant deepwater coral reefs within 200 m at both the planned Morvin production templates, it was not possible to use the normal practice without incurring damage to some of the reefs. It was therefore decided to seek authorities' permission for discharging top-hole cuttings some distance away from the templates, at locations deemed safe for the coral reefs. The chosen technical solution was to employ a "Cutting Transportation System" (CTS), consisting of two 600 m long flexible tubes installed on the seabed. They were attached to five gravel filled bags ("bigbags") to prevent them from moving on the seafloor and were attached to a manifold placed on the seafloor. This manifold had one exhaust pipe leading from the wellhead where the cuttings were collected and pumped to the CTS. Currents were predicted and numerically modeled, and the optimal discharge locations were found where cuttings would not do harm to the surrounding significant coral reefs. The cuttings were discharged at relatively high speed through recoil dampers placed about 1 m proud of the seafloor at the end of the CTS-hoses.

The cuttings emitting from the recoil damper were found to be heavier and less fine-grained than predicted. This resulted in the discharge plume emitting as a heavy fluid, and spreading along the seafloor as a turbid and heavy cloud. Because this cloud did not spread high up into the water mass most of it accumulated near the end of the CTS hoses without any damage to the corals. Continuous visual monitoring by ROVs of the CTS and the nearest coral reefs was done during the whole top-hole drilling operation at Morvin. This ROV-work also included lifting and repositioning (adjusting) of the recoil dampers, as the heap of cuttings grew to heights of up to 1 m. A total of three sediment traps were also placed on the seafloor in order of documenting the final distribution of any resuspended cuttings. Three of the downstream coral reefs were continuously monitored by using an automatic time-lapse "satellite photo rig" and by ROV-monitoring. This Morvin experience proved that production drilling can also be done near scattered coral reefs without harming them.

Observations Along Pipelines

Prior to the laying of trunk pipelines on the seafloor, detailed visual surveys are performed. The objectives of these "pre-lay" surveys are to make sure the pipeline does not cross any dangerous areas or any features that can hinder the laying. In addition, all features and marine life are also recorded along the pipeline route. Because underwater visibility is restricted, the visually documented corridor is most often no wider than about 10 m. However, the length of the surveyed corridor, which depends on the length of the pipeline to be laid is anything up to several hundred kilometers. During the last quarter century, 1984–2010, several thousands of km long transects were surveyed by Statoil. Thus, trunk gas and oil pipelines have been laid from the Snehvit field in the southern Barents Sea and from numerous fields offshore Mid- and Southern-Norway. Detailed knowledge about the seafloor and marine life has, therefore, been gathered along transects crossing all ranges of water depths of the North Sea, stretching from Norway to the UK, and from Norway to Germany, Belgium, and France (Dunkerque). In this way, the general seafloor has been imaged, meter for meter, across 50 m iceberg ploughmarks in the Barents and Norwegian Seas, to typically 100 m wide, 5 m deep pockmark craters in the northern North Sea, to wide stretches of rugged sandwave fields off Netherlands and Belgium, to the normal, even, and uneventful seabed in all of the seas. The seabed life pattern always seems to be the same: there is apparently very little variation in macrofaunal life except for when there are special features on the seafloor, such as boulders and rock outcrops, besides the special fluid-flow-related features: pockmark craters and other microseeps, and the eventual macro-seeps, which are very rare.

The fact that marine life proliferates wherever there are "special features" on the seafloor means that wherever a new man-made structure is installed on the seafloor, there may be a dramatic impact on the visible macrofaunal life pattern (Fig. 6). On all concrete-coated trunk pipelines laid across the pre-surveyed sections, it is clearly noticed how the pipeline structure introduces new opportunities for benthic life. For example, it only took 1 year before thousands of *Nephrops norvegicus* crustaceans were established along kilometer long stretches of the "Statpipe" pipeline, the first 3600 trunk pipeline laid across the 300 m deep Norwegian Trench (North Sea), in 1985. They were clearly colonizing small sections of the seafloor along the outer, curved wall of the pipeline. Their holes were located snugly inside the sediments nearest to the pipeline wall. The clawed animals had their bodies halfway out of their openings, easily visible during the first thorough visual inspection of the pipe, the so-called as-laid survey, conducted about 1 year after the pipeline was installed. Over time, the annual pipeline inspections, which cover long sections of both sides and the top of pipe, every other year, document not only how the trunk pipelines gradually bury themselves into the seabed, but also how they are colonized by a variety of different invertebrates. Unfortunately, this information is seldom used for quantifiable studies by biologists or marine ecologists, although the data should be fully available for such work. The reason is probably, that this type of visual information is unknown and that it perhaps does not belong in the scientific "vocabulary" of traditional marine biology or ecology.

To illustrate some of the important information content in this unique pipeline inspection data set, there are three particularly noticeable occasions, episodes, and occurrences that are worthy of further description: (1) A temporary anoxic condition recorded in the central North Sea, (2) "Drifts" of dead fish (mackerel) along a pipeline section, and (3) Fish protected by OHI-pipelines.

Temporary Anoxic Conditions

During the annual inspection in 2004, of the Europipe 1 trunk gas pipeline from Kårstø, SW Norway, to Northern Germany, a few dead fish and invertebrates were seen lying on the seafloor along the pipeline about 50 km south of the Draupner platform, central North Sea. There was also some white material, believed to be bacterial mats. This section of the pipe is partly buried into the seafloor and is about 40 cm proud of the smooth seafloor surface. Both sides of the pipeline were affected and, there were also dead invertebrates on top of the pipe itself. Although this anomalous observation could easily have been interpreted as something to do with the pipeline, some further investigations and measurements precluded such a conclusion. Anoxia was obviously the cause of this "mass extinction," - but, what had caused it? Along one short section, it was only the southern side of the pipeline that was affected. This annual inspection was conducted during May, when the expected ambient water temperature should be around 6 °C. But, actual measurements showed an ambient water temperature of only 1.1 °C near the seafloor. Measurements through the water column showed the lower 20 m of seawater only had between 1.1 °C and 2.0 °C. This very cold, nearly freezing water apparently occurred as a dome-shaped, dynamically stable water mass occupying the general depression in the seafloor, where the pipeline had been laid. During installation of the Europipe 1 pipeline some years earlier, portions of it were laid into the meandering relict "Elbe valley" on the seafloor of the mid- and southern North Sea. This was obviously where the cold water had accumulated. Investigations of recorded water temperatures further southeast in the German Bight showed that very cold water had formed during the winter months, January and February when there was a long cold spell. Because there had been no subsequent storms, the cold, dense water mass had since followed the deepest local portions of the seafloor, and gradually flowed toward the central portion of the North Sea, where it engulfed parts of the Europipe 1 pipeline and where it became stagnant and anoxic. During the subsequent annual pipeline inspection, 2 years later, there was neither sign of dead animals nor any other evidence of this temporary "extinction event." These observations clearly demonstrate that great variations in the natural seafloor environment, even in the relatively busy central North Sea can occur without being noticed on the sea surface.

"Drifts" of Dead Fish Along a Pipeline

The fact that large trunk pipelines laid onto the even, flat, featureless seafloor act as barriers for objects drifting across the seafloor is well known. During the early years of trunk pipeline construction in Norwegian waters (1983–1997), quite a lot of garbage, that is, paper, plastic bags, bottles, cans, etc., together with natural debris, such as kelp and sponges, etc., was noted to accumulate along sections of the newly laid pipes. During the last decennium, however, the amount of garbage has decreased markedly, probably as a result of stricter garbage handling rules in general, and Norway in particular. However, such rules are not always obeyed, as the following example proves. During the annual inspection of a section of the 36″ trunk gas pipeline, Europipe 2, in the Norwegian sector of the North Sea, a huge and elongated pileup of dead fish (mackerel) was visually recorded. It formed a ca. 3 km long and nearly 1 m high "drift" along the western side of the pipeline. The volume of dead fish was estimated to about 10,000 m^3. Because this freshly killed mackerel was obviously dumped illegally by one or several of about seven nearby fishing vessels, the Norwegian Fishery Directory was contacted, and on the following evening, live video footage of the drift was broadcast as a news item by the directory on national TV. Although this live footage, from one of OHI's annual pipe inspections surely made a public impression, it is very rare that such information is used for public purposes.

Fish Protected by OHI-Pipelines

The largest pipeline constructed to date, in the North Sea and Norwegian Sea, is the 1,200 km long 40″ and 44″ Langeled pipeline, transporting gas from the giant Ormen Lange field off Mid-Norway, via Nyhavna in Mid-Norway and the Sleipner field in the central North Sea to Immingham on the east coast of the UK. During the initial as-laid survey, immediately after laying, in 2006, a large school of newly spawned fish was documented along a 50–100 km long section of the pipe, in UK waters. This school was also observed to swim along the pipeline about 3 months later, after the individual fish had grown significantly larger in size. Trawling is routinely practiced along exposed trunk pipelines. The intent is to catch fish that swim along and hide beneath free-spanning sections of the pipelines. Because this large school of fish had obviously been spawned near the newly laid pipeline, and because it utilized the opportunity for refuge along the pipeline, it was suggested by OHI-personnel that measures to protect the school should perhaps be taken by British Authorities. To protect a corridor of, say 3 km width, along a 100–200 km long section of the Langeled pipeline could have been feasible. The following annual inspections could also have visually documented the fate of such a school of fish. However, it was soon realized that suggestions of this kind were premature, in 2006, as it was not expected that anybody would actually understand what the problem may be, or what the suggestion was really about. This demonstrates clearly how little is known and currently understood about the subsea world and how difficult it is to convey ideas and impressions from this world, unless the receivers have similar or comparable experience.

The higher abundance of fish sightings (of fish >0.5 m in length) along trunk pipelines was already noticed after the installation of Statoil's first trunk gas pipeline, the over 250 km long and 36″ Statpipe pipeline from the Statfjord field to Kalstø on the SW coast of Norway. This pipeline crosses the 300 m deep Norwegian Trough, where there is a varying density of pockmark craters in the seafloor. During the very first annual survey of this pipeline, a strong correlation between number of fish sightings and the density of pockmarks was noticed. The fish counted in 1987 were either seen swimming along the sides of the pipeline, or occupying some of the space underneath the pipeline where it had intermittent spans over irregular terrain in the seafloor. This comparison was made with fish sightings in 1984, over the same, undisturbed section of seafloor, before the pipeline was installed. The increased number of fish seen, after the pipeline was installed was dramatic and close to tenfold along some of the seafloor sections. Also before the pipeline was installed, there was evidently many more fish in densely pockmarked areas than in areas without pockmarks. This example shows that trunk pipelines attract fish and also likely protects them to some degree.

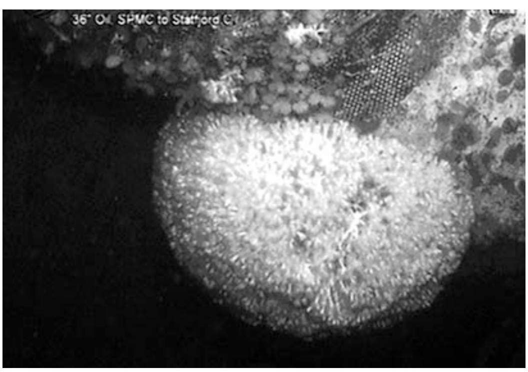
The question of protection of trunk pipelines against damage by trawl-board impact was assessed prior to installation of the Statpipe pipeline. Carefully designed laboratory studies in test tanks and rigs, however, soon proved that no serious damage would possibly be inflicted on the concrete coated steel trunk pipelines by normal trawl-boards. Therefore, most such pipelines are left on the seafloor without being actively protected by trenching or gravel cover. In some areas, where the seafloor relief is varied, by numerous pockmarks or iceberg ploughmarks, there may be a need to modify the seafloor topography to avoid long free spans in the pipe. Spans longer than about 80 m are unacceptable for two reasons: (1) motions in the pipe caused by water currents may inflict material damage to the pipe, and (2) trawl-boards may snag underneath the pipe, during fishing along the pipeline. Seafloor preparation and intervention is then performed, either by trenching the seafloor highs or by dumping gravel inside seafloor troughs. This type of intervention is performed before installation of pipeline or immediately after installation. In the German sector of the North Sea, where the water depth is not greater than 150 m, all trunk pipelines have to be buried below the seafloor surface. The annual inspection of buried pipelines is done by the use of induced electric currents and magnetic detection of the buried pipe, combined with visual inspection of the seafloor above the buried pipe. During the last 10–15 years, the annual surveys in these waters have not, unfortunately produced many fish sightings, probably as a consequence of lack of protection and intensive fishing activity.

The Impact of Platforms and Other Fixed OHI-Structures

The six giant concrete platforms Statfjord, A, B, C and Gullfaks A, B, C were installed on the seafloor at the Statfjord and Gullfaks fields in the northern North Sea during a 15 year period in the 1980s and 1990s. Apart from the seafloor at the deepest site, Gullfaks C, the seafloor is dominated by sand, gravel, and patches of boulders. Prior to installation, large fish, such as *Brosme* sp., Ling, Cod, and Seithe, were only occasionally seen swimming around. However, during a detailed preinstallation site survey of the deeper Gullfaks C site, a more complex situation was found. The seafloor at Gullfaks C contrasts with the other sites by having a seafloor covered by soft, fine-grained clay-dominated sediments. Furthermore, there are normal-pockmarks in this area. A dedicated visual survey of pockmark craters here revealed that larger fish, most often the *Brosme* sp., were located inside many of these pock-marks. And inside one particular pockmark, of 8 m depth and 120 m length, there were up to 20 large fish located in the deepest end of the pockmark crater [10]. About fifteen of the large fish (Ling, Cod, and Brosme) were swimming against the southward flowing current, whereas about five fish were found inside a tunnel at the deepest portion of the pockmark [10]. This example clearly demonstrates some of the unexpected heterogeneity of the seafloor and that it takes dedicated visual surveying to unravel the true nature of marine life on the undisturbed seafloor.

After about 30 years of operating large fixed structures, it is a well-known fact within the OHI that there are plenty of large fish living near the structures. At some of the hydrocarbon fields in the North Sea special trawl-fenders have been installed on the seafloor to prevent trawlers from fishing within the 500 m forbidden safety zone surrounding all fixed platforms which are over the sea surface. These trawl-fenders are robust, up to 3 m high steel poles driven into the seafloor. They are intended to snag the trawl equipment before it can damage any of the production gear and infrastructure placed on the seafloor within the safety zone. It is also a well-known fact that filter-feeders, such as sea anemones and other sessile animals colonize the legs of both steel and concrete platforms (Fig. 7). However, until for about 15 years ago it was not known that colonizing ahermatypic corals also lived on these structures (Fig. 8). Thus, OHI-related fixed structures make ideal artificial reefs on the seafloor.

By law, all the OHI-related structures placed on the seafloor shall be decommissioned after use, that is, they shall be dismantled and disposed of in



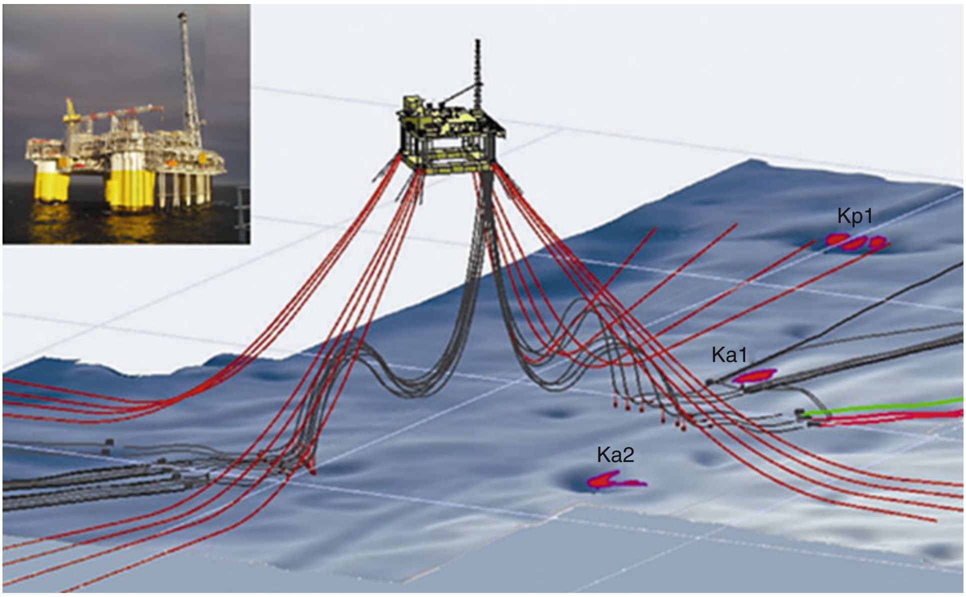
Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 8 A relatively small Lophelia colony seen on a man-made structure at Statfjord, Norwegian sector of the North Sea. The colony cannot be older than about 22 years (age of structure)

a safe and environmental friendly manner. However, judging from experiences in the North Sea, even some of the large trunk pipelines could perhaps remain until they either buried themselves or corroded into destruction. In the meantime, they would undoubtedly represent an added biologically friendly asset to the normal seafloor, at least in some places (Fig. 9). For the giant concrete structures, the case is also one of cost benefit. These structures are placed firmly into the ground and will require a lot of energy to remove. Because concrete is a very environmental friendly substance and has a high longevity, it should be seriously debated if they could remain in place for 'ever.' They could act as home for plenty of fish, and could even be used by professional fishermen as fish nurseries and fishing grounds.

Future Directions

More dedicated detailed seafloor mapping, combined with visual documentation is undoubtedly needed for improving the understanding and knowledge of life on the seafloor in general. Through the Norwegian "Mareano" project, started by the government in 2005, this type of work has actually started in Norway. Swathe multi-beam detailed seafloor mapping is here combined with up to kilometer long video (visual) transects. In addition to sampling, the seafloor is mapped for biotype and many other parameters. In the long run, such mapping and documentation is necessary for adequate management of the resources in the sea. However, it may take another 100 years for one vessel to map out the whole Norwegian EEZ (Exclusive Economic Zone) in this fashion. This therefore, calls for more technological development and probably also for the use of autonomous robotic systems in the future.

With regard to the OHI-related structures placed on the seafloor, such as trunk pipelines, steel templates, and giant platforms, there needs to be a thorough discussion as to which of them should be relinquished (removed) and which structures could serve to stimulate and protect the marine bio-environment. Judging from the



Marine Life Associated with Offshore Drilling, Pipelines, and Platforms, Fig. 9 A perspective image of how the Kristin semi-floating platform is located above the seafloor. The water depth below the platform is 305 m. *Red lines* indicate anchor chains and wires. The *black lines* indicate flowlines and risers piping products (gas, condensates, etc.). The chains and wires that hold Kristin in place are attached to "suction anchors," inverted domed steel cylinders sucked into the seafloor clays. The coral reefs, Ka2, Ka1, and Kp1 (in *red*), have been inspected twice before installation and several times after installation of the platform and the other infrastructure on the seafloor. No damage has been found on the reefs. Image is based on true multibeam mapping of the seafloor and digital technical drawings of the platform, shown in a photo *upper left*. This image shows that the platform and infrastructure can be relinquished after use, without any damage to sensitive biology on the seafloor. (Courtesy of Statoil ASA and Leslie Austdal)

visual documentation over the last 30 years in the North Sea, it can be deemed favorable to let some of the installations remain on the seafloor as artificial reefs and for the future benefit of both marine life and humans.

Bibliography

1. Martin TR, Olsen KR, Cahill MM (2010) Artificial reefs – an important tool for mitigation and restoration. Ocean News Technol 16(4):28–29
2. Maurer BA, McGill BJ (2004) Neutral and non-neutral macroecology. Basic Appl Ecol 5:413–422
3. Humphris SE, Zierenberg RA, Mullineaux LS, Thomsen RE (eds) (1995) Seafloor hydrothermal systems: physical, chemical, biological, and geological interactions. Geophysical monograph 91American Geophysical Union, Washington, DC, 510 pp
4. Etter RJ, Rex MA(1990) Population differentiation decrease with depth in deep-sea gastropods. Deep-Sea Res 37:1251–1261
5. Rex MA (1983) Geographic patterns of species diversity in the deep-sea benthos. In: Rowe GT (ed) The sea. Wiley, New York
6. Piepenburg I et al (2001) In: Schãfer S et al (eds) The northern north Atlantic – a changing environment. Springer, Berlin
7. Schäfer P, Ritzrau W, Schlüter M, Thiede J (2001) The northern north Atlantic – a changing environment. Springer, Berlin, 500 pp
8. Judd AG, Hovland M (2007) Submarine fluid flow, the impact on geology, biology, and the marine environment. Cambridge University Press, Cambridge, 475 pp
9. Suess E (2010) Marine cold seeps. In: Timmis KN (ed) Handbook of hydrocarbon and lipid microbiology, 1. Springer, Berlin, pp 187–203. (Part 3)
10. Hovland M, Judd AG (1988) Seabed Pockmarks and Seepages. Impact on geology, biology and the marine environment. Graham & Trotman, London, 293 pp
11. Hovland M, Thomsen E (1989) Hydrocarbon-based communities in the North Sea? Sarsia 74:29–42
12. Hovland M (2002) On the self-sealing nature of marine seeps. Cont Shelf Res 22:2387–2394
13. Wegener G, Shovitri M, Knittel K, Niemann H, Hovland M, Boetius A (2008) Biogeochemical

processes and microbial diversity of the ullfaks and Tommeliten methane seeps (northern North Sea). Biogeosciences 5(4):1127–1144
14. Hovland M (2007) Discovery of prolific natural methane seeps at Gullfaks, northern North Sea. Geo-Mar Lett. https://doi.org/10.1007/ s00367-007-0070-6
15. Seibel BA, Dierssen HM (2009) Animal function at the heart (and gut) of oceanography. Science 323: 343–344
16. Moore CJ (1999) Seeps give a peek into plumbing, explorer. Am Assoc Pet Geol Bull 99:22–23
17. Dimitrov LL (2002) Mud volcanoes – the most important pathway for degassing deeply buried sediments. Earth Sci Rev 59:49–76
18. Dupré S, Woodside J, Klaucke I, Mascle J, Foucher J-P (2010) Widespread active seepage on the Nile Deep Sea Fan (offshore Egypt) revealed by high-definition geophysical imagery. Mar Geol 275:1–19
19. Vogt PR, Crane K, Pfirman S, Sundvor E, Cherkis N, Flemming H, Nishimura C, Shor A (1991) SeaMarc II sidescan sonar imagery and swath bathymetry in the Nordic basin. EOS Trans 72:486
20. Hovland M, Hill A, Stokes D (1997) The structure and geomor-phology of the Dashgil mud volcano. Azerbaijan Geomorphol 21:1–15
21. King LH, MacLean B (1970) Pockmarks on the Scotian shelf. Geol Soc Am Bull 81:3142–3148
22. Cathles LM, Su Z, Chen D (2010) The physics of gas chimney and pockmark formation, with implications for assessment of seafloor hazards and gas sequestration. Mar Pet Geol 27:82–91
23. Hovland M, Heggland R, de Vries MH, Tjelta TI (2010) Unit pockmarks and their potential significance for prediction of fluid flow. J Mar Petrol Geol 27:1190–1199
24. Buhl-Mortensen L, Vanreusel A, Gooday AJ, Levin LA, Priede IG, Buhl-Mortensen P, Gheerardyn H, King NJ, Raes M (2010) Biological structures as a source of habit heterogeneity and biodiversity on the deep ocean margins. Mar Ecol 31: 21–50
25. Gunnerus JE (1768) Om nogle Norske coraller. In: Kongelige Norske Videnskabers Selskabs Skrifter, vol 4, pp 38–73
26. Dons C (1944) Norges korallrev. Norsk Vidensk Selsk Trond-heim Forh 16A:37–82
27. Wilson JB (1979) The distribution of the coral *Lophelia pertusa* (L.) [*L. Prolifera* (Pallas)] in the north-east Atlantic. J Mar Biol Assoc UK 59:149–164
28. Wilson JB (1979) “Patch” development of the deep-water coral *Lophelia pertusa* (L.) on rockall bank. J Mar Biol Assoc UK 59:165–177
29. Hecker B, Blechschmidt G, Gibson P (1980) Epifaunal zonation and community structure in three Mid- and North Atlantic Canyons. Contract report BLM AA551-CT8-49 prepared by Lamont-Doherty for US Department of Interior
30. Zibrowius H (1980) Les Scle’ractiniaires de la Me’diterrane’e et de l’Atlantique nord-oriental. Memoires de l’Institute Oceanographique 11:247
31. Hovland M (1990) Do carbonate reefs form due to fluid seepage? Terra Nova 2:8–18
32. Hovland M, Croker PF (1993) Fault-associated seabed mounds in the Porcupine Basin, offshore Ireland. Expanded abstract. In: Proceedings of the 55th EAEG Ann. Mtg., Stavanger, Norway
33. Hovland M, Croker P, Martin M (1994) Fault-associated seabed mounds (carbonate knolls?) off western Ireland and NorthWest Australia. Mar Pet Geol 11:232–246
34. Freiwald A, Henrich R, Pätzold J (1997) Anatomy of a deep-water coral reef mound from Stjernsund, west Finnmark, northern Norway. In: James NP, Clarke JAD (eds) Cool-water carbonates. Soc Sediment Geol (SEPM), Special Publ 56. SEPM, Tulsa, pp 140–161
35. Henriet J-P, De Mol B, Pillen S, Vanneste M, Van Rooij D, Versteeg W, Croker PF, Shannon PM, Unninthan V, Bouriak S, Chachkine P (1998) Gas hydrate crystals may help build reefs. Nature 391:648–649. (Porcupine-Belgica Shipboard Party)
36. Hovland M, Mortensen PB (1999) Norske korallrev og prosesser I havbunnen (Norwegian coral reefs and seabed processes), John Grieg, Bergen, Norway, 167 pp (in Norwegian with English summary)
37. Armstrong CW, van der Hove S (2007) The formation of policy for protection of cold-water coral off the coast of Norway. Internal report, University of Tromsø
38. Fosså JH, Mortensen PB (1998) Artsmangfoldet på *Lophelia*-korallrev og metoder for kartlegging og overvåkning. The biodiversity on Lophelia-reefs and methods for mapping and monitoring. Fisken og Havet 17:1–95. (in Norwegian)
39. Willison JHM, Hall J, Gass SE, Kenchington ELR, Butler M, Doherty P (eds) (2001) In: Proceedings of the first international symposium on deep-sea corals, Ecology Action Centre and Nova Scotia Museum. Halifax, Canada
40. Hovland M, Risk M (2003) Do Norwegian deep-water coral reefs rely on seeping fluids? Mar Geol 198: 83–96
41. Jensen S, Neufeld JD, Birkeland N-K, Hovland M, Murrell JC (2008) Insight into the microbial community structure of a deepwater coral reef environment. Deep-Sea Res I 55:1554–1563
42. Yakimov MM, Cappello S, Crisafi E, Tursi A, Savini A, Corselli C, Scarfi S, Giuliano L (2006) Phylogenic survey of metabolically active microbial communities associated with the deep-sea coral *Lophelia pertusa* from the Apulian plateau, Central Mediterranean Sea. Deep-Sea Res I 53:62–75
43. Sorokin YuI, Sorokin Yu P (2009) Analysis of plankton in the southern Great Barrier Reef: abundance and roles in throphodynamics. J Mar Biol Assoc UK 89:235–241
44. Neulinger SC, Gärtner A, Järnegren J, Ludvigsen M, Lochte K, Dullo W-C (2008) Tissue-associated “*Candidatus Mycoplasma corallicola*” and filamentous bacteria on the cold-water coral *Lophelia pertusa* (Schleractinia). Appl Environ Microbiol 75:1437–1444

45. Tavormina PL, Ussler W, Orphan VJ (2008) Planktonic and sediment-associated aerobic methanotrophs in two seep systems along the North American margin. Appl Environ Microbiol 74:3985–3995
46. Penn K, Wu D, Eisen JA, Ward N (2006) Characteristics of bacterial communities associated with deep-sea corals on Gulf of Alaska seamounts. Appl Environ Microbiol 72:1680–1683
47. Hovland M (2008) Deep-water coral reefs – unique biodiversity hot-spots. Springer Praxis, Chichester, 278 pp
48. Costello MJ, McRea M, Freiwald A, Lundälv T, Jonsson L, Bett BJ, Van Weering TCE, de Haas H, Roberts MJ, Allen D (2005) Role of cold-water coral *Lophelia pertusa* coral reefs as fish habitat in the North East Atlantic. In: Freiwald A, Roberts M (eds) Cold-water corals and ecosystems. Springer, Heidelberg, pp 771–805
49. Furevik D, Nøttestad L, Fosså JH, Husebø A, Jørgensen S (1999) Fiskefordeling i og utenfor korallområder på Søregga. Fisken og Havet no 15, 33 pp
50. Husebø A, Nøttestad L, Fosså JH, Furevik D, Jørgensen SB (2002) Distribution and abundance of fish in deep-sea coral habitats. Hydrobiologia 471: 91–99
51. Roberts JM, Wheeler AJ, Freiwald A (2006) Reefs of the deep: the biology and geology of cold-water coral ecosystems. Science 312:543–547
52. Turley C, Blackford J, Widdicombe S, Lowe D, Nightingale PD, Rees AP (2006) Reviewing the impact of increased atmospheric CO_2 on oceanic pH and the marine ecosystem. In: Schnellnhuber HJ, Cramer W, Nakicenovic N, Wigley T, Yohe G (eds) Avoiding dangerous climate change. Cambridge University Press, Cambridge, pp 65–70
53. Turley CM, Roberts JM, Guinotte JM (2007) Corals in deep-water: will the unseen hand of ocean acidification destroy cold-water ecosystems? Coral Reefs. https://doi.org/10.1007/ s00338-007-0247-5
54. Gass SE, Roberts JM (2006) The occurrence of the cold-water *Lophelia pertusa* (Scleractinian) on oil and gas platforms in the North Sea: colony growth, recruitment and environmental controls on distribution. Mar Pollut Bull 52:549–559
55. Roberts JM, Long D, Wilson JB, Mortensen PB, Gage JD (2003) The cold-water coral *Lophelia pertusa* (Scleractinia) and enigmatic seabed mounds along the north-east Atlantic margin: are they related? Mar Pollut Bull 46:7–20
56. Koslow JA, Gowlett-Holmes K, Lowry JK, O'Hara T, Poore GCB, Willimams A (2001) Seamount benthic microfauna off southern Tasmania: community structure and impacts of trawling. Mar Ecol Prog Ser 213: 111–125
57. Fosså JH, Mortensen PB, Furevik DM (2000) *Lophelia*-korallrev langs norskekysten. Forekomst og tilstand. Lophelia coral reefs along the Norwegian coast. Occurrence and conditions. Fisken og havet, 2, 94 pp (in Norwegian)
58. Rogers AD (1999) The biology of *Lophelia pertusa* (Linnaeus 1758) and other deep-water reef-forming corals and impacts from human activities. Int Rev Hydrobiol 84:315–406
59. Mortensen PB, Fosså JH (2006) Species dieversity and spatial distribution of invertebrates on *Lophelia* reefs in Norway. In: Proceedings of the 10th international coral reef symposium. Okinawa, Japan, pp 1849–1868
60. Hall-Spencer J, Allain V, Fosså JH (2002) Trawling damage to Northeast Atlantic ancient coral reefs. Proc R Soc Lond Ser B Biol Sci 269:507–511
61. Myhrvold A, Hovland M, Nøland S-A (2004) Baseline and environmental monitoring in deep water – a new approach. In: Seventh international SPE conference on health, safety, and environment, Calgary, 29–31 Mar 2004. Paper no. SPE 86776
62. Etiope G, Feyzullayev A, Baciu CL (2009) Terrestrial methane seeps and mud volcanoes: a global perspective of gas origin. Mar Pet Geol 26:333–344
63. Mikkelsen N, Erlenkauser H, Killingley JS, Berger WH (1982) Norwegian corals: radiocarbon and stable isotopes in *Lophelia pertusa*. Boreas 5:163–171
64. Mortensen PB, Rapp HT (1998) Oxygen- and carbon isotope ratios related to growth line patterns in skeletons of *Lophelia pertusa* (L) (Anthozoa: Scleractinia): implications for determination of linear extention rates. Sarsia 83: 433–446

Web Sites

www.mareano.no
www.npd.no
www.nerc.ac.uk
www.iodp.org
www.deepseadrilling.org
www.offshoremagazine.com
www.serpentproject.com
www.diverdiscover.whoi.edu
www.statoil.com
www.eu-hermes.net

Natural Resource Flows and Sustainability in Urban Areas

Stefan Anderberg
Lund University Centre for Sustainability Studies, Lund University, Lund, Sweden

Article Outline

Glossary

Ecological footprint A measure of the resource consumption of a given unit (e.g., a city and its population) that represents the area of ecologically productive land and water needed for the production and assimilation of wastes generated.

Emergy analysis Emergy is the amount of energy, usually solar, that is directly or indirectly required to generate a given output flow. Emergy analysis allows analysis of the whole embodied environmental work expressed in a single unit and has been applied to flows in cities and other regions as well as different types of products and production processes.

Urban ecology Research on urban ecosystems and applications of ecological principles in connection with urban planning.

Urban metabolism Involves conceptualizing a city as an organism or ecosystem and tracking resources that go into the system and products and wastes that leave it.

Urban mining Recovery of metals and other valuable materials in urban areas. Studies that have identified metal concentrations comparable to mines in cities and urban waste dumps suggested that an increasing share of the raw materials for metal production in the future will be recovered from wastes in urban areas.

Definition of the Subject and Its Importance

Cities and urban areas are of central importance for global sustainable development both as dominating sites of production and consumption and as contexts for developing new, more sustainable practices. All cities depend on large imports of energy and other natural resources to satisfy consumption of their inhabitants as well as local production, trade, and services. These resource flows are closely linked to global and regional sustainability issues, such as resource scarcity, pollution, and competition for land or water, as well as local health, environmental, and distributional issues. How resources are consumed and managed in the expanding cities has strong implications for the global resource flows and related pressures in different scales. The goal of "sustainable development" puts into question many traits of current urbanization and city development and calls for a sustainable urban transformation. Such a transformation needs to rely on sound, efficient, and sustainable resource management and sustainable urban structures in terms of the built environment, transport systems, and green and blue structures. The handling of resource flows in cities and urban areas is the core of this challenge. For increasing the capacity of local strategic action toward a more sustainable use of resources, it is essential to develop improved understanding of the urban resource flows. Studies of urban ecosystems and resource flows as well development of ecological planning approaches connect to this challenge.

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_441

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC, 2012
https://doi.org/10.1007/978-1-4419-0851-3_441

Introduction

After centuries of intense urbanization, more than half of the world's population lives in urban areas. The global urban population is expected to grow by more than 50%, or 1.7 billion people, between 2008 and 2030 [1]. Therefore, the cities are increasingly recognized as the most important global sustainability challenge of this century [2]. The urban ecological sustainability challenges are both linked to global resource scarcity, pollution and biodiversity issues, and local health and environmental problems as well as resource access and distribution among different groups of people. With their central role in connection with manufacturing, trade, and consumption, big cities since long dominate global energy use and material flows [3]. How resource use develops in cities have strong implications for resource flows in different scales and related pressures. How cities are built, heated and cooled; how efficient their infrastructures are, and what urban citizens consume or how they travel in their everyday lives have great influence on greenhouse emissions, and exploitation of scarce land, water and mineral resources in different parts of the world. For increasing the capacity of local action toward a more sustainable use of resources, it is essential to develop improved analysis tools that allow a better guidance for strategic action [4, 5]. A key challenge is to improve the overview and understanding of the environmental linkages of the urban resource flows in different scales and develop urban governance that contributes to a more sustainable use of resources [6, 7]. Studies of urban ecosystems and resource flows as well as development of sustainability planning connect to this challenge. Urban metabolism studies explore the interactions among resource flows, urban transformation processes, waste streams, and quality of life and connect to the challenge of combining urban planning with sustainable resource management. This entry focuses particularly on natural resource flows in urban areas and urban metabolism studies, their results and challenges, and their potential contributions to sustainable urban planning and governance.

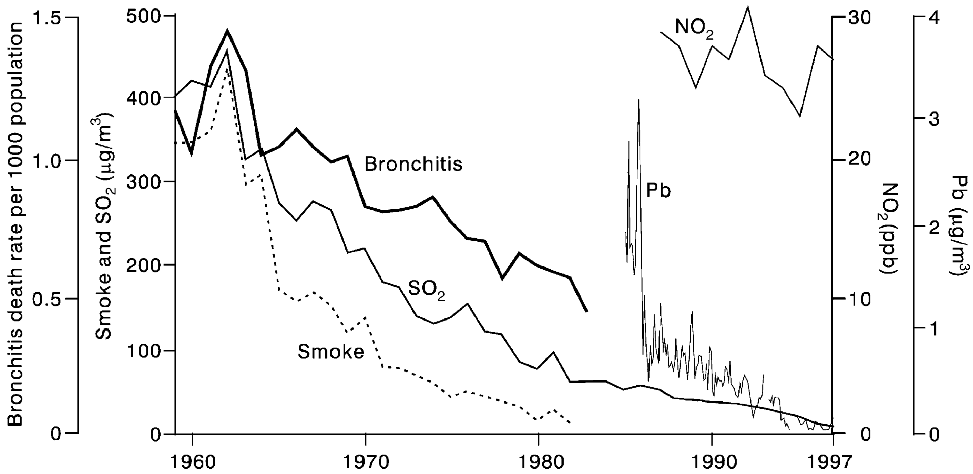
Urban Sustainability Challenges

Cities alter their local environments and influence the global environment in several ways. They occupy and modify space and are characterized by intense competition for space between different uses of land [8]. They import vast quantities of food, water, and energy and export emissions and waste, and they depend in various ways on local and global ecological services. Cities have always been a scene for unsafe living conditions and ecological problems [9]. Historically, it is possible to distinguish three different but often overlapping types of urban sustainability crises during the industrial era:

- A *hygienic* crisis connected to unsafe drinking and inadequate sanitation and waste management, and related epidemics in fast growing cities
- A *social* crisis connected to poverty; insufficient, poor, and unhealthy housing; and lacking access to light, fresh air, and green spaces
- An *environmental* crisis particularly connected to severe pollution and health hazards for local people

In cities in industrialized regions, these crises have generally been successfully addressed via development of urban infrastructures, planning, organization, and control, and the most urgent local problems have disappeared [10]. Local pollution problems still remain in most cities, mostly connected to traffic, waste, and diffuse pollution. Figure 1 shows the development in Manchester as an example of the improvement of environmental quality and health in European cities toward the end of the twentieth century. By the end of the nineteenth century, the bronchitis death rate was above 3 per 1,000 inhabitants compared with less than 0.5 in the 1980s.

In fast-growing cities in the developing world, these three types of problems still remain, overlap, and often get worse. In many cities in Africa, Asia, and Latin America, large parts of the inhabitants live under unhealthy conditions in poor housing without access to safe drinking water,



Natural Resource Flows and Sustainability in Urban Areas, Fig. 1 Development of air pollution (smoke, SO_2, lead, and NOx) and bronchitis death rate (per 1,000 inhabitants) in Manchester. (Source: [10])

sanitation, or other basic services [11, 12]. In great contrast to the desperate needs for basic infrastructure in Third World cities, the emerging sustainable city efforts in Europe, North America, and Australia during recent decades have, to a large extent, been concerned with the unsustainability of the cities in a global context. Global change problems, such as resource scarcity, climate change, land use competition, or biodiversity, put into question the energy use, spatial structures, and infrastructural arrangements of modern society in general and big city regions in particular. In this context, promoting efficient and sustainable management of resource flows in the city has become of central importance.

Urban Planning

Societal development is in different ways reflected in changed flow and flow patterns. Industrialization, increasing wealth, expansion of the city, changed design of buildings and city areas, and changed patterns of settlement, work, commuting, production, consumption, trade, leisure activities, and use of the urban landscape of the big city all result in changed flows. Resource flows are thus strongly influenced by urban planning and design of urban areas. Urban planning has developed in interaction with different social welfare issues, e.g., local environmental problems. It has traditionally primarily focused on regulating the location and intensity of activities to avoid health problems and environmental degradation and influenced resource flows, but planning has seldom comprehensively and explicitly addressed resource management [13]. Urban planning emerged as a response to the unhealthy and polluted conditions in the fast-growing industrializing European cities in the nineteenth century. The overwhelming problems were gradually overcome by investments in infrastructure and introduction of a system of laws and authorities to take care of waste collection, water, hygienic control, and worker protection. Society took on a growing responsibility for the health and well-being of its citizens, and the major tools have been infrastructural investments and an increasingly ambitious spatial planning. During the twentieth century, the social crisis connected to poor, unhealthy housing conditions was, in European cities, addressed by demolition, renovation and social housing projects, and establishment of parks, green and recreational areas, and separation of industry and housing. Cities were, in the twentieth century, transformed by motorized road traffic and the breakthrough of the private car, but also visions of more human and ecologically sound urban environments influenced the urban development. Both decentralized city visions, such as Howard's "Garden City," and more centralized visions, such as Le Corbusier's "Radiant City," have influenced urban and suburban developments.

With the growing industrial production, increasing use of fossil fuels, cities in general were severely polluted until the mid-twentieth century. From the 1960s to the 1990s, increasing concern for the environment stimulated important infrastructural investments in wastewater treatment, district heating, energy source substitution, which, together with improved pollution control and industrial restructuring, decreased the industrial pollution and health hazards for urban inhabitants in Europe and North America.

Planning for Urban Sustainability

From the 1970s, growing concern about environment and resource use also influenced urban planning. New urban forms, such as the compact city with high-density settlement and mixed uses, were promoted as a response to urban sprawl and high dependency on cars and fossil fuel consumption after the oil crisis. The real breakthrough for environmental and ecological planning and interest for the relationship between planning and resource management came with the goal of "sustainable development" that became influential from the late 1980s when ecocity or sustainable city visions and ambitions spread rapidly through national, international, and intercity initiatives. UN Habitat, the World Bank, and the EU have launched sustainable urban development programs that, in different ways, support cities in their strive for a more sustainable development. In Europe, more than 2,500 local and regional governments have signed the Aalborg Charter from 1994 and joined the European Sustainable Cities and Towns Campaign. The most ambitious of these local communities have signed the Aalborg Commitments from 2004.

Sustainable urban development ambitions have brought both broader concerns and new ambitions to urban governance and planning. For increasing sustainability, waste and pollution must be minimized, natural resources conserved, and the carrying capacity of ecosystems respected. The sustainable city agenda includes often compact, efficient land use; reduced car use, efficient public transport, efficient resource use, and reduced pollution and waste; restoration of natural systems; good housing and living environments, community participation and involvement, social equity and inclusion, and a sustainable economy [14]. Sustainability planning requires a holistic view of a city or region that includes equal concern for environmental, economic, and social sustainability [15]. It is conceived as a process, and sustainability, as a goal to work toward rather than something soon to be achieved. In order to achieve such goals, comprehensive plans, zoning ordinances, building codes, and planning processes must be revised for incorporating sustainability principles into urban environmental planning [15–17]. This does not only imply that environmental quality or resource conservation goals must be as important as economic but also that other goals, such as social justice and equity, must be addressed in relation to environmental efforts. For realizing such integrative ambitions, it is often stressed that planning and decision processes must become more open and integrated: all levels of government, business, civil society, and individual must work together [17].

Sustainability is a long-term goal, which can only be achieved through consistent strategic action over the long term. Sustainable urban development needs to rely on sound resource management and sustainable urban structures. To realize this and to reduce the demand of resources and flows of energy, water, waste, and materials, it is necessary to develop a more effective, integrated, and flexible infrastructure for water supply, waste, sanitation, heating, electricity, and fuels that is resilient to changes in population, economy, and climate and the hydrological regime. Urban systems should be multifunctional, serving more than one technical purpose and be able to integrate ecological, recreational, and aesthetic values. For increasing the capacity of municipalities to act strategically, an important challenge is to develop systems-based analysis tools in collaboration with stakeholders that allow a better overview and guidance for strategic action [4]. This is the challenge that has inspired many studies of urban ecosystems and urban metabolism.

Cities as Ecosystems

The concept of "urban ecosystems" has proved useful for investigating the relationship between urban areas and the environment [18]. Urban ecosystems are defined by Pickett [19] as "those in which people live at high densities, or where the built infrastructure covers a large proportion of the land surface." Research on urban ecosystems, often called urban ecology research, has grown and broadened during recent decades. This research can be divided into research on *ecological systems in the city* and on *the city as an ecological system* [20]. The former is dominating and includes many specialized study areas that focus on particular ecological aspects in urban areas [19], e.g., urban climates, soils, hydrology, vegetation, animals, nutrient and carbon flows, as well as different human and social urban systems. The latter analyzes the city as a whole ecosystem based on different ecosystem approaches, human ecosystem or sustainability assessment frameworks or watersheds as integrative tools.

Natural resource flows and their transformations are essential for the functional structure and dynamics of both natural and human ecosystems. From such a flow-based systems perspective, cities can be described as large stocks of accumulated materials stored in buildings and infrastructure, and they are characterized by large and concentrated flows of raw materials, energy, goods, waste, people, and information. In thermodynamic terms, cities are open systems that depend on imported energy and matter. The cities require vast amounts of resources to function. In the past, the amount of accessible resources in the nearest hinterland was a constraint to the growth of cities [21], but with improved transport and development of international trade, cities no longer rely on their immediate hinterlands for a growing share of resources, including food. Most environmental problems are connected to these natural resource flows.

Urban Metabolism

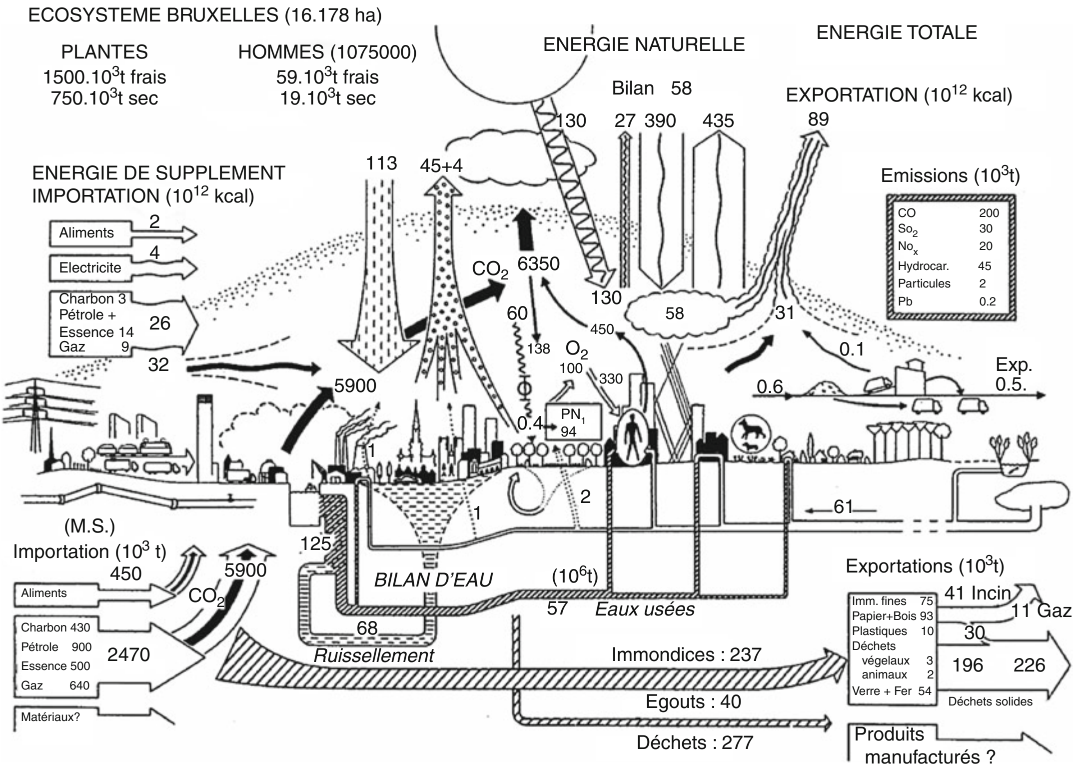
Urban flows and resource management in different cities have been analyzed in various studies since the 1970s. Figure 2 shows the urban metabolism of Brussels from the early 1970s from the study of Duvigneaud and Denayeyer-De Smet [22]. The term "urban metabolism" was launched by Abel Wolman [23] in his study on generalized flows in a "typical" city in the USA. Kennedy et al. [24] defines "urban metabolism" as "the sum total of the technical and socio-economic processes that occur in cities, resulting in growth, production of energy, and elimination of waste." The notion of "metabolism" is based upon analogy with the metabolism of organisms, which is one example of parallels between cities and ecosystems, which are central in urban ecology. The term "metabolism" has also become established in connection with material and substance flow analysis in other kinds of man-dominated systems, e.g., "industrial metabolism" studies [25] have analyzed material and substance flows in different regions.

The study of the urban metabolism involves in practice a holistic conceptualization of a city as a system of flows and inputs, outputs, and storage of focused flows, e.g., energy, water, nutrients, materials, and wastes that are then quantified for a certain time period. Figure 3 shows an example of a summary of results in terms of inflows and outflows in the metabolism of Sydney in 1990. By looking at the city as a whole system and by analyzing the pathways along which energy and materials, including pollutants, move, the goal is to improve resource management systems, allow assessment of impacts, and identify technological and organizational opportunities to increase the efficiency of resource use through, e.g., recycling of wastes as valuable materials and energy conservation [7].

Different Urban Metabolism Studies

Kennedy et al. [26] identified almost 50 "urban metabolism" studies on more than 30 different cities since the 1970s. After a few forerunners in the 1970s, urban metabolism studies reappeared in the 1990s, and the great majority of studies have been published since the year 2000. Hong Kong, Vienna, Stockholm, Toronto, London, and Paris have been subject for several studies.

Most "urban metabolism" studies have focused on energy, water, food (nutrients), materials

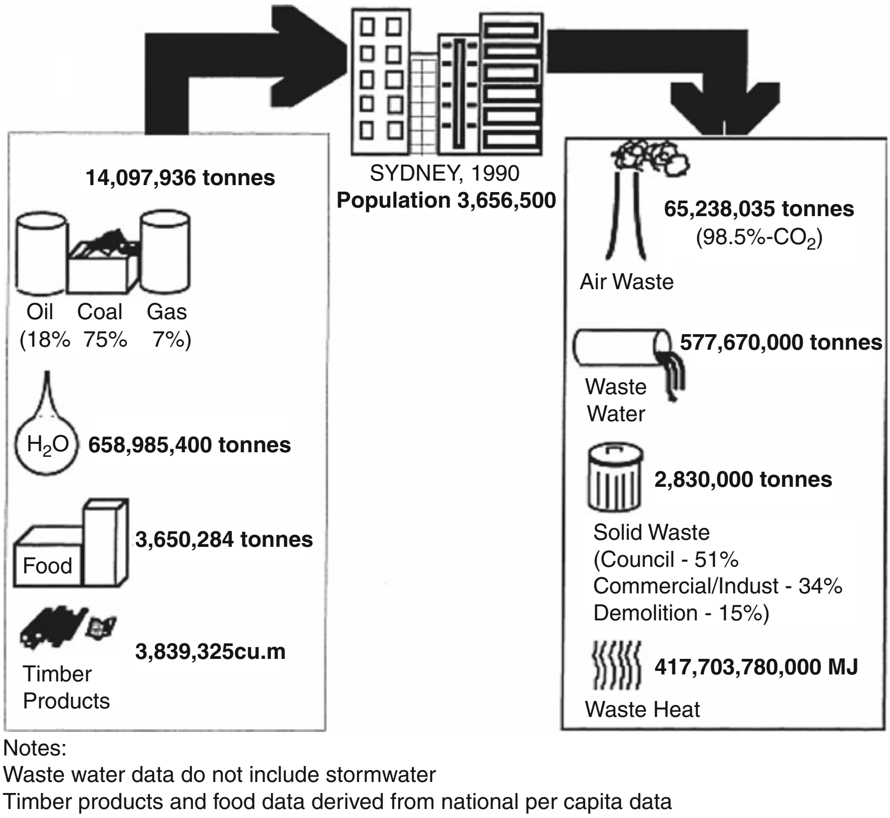


Natural Resource Flows and Sustainability in Urban Areas, Fig. 2 The urban metabolism of Brussels, Belgium, in the early 1970s. (From: [22])

(metals), and resulting pollution (typically CO_2, SO_2, NO_x, nutrient releases, or heavy metals) and waste in a limited time period. There are also examples of analyses of the historical development, such as for heavy metals in Stockholm [27] and nitrogen in food metabolism in Paris [28] and total material flows in Limerick [29]. In some other cities, such as Hong Kong and Toronto, repeated studies have provided some basis for comparison over time [26].

Most studies have been based on mass or energy measurement for different materials and energy flows, respectively. But some studies have sought to use agglomerated measurement units. Several of the earliest studies in the 1970s e.g., [30] described the metabolism in terms of emergy, an approach which more recently has been applied in connection with analyses of Taipei [31], Beijing [32], and Rome [33]. Related to the urban metabolism concept is also the application of the ecological footprint technique to cities [34]. The studies of Vancouver [35], Santiago [36], and cities of the Baltic region of Europe [37] in the 1990s have been followed by many studies. It has also become increasingly popular to describe the urban metabolism in terms of ecological footprints e.g., [38–41].

Many urban metabolism studies have focused primarily on urban material stocks and flows, e.g., of Lisbon [42], Singapore [43], and York [38]. In connection with heavy metal analyses, the urban stocks have been recognized as future environmental risks or potential raw material sources [44–47]. There has also been a growing interest for food flows and related nutrient flows [48–51]. Urban water metabolism has been addressed in several studies [52–56]. In recent years, urban CO_2 emissions in different cities have received growing interest, and many estimations are built on urban metabolism approaches [57]. There are also many broader studies that in different ways connect to urban metabolism,



Natural Resource Flows and Sustainability in Urban Areas, Fig. 3 Resource inputs and waste outputs in Sydney, 1990. (From: [7])

e.g., transportation of materials [58], and urban infrastructure systems [59]. An indication for the renewed and increasing interest in urban metabolism studies is that two projects have recently started under the EU Seventh Framework Programme: SUME [60] and BRIDGE [61].

Contributions by Urban Metabolism Studies

The urban metabolism studies have contributed to the overview and understanding of the flows and stocks of various materials in urban areas, and of pollution risks. They confirm the importance of cities for energy and material flows in different scales. Furthermore, they have shown that large parts of the resource flows accumulate in buildings, wastes, and pollution in urban areas. In some cases, they have indicated that accumulation results in surprisingly high concentrations of hazardous materials "hibernating" in cities, which bring future risks and challenges in connection with demolition and waste, diffuse pollution, and pollution of groundwater (e.g., [27]. Brunner and Rechberger [46] have inspired a growing research on the potentials for "urban mining."

The studies have also contributed to the increase in the understanding of the development of the urban flows in different cities, which often have changed dramatically during the twentieth century. Where city development has been characterized by increasing welfare, consumption, and mobility, the societal flows have grown dramatically and become more diversified, and spatial patterns have changed with the development and expansion of city regions. Most studied cities have, in recent decades, become more energy and material intensive, while other flows, such as water and wastewater, as well as different types of pollution show varying trends in different cities

[26]. While Asian cities may still be characterized by increasing levels of traditional pollutants, such as SO_2 and particulates, European cities have generally experienced a dramatic trend break of most of such traditional air pollutants, while emissions of NO_x and CO_2 may remain stable [10]. In parts of East, South, and Central Asia, urbanization has dramatically increased water consumption and water pollution, which adds pressure on scarce water resources in the regions.

Urban Metabolism Studies and Planning

Urban metabolism is often described as "a holistic approach to urban planning" [5] and is fundamental for developing sustainable cities and communities. Such overviews of flows in different cities have, in many cases, been important for informing and stimulating strategic discussions concerning resource provision, resource efficiency, and pollution. Metabolism studies contributed importantly to the development of flow-based ecocycle strategies and related introduction and campaigns for recycling in the 1990s. There are examples of collaboration between authorities and researchers in some studies e.g., [27] and that studies have been used for developing local sustainability indicators connected to energy or material consumption or ecological footprints. There are also examples of policy frameworks that have been developed for supporting sustainable urban policy development, e.g., the extended metabolism for human settlements from Australia that link the urban metabolism to both socio-economic dynamics and quality of life (Fig. 4). However, it is still rare that urban metabolism studies and perspectives have been fully integrated into local strategies or planning processes.

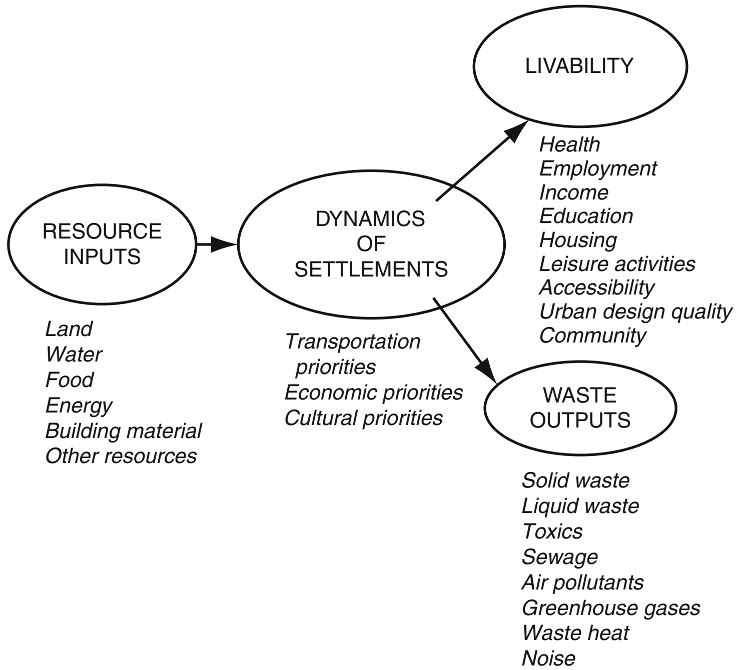
Kennedy et al. [26] suggest that particularly relevant practical applications of urban metabolism for urban planners and designers include sustainability reporting, urban greenhouse gas accounting, mathematical modeling for policy analysis, and urban design. Urban metabolism studies have often contributed to sustainability reporting and indicators. Indicators based on metabolism studies have many advantages: they are not only indicative of a city's sustainability but also are scientifically valid, relevant to urban planners and dwellers, comparable over time, and understandable. With the growing interest for setting local goals for emissions of greenhouse gases and ecological footprints, metabolism approaches are needed for estimating the emissions in different cities. An increasing number of cities – from the planned city of Masdar in Abu Dhabi to big cities such as Copenhagen and Seattle – have, in the last few years, declared ambitions to become carbon neutral within the next few decades. Partly connected is a trend among "Sustainable Cities" to set targets in terms of decreasing their ecological footprints.

The ambitions of metabolism studies to contribute to regional assessment, modeling, and scenario construction have only seldom been realized, though there are exceptions, such as City Region 2020 [4] and Maastricht 2030 [5], assessment and scenario studies based on metabolism approaches. There are also examples of mathematical models of metal and nutrient flows, such as SIMBOX [62] and STAN [46, 63], that can be used for simulating future changes and identify solutions to perceived problems. Urban metabolism has furthermore been used for development of tools for designing and redesigning urban areas and their infrastructure [26, 64].

Limitations and Challenges

The metabolism studies have traditionally limitations that can be connected to data availability, perspectives, and available resources and priority setting in individual projects. Observations and discussion about the difficulties of the studies in terms of data accessibility, missing information, and uncertainties are often neglected in presentations of various studies. Statistics do not adapt very well for following all types of flows and particularly not in a local area. Many types of statistical data and inventories are often only available at the national level, and these data are often used as basis for estimating local flows. Data on water, energy, and waste flows handled by local public organizations are more accessible, which explains the dominance of these in studies. With the exception of very few city areas, such as Hong Kong, which has been subject to repeated studies [65–67], cities are difficult to delimit,

Natural Resource Flows and Sustainability in Urban Areas, Fig. 4 Extended metabolism model of human settlements. (From: [7])



which creates problems to estimate inflows to the urban system as well as their international and local exchanges.

Integrating nature and socioeconomic aspects is still an important challenge in connection with ecosystem analysis. The analysis of urban metabolism has largely been dominated by quantification of material inflows and outflows, and the connections to social aspects have often been fairly weak. Only few studies have made efforts to analyze flows in their broader social, economic, and geographical setting, and studies have seldom been able to connect to decision-making or involve stakeholders. Despite increased interest in international aspects of material flows, only limited systematic efforts to evaluate ecological linkages of various cities in more detail than standardized ecological footprints have been made. It is also rare with several levels of scale in the analysis. Summarized flows contribute to an overview but are not adequate for making connections to decision points in society or say much about their qualitative usefulness to society [3].

In making connections to decision-making, planning, and strategy development, it is necessary with closer connections to decision points and institutional structures, and the analysis must become more flexible in scale and detail. A more systematic multilevel analysis and connections between different levels of scale and governance are needed. In many cities, there is a lack of evaluation of local ecocity projects in relation to the whole city and its development. In a discussion paper on how cities can integrate global concerns, Xia Bai [68] presented some thoughts on how scale considerations should be taken into account more systematically when addressing city strategies, which can inspire urban metabolism studies to become more policy relevant.

Future Directions

Many innovative ideas regarding how a city could be planned or developed to minimize its resource consumption have been proposed in recent decades. There are numerous ecocity or sustainable city initiatives in different cities around the world. Cities such as Curitiba, Freiburg, London, Vancouver, Stockholm, Tianjin, and Melbourne can show inspiring examples of initiatives and positive developments. Public planning authorities have developed growing awareness of sustainability issues, and it is difficult to find any comprehensive urban plan that lacks sustainability goals. In local ecocity projects, flow-based initiatives are common in relation to waste, energy, and traffic. Despite this marked increase in awareness and ambitions, there is still a lack of powerful initiatives that decisively shift urban development in a sustainable direction. Even cities that have a high profile as sustainability forerunners, such as Copenhagen and Malmö [69] or Portland [70, 71], may show developments that are fairly inconsistent with declared sustainability ambitions, particularly in terms of road traffic and urban sprawl.

There is generally a need for more comprehensive analysis at the city and big city region level by revitalized urban metabolism efforts. It is, however, possible that the increasing interest among cities concerning climate change mitigation and ecological footprints will stimulate further analysis of the urban metabolism at the city level.

Urban metabolism and ecological footprints have become influential concepts. But it is a bit paradoxical that while urban ecology and sustainable urban development research and actions have grown dramatically, the analysis of urban systems at the city level or big city regional level has played only a marginal role. Efficiency of energy, problems of traffic flows, and consumption are addressed in an urban setting, but the scale is most often far below the city or city region level (e.g., house, household, city district, infrastructural sector). There exists seldom a sufficiently good information basis for assessing the development of the city and its region and their external exchanges with the surrounding world and for developing comprehensive strategies at the city level.

There are many intriguing questions concerning the development of big city regions, e.g., connected to urban form and location policies that should be systematically analyzed from a metabolic perspective. From studies primarily in the UK and the Netherlands, there is evidence of positive effects in terms of provisions of ecoservices and minimized loss of natural environment and biodiversity from promotion of more compact urban development [72, 73]. However, there is still debate about the desirability and challenges of such compact city ambitions e.g., [74] and need for more empirical evidence [3].

In dynamic big city regions that are characterized by regional enlargement and polycentric developments, such issues are very complex and challenging but need to be analyzed further. In national planning systems all over the EU, the concept of polycentricity is implemented [75]. The European Spatial Development Perspective [76] strongly argued that polycentric urban systems are more equitable, efficient, and sustainable than monocentric or dispersed urban systems and should therefore be actively encouraged, but this hypothesis remains to be tested, and even if polycentric regions have certain potential advantages, these opportunities still need to be identified and used. Major research questions in both the large urban metabolism projects under EU-FP7 connect to these issues. They both focus on urban form, traffic, and land use, and their ambitions to connect to planning and institutional structures are higher and more elaborated than most past urban metabolism projects.

There are many challenges for future urban metabolism studies in relation to developing the overview and understanding of the urban flow systems and to become more policy relevant. Only about 30 cities have been so far analyzed, and there have not been much comparative analysis, generalization, and theoretical development. Kennedy et al. [26] suggested that identification of various classes and comparative analysis of various factors, such as climate, age of city, and development history, should be further focused. Comparative studies are difficult due to very different preconditions in various cities, and comparisons between existing studies are problematic due to inconsistent basic data and lack of standardization of procedures. Even though there has been standardization effort in some study areas (e.g., metal metabolism studies), such challenges are likely to remain. Despite this, comparative studies and comparisons with other cities may still be valuable and almost necessary for developing perspectives on different resource flows and their efficiency in individual cities. Comparative studies can also form a base for looking closer into driving forces of the urban metabolism and effects of different developments or strategies. Besides improving the assessment of linkages of the urban metabolism in different scales, future studies should also aim to improve the overview of storage and accumulation processes in the urban environment. Analysis of urban natural flow systems is of central importance for a more comprehensive understanding of the resource use in the city and for assisting cities to contribute more forcefully to global sustainable development.

Bibliography

Primary Literature

1. UNFPA (2007) State of world population 2007. Unleashing the potential of urban growth. United Nations Population Fund, New York
2. Grimm NB, Faeth SH, Golubiewski NE, Redman CL, Wu J, Bai X, Briggs JM (2008) Global change and the ecology of cities. Science 319:756–760
3. Huang SL, Yeh CT, Chang LF (2010) The transition to an urbanizing world and the demand for natural resources. Curr Opin Environ Sustain 2:136–143
4. Ravetz J (2000) Integrated assessment for sustainability appraisal in cities and regions. Environ Impact Assess Rev 20:31–64
5. Rotmans J, van Asselt MBA, Vellinga P (2000) An integrated framework for sustainable cities. Environ Impact Assess Rev 20:265–276
6. Alberti M (1996) Measuring urban sustainability. Environ Impact Assess Rev 16(4–6):381–423
7. Newman PWG (1999) Sustainability and cities: extending the metabolism model. Landsc Urban Plan 44:219–226
8. Batty M (2008) The size, scale, and shape of cities. Science 319:769–771
9. Luckin B, Massard-Guilbaud G, Schott D (eds) (2005) Environment and the City. Modern European cities and the management of their resources. Ashgate, Aldershot
10. Douglas I, Hodgson R, Lawson N (2002) Industry, environment and health through 200 years in Manchester. Ecol Econ 41(2):235–255
11. Keiner M, Koll-Schretzenmayr M, Schmid WA (2005) Managing urban futures: sustainability and urban growth in developing countries. Ashgate Publishers, Aldershot
12. Ooi GL (2009) Challenges of sustainability for Asian urbanization. Curr Opin Environ Sustain 2009(1): 187–191
13. Agudelo-Vera CM, Mels AR, Keesman KJ, Rijnaarts HH (2011) Resource management as a key factor for sustainable urban planning. J Environ Manag 92: 2295–2303
14. Roseland M (2005) Toward sustainable communities: resources for citizens and their governments. New Society, Gabriola Island
15. Newman P, Jennings I (2008) Cities as sustainable ecosystems: principles and practices. Island Press, Washington, DC
16. Girardet H (1999) Creating sustainable cities. Green Books/The Schumacher Society, Totnes
17. Register R (2006) Ecocities: rebuilding cities in balance with nature. New Society, Gabriola Island
18. McIntyre NE, Knowles-Yánez K, Hope M (2008) Urban ecology as an interdisciplinary field: differences in the use of “Urban” between the social and natural sciences. In: Marzluff JM (ed) Urban ecology. An international perspective on the interaction between humans and nature. Springer, New york, pp 49–65
19. Pickett STA, Cadenasso ML, Grove JM, Boone CG, Groffman PM, Irwin E, Kaushal SS, Marshall V, McGrath BP, Nilon CH, Pouyat RV, Szlavecz K, Troy A, Warren P (2011) Urban ecological systems: scientific foundations and a decade of progress. J Environ Manag 92:331–362
20. Hultman J (1993) Approaches and methods in urban ecology. Geograf Ann Ser B Hum Geogr 75(1):41–49
21. McGranahan G, Marcotullio P, Bai X, Balk D, Braga T, Douglas I, Elmqvist T, Rees W, Satterthwaite D, Songsore J, Zlotnick H, Eades J, Ezcurra E, Whyte A (2005) Urban systems. In: Millennium ecosystem assessment: conditions and trends assessment. Island Press, Washington, DC
22. Duvigneaud P, Denayeyer-De Smet S (1977) L’Ecosystéme Urbs, in L’Ecosystéme Urbain Bruxellois. In: Duvigneaud P, Kestemont P (eds) Productivité en Belgique. Traveaux de la Section Belge du Programme Biologique International. Bruxelles, pp 581–597
23. Wolman A (1965) The metabolism of cities. Sci Am 213: 179–190
24. Kennedy CA, Cuddihy J, Engel Yan J (2007) The changing metabolism of cities. J Ind Ecol 11(2):43–59
25. Ayres RU, Simonis UK (eds) (1994) Industrial metabolism. United Nations University Press, Tokyo
26. Kennedy C et al (2011) The study of urban metabolism and its applications to urban planning and design. Environ Pollut 159:1965–1973
27. Bergbäck B, Johansson K, Mohlander U (2001) Urban metal flows – a case study of Stockholm, review and conclusions. Water Air Soil Pollut Focus 1(3/4):3–24
28. Barles S (2007) Feeding the city: food consumption and flow of nitrogen, Paris, 1801–1914. Sci Total Environ 375:48–58
29. Browne D, O’Regan B, Moles R (2009) Assessment of total urban metabolism and metabolic inefficiency in an Irish city-region. Waste Manag 29(10):2765–2771
30. Zucchetto J (1975) Energy, economic theory and mathematical models for combining the systems of man and nature. Case study, the urban region of Miami. Ecol Model 1:241–268
31. Huang SL, Hsu WL (2003) Materials flow analysis and energy evaluation of Taipei’s urban construction. Landsc Urban Plan 63(2):61–74

32. Zhang Y, Yang Z, Yu X (2009) Evaluation of urban metabolism based on energy synthesis: a case study for Beijing. Ecol Model 220(13–14):1690–1696
33. Ascione M, Campanella L, Cherubini F, Ulgiati S (2009) Environmental driving forces of urban growth and development. An emergy-based assessment of the city of Rome, Italy. Landsc Urban Plan 93:238–249
34. Wackernagel M, Kitzes J, Moran D, Goldfinger S, Thomas M (2006) The ecological footprint of cities and regions: comparing resource availability with resource demand. Environ Urban 18(1):103–112
35. Wackernagel M, Rees W (1995) Our ecological footprint: reducing human impact on the earth. New Society, Gabriola Island
36. Wackernagel M (1998) The ecological footprint of Santiago de Chile. Local Environ 3(1):7–25
37. Folke C, Jansson Å, Larsson J, Costanza R (1997) Ecosystem appropriation by cities. Ambio 26:167–172
38. Barrett J, Vallack H, Jones A, Haq G (2002) A material flow analysis and ecological footprint of York. Technical report. Stockholm Environment Institute, Stockholm
39. Collins A, Flynn A, Weidmann T, Barrett J (2006) The environmental impacts of consumption at a subnational level: the ecological footprint of Cardiff. J Ind Ecol 10(3):9–24
40. Chartered Institute of Wastes Management (2002) A resource flow and ecological footprint analysis of greater London. Best Foot Forward, London
41. Eaton RL, Hammond GP, Laurie J (2007) Footprints on the landscape: an environmental appraisal of urban and rural living in the developed world. Landsc Urban Plan 83:13–28
42. Niza S, Rosado L, Ferrão P (2009) Urban metabolism: methodological advances in urban material flow accounting based on the Lisbon case. J Ind Ecol 13(3):384–405
43. Schulz NB (2007) The direct material inputs into Singapore's development. J Ind Ecol 11(2):117–131
44. Sörme L, Bergbäck B, Lohm U (2001) Century perspective of heavy metal use in urban areas. A case study in Stockholm. J Water Air Soil Pollut Focus 1(3–4):197–211
45. Svidén J, Jonsson A (2001) Urban metabolism of mercury turnover, emissions and stock in Stockholm 1795–1995. J Water Air Soil Pollut Focus 1(3–4): 79–196
46. Brunner PH, Rechberger H (2004) Practical handbook of material flow analysis. CRC Press, Boca Raton
47. Obernosterer R, Brunner PH (2001) Urban metal management the example of lead. J Water Air Soil Pollut Focus 1(3–4): 241–253
48. Baker LA, Hope D, Xu Y, Edmonds J, Lauver L (2001) Nitrogen balance for the Central Arizona Phoenix (CAP) ecosystem. Ecosystems 4:582–602
49. Færge J, Magid J, Penning de Vries FWT (2001) Urban nutrient balance for Bangkok. Ecol Model 139:63–74
50. Forkes J (2007) Nitrogen balance for the urban food metabolism of Toronto Canada. Resour Conserv Recycl 52:74–94
51. Neset TSS, Lohm U (2005) Spatial imprint of food consumption. A historical analysis for Sweden, 1870–2000. Hum Ecol 33(4):565–580
52. Hermanowicz SW, Asano T (1999) Abel Wolman's the metabolism of cities' revisited: a case for water recycling. Water Sci Technol 40(4):29–36
53. Gandy M (2004) Rethinking urban metabolism: water, space and the modern city. City 8(3):363–379
54. Thériault J, Laroche AM (2009) Evaluation of the urban hydrologic metabolism of the Greater Moncton region, New Brunswick. Can Water Resour J 34(3):255–268
55. Sahely HR, Kennedy CA (2007) Integrated systems flow model for quantifying environmental and economic sustain-ability indicators: case study of the City of Toronto urban water system. ASCE J Water Resour Plann Manage 133(6):550–559
56. Baker LA (ed) (2009) The water environment of cities. Springer, New York
57. Kennedy C, Steinberger J, Gasson B, Hillman T, Havránek M, Hansen Y, Pataki D, Phdungsilp A, Ramaswami A, Villalba Mendez G (2009) Greenhouse gas emissions from global cities. Environ Sci Technol 43:7297–7302
58. Fischer-Kowalski M, Krausmann F, Smetschka B (2004) Modelling scenarios of transport across history from a socio-metabolic perspective. Rev – Fernand Braudel Center 27:307–342
59. Bettencourt L, Lobo J, Helbing D, Kuhnert C, West GB (2007) Growth innovation scaling and the pace of life in cities. Proc Natl Acad Sci U S A 104:7301–7306
60. SUME (Sustainable Urban Metabolism for Europe) FP7 Collaborative Research Project. http://www.sume.at. Accessed 1 July 2011
61. BRIDGE programme (sustainaBle uRban plannIng Decision support accountinG for urban mEtabolism). http://www.bridge-fp7.eu. Accessed 1 July 2011
62. Baccini P, Bader HP (1996) Regionaler Stoffhaushalt. Spektrum Akad. Verlag, Heidelberg
63. Cencic O, Rechberger H (2008) Material flow analysis with software STAN. J Environ Eng Manage 18(1):3–7
64. Oswald F, Baccini P (2003) Netzstadt: designing the urban. Birkhäuser, Basel
65. Newcombe K, Kalma J, Aston A (1978) The metabolism of a city: the case of Hong Kong. Ambio 7:3–15
66. Boyden S, Millar S, Newcombe K, O'Neill B (1981) The ecology of a city and its people: the case of Hong Kong. Australian National University Press, Canberra
67. Warren-Rhodes K, Koenig A (2001) Escalating trends in the urban metabolism of Hong Kong: 1971–1997. Ambio 30(7):429–438
68. Bai X (2007) Integrating global environmental concerns into urban management – the scale and readiness arguments. J Ind Ecol 2(11):15–29
69. Anderberg S, Clark E (2011) The green and sustainable Øresund region: eco-branding Copenhagen and Malmö. In: Vojnovic I (ed) Sustainability: a global

urban context. Michigan State University Press, Lansing. (in press)
70. Redkin A. Portland again tops a sustainable cities list, dot earth, the opinion pages, New York Times. http://dotearth.blogs.nytimes.com. Accessed 22 June 2011
71. Sustainable Business Oregon Portland's traffic is 16th worst. http://www.sustainablebusinessoregon.com. Accessed 22 June 2011
72. Tratalos J, Fuller R, Warrren P, Davies R, Gaston K (2007) Urban form, biodiversity potential and ecosystem services. Landsc Urban Plan 83:308–317
73. Geurs KT, van Wee B (2006) Ex-post evaluation of thirty years of compact urban development in the Netherlands. Urban Stud 43(1):139–160
74. Jenks M, Burton E, Williams K (eds) (2000) The compact city: a sustainable urban form? Oxford Brookes University, Oxford
75. Halbert L, Pain K, Thierstein A (2006) European polycentricity and emerging Mega-City-Regions – "one size fits all" policy? Built Environ 32(2):194–218
76. European Commission (1999) European spatial development perspective. Towards balanced and sustainable development of the territory of the European Union

Books and Reviews

Anderberg S, Prieler S, Olendrzynski K, de Bruyn S (2000) Old sins – industrial metabolism, heavy metal pollution and environmental transition in Central Europe. UN University Press, Tokyo

Baccini P, Brunner PH (1991) Metabolism of the anthroposphere. Springer, Berlin

Bernhardt C (ed) (2004) Environmental problems in European cities in the nineteenth and twentieth centuries. Waxmann, Münster\New York

Odum HT (1996) Environmental accounting: emergy and environmental policy making. Wiley, New York

Tarr JA (2002) The metabolism of the industrial city: the case of Pittsburgh. J Urban Hist 28(5):511–545

Remediation in Karst

Petar T. Milanović
Belgrade, Serbia

Article Outline

Glossary

Collapse Abrupt breakdown of unconsolidated sediments deposited above the cave or karst channel or breakdown of the cavern overburden. Synonym – subsidence

Cutoff A construction below ground level intended to reduce water seepage (grout curtain or cutoff wall)

Cutoff wall A watertight wall of clay or concrete which is built up by applying structure in the form of trench or overlapping piles

Dam failure Collapse or movement of part of a dam or its foundation, so that the dam cannot retain water

Grouting Is a procedure of which grout mix is injected into the fissures, cavities, and voids in rock formations in order to reduce its permeability and to improve geotechnical properties of rock mass

Induces collapse Breakdown of unconsolidated sediments as the result of human activities (groundwater pumping, mining, tunneling, or man-made reservoir influence)

Karst Terrain composed of high soluble rocks (limestone, dolomite, gypsum, salt, and conglomerate with carbonate matrix) and very risky environment for any surface and underground structure construction

Definition of Subject

Karst areas are extremely complex and produce a great variety of geomorphological, hydrogeological, hydrological, and engineering geological conditions. As a consequence, a number of structures in karst need remediation during operation despite extensive investigations and extensive protective measures applied during design stage and construction. The risk of constructing in karst cannot be eliminated completely due to the hydrogeological and geotechnical complexity of karst, even when best engineering practices are followed. Due to karst complexity, conventional geotechnical measures have limitations that, sometimes, are useless. To deal with karst successfully, innovation, engineering practice, execution feasibility, and commercial understanding are required.

The most frequent geotechnical difficulties in karst engineering are presence of cavern at foundation of any building or along the grout curtain and tunnel routes, leakage from reservoirs, groundwater bursts during underground excavations, provoking induced subsidence, and number of natural and anthropogenic impacts. In karstified rock masses, a number of karst features such as caverns and karst conduits are hardly predictable from surface observations or during construction. In many cases, these features become active during operation. To prevent failures, remedial works become necessary.

The feasible remediation alternatives are different kinds of geotechnical treatment: surface treatment and sealing in underground. Commonly applied surface methods are compaction of surface clayey layer (clay lining), different kinds of geomembrane, shotcrete lining, shallow grouting

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_975

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_975-1

blanket, dental plugging, heavy reinforced concrete slabs, and construction of cylindrical dams or dikes to isolate large ponors and estavelles. Common underground watertight structures are different kinds of cutoffs, concrete plugs and different drainage structures, and conical grouting ahead in the case of tunnel driving. Frequently applied grouting materials are cement, clay, asphalt, or hot bitumen. Particular approaches and materials are necessary for grouting the cavernous zones (rock blocks, crushed aggregate, sand, gravel, chemicals, synthetic sponge, and polyurethane foam). In many cases, the successful solution needs a combination of surface and underground structures and combination of different methods and technologies.

Introduction

In the first half of the twentieth century, the understanding of karst and karstification from the engineering viewpoint was in its infancy. During that period, a significant number of different structures failed in karst, particularly dam and reservoirs. In many of these cases, the reservoir never filled up despite intensive remedial works, ending in abandonment. Different technologies and various structures adapted to remediation in karst were not developed. For a long time, the term karst had a bad meaning for engineers, suggesting failure. By increasing the knowledge in karstology and development of remediation technologies, the number of failures has considerably decreased. However, karst is still far from being accepted as a friendly environment for construction of any large structure: dam, reservoir, tunnel, railroad, roads, and buildings. The risk in karst cannot be eliminated and, however, can be minimized to the technically and economically acceptable level. These general premises are well known to anybody working in karst.

Surface Remediation Measures for Dams and Reservoirs

Approximately 80% of dams in karst that needs additional watertightness treatment after construction have been successfully finalized. In most cases, the leakage occurred during the first filling. In more than 30 cases, a leakage rate between 2 and 50 m^3/s was registered. More than 60 dams in evaporites have needed rehabilitation. In a majority of these cases, rehabilitation was successful. However, some dams were abandoned because problems were too complicated for the available sealing technology: Montejaque (Spain), Hales Bar (USA), Vrtac (Montenegro).

Many attempts have been made at the surface to stop water from sinking underground either coming from riverbeds or from reservoirs. To reduce water losses from the reservoirs, in present practice, the following methods are being applied:

- Protection of alluvial reservoir bottom by surface compacting and construction of impervious clay blankets
- Covering karstified limestone (river bottom or reservoir banks) by shotcrete
- Protection of sinking zones in alluvial overburden by geosynthetics for soil reinforcement
- Plugging of ponor (swallow holes) inlet channels by SCC
- Construction of grouting carpets (consolidation treatment of karstified surface or karstified paleo-refief)
- Encompassing large swallow holes or estavelles by cylindrical concrete dams
- Isolating large ponor (swallow hole) zones by dykes
- Closing estavelles by concrete plugs equipped with non-return valves
- Large cracks plugging by grouting
- Reservoir bottom protection by aeration tubes
- Dental treatment in the vicinity of different structures or in the reservoir limestone banks

Subsidence development is a very common process which endangers safety and integrity of reservoirs. However, if a reservoir is situated within a hydrogeological closed regional structure, subsidence does not provoke seepage out from the reservoir.

Surface treatment, particularly construction of clay blankets, is commonly used to protect reservoir watertightness if the groundwater table is deep below the reservoir bottom and dam sites. In the case of the presence of estavelles (intermittent openings active only in wet seasons) at the reservoir banks and bottoms, the surface treatment is very complicated [25]. An impervious surface blanket is commonly applied for shallow reservoirs Hutovo Reservoir (Herzegovina), Hammam Grouz, and Ourkis (Algeria).

If the consequence of alluvial ponors (subsidences) at the bottom of reservoirs is seepage, their activity may be stopped by using different kind of geotextile (geomembranes). In this case, the ponor opening (collapse funnel at the surface) and channel through the alluvial section are to be filled with crushed stone according to the principle of inverted filter. An inverted filter consists of layered crushed stones in such a way that each layer is progressively finer-grained up toward the surface, capped with a layer of impervious compacted clay, and, finally, covered with a geomembrane.

Geomembranes are frequently applied as waterproofing technology for reservoir and tailings bottom and banks (Fig. 1a).

In some cases, heavily reinforced concrete slabs are applied as a conservative and less risky solution (Fig. 1b).

In solving the problem of watertightness of the reservoirs, the bottom of which is being covered with alluvial sediments, in many cases, great attention should be paid to the captured air in the karstified bedrock and intergranular porosity in the alluvial deposits. During saturation, the underground water moves (presses) air out from the aeration zone, producing air bubbles below the geomembrane. If the air pressure reaches the limit of the plastic deformations, the geomembrane explodes.

To solve this problem, an aeration pipe (∅ 200–600 mm) has to be installed in that opening, and the cleaned part of the rock around the pipe has to be cemented. The end of pipe must be above the reservoir level (Fig. 2).

Shotcrete represents an efficient technology to achieve the required level of watertightness for reservoirs, riverbeds, and tunnels located in karstified carbonate rocks. It is placed on the rock surface under pressure so that it penetrates in fractures, caverns, and cavities. It has adhesive capacity and sealing action in the same time, preventing water losses through the karstified rock. It is also used for dental sealing as well as for blanketing of large surfaces of karstified rocks (reservoir flanks, riverbeds).

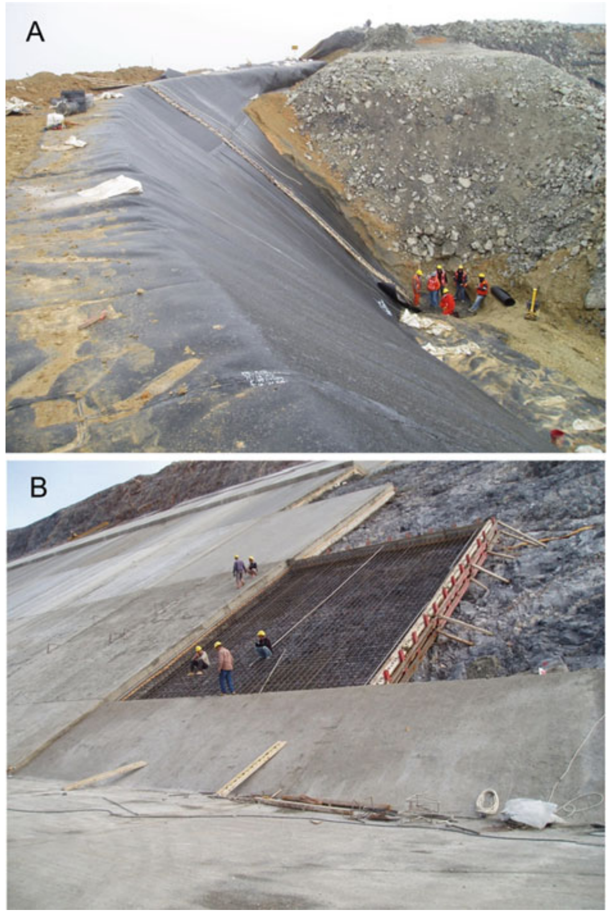
The thickness of the shotcrete for lining the shallow reservoir or riverbed flanks is 5 cm on average. To prevent the occurrence of micro-cracks, steel reinforcement meshes are used (∅ 3 mm) or different fiber types (steel, plastic, or glass). The length of the fibers is from 3 to 30 mm, i.e., at least three times the diameter of the largest grain in the aggregate. In order to assure satisfactory anchoring, ends of steel fibers should be thickened or bent.

Important experience in shotcrete application in karsts was acquired during the waterproofing of the largest European lost river Trebišnjica (Herzegovina). The water losses through the ponors and widened cracks along the 65 km of the riverbed amounted to 75 m^3/s under natural conditions. To secure the impermeability of the extremely karstified zone, the bed and flanks of the upper storage reservoir of Reversible PP Čapljina, an area of 2.2×10^6 m^2, were blanketed with shotcrete (Fig. 3).

During the dry period, the groundwater levels are deep beneath the bottom of the riverbed, lasting several months in a year. In the wet period of the year, groundwater rises extremely fast up to the surface, creating strong and uplift to the shotcrete. Usually, uplift is concentrated, and the demolished part of shotcrete is localized to a range from 0.5 to a few m^2, (Fig. 4) [21].

Rarely, strong uplift demolished a few hundreds of square meters of shotcrete. In these cases, the pressure releasing valves are required to eliminate uplift pressure.

One of the ways to solve problems of estavelles in the storage reservoirs is by closing its opening with concrete plugs equipped with non-return valves. These structures should be built at the opening of estavelle, and the surrounding rock should be blanketed with concrete. Under the reservoir water head, these valves prevent losses. When estavelle discharges water, the valves are opened, and water flows through them in the



Remediation in Karst, Fig. 1 (**a**) Tailings bank waterproofing by geomembrane; (**b**) heavy reinforced concrete slab to prevent seepage through the karstified reservoir bottom and bank

reservoir as long as the underground water pressure exceeds the water pressure in the reservoir. These constructions are not always successful because new estavelles can be created in the surrounding bottom area.

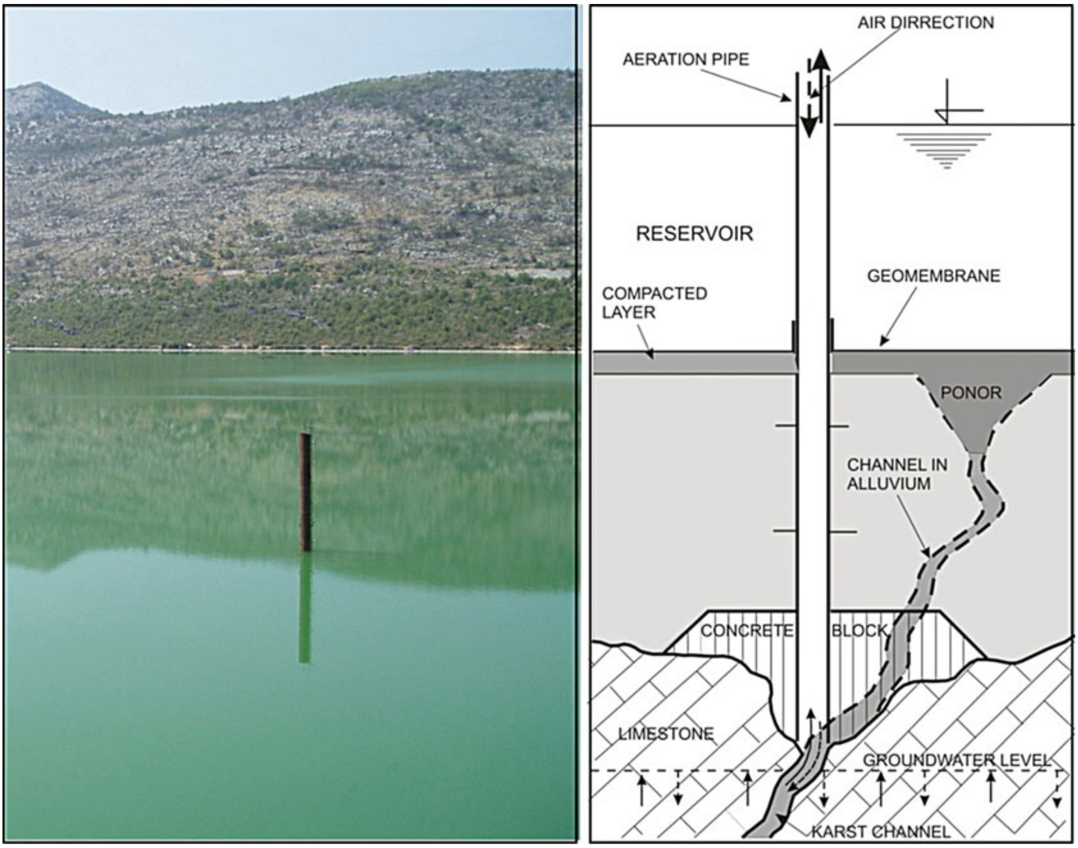
To prevent losses from the Nikšićko Polje (Montenegro), a large ponor Slivlje (swallowing capacity of around 120 m^3/s) was isolated by a cylindrical dam but without great success. The water losses were decreased by approximately 10%. A similar structure was constructed around the large estavelle Opačica at the rim of storage reservoir in the same polje, which prevented the loss of water from the reservoir. Similar structures were applied in Chinese karst to prevent seepage from reservoirs (Fig. 5).

The rock-filled or earth-filled dikes can be used to isolate the broad ponor zone at the edges of the storage reservoirs, the length of which can reach hundreds of meters. Foundation of those dikes is in the nonkarstified or partially karstified rock mass (Buško Blaro Reservoir, Bosnia and Herzegovina, Mavrovo Reservoir, FYUR Macedonia).

Underground Remediation Measures

Underground sealing is a geotechnical protection measure applied either in the karst aquifer or above it in the aeration zone.

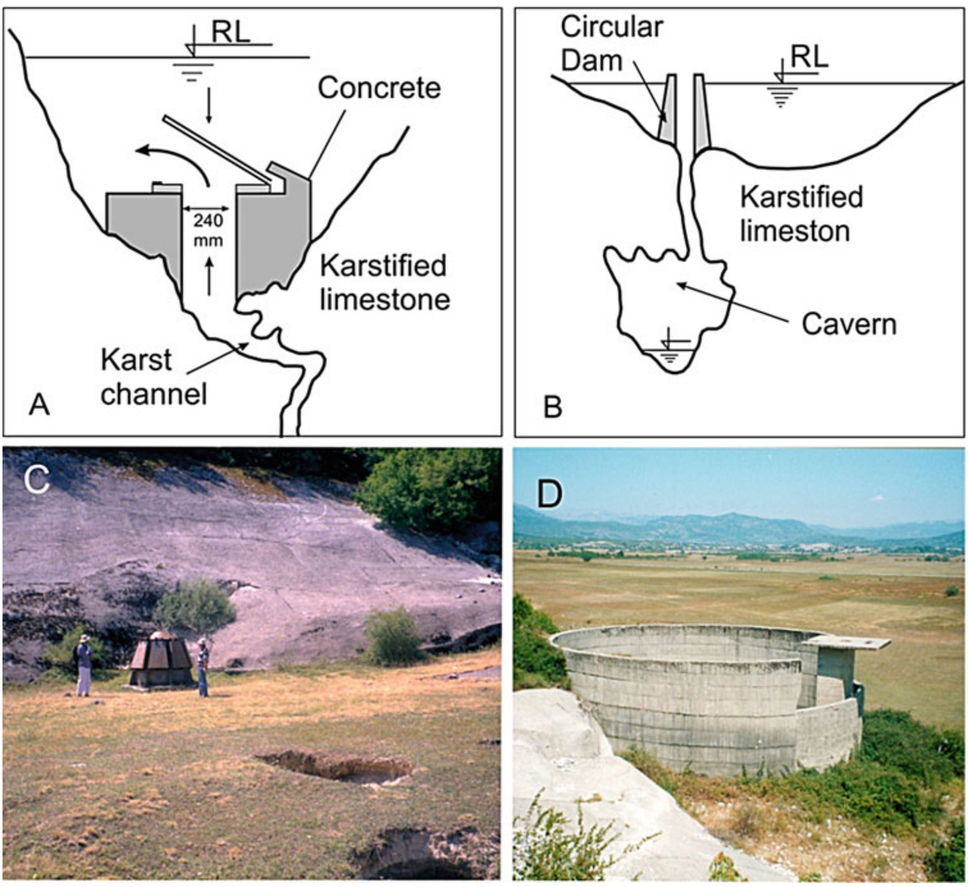
The most common underground protective and remediation measure are grout curtains. Due to specificity of the karst hydrogeological nature, the grout curtains executed in karstified rock mass are more complex and much larger than curtains in other geological formations. Following examples confirm this conclusion:



Remediation in Karst, Fig. 2 Aeration pipe installed to protect waterproofing blanket at reservoir bottom against destructive role of air under pressure

Remediation in Karst, Fig. 3 Shotcrete blanket to prevent seepage from karstified and porous riverbed

Remediation in Karst, Fig. 4 Shotcrete at reservoir bank demolished by strong and concentrated groundwater uplift



Remediation in Karst, Fig. 5 One-way valve (**a**) reservoir in China and (**c**) reservoir in Montenegro; (**b**) cylindrical dam in China and (**d**) cylindrical dam to prevent seepage from reservoir in Montenegro

– Ataturk (Turkey), surface 1,200,000 m^2, length 5.5 km, depth up to 300 m
– El Cajon (Honduras), surface 610,000 m^2
– Berke (Turkey), surface 533,000 m^2, depth up to 235 m
– Dokan (Iraq), 471,000 m^2
– Khao Laem Dam (Thailand), 437,000 m^2
– Slano (Yugoslavia), surface 404,224 m^2, length 7.011 m
– Keban (Turkey), 338,000 m^2
– Salman Farsi (Iran), 250,000 m^2

The design of the grout curtain is tentative and depends on the findings compiled during execution. Modifications and adaptations of curtain route, depth, number of grouting rows, and grouting technology, on the basis of the geological findings in subsurface, are calculated from experiences learned from many projects in karst.

The possible appearance of karst phenomena at grout curtain route might call for special and time-consuming treatment performing from adits or shafts, which have to be excavated from grouting galleries. Because of that, suggested distance between grouting galleries is not more than 30 m. A few typical examples are presented below.

Extreme nonhomogeneity of the karstified rock mass leads to great variability of grout mass consumption. During execution of grout curtain in karst, the common case are boreholes with very low or low grout mass consumption beside a borehole with consumption of a few thousand kg/m'.

The Slano Reservoir (Montenegro) is protected by a 7011 m long grout curtain. At only one section of this curtain at a depth 100–110 m, the following amounts were used: 4,110 m^3 of aggregate and 2300 t of grout mix (cement, clay, sand, and calcium soda). This section was situated 40 m below minimal groundwater level. Totally, this curtain required 91,531 t of material (27,826 t cement, 48,666 t clay, 307 t bentonite, 1766 t calcium soda, 4873 t sand, 8050 aggregate (gravel), and 43 t silicate). By primary grouting work, the initial seepage of 34 m^3/s was reduced to 3.5 m^3/s [27]. Due to degradation (erosion) of the curtain during the past 25 years, the seepage gradually increased to $\approx$ 7 m^3/s in the case of a full reservoir.

In some cases of heavily karstified limestone, a triple-row curtain may be necessary. In the case of Berke Dam [1] quinary holes (B) at a spacing of 0.75 m were planned for triple rows in heavily karstified limestone. The effectiveness of grouting is obvious as demonstrated by the comparative average grout mass consumption: P holes, 3270 kg/m'; S, 632.9 kg/m'; Q, 142.6; and B, 71 kg/m'. Average consumption was 518.1 kg/m'. Maximum grout mass consumption in the highly karstified area varied from 500 to 700 t/m'.

The grout curtain at Salman Farsi dam site was constructed from galleries at five levels. The vertical distance between galleries varies from 18 to 36 m in order to intersect as many karst caverns by galleries as possible. Total length of grouting galleries is 3826 m. In the limestone without cavities, the grout mass consumption was less than 100 kg/m'. However, in a number of stages, the grout mass consumption was between 10 t/m' and 160 t/m'.

Large and empty caverns and systems of caves, especially in the dam site abutments, should be avoided by the modification of grout curtain alignment. The huge cavern can be isolated by bypassing the cave upstream and downstream sides. In the case of downstream alignment of grout curtain routes, the empty cavities will be filled by water.

Plugging of Concentrated Underground Flows

Plugging an active karst flow, deep in underground, is an extremely uncertain, dubious, and risky task. Quite often, the result is failure. Particularly complicated and difficult to plug are deep and siphonal flows. Practical solutions for this kind of problem are extremely complex and require lots of time and significant financial resources. Every problem is unique, and past solutions may almost never be repeated. Particular problems appear due to concentrated and very fast underground flows. When the flow is deeper, the water pressure is higher, and possibility for a successful grouting considerably decreases.

The main problem of this type of sealing is the determination of the karst channel position, particularly the location where it crosses the grout curtain. The successful sealing of underground conduit flows depends upon the accurate determination of:

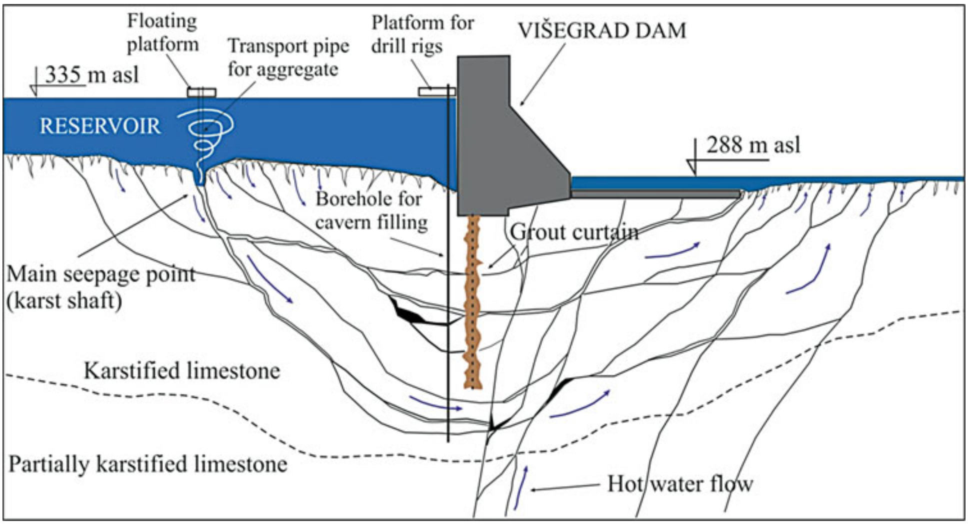
- Position of karst channel or concentrated zone flow
- Channel dimensions (cross-sectional area)
- Underground flow, characteristics in the channel (permanent or intermittent flow under pressure or with free surface, direction, and velocity)
- Possible presence of clay and silty cave deposits and their thickness

Treatment of underground flows requires different technologies to be applied: grouting by different types of cement mortar; grouting the aggregate which is first filled into the karst channel; filling of karst channel or cavern by rock blocks; polyurethane foam grouting; and grouting with molted asphalt, bitumen, cotton flocks, synthetic sponge, and different chemicals [17]. A few examples are presented to confirm the necessity of various methodologies.

During the first filling of the Višegrad Reservoir, Bosnia and Herzegovina (1986) underseepage below the suspended grout curtain was 1.4 m^3/s. The seepage was followed by massive washing of clay from caverns situated below the grout curtain. In spite of intensive re-grouting and having the local curtain deepened to110 m, the leakage had increased steadily from 6.5 m^3/s in 1996 to 9.4 m^3/s in 2003 and 13.92 m^3/s in 2008. Karst channels filled with clay at a depth of more than 130 m below the dam foundation were reactivated. A study of geological data from historical records, particularly geological mapping data of dam foundation surface, directed an intense site investigation program via boreholes (including TV logging), tracer testing, and monitoring of seepage outflow points downstream from the dam. Particularly successful was prospection of reservoir bottom by divers. During investigations in 2009 and 2010, a large karst opening (ponor – diameter 3.5 m) with massive water inflow was discovered by divers at the reservoir bottom, approximately 160 m upstream from the dam [23]. One important limitation was that the reservoir could not be drawn to a level which would minimize seepage and permit the treatment of deep seepage zone under lower pressure. In this case, the reservoir level created a maximum differential head between seepage zone(s) and downstream outlet points, meaning the largest pressure and fastest flows were through the deep karstified zone. To organize intensive treatment of the incipient karst features (ponor at bottom and large karst channels beneath the dam), two floating structures were constructed: (a) structure to plug ponor at the reservoir bottom and (b) float platform near the upstream dam face for deep boreholes (Ø 122.7) to fill caverns at depths of 100–150 m below the reservoir bottom (Fig. 6). Ponor at reservoir bottom was filled with 1286 m^3 of aggregate (1–4 mm). Crushed stone (aggregate) was introduced unto the empty cave-like space of the following sizes and proportions: 4–8 mm 20%; 8–16 mm 30%, and 16–32 mm 50%. After 38,000 m^3 of aggregate was inserted, the seepage was recorded as having been reduced to about 4.5 m^3/s.

In the case of Keban Dam (Turkey), the huge Petek Cave was filled with about 605,000 m^3 of limestone blocks, sand, and clay to reduce water losses from 26 to 8–9 m^3/s.

To plug the underground flow (4 m^3/s) at the Krupac ponor (Nikšić Polje, Montenegro) 750 m^3 of crushed stone and a considerable amount of sawdust and grouting mixture (70 percent sand and 30 percent grout mix) were used. Below the Vrtac Dam (same area), a cavern was detected in the area of large fault zone. Rod fall at depth of 85.5 m was only 0.5 m. Nine additional boreholes were drilled in the surrounding area, but exact contours were not identified. To block the underground flow, 1767 m^3 of aggregate (16–40 mm) and 1638 t of grout mix (dry component) were inserted through the boreholes. Temporary groundwater level behind the filled section increased to 30 m. But after a few months, groundwater decreased to the same level as it was before remediation.



Remediation in Karst, Fig. 6 Simplified cross section of main seepage flows beneath the dam body including general concept of remedial works [23]

During the plugging and grouting work at the Charmine Reservoir (France), a karst channel with a diameter of 80 cm was grouted. Material required for this plugging operation was 2000 m^3 of sand and 540 t of grout mixture. As a result, water loss from the reservoir was reduced from 800 to 10 l/s.

To block underground flows at El Cajon Dam site, 8000 wooden balls 7 cm in diameter, 25,000 polypropylene bags, and 650 mortar balls 6 cm in diameter were inserted into the cavities using 100 mm diameter boreholes. Accessible caverns filled with clay were cleaned up and plugged with concrete.

At the Lar Dam (Iran), the leakage problem is one of the most complicated problems related to plugging of active karst flow. During the first filling of reservoir, a very huge leakage (7.7–10.8 m^3/s) was measured. The cavern discovered at a depth of 210 m below the riverbed had a volume of 90,000 m^3 [5]. Through boreholes drilled from a gallery and having a diameter of 214 mm, the downstream section of the cavern was filled with about 28,000 m^3 of gravel, grading 5–50 mm, and 34,000 m^3 of crushed rock. The upper part of the cavern above the gravel cones up to the cave roof was filled with cement mortar totaling 13,000 m^3. However, in spite of the extensive grouting and filling of deeper channels with active flows, water losses are still very high. The flows are located at the depths of 250 to 430 m below the riverbed. About 42,000 m^3 of material (grout mix, sand, pebble, and aggregate) was not sufficient. Part of this material was transported by underground flow and deposited in front of large spring 12 km downstream.

In the case of Salman Farsi Dam (Iran), during the investigation drilling, no caverns were detected on either side. However, during grout curtain execution large-diameter karst channels and caverns were discovered. All accessible caverns were speleologically investigated. On the basis of results of speleological investigations, different plugging approaches were applied. The volume of the largest one was more than 150,000 m^3. This cave was avoided by modification of the grout curtain route. The caverns in the path were filled with clay and were washed and filled with self-compacting concrete (SCC), as were the large open caverns. For filling of openings and channels with apertures less than 20 cm, grout mix and grout mortar (up to 4 mm) were used.

Summarizing the results of the average grout mass consumption for 84 curtains from different karst regions of the world, the following classification can be established – Table 1.

Remediation in Karst, Table 1 Classification of grout mix consumption

Classification of grout mass consumption	Quantity of dry mass (kg/m’)	Distribution (%)
Very low	Up to 50	7
Low	50–100	24
Most frequent value (medium)	100–400	44
High	400–1000	20
Very high	>1000	5
Undefined high		

- Very low and low grout mass consumption

  These consumptions can be expected in poorly fractured and mostly thick bedded and slightly karstified limestone and dolomite. Very low consumption is rare. In 7% of analyzed examples, consumption was lower than 50 kg/m’ and in 20% of cases was between 50 and 100 kg/m’.
- Medium (most frequent value)

  The consumption rate in intensively fractured and karstified limestone, with open cracks and joints, enlarged by karstification process, including presence of cavities, can vary between 100 and 400 kg/m’. In some cases, water circulates through it. About 44% of the analyzed examples belong to that category. It means that the average and most frequent consumption of dry mass can be expected in the range of that category.
- The category of high and very high consumption belongs to extremely fractured limestone. The karstification process enlarges dense fractured networks. The active groundwater flows are the consequence of the karstification porosity and great transmissivity. About 25% of analyzed curtains have an average consumption in range of these two categories (high 400–1000 kg/m’ and very high >1000 kg/m’).

In the above-presented five categories, the consumption of grout mass used for filling and plugging the large karst conduits and caves is not included, i.e., the cases where grout takes more than 1000 kg/m’. This consumption is categorized as undefined high and cannot be expressed in kg/m’.

For sealing the very pervious fissured zones or small-diameter karst conduits, the following materials are commonly applied:

– Cement-based grout mass (slurry, grout mix, suspension) is commonly applied in recent grouting practice. Theoretical and practical aspects of grouting procedure using different types of cement grout mix are widely discussed in detail by various authors.
– Clay-cement stable grout mass is based on clay as basic component. Clay is cheaper than cement, reduces the permeability of the injected compound, and is resistant against aggressive water action. Mostly, the kaolinite and montmorillonite (bentonite) clays are used as components for grout mass. The flaky crystalline particles of clay are smaller than 3 μm (regular cement 100 μm and microfine cement 10 μm) and easily penetrable into tight joints. Clay-cement stable grout mass was used at many grout curtains in Dinaric karst region: Peruća Dam, 75% of clay; Buško Blato, 55%–75%; and Rama, 45%–57%. In the case of 7 km long grout curtain along the Slano Reservoir in Nikšićko Polje, the clay was a dominating component: 50%–65% of clay, 30%–40% of cement, and 1%–4% of Na_2CO_3. Another good example is the grout curtain below the Grančarevo Dam (Herzegovina) where the common grout mass was clay 66% and bentonite 33%.
– For plugging the large open cracks and small cavities, different types of cement mortar with additives and fillings such as sand, gravel, chemicals, etc., are routinely used.
– Polyurethane foams are a gaseous emulsion of uniform bubbles having very small diameter. These are characterized by a very high expansion coefficient. As compared to the initial volume, final volume can be more than 20 times larger after the foam is mixed.
– To plug the main leakage zone in the case of Dokan Dam, diesel oil and cotton flocks were used.
– Asphalt grouting technology has been used for leakage sealing at Hales Bar Dam site (1919) and Great Falls Reservoir, both in the USA.

– In modern grouting practice, asphalt is rarely used. Within the last few years, hot bitumen has been used in the USA for plugging concentrated underground flows [4]. The hot bitumen is pumped continuously down, through specially installed pipes, to the karst channel with the massive flow (injected at temperatures of 200 °C and higher).

For the large underground spaces treatment (channels and caverns), the most common technologies are:

– If the cavern is above the water table, the best solution is manual cleaning, washing, and filling by self-compacting concrete. Auxiliary adit or shaft from grouting gallery should be connected with cavern for easy communication with cavern.
– In the case of submerged caverns, large-diameter boreholes are required for filling the cavern with aggregate and grouting by common grout mix (prepacted concrete).
– When the size of cavern is enormous, a large shaft should be excavated for filling the cavern by large rock blocks, aggregate, sand, and other material (e.g., Keban Dam (Turkey)).
– For large caverns or wide tectonized and karstified zones, construction of cutoff (diaphragm) walls is a frequently applied technology.

The decision on which remedial method and material should be applied depends on many specific properties. Many of these features can be defined during constructions only, i.e., when direct observation in the underground is possible.

Remedial Works During Underground Excavations

Construction of underground structures in karst, particularly excavations of tunnels, is extremely risky and dangerous geotechnical work. During tunnel excavation, the crucial problems are the presence of cavern at tunnel route and groundwater intrusion. In spite of application of new sophisticated excavation and remediation technologies, construction time in tunneling is usually much longer than the projected time during design.

A number of examples of problems during excavation and tunnel operation are presented by Milanović [18], Marinos [16], and many other authors: Ghion and Dodoni tunnels, Greece; Montelungo, Italy; Steinbühltunnel, Germany; Sozina Tunnel, Montenegro; head race tunnels for power plants, Dubrovnik and Čapljina, Herzegovina; water transfer tunnel, Fatnica, Bileća, Herzegovina; Yellow River diversion tunnels, China; Karawanken Tunnel, Slovenia to Austria; SMART, road, and stormwater tunnel, Malaysia; Bordeaux Tunnel, France; tunnel in Ontario, Canada; tunnel in Ohio, USA; Lötschberg Base Tunnel, Switzerland; Gotthard Base Tunnel, Switzerland; Zakučac I and Zakučac II, Croatia; and Kuhrang III and Nowsoud diversion tunnels, Iran.

If a cavern happens to be encountered by the tunnel route or water inflow suddenly increases during the excavation, progress can be retarded because of many different difficulties such as:

- Excessive free space problem which must be overcome by filling, bridging, or bypassing.
- Cavity which is masked by cave deposits so that dimensions and geotechnical features are not known. This situation requires serious investigations and sometimes changing of excavation technology, thus making complicated remedial work indispensable.
- Source of groundwater discharges into the tunnel causing flows of hundreds, even thousands of liters per second.
- Undiscovered caverns very close to the tunnel or beneath the tunnel invert can provoke serious problems during the excavation and also during the tunnel operation.
- Concentrated external groundwater pressure frequently provokes large destruction in the tunnel lining.
- In the case of shallow overburden, collapse in the tunnel can provoke collapse at the surface.
- Because each tunnel has a role of regional drainage for the surrounding aquifer, the springs above the tunnel route and its close vicinity frequently dry up.

Consequently, all karst channels are subject to high pressure, destructive effects of turbulent inflows, and enormous amounts of water and sediments. The tunnel drainage is thereby involved in aggravating tunneling conditions. The destructive effect can be multiplied if the excavated tunnel tube connects active and nonactive (partially or fully filled) karst channels.

Any inflow in the tunnel of more than 100 l/s, with no possibility for dewatering by gravity, is an extremely serious and dangerous problem. In some cases, if the water pressure exceeds 10 bars, a small discharge (30–50 l/s) can appear as very complex tightening problem. To solve this problem, special and expensive treatment technologies have to be applied.

In spite of detailed investigations and determination of hydrogeological and engineering geological characteristics, detecting caverns in the front of tunnel head is questionable. Many different methods are applied; however, the results are often poor. Horizontal drilling ahead of the tunnel face is a common and most effective investigation method. Particularly sensitive is the usage of a tunnel-boring machine (TBM). To pass through a large cavern, either empty or filled with clay, application of the conventional excavation method in front of the TBM head is required. Even if the cavern is relatively small, remedial procedures can be complicated and time-consuming (Fig. 7). Almost all procedures need about 70 days.

Due to frequent floods of the Kuhrang III tunnel in Iran (23 km long), excavation time (proposed by design) was more than doubled. In a few cases, the huge water inflow in front of the TBM head increased abruptly up to a peak of 1.2 m^3/s in only 4 h. Reactivating the old drainage systems results in additional water and sediments inflowing into the tunnel. From one channel only for 24 h, more than 1000 $m^3$of boulders, gravel, sand, and 500 t of suspended sediments were transported from underground cavities into the tunnel. The hydraulic pressure of 26 bars was measured.

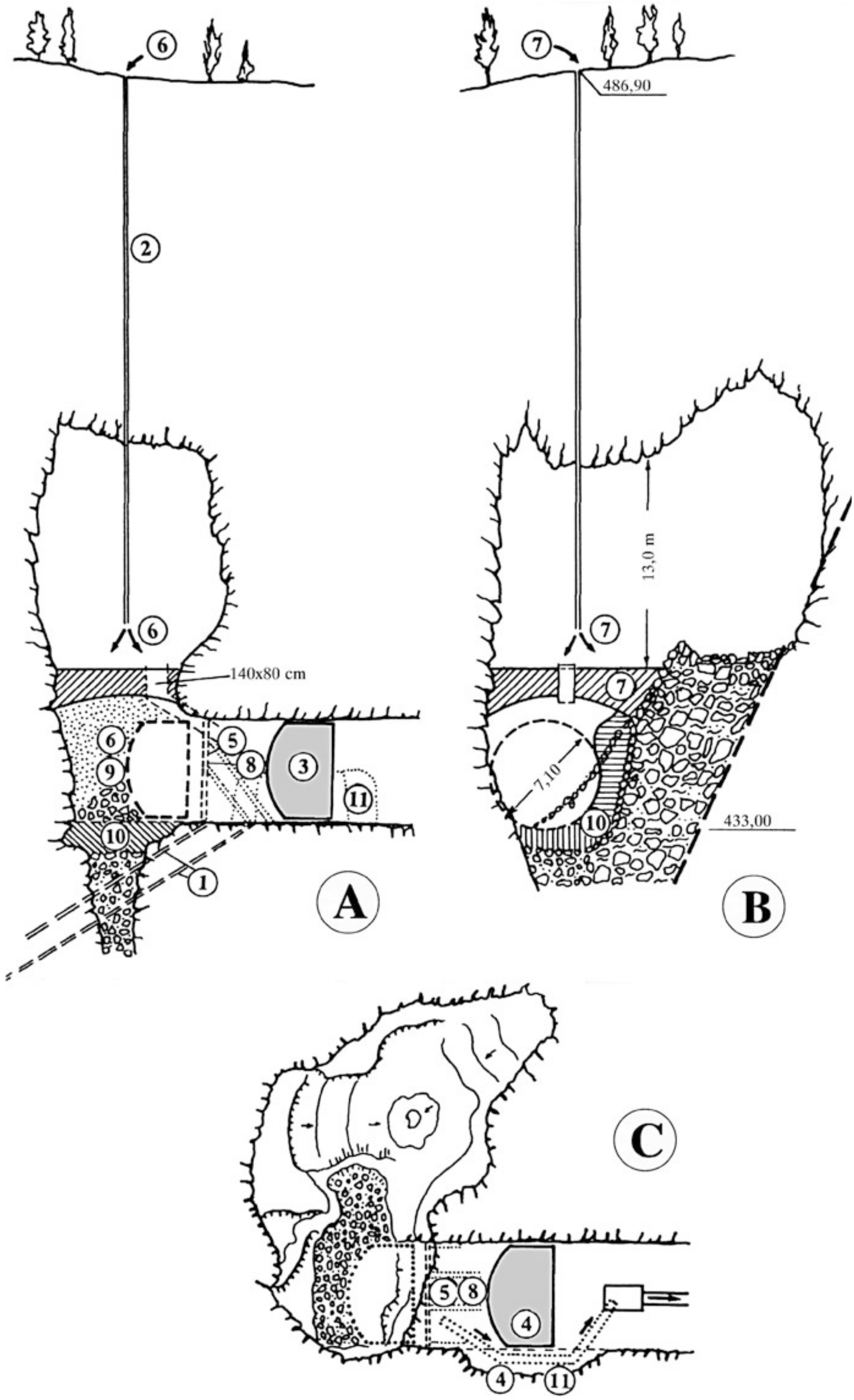
Due to huge water pressure, the construction of a concrete plug between the tunnel face and TBM head was the only solution. This procedure is very complex and risky for the people involved. Three key remediation phases (consisting of six steps) are presented in Fig. 8:

1. TBM head has to be shifted back 8–10 m and a bypass excavated around the head.
2. Insert the large-diameter pipes into the karst channel for water drainage.
3. Construction of the partition wall and filling the space between the wall and tunnel face by a concrete.
4. After construction of massive (4 m thick) concrete plug in front of tunnel face to stop the water inflow, drainage pipes were closed to block water flow in karst channel.
5. Next step is grouting the rock mass in front of the tunnel face. After injection of 446 t of dry cement, the total inflow decreased to Q ~ 270 *l*/s.
6. Tunnel driving continues through the concrete plug and grouted rock mass.

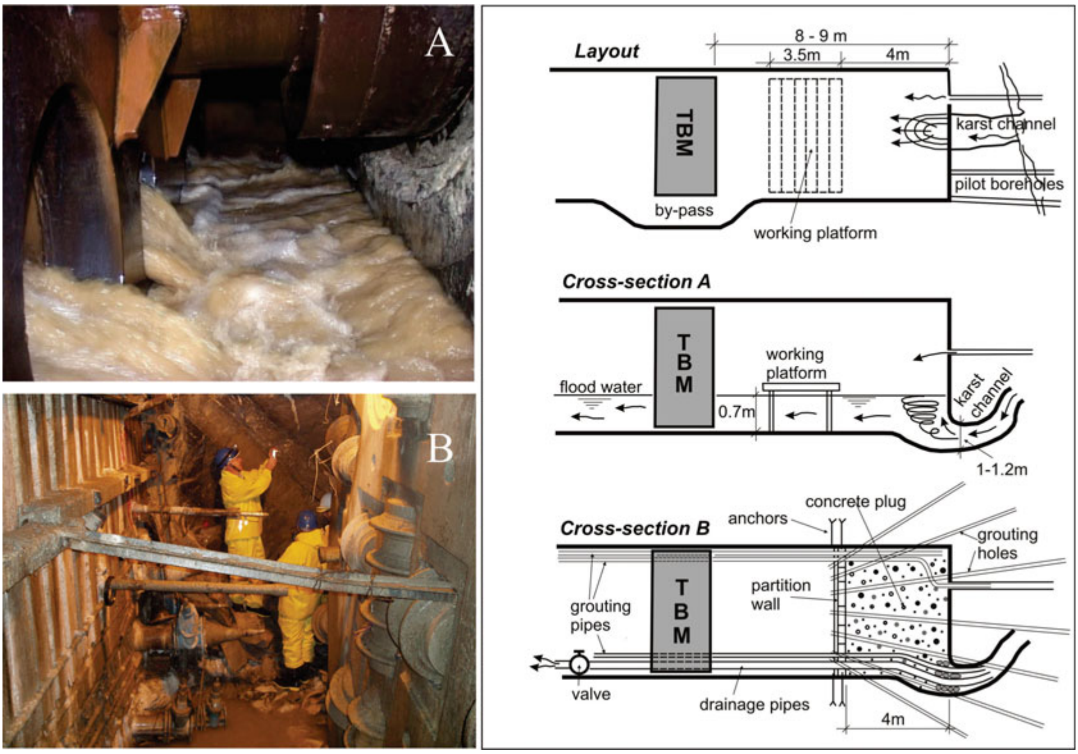
The large inflow of underground water has considerably retarded excavation in a number of tunnels in karstified rocks. During excavation of the headrace tunnel for Power Plant Dubrovnik (16.5 km long and 6 m diameter), excavation was stopped 25 times due to flood as consequence of heavy rains. In spite of the water table being deep below the tunnel level, the water inflow frequently exceeded 2.5 m^3/s. Intensive rain and fast vertical infiltration were the cause of such a massive water burst into the tunnel. Tunnel driving was suspended for 160 working days.

During excavation of upstream section of the conveyance tunnel from Fatničko Polje to the Bileća Reservoir (length 15.6 km, diameter 7.1 m), the tunnel was flooded 120 days every year.

Due to the possibility of an abrupt water burst during excavation, a monitoring system of precipitation and groundwater level fluctuation has to be established at the broad catchment area. Detailed geological mapping, geophysical investigations, drilling from surface along the tunnel route, and speleological investigations are basic investigations for the final selection of tunnel route and technology of excavation. Bypasses around cavernous zones or complete rerouting due to huge caverns are common solutions. Applying conical grouting ahead of the tunnel, step by step, in many

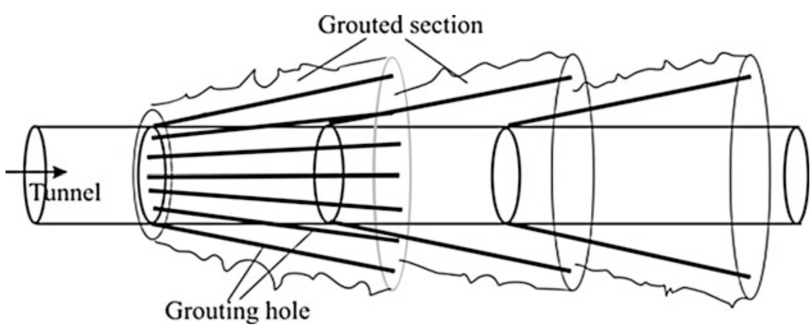


Remediation in Karst, Fig. 7 Overcoming partially filled cavern: (**a**) cross section along the tunnel route; (**b**) cross section perpendicular to the tunnel trace; (**c**) horizontal cross section



Remediation in Karst, Fig. 8 Plugging the water burst into the tunnel. (**a**) Water burst at tunnel face and (**b**) remediation works between partition wall and TBM head. Right graphic: presentation of three key remediation work stages

Remediation in Karst, Fig. 9 Tunnel driving by grouting the rock mass ahead of tunnel



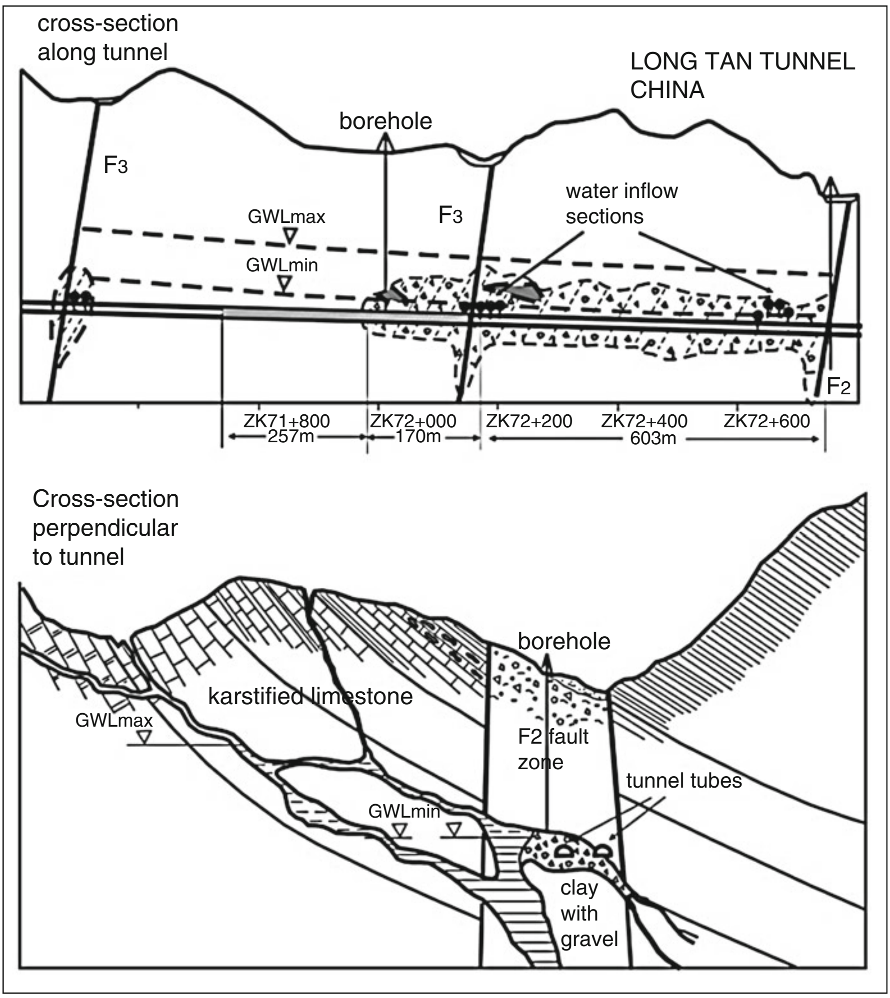
cases is the only possibility to cross through a heavily karstified zone saturated with water under pressure (Fig. 9).

A number of different treatment technologies were used to stop groundwater intrusion into the tunnel: drainage in front of the tunnel face; fan consolidation grouting; cone grouting ahead; heavy concrete plugs in the main tunnel to isolate flooded section; heavy reinforced and anchored concrete slabs to plug karst conduits with water under pressure; and grouting of individual karst channels. For material and energy transport, a large-diameter borehole has been drilled from the surface.

In Han Xingrui's book *Prediction and Engineering Treatment of Water Gushing and caves for Tunnelling in Karst* [10], the author presents 12 case studies of tunnels excavated in Chinese karst. Each of the examples presents problems with caverns (filled or empty) and huge groundwater inflow, as well as remedial works needed to pass through the weak and dangerous sections. One typical example is the Long Tan Tunnel (8693 m long) presented in Fig. 10. The key problem is the Tanchunguan underground river in the route of tunnel.

Drainage prior to tunnel excavation is the most frequent and most effective method. Horizontal boreholes from the tunnel face (pilot holes 25–40 m long) have a drainage role and an investigation function to detect caverns and cavern zones. In some cases, the groundwater inflow is too high, and drainage holes are not effective enough. In that case, a drainage gallery beneath the tunnel route is required. After 6 years of operation of the 6170 m long Sozina railway tunnel (Montenegro), it was flooded by an abrupt groundwater intrusion through the karst channel at the rate of 6.5 m^3/s. A 1750 m long drainage tunnel, 2 m lower and 15 m from the main tunnel, was excavated to resolve the problem.

Defects during tunnel operation are also very common in karst. Particularly vulnerable are derivation tunnels. These tunnels are exposed to high pressure from inside and from outside. Even if the tunnel is situated in an aeration zone, any high precipitation can provoke fast groundwater levels rising and unpredictable concentrated outside water pressure, or water pressurized air, on the lining. Demolishing of the tunnel lining frequently occurs (Fig. 11). An artificial bypass



Remediation in Karst, Fig. 10 The Long Tan Tunnel, China

Remediation in Karst, Fig. 11 Tunnel lining damaged due to external concentrated water pressure (right) and remedial solution (bypass) to drain water around tunnel lining

connection for water around the tunnel lining is required [19]. This connection prevents the tunnel lining destruction due to concentrated water pressure from above during periods of heavy rain. The lining is also protected against air or water pressure from below in the case of rapidly rising groundwater levels.

Collapses (Subsidences)

At numerous areas in karst regions of Dinarides, Helenides, and Taurides, the century-old problem for local inhabitants from Roman until the present time is the natural phenomena known as collapses (subsidences, alluvial ponors). In order to prevent erosion of agricultural soil, great attempts have been made to plug subsidence. The following natural materials have been used: stone, wood, and clay (Fig. 12a). In some cases, collapses are protected by simple stone walls (Fig. 12). All these structures were efficient for a short time. After one or few floods, they were damaged and needed to be rebuilt.

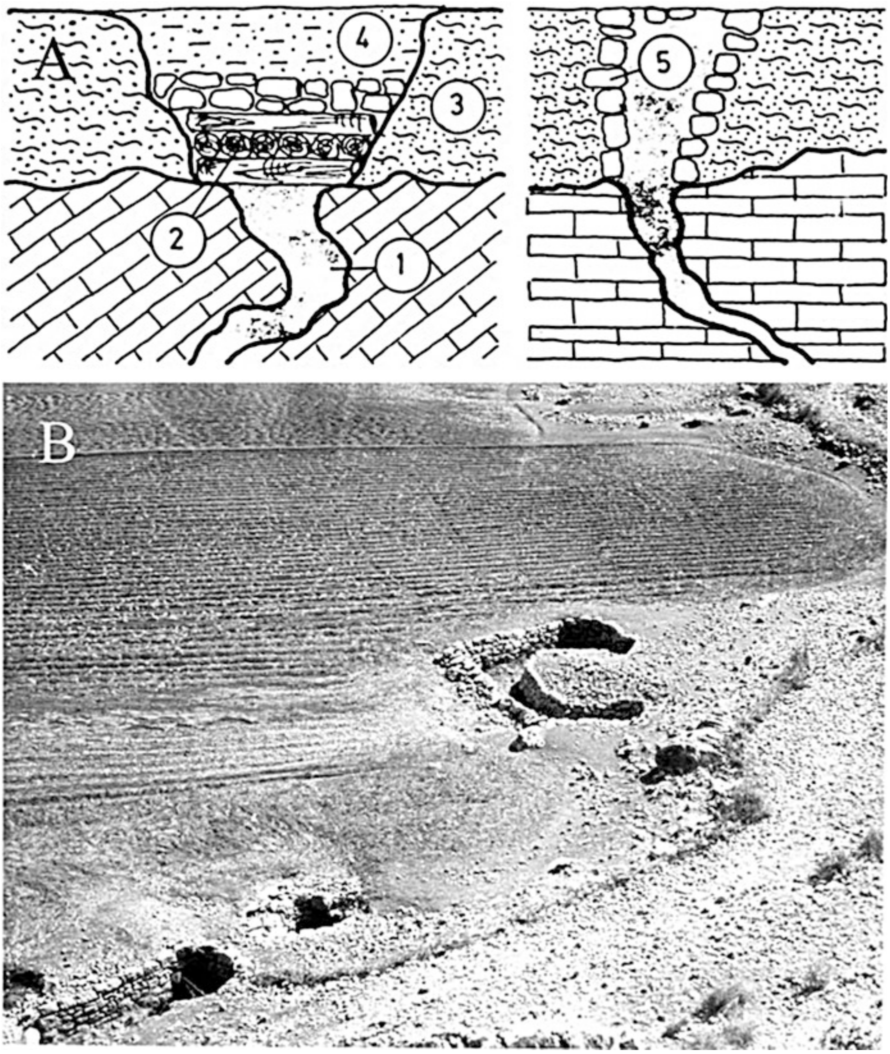
Genesis and mechanism of karst collapses and measures for prevention and remediation have been analyzed by many authors. Some of them are LaMoreaux and Newton [15], Beck [2], Waltham [28], White and White [29], Milanović (1981) [20, 21], Klimchouck and Andrejchuk [14], Ford and Williams [7], Gutierrez et al. [8], and Lu (2009).

Overburden thickness in subsidence-prone areas varies from a few meters to greater than 70 m. The key process of collapse creation is suffosion. Suffosion acts from the surface as flood water or from below as underground water. After an initial channel is formed as consequence of flood water suffosion, the erosion processes occur. Large volumes of eroded material are transported through the karst channels. Subsidence in the form of funnel-like shafts origins at surface. In the places where the karst channel is covered with unconsolidated sediments, the strong upward flow can provoke fluidization and piping process. Final result is collapse of the upper alluvial layers, i.e., the subsidence origin.

Subsidence in reservoir bottom and banks occurs under the influence of water (groundwater, flood water, and pore water) as erosion and piping action breaks down the support of poorly consolidated sediments. In some cases, the water under pressure (water hummer effect) or water pressurizing air in the aeration zone (air hammer effect) has triggered blowouts through the overlaying sediments.

Induced collapses (subsidence) have developed as consequence of urbanization, mining, man-made reservoir construction, and enormous groundwater extraction. Such incidents dramatically increased in the second part of the twentieth century. According to Ford and Williams [7] in recent decades, hundreds of thousands of new induced subsidence cases have been reported worldwide.

In the area of Dzerzhinsk, N. Novgorod (Russia), more than 5000 induced collapses occurred in the past 50 years. Some of them were catastrophic, causing damages to buildings, factories, and other constructions [26]. The area of Kungur town registered 415 subsidences created as consequence of urbanization. More than 80% of these subsidences occurred during the period 1971–1975 [6]. From



Remediation in Karst, Fig. 12 Ancient technology of protection of collapses: (**a**) 1. Karst channel; 2. protective structure, wood, and stone; 3. alluvial deposits; 4. final cover by soil; 5. stone wall. (**b**) Protection of collapses by stone walls

1947 until recently, a number of conferences that focused on the problem of subsidences were organized in the former Soviet Union and Russia.

According to Hua [11], collapses as consequence of groundwater pumping were reported from 25 areas in China (1974–1986). In one extreme case, 600 collapse events were recorded in an area of 5 km^2. Yuan [30] presents a list of collapses in 11 regions in southern China with 14,932 registered collapses. The majority of collapses occurred due to dewatering of coal mines – 12,750. Analysis of these collapses shows two main mechanisms of creation: soil suffosion and vacuum suction erosion. However, some of collapses were created due to water and gas explosions in underground karst conditions, soil liquefaction caused by vibration, soil erosion along decayed plant roots, etc.

In the USA alone, more than 25,000 km^2 have been directly affected by induced subsidence. According to LaMoreaux and Newton [15], thousands have been formed in the USA since 1950.

Due to the abrupt appearance sinkholes in Florida, the Florida Sinkhole Research Institute was founded in 1981. In this period, 14 conferences on sinkholes were organized in different karst regions of the USA. For over 30 years, the Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst is one of the premier conference series for all engineering aspects of karst and is particularly focused on subsidence problems.

According to Brink [3] after 25 years pumping in a gold mine district in South Africa, 38 people have lost their lives in collapse.

The solution process in carbonate rock produces a variety of different cavities and conduits that are covered by different poorly consolidated sediments. Because of that, there are many variations and combinations of different cases. Possibility of collapse classification according to geotechnical properties and expected levels of hazard can be approximated. Different authors suggest remedial approaches based on experience. According to Sowers [24], a general approach to overcoming hazards in the case of karstified rocks covered with unconsolidated sediments includes “avoiding areas of concentrated hazards, correcting the hazards by filling them or collapsing them, bridging over small hazards, reinforcing the rock, bypassing shallow hazards with deeper foundations, and minimizing activation of the hazard-forming processes.” Sowers presented 23 examples of technology applied to correct or minimize or mitigate defects during the foundation of different structures.

Man-made reservoirs change the regime of underground and surface water, provoking many different processes. Some of them are destructive, resulting in collapses. Induced collapses resulting from extensive water level fluctuation in man-made reservoirs are a common failure. Process of piping together with air pressure (air hammer) and water uplift (water hammer) triggers collapses, provoking huge leakage from reservoirs through the reactivated karst channels.

Subsidence development is a very common process which endangers the safety and integrity of reservoirs. Subsidence induced by filling or extensive water level fluctuation in man-made reservoirs has resulted in much leakage from certain reservoirs (Lar, Iran 10.8 m^3/s; Hutovo Reservoir, Herzegovina 3 m^3/s; Vrtac, Montenegro, 25 m^3/s; Keban, Turkey, 26 m^3/s; Mavrovo Reservoir, FYUR Macedonia, 7 m^3/s; Perdika Reservoir, Greece; Samanalawewa, Sri Lanka; Hammam Grouz, Algeria; Mosul, Iraq; and many others). Figure 13 presents some examples of subsidences that are formed during different reservoirs operation.

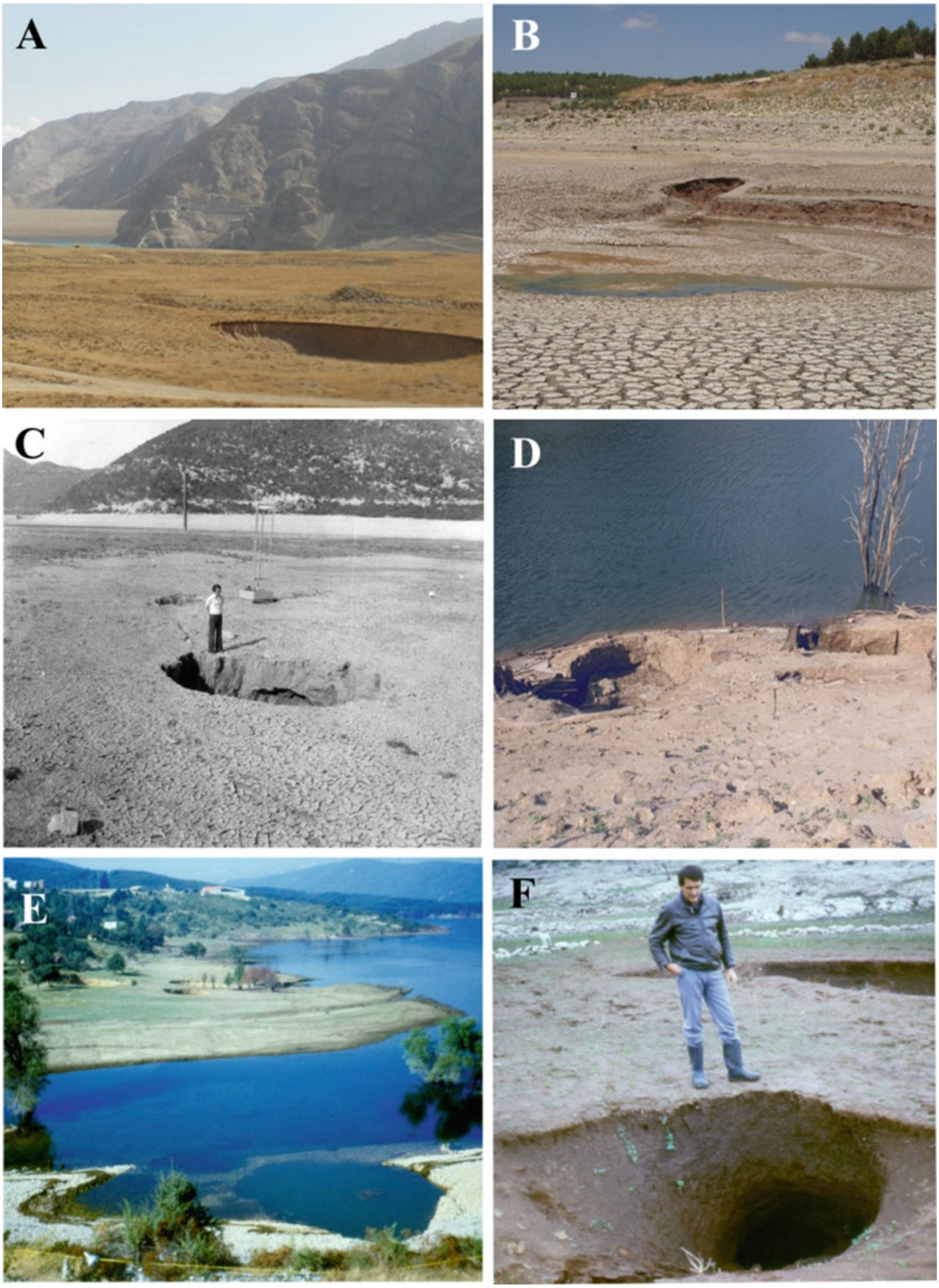
Collapses at the reservoir bottom are a source of leakage [22]. In some cases, the reservoir cannot operate at full capacity. Different waterproofing approaches and technologies should be applied:

- Dental treatment, i.e., construction of inverted filter, including concrete cap at the top
- Clay blankets
- Application of different geomembranes
- Grouting the karstified rock mass below the sinkhole
- Digging out unconsolidated sediments and filling karst conduit openings
- Construction of reinforced concrete slab
- Construction of a concrete plug in the conduit connected to the collapse

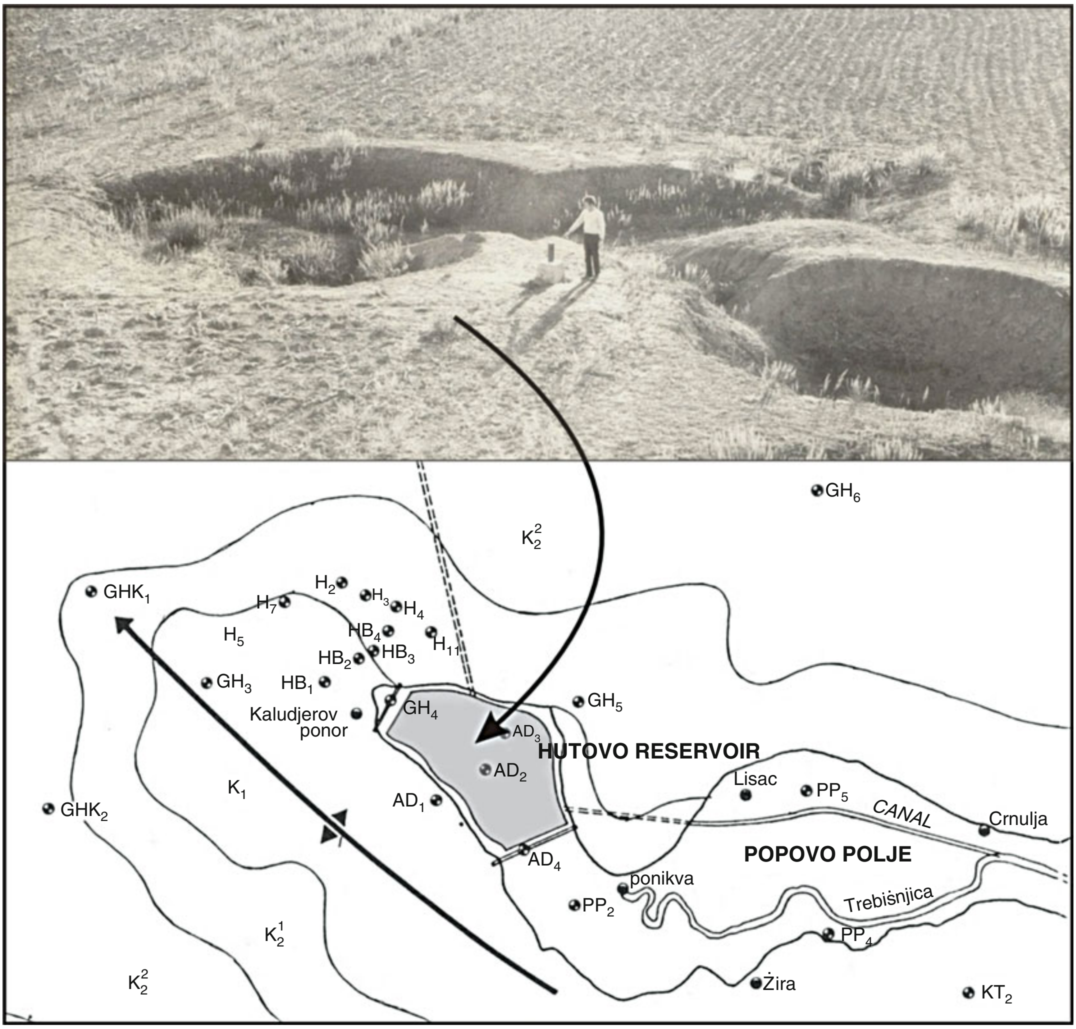
In general, remediation approaches are same or similar to those presented in the section on “Surface Remediation Measures for Dams and Reservoirs.”

Reservoir Hutovo, located at the lowest part of large karst depression Popovo Polje (Herzegovina), is a good example of induced subsidence occurrence. The reservoir bottom is covered with alluvial deposits. Their thickness increases from the flanks toward the middle part of the polje, where thickness is about 30 m. Topography of bedrock (Cretaceous limestone) is typical for karst. The area of the reservoir was losing water under natural conditions through 75 registered ponors in the shape of different size subsidences [20]. The largest one was formed above the fossilized ponor of river Trebišnjica (Fig. 14).

Very complex remedial works were applied to prevent seepage through the deep collapse at the bottom of Hutovo Reservoir: (1) filling collapse space with clay/cement grout mixture, (2) grouting



Remediation in Karst, Fig. 13 Subsidences created during operation of reservoirs: (**a**) Lar, Iran; (**b**) Hammam Grouz, Algeria; (**c**) Hutovo, Herzegovina; (**d**) Samanalawewa, Sri Lanka; (**e**) Mavrovo, FYUR Macedonia; and (**f**) Bileća, Herzegovina

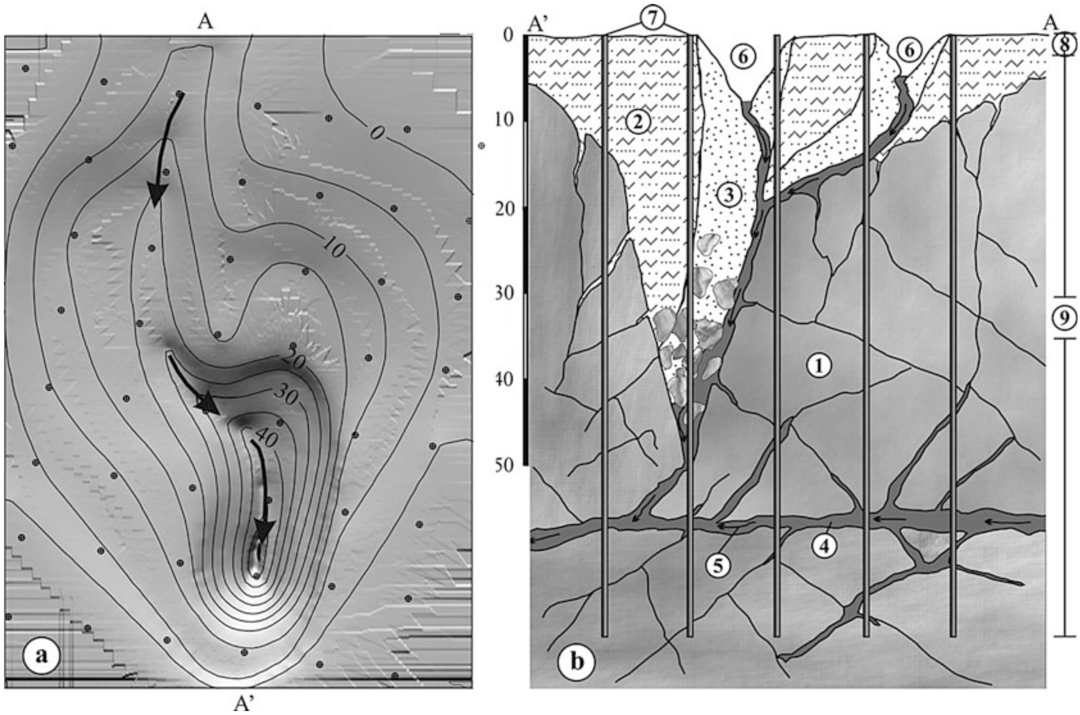


Remediation in Karst, Fig. 14 Natural collapse at future reservoir bottom

the rock mass at the area of collapse, (3) covering the surface above the collapse with compacted clay, and finally, (4) covering the entire area by geomembrane (Fig. 15).

Human activities can also play a special role in inducing or enhancing karst processes in evaporite rocks, and the results can be catastrophic [12]. Owing to the extremely high and rapid dissolution process, the formation of initial cavity (collapse) deep below the surface, and its migration upward to the land surface, occurs much faster than the same process in carbonate rocks. According to Johnson, gypsum-karst problems are caused by the following human activities [12, 15]:

- Building structures that induce differential compaction of soils above an irregular gypsum-bedrock surface
- Building structures directly upon gypsum-collapse features
- Impounding water above, or directing water into, a gypsum unit where soil piping can divert water (and soil) into underground gypsum cavities



Remediation in Karst, Fig. 15 Remediation of deep sinkhole at the reservoir bottom. (**a**) Topography of limestone bedrock with rate of grouting holes. (**b**) 1. Karstified limestone; 2. alluvial deposits; 3. sandy deposits with boulders; 4. karst channel; 5. direction of underground flow; 6. collapse before remedial works; 7. grouting boreholes; 8. surface zone compacted and covered with geomembrane; 9. depth of grouted rock mass

Induced collapses in salt deposits mostly are associated with solution mining and petroleum industries. The size of collapse varies between 10 and 100 m in diameter and is 10–600 m in depth.

Huge subsidences, known as *obruks*, are very frequent at broad Konya Plain (Turkey). Some of the very recent collapses have depths of more than 70 m and occurred instantaneously (Fig. 16).

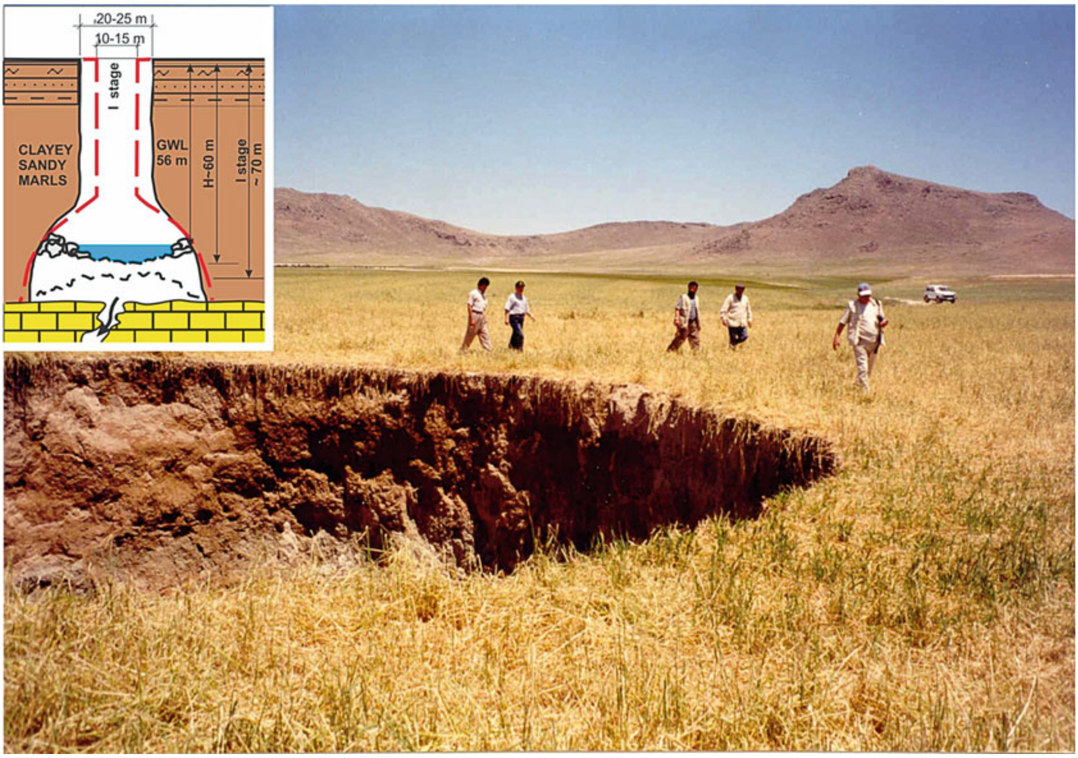
Further Directions

In spite of very serious and complex investigation programs, including all available investigation methods, remediation is practically unavoidable for any kind of human activities in karstified rocks. The necessity for remediation cannot be totally eliminated by increasing the investigation programs. It can be minimized at an acceptable level but never absolutely eliminated.

Difficulties and failures, generally, can be classified as technical and ecological, sometimes archaeological [13]. The technical difficulties and failures are connected with various man-made structures, particularly in the case of dams, reservoirs, and tunnels.

The most frequent technical difficulties are land subsidence at land surface and at the bottom of the reservoirs, water leakage at dam sites and from reservoirs, subsidence around the tunnel tubes, break-in of water and mud during the underground excavation, induced seismicity as a consequence of artificial storage, floods as a consequence of surface water regime changing, global groundwater balance change, and unproductive deep-well drilling.

The reservoir in karst may fail to fill despite an extensive sealing treatment. In spite of a serious and exhaustive investigation program and large-scale sealing treatment, a number of examples confirm this. In many examples, the designed



Remediation in Karst, Fig. 16 Large recent collapse at Konya Plain, Turkey

and executed watertightness treatment was only partially successful. In such cases, tremendous additional remedial works have been carried out to stop or minimize leakage. It requires a lot of patience and perseverance, and adequate funds. The practical solutions of this kind are extremely complex and require close collaboration between experienced geologists and civil engineers.

As a result of persistent, long-term, and step-by-step sealing treatment, in some cases, the results justify invested money. But in some cases, owing to the high cost or lack of sufficient technology, remediation of damage is not feasible.

Dam sites and reservoirs located in the karstic river valleys need more care than in the case of canyon sites. The most unfavorable conditions, from a watertightness point of view, occur in valleys with a hanging riverbed situated at high elevation. Deep and expensive grouting treatment and filling of empty caverns by crushed material and prepack should be expected. Generally, hanging river valleys should be avoided for surface water storage. But in the case of the normal valley (valley is erosional level, i.e., discharge zone for adjacent karst aquifer), the chances for safe water storage are realistic. Possible leakage through the dam sites or reservoir banks can be limited or eliminated in a technically and economically acceptable way.

In the case of the dam and reservoir construction in karst, the leakage risk is dominant. Dam stability risk is much less exposed. Risk related to the underground structures particularly during tunnel driving is permanently present due to abrupt floods and huge caverns at tunnel route.

Induced sinkholes are a common and serious natural problem, which increases with urbanization and industry development. They are separated into those caused by frequent and long-term water table decline and those caused by different construction. Induced sinkholes are very common in the reservoirs and in the areas of extensive groundwater extraction. Sometimes the magnitude of subsidence is catastrophic. The sinkholes are spatially independent random events; in such cases, the possibility for prediction is limited.

In a karstic environment with a highly random distribution of karstic features, always some uncertainties remain. Every problem with construction in karst has its own individual circumstances, and no case or situation is ever repeated. Modification and adaptation to real-life situations in the underground are almost the rule, not an exception. Because of that, the necessity for remediation in karst should be expected during all times of structure operation. Geological and hydrogeological diagnoses must be accomplished with the support of new investigation technologies. During the entire functional life of any structure built in karst, operators should be prepared with accessible equipment, personal and materials necessary to minimize local damages, and the possibility of remediation work.

Bibliography

1. Altug S, Saticiogly Z (2001) Berke arch dam, Turkey: Hydrogeology, karstification and treatment of limestone foundation. In: Proceedings of the 6th international symposium and field seminar, Marmaris, Turkey, IHP-V Technical documents in Hydrology, No. 49. Vol. I UNESCO, Paris, p 315
2. Beck BF (1991) On calculating the risk of sinkhole collapse. In: Kastning EH, Kastning KM (eds) Appalacian karst. Proceedings of the Appalachian Karst Symposium. National Speleiological Society, Radford, pp 231–236
3. Brink ABA (1984) A brief review of the South African sinkhole problem. In: Beck F (ed) Sinkholes: their geology, engineering and environmental impact. A.A. Balkema, Rotterdam/Boston, pp 123–127
4. Bruce DA (2003) Sealing of massive water inflows though karst by grouting: principles and practice. In: Back BF (ed) Sinkholes and the engineering and environmental impacts of Karst. Geotechnical Special Publication No. 122. American Society of Civil Engineers, Reston
5. Djalaly H (1988) Remedial and watertightening works of Lar Dam, Iran. Paper presented at Seizieme Congres das Grandes Barages, San Francisco
6. Dublansky GN, Shilova EV, Kadebski YV (2004) Imperilment of the Kungur town by subsidence. In: Russian. Karstology – XXI century: theoretical and practical significance. Proceedings of the international symposium, Perm
7. Ford D, Williams P (2007) Karst hydrogeology and geomorphology. Wiley, Boston
8. Gutierrez F, Guerrero J, Lucha P (2008) Quantitative sinkhole hazard assessment. A case study from the Erbo Valley evaporate alluvial karst (NE Spain). Nat Hazards 45:221–233
9. Günay G, Milanović P (2005) Karst engineering studies at the Akkopru Reservoir area, SW of Turkey. In: Stevanović Z, Milanović PT (eds) Water resources and environmental problems in Karst. Proceedings of the international conference, Belgrade/Kotor, Sept 2005, pp 651–658
10. Han X (2010) Prediction and engineering treatment of water gushing and caves for tunnelling in Karst, BBT Press, Guangxi
11. Hua SL (1984) Pumping subsidence of surface in some karst areas of China. In: Symposium on human influence on Karst, Postojna
12. Johnson KS (1997) Evaporate Karst in the United States. Carbonates Evaporates 12(1)
13. Kiernian K (1988) Human impact and management responses in the karsts of Tasmania. In: Proceedings of the international geographical union, Study Group Man's Impact on Karst, Sydney
14. Klimchouck A, Andrejchuk V (2005) Collapse and breakdown mechanisms from observations in the gypsum caves of Western Ukraine: implications for subsidence hazard zonation. Environ Geol 48:370–383
15. LaMoreaux PE, Newton JG (1986) Catastrophic subsidences: an environmental hazard, Shelby County, Alabama. Environ Geol Water Sci 8(1/2)
16. Marinos PG (2005) Experiences in Tunnelling through Karstic rocks. In Stevanović Z, Milanović PT (eds) Water resources and environmental problems in Karst. Proceedings of the international conference, Belgrade/Kotor, Sept 2005, pp 651 – 658
17. Maximovich NG (2006) Safety of dams on soluble rock (The Kama hydroelectric power station as an example). The Russian Federal Agency for Science and Innovations, Perm
18. Milanović P (1997) Tunneling in Karst: common engineering-geology problems. In: Marinos PG, Koukis GC, Tsiambaos GC & Stournaras GC (eds) Engineering geology and the environment. Balkema, Roterdam
19. Milanović P (2000) Geological engineering in karst – dams, reservoirs, grouting, groundwater protection, water tapping, tunneling. Zebra Publishing, Belgrade
20. Milanović P (2002) Subsidence hazards as a consequence of dam, reservoir and tunnel construction. Int J Speleol 31(1/4):169–180
21. Milanović P (2002) The environmental impacts of human activities and engineering constructions in karst regions. Epizodes 25(1):13–21
22. Milanović P (2011) Groundwater, impacts of infrastructure construction. In: Proceedings of the 9th conference on limestone hydrogeology, Besancon
23. Milanović S (2015) Prevent leakage and mixture of Karst groundwater. In: Stevanović, Z (ed) Karst aquifers – characterisation and engineering. Springer Heidelberg

24. Sowers GF (1984) Correction and protection in limestone terrane. In: Back B (ed) Sinkhole: their geology, engineering & environmental impact. Balkema, Rotterdam, pp 373–378
25. Stevanović Z, Milanović P (2015) Engineering challenges in Karst. In: Gabrovšek F, Ravbar N (ed) Acta Carsologica. 44/3, Academia Sciantiarum et artum Slovenica, Ljubljana
26. Tolmachev V, Leonenko M (2011) Experience in collapse risk assessment of building on covered Karst landscapes in Russia. In: Van Beynen PE (ed) Karst management. Springer, Dordrecht
27. Vlahović V (1981) Karst Reservoir Slano. The Montenegrian Academy of Sciences and Arts, Special Edition, Podgorica
28. Waltham AC (1996) Ground subsidence over underground cavities. J Geol Soc China 39(4):605–626
29. White EL, White WB (1969) Processes of cavern breakdown. Natl Spelol Soc Bull 31:83–96
30. Yuan D (1991) Karst of China. Geological Publishing House, Beijing

Book and Reviews

Andreo B, Carrasco F, Duran JJ, LaMoreaux JW (2010) Advances in research in Karst media. Springer, Berlin

Back BF (ed) (2003) Sinkholes and the engineering and environmental impacts of Karst. Edited by Geotechnical Special Publication No. 122. American Society of Civil Engineers

Bonacci O (1987) Karst Hydrology, With Special References to the Dinaric Karst. Springer, Berlin/Heidelberg

Engineering geological problems of construction on soluble rocks (1981) International engineering geological symposium, Istanbul

Ford D, Williams P (2007) Karst hydrogeology and geomorphology. Wiley, Chichester

Gabrovšek F, Ravbar N (eds) (2015) Acta Karsologica 44-3. Academia Scientorum et Artum Slovenica, Ljubljana

Gutierrez F, Desir G, Gutierrez M (2003) Causes of the catastrophic failure of an earth dam built on gypsiferous alluvium and dispersive clays (Altorricon, Huesca Province, NE Spain). Environ Geol 43:842–851

Gutierrez F, Orti F, Gutierrez M, Perez-Gonzalez A, Benito G, Prieto JG, Valsero JJD (2001) The stratigraphical record and activity of evaporate dissolution subsidence in Spain. In: Gutierrez F, Orti F, Gutierrez M, Perez-Gonzales A, Benito G, Prieto JG, Valasero JJD (eds) Carbonates and Evaporites 16(1). Springer, Heidelberg, pp 46–70

Guzina B, Sarić M, Petrović N (1991) Seepage and dissolution at foundations of a dam during the first impounding of the reservoir. In: Congres des Grandes Barrages, Q66 Vienne

Günay G, Johnson IA, Back W (eds) (1990) Hydrogeological processes in Karst Terranes. IAHS Press, Wallingford

Günay G, Johnson IA (eds) (1997) Karst water environmental impact. A.A. Balkema

Karst in Carbonate Rocks (1972) Publishing house of the Moscow University. In Russian, Moskwa

Karstology – XXI Century (2004) In: Proceedings of the international symposium. Perm State University, Perm

Krešić N, Stevanović, Z (eds) (2010) Groundwater hydrogeology of springs, engineering, theory, management, and sustainability. Elsevier, Burlington

LaMoreaux PL (1984) Catastrophic subsidence, Shelby County, Alabama. In Back BF (ed) Sinkholes; Their geology, engineering and environmental impact; the first multidisciplinary conference on Sinkholes, Orlando, 15–17 October. A. A. Balkema, Netherlands pp 131–136

Lu Y (2009) Karst in China, a World of improbable peaks and wonderful caves. Ministry of Land and Resources, Peking

Milanović P (1981) Karst hydrogeology. Water Resources Publications, Littleton

Milanović S, Vasić LJ (2015) Monitoring of Karst groundwater. In: Stevanović Z (ed) Karst aquifers – characterisation and engineering. Springer International Publishing Switzerland

Newton JG(1976) Early detection and correction of sinkhole problems in Alabama, with a preliminary evaluation of remote sensing application. Alabama highway research, HPR no. 76, research project 930-070

Stevanović Z (ed) (2015) Karst aquifers – characterisation and engineering. Springer

Stevanović Z, Milanović P (2005) Water resources & environmental problems in Karst. In: Proceedings of the international conference and field seminars. Faculty of Mining and Geology, Belgrade/Kotor

Van Beynen PE (ed) (2011) Karst management. Springer, Dordrecht

Part V

Construction and the Environment

Construction Planning: Environmental Impact of Foundation Studies and Earthquake Issues

Asadullah Kazi[1] and Bashir Memon[2]
[1]Isra University, Hyderabad, Pakistan
[2]P.E. LaMoreaux & Associates, Tuscaloosa, AL, USA

Article Outline

Glossary

Aftershocks Smaller earthquakes following the largest one (the main shock).

Amplification Increase in the amplitude of earthquake wave depending on the nature of ground.

Amplitude The maximum height of a wave crest or depth of a trough.

Attenuation Decrease in the amplitude of earthquake waves.

Earthquake The vibration of the Earth caused by the passage of seismic waves.

Epicenter The point on ground surface directly above the focus (hypocenter) of an earthquake.

Focus The place at which an earthquake is initiated.

Frequency Number of vibrations per unit time.

Fundamental period It is the period of an oscillating structure when subjected to shaking.

Ground motion Movements of ground caused by the vibrations of earthquake.

Intensity A measure of ground shaking obtained from damage to man-made structures (Modified Mercalli scale).

Isoseismal map Contour lines drawn to separate one level of seismic intensity from another for a given earthquake or maximum intensity expected of earthquakes in a given area.

Magnitude A measure of energy released by an earthquake (Richter scale).

Main shock The main earthquake event, which may be followed by aftershocks.

Mitigation To lessen or alleviate the damage likely to be caused by an earthquake.

Period The time interval between successive crests or troughs of an earthquake wave.

Resonance Increase in the amplitude of an earthquake caused by an overlap between the frequency of ground and that of the structure resting on it.

Retrofit To bring back to stable condition after having been damaged by an earthquake.

Return period Recurring period of an earthquake of a given magnitude or intensity.

Seismic wave An elastic wave generated by natural or man-made causes.

Seismic zoning A map based on the distribution of earthquakes, classified according to expected intensity translated into acceleration as the percentage of gravity.

Seismicity The occurrence of earthquakes in space and time.

Seismology The study of earthquakes.

Tectonic plates Major regions of the Earth separated by geological faults along which these regions (plates) move relative to each other.

Definition of the Subject

Earthquakes are simply the shaking of the ground initiated by natural or man-made causes. The waves generated by earthquakes are called seismic waves, and they are broadly classified into

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_442

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_442-3

body waves and surface waves. A point from which the earthquake waves first emanate is called the earthquake focus. Moreover, the point directly above the focus is called the epicenter of an earthquake. Natural earthquakes can be caused by tectonic, volcanic (related to movement of tectonic plates), or collapse-related behavior of rock/soil. Man-made shaking, on the other hand, can be initiated by a variety of activities such as blasting, underground nuclear explosions, movement of heavy vehicles, and so forth.

Natural earthquakes are known to have occurred throughout history [1, 2]. They have left their traces as manifestations of geological processes at work on the geological time scale. These traces are generally evident in places where people lived, and no such records are available for places which were uninhabited. Most of these earthquakes are located along tectonic plate boundaries, while some are located in the intraplate regions of the world [3], as reflected in global hazard maps published by the United States Geological Survey.

Many lessons were learned from the damage caused by the Great California earthquake of 1857 and subsequent earthquakes associated with movements along the San Andreas Fault [4]. The Great Alaskan earthquake of 1964 [5] and a history of earthquakes in Lima, Peru, and elsewhere have helped in developing guidelines for the design of earthquake-resistant small- and higher-rise buildings, under different ground conditions [6].

The Engineering Research Institute in California has recently prepared a report dealing with contributions of earthquake engineering to protecting communities and critical infrastructure from multi-hazards. Both the 2009 International Building Code in general and the Seismic Design Maps and Tools for engineering in particular give guidance to these questions.

Introduction

Earthquakes are one of the most destructive environmental hazards. The power of an earthquake is expressed according to its intensity or magnitude. The effect of both the parameters generally decreases as the distance of the affected place increases from the epicenter of the earthquake.

Earthquakes occur at different depths below the ground surface. The ones that take place between 0 and 70 km below the surface are often classified as shallow. However, if they occur between 70 and 300 km, they are designated as intermediate focus, while the deep earthquakes occur between 300 and 800 km.

Intensity and Magnitude

The intensity is a noninstrumental quantity assigned according to observed geological effects, damage to structures, and the perception of shaking by people and animals at a given place. It is not a fixed quantity but varies from place to place. There are many scales of intensity, but the most common, as illustrated in Table 1, is the modified Mercalli intensity scale [7]. The scale ranges from the lowest of 1 (not felt except by a few) to a maximum of 12 (total damage, where objects are thrown into the air).

The magnitude of an earthquake, on the other hand, is a fixed instrumental quantity and is a measure of the energy released by an earthquake [8]. They are generally classified according to their magnitude, as illustrated in Table 2. For example, the magnitude of 6 is approximately equal to the energy released by the atomic bomb used at Hiroshima during the Second World War and is classified as a large earthquake. Earthquakes stronger than magnitude 8 are classified as great earthquakes.

Ground Motion Considerations

The outcome of earthquakes on buildings depends generally more on the severity of ground shaking than on any other single factor. Two earthquakes of the same magnitude, but located in two different areas, may produce quite different effects on buildings. The depth of an earthquake from its focus is an important factor in this regard. Generally, the shallower the earthquake, the more severe the effect. The energy of deeper earthquakes is dissipated over a larger volume of rock, while that

Construction Planning: Environmental Impact of Foundation Studies and Earthquake Issues, Table 1 Abridged modified Mercalli intensity scale [1] (1956 version)

Description	
I.	Not felt except by very few under favorable circumstances (I Rossi-Forel scale)
II.	Felt only by a few people at rest, especially on upper floors of buildings
	Delicately suspended objects may swing (I–II on Rossi-Forel scale)
III.	Felt quite noticeably indoors, especially on upper floors of buildings, but many people do not recognize it as an earthquake. Standing automobiles may rock. Vibration like passing of truck (III on Rossi-Forel scale)
IV.	During the day felt indoors by many, outdoors by few. At night, some awakened. Dishes, windows, doors disturbed; walls make creaking sound. Sensation like heavy truck striking building. Standing automobiles rocked noticeably (IV–V or Rossi-Forel scale)
V.	Felt by nearly everyone, many awakened. Some dishes, windows, and so on broken; cracked plaster in a few places; unstable objects overturned. Disturbance of trees, poles, and other tall objects sometimes noticed. Pendulum clocks may stop (V–VI Rossi-Forel scale)
VI.	Felt by all, many frightened and run outdoors. Some heavy furniture moved; a few instances of falling plaster and damaged chimneys. Damage slight (VI–VII Rossi-Forel scale)
VII.	Everybody runs outdoors. Damage negligible in buildings of good design and construction; slight to moderate in well-built ordinary structures; considerable in poorly built or badly designed structures; some chimneys broken. Noticed by people driving cars (VIII Rossi-Forel scale)
VIII.	Damage slight in specially designed structures; considerable in ordinary substantial buildings, with partial collapse; great in poorly built structures. Panel walls thrown out of frame structures. Fall of chimneys, factory stacks, columns, monuments, and walls. Heavy furniture overturned. Sand and mud ejected in small amounts. Changes in well water. People driving cars disturbed (VIII + to IX Rossi-Forel scale)
IX.	Damage considerable in specially designed structures; well-designed frame structures thrown out of plumb; great in substantial buildings, with partial collapse. Buildings shifted off foundations. Ground cracked conspicuously. Underground pipes broken (IX+ Rossi-Forel scale)
X.	Some well-built wooden structures destroyed; most masonry and frame structures destroyed with foundations; ground badly cracked. Rails bent. Landslides considerable from river banks and steep slopes. Shifted sand and mud. Water splashed, slopped over banks (X Rossi-Forel scale)
XI.	Few, if any, (masonry) structures remain standing. Bridges destroyed. Broad fissures in ground. Underground pipelines completely out of service. Earth slumps and land slips in soft ground. Rails bent greatly
XII.	Damage total. Waves seen on ground surface. Lines of sight and level distorted. Objects thrown into the air

Masonry A, B, C, and D. To avoid ambiguity of language, the quality of masonry, brick, or otherwise is specified by the following lettering:

Masonry A. Good workmanship, mortar, and design; reinforced, especially laterally, and bound together by using steel, concrete, etc.; designed to resist lateral forces

Masonry B. Good workmanship and mortar; reinforced, but not designed in detail to resist lateral forces

Masonry C. Ordinary workmanship and mortar; no extreme weaknesses like failing to tie in at corners, but neither reinforced nor designed against horizontal forces

Masonry D. Weak materials, such as adobe; poor mortar; low standards of workmanship; weak horizontally

of the shallower earthquake is spread over a smaller volume of rock.

Engineers use strong-motion seismographs to record how the ground shakes during an earthquake. The most important parameters of ground motion (free-field motion) obtained from these instruments are the peak velocity, the peak acceleration, the frequency or period of waves, and the duration of strong motion. The peak velocity determines how fast the ground is shaking, whereas the peak acceleration indicates how quickly the velocity is changing, and the latter is more often used in the engineering design of structures. It may be borne in mind that the force generated by an earthquake is a product of the mass of a structure and the ground acceleration.

Ground Condition Considerations

Earthquake shaking is attenuated or reduced with distance from the epicenter. Nevertheless, for a

Construction Planning: Environmental Impact of Foundation Studies and Earthquake Issues, Table 2 Energy release of an earthquake. Magnitude (M)-energy (E) relation can be given the following formula: Log E = 11.8 + 1.5 M. If M is increased by 1.0, E is magnified by a factor of $10^{1.5}$, i.e., approximately 32. Therefore, the seismic energy of an M = 6 earthquake is about 32 times as large as that of an M = 5 earthquake and is about 1,000 times that of an M = 4 earthquake. Magnitude and energy of earthquakes (Modified from Earthquake Mechanics, [8])

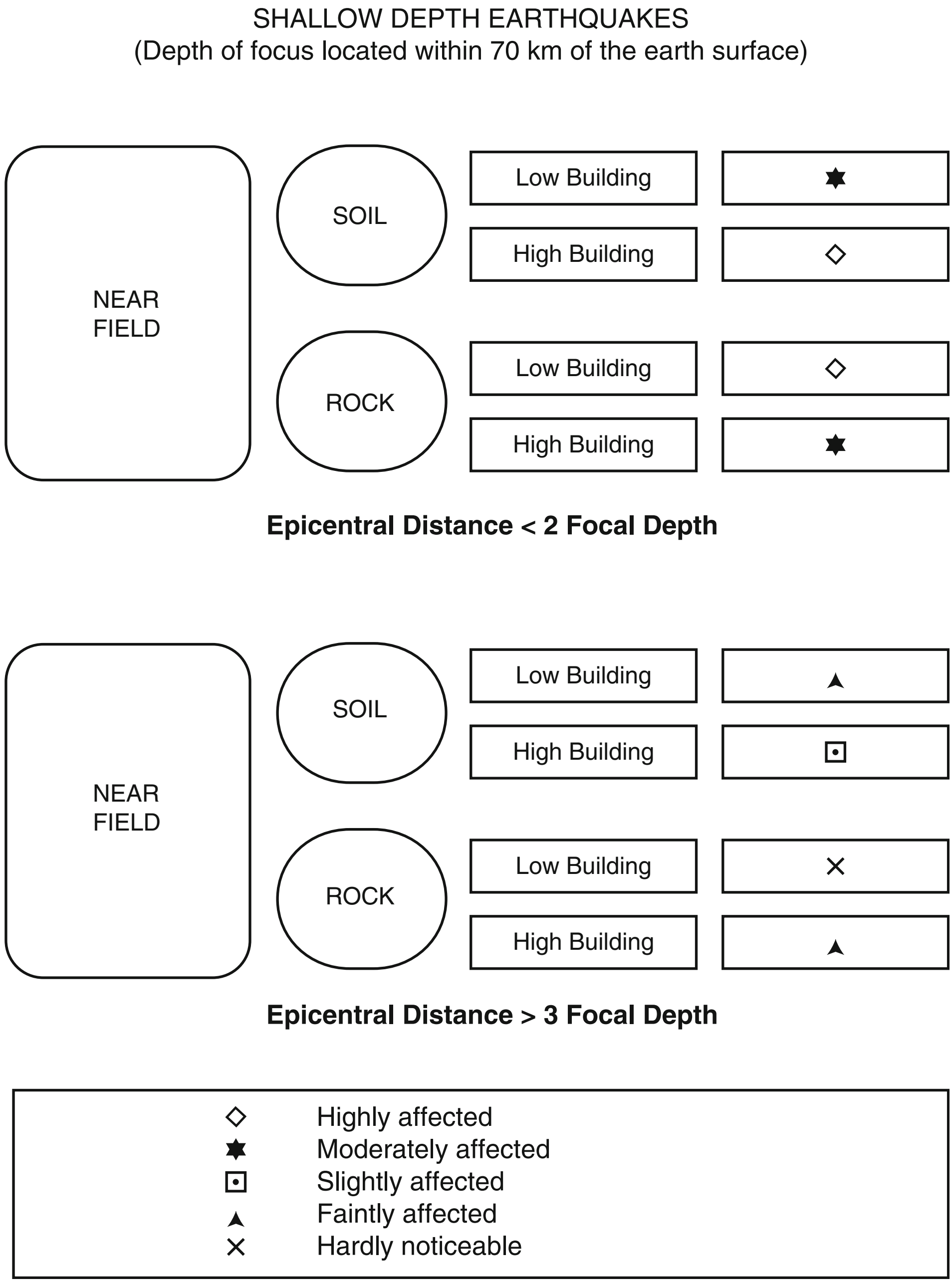
Magnitude (M)	Energy (Ions of TNT)	Nuclear bomb equivalence	Class
4.0	15	0.001 × Hiroshima	Small
		0.0008 × Nagasaki	
4.5	86	0.006 × Hiroshima	
		0.004 × Nagasaki	
5.0	478	0.03 × Hiroshima	Moderate
		0.02 × Nagasaki	
5.5	2,629	0.17 × Hiroshima	
		0.13 × Nagasaki	
6.0	15,057	1 × Hiroshima	Large
		0.75 × Nagasaki	
6.5	86,042	5.7 × Hiroshima	
		4.3 × Nagasaki	
7.0	478,011	31.7 × Hiroshima	Major
		23.8 × Nagasaki	
7.5	2,639,063	175 × Hiroshima	
		131 × Nagasaki	
8.0	15,057,361	1,000 × Hiroshima	Great
		750 × Nagasaki	
8.5	860,420,065	57,514 × Hiroshima	
		4,285 × Nagasaki	

TNT (trinitrotoluene) is a chemical compound
Dynamite has 60% greater energy than TNT of the same weight
Hiroshima atom bomb ~15,000 tons TNT
Nagasaki atom bomb ~20,000 tons TNT

fixed distance, the type of geological material has also a strong influence on the nature and magnitude of ground motion. It may be emphasized that the high-frequency (short-period) earthquake waves, transmitted by strong rocks, attenuate more rapidly with distance than the low-frequency (long-period) waves transmitted by soils (Fig. 1). Furthermore, earthquake waves change their characteristics as they encounter a material different from the one through which they were initially propagating. For instance, the high-frequency waves, while passing through rocks, are slowed down as they enter the overlying soil cover, but their period is increased instead. A clear distinction must, however, be made between attenuation described above and amplification. Although the ground motion of soils is attenuated more than that of rocks, the reverse is true as far as amplification of ground motions in soil. For the same epicenter distance, the thicker the soil cover, the greater the amplification factor, and the longer the duration of ground shaking. Topographic variations can also cause amplification of seismic waves. Summits of isolated hills and edges of plateaus and cliffs are sites of large amplification within a fairly wide frequency range.

Ground and Structure Interaction

Fundamental periods or frequencies of buildings vary according to the height and nature of buildings. When such structures are subjected to free-field ground motions, they respond differently to



Construction Planning: Environmental Impact of Foundation Studies and Earthquake Issues, Fig. 1 Effect of earthquakes on low- and high-rise buildings founded on soil and rock

different frequencies or periods of ground motion. For the same amplitude of ground motion, it is the resonance of frequencies (or periods) of ground motion and that of the building which is the cause of maximum damage. Thus, in terms of the frequency content of an earthquake wave, low-rise buildings (such as ordinary one- or two-story dwellings and similar structures) are more affected by high-frequency or short-period ground motions, as compared to low-frequency or long-period motions of the ground. The reverse is true for the high-rise buildings, which are less affected by high-frequency or short-period motions of the ground. This change in behavior of buildings is mainly attributed to resonance of soil-structure interaction, which can cause more damage when the ground motion frequency of earthquake waves and that of the building structure is the same [9]. It is, therefore, advisable to avoid high-rise buildings (low frequency) on low-frequency soils and, similarly, avoid low-rise buildings (high frequency) on high-frequency grounds.

It must also be remembered that the frequency of earthquake waves decreases as the distance of the site in question increases. Therefore, it is not surprising to note that sometimes high-rise buildings located farther from the epicenter suffer more than high-rise buildings located closer to the epicenter. This is particularly noticeable in situations where the buildings are underlain by rock. Low-rise buildings, founded on rock and close to the epicenter, are more vulnerable to damage than the high-rise buildings situated in the same area.

It has been shown from the foregoing that the damage to structures in earthquake-prone areas is largely dependent on ground acceleration; depth of earthquake focus; distance of the epicenter from the structure; site conditions including the nature of the ground, its thickness, and topography of the site; duration of ground shaking; and last but not the least, the type of structure. It must be emphasized that for engineering purposes, it is the intensity or resulting acceleration of the ground motion that is important and not the magnitude of the earthquake.

Mitigation, Construction Planning, and Management

Earthquake risk varies from location to location, from structure to structure, and from person to person. Damage due to earthquakes is particularly great in certain locations, and it is very difficult to prevent the earthquakes from occurring. Nevertheless, precautions must be taken to mitigate the impacts of earthquake in the vulnerable area. Earthquakes are often followed by aftershocks, which can be more detrimental to the main earthquake shock, because the structures may have weakened by the main shock. The magnitude of the main shock is generally greater than the aftershocks, the magnitude of which generally decreases with time and can be empirically determined as to how long the aftershocks will continue.

Earthquakes are inevitable, but the damage from earthquakes is not. Safety and survival in an earthquake is very important. It is important to be prepared and take necessary precaution to reduce losses; in the event of an earthquake, it is advisable to protect belongings and persons from falling objects. The instructions published by the United States Geological Survey, Department of the Interior, are very helpful in this regard.

If need be, vacate or demolish hazardous old buildings, and retrofit by strengthening the newly built structure, as well as regulating the construction of new buildings by adopting methods that cater to ground-structure interaction principles. Instances of fire, breaking of pipelines, disruption of electricity and water supply, and outbreak of disease are common following earthquakes, and among other things, they call for rescue, relief, rehabilitation of displaced people, relocation, and reconstruction of structures at safer places. The migration of displaced people from one area to another creates problems of adaptability and issues raised by the native community in accepting people from different social backgrounds. The burden on the native community can be reduced by the construction of additional infrastructure such as hospitals, schools, community centers, market places, and so on.

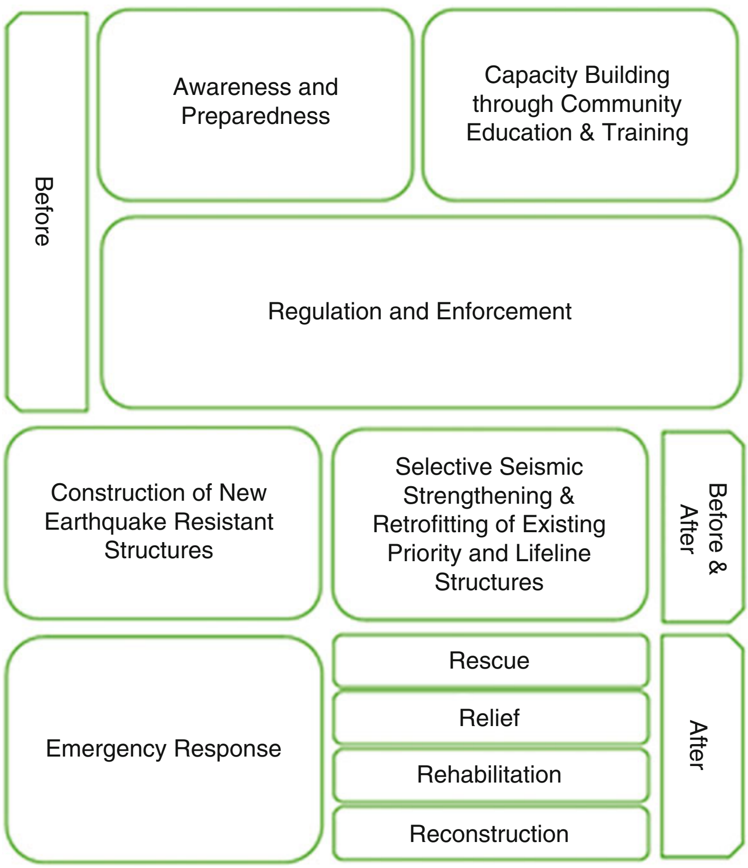
In an attempt to avoid adverse environmental impact to buildings, it is important to prepare isoseismic maps of major earthquakes, which may have occurred in the area [10, 11]. These maps delineate areas of equal damage or of equal intensity of the earthquake in the past. A similar exercise can be initiated to delineate areas in terms of return periods of earthquakes of the same intensity. Furthermore, seismic zoning maps can be prepared to define areas which are extremely, highly, moderately, and slightly susceptible to the likely damage, which may be caused by an earthquake of a specified return period. Building codes should be prepared and enforced to avoid damage to high-risk installation as well as buildings in different seismic zones.

Figure 2 illustrates management of earthquake-prone areas and after an earthquake. It includes measures which are necessary before an earthquake, before and after an earthquake, and after an earthquake. Before an earthquake, emphasis is placed on awareness and preparedness, capacity building through community education and training, as well as enforcement of regulations. There are measures, which can be adopted before and after an earthquake, such as construction of new earthquake-resistant structures, together with selective seismic strengthening and retrofitting of existing priority and lifeline structures. The emergency measures (4 Rs), which essentially apply after an earthquake has devastated the area, sequentially include rescue, relief, rehabilitation, and reconstruction.

Future Directions

Natural earthquakes are a major threat to society. Therefore, monitoring earthquakes is essential for providing scientific data to investigate complex earthquake phenomena and to mitigate seismic hazards [12].

The earthquake early warning systems offer practical information for reducing seismic hazards in earthquake-prone regions. Attempts will be



Construction Planning: Environmental Impact of Foundation Studies and Earthquake Issues, Fig. 2 Management of earthquake-prone areas before and after an earthquake made here to mention about some recent developments followed by suggestions for future work, in this direction.

Over the past few decades, notable progress has been made in the understanding of some of the factors and processes that govern the occurrence of earthquakes. For instance, it is known [13] that the amount of damage caused by an earthquake is not only related to its magnitude, depth, and geographical location of the site but also factors such as the type and quality of affected buildings, as well as the nature of the ground on which such buildings are constructed.

Building codes, relevant to different regions of the world, defined by earthquake zoning, earthquake recurrence interval, and isoseismic maps help in providing guidelines for the design, evaluation, and in some instances rehabilitation or retrofitting of building structures, in a given area.

The advent of Global Positioning Systems (GPS) opened a new avenue of opportunity to seismologists for monitoring movements in geographical positions of marked stations and allows better long-term forecasts of areas likely to be affected by earthquakes. This development together with advances in telemetry and invention of earthquake early warning systems linked with real-time seismology offers practical information for reducing seismic hazards in earthquake-prone regions. Such systems are already successfully deployed in Mexico, Japan, and Taiwan [14]. However, there is no general method for predicting earthquakes in the actual time domain, and only probable estimates of recurrence of earthquakes of a given magnitude are available. Moreover, as described above, there are techniques that can detect an earthquake in progress and provide notice of seconds to tens of seconds, prior to actual ground shaking.

Real-time seismology refers to a practice in which seismic data are collected and analyzed quickly after a significant seismic event. The results thus obtained can be effectively used for post-earthquake emergency response and early warning. Such systems, on their own, can generally detect strong shaking at an earthquake's epicenter and transmit alerts ahead of the damaging earthquake waves. The speed of an electronic warning signal, in time domain, is faster than the speed of earthquake waves traveling through the earth.

Unfortunately, there is no worldwide public early warning system for earthquakes. It is important to use this relatively new science and modern technology to rapidly detect the beginning of an earthquake, assess the hazard that the earthquake poses, and then provide a warning to people if they are in any harmful situation. Such early information about past earthquake occurrences is particularly useful in the discipline of disaster management. It calls for hazard mitigation plans from on-site and regional bases to consider post-earthquake emergency response and early warning.

There is a real need for a radically different design of seismographic networks employing earthquake early warning system. The importance of actively exploring the potential of wireless sensor networks cannot be overemphasized. However, building a radically different design of high-density seismic networks at national and international levels is not economical using existing seismic technology. Innovative approaches must be developed, and perseverance is needed.

Reference may be made to, "The November 7, 2012, M7.4 Guatemala Earthquake and its Implications for Disaster Reduction and Mitigation," a study jointly conducted by Earthquake Engineering Research Institute (EERI), Asociación Guatemalteca de Ingeniería Estructural y Sísmica (AGIES), and the World Bank [15]. The report presents a noteworthy practice for guidance pertaining to understanding the causes of structural damage to structures as well as the improvement of building codes and standards.

Bibliography

Primary Literature

1. Ambraseys NN (1971) Value of historical records of earthquakes. Nature 232:375–379
2. Poirier JP, Taher MA (1980) Historical seismicity in the near and Middle East, North Africa, and Spain from

Arabic documents (VIIth–XVIIIth century). Bull Seismol Soc Am 70(6):2185–2201
3. Press F, Sielver R (1982) Earth. W. H. Freeman, New York
4. Wesson RL, Wallace RE (1985) Predicting the next great earthquake in California. Sci Am 252:23–31
5. Mavroeidis GP (2008) Estimation of strong ground motion from the great 1964 Mw 9.2 Prince William Sound, Alaska earthquake. Bull Seismol Soc Am 98: 2303–2324
6. Chopra AK (1995) Dynamics of structures. Prentice Hall, Englewood Cliffs
7. Richter CF (1958) Elementary seismology. W. H. Freeman, San Francisco
8. Kasahara K (1981) Earthquake mechanics. Cambridge University Press, Cambridge
9. Newmark NM (1970) Current trends in seismic analysis and design of high-risk structure. In: Weigel RL (ed) Earthquake engineering. Prentice Hall, Englewood Cliffs
10. Deputy Ministry for Mineral Resources (1996) Gulf of Aqba earthquake of November 22, 1995 with emphasis on its effects in Saudi Arabia. Special report: Ministry of Petroleum and Mineral Resources, Jeddah
11. Shehata WM, Kazi A, Zakir FA, Allam AM, Sabtan AA (1983) Preliminary investigations on Dhamar earthquake, North Yemen of December 13, 1982. King Abdulaziz University, Jeddah. Bull Fac Earth Sci 5:23–52
12. Kanamori H (2005) Real-time seismology and earthquake damage mitigation. Annu Rev Earth Planet Sci 5:195–214
13. Green AG (1999) Mitigating the effects of earthquakes: problems, progress and future trends. In: Proceedings of the WMO/UNESCO sub-forum on science and technology in support of natural disaster reduction, WMO-914, pp 74–82
14. Nishenko SP, Savage WU, Johnson T (2009) Earthquake early warning: a prospective user's perspective (Invited), abstract #S21C-04. American Geophysical Union, Fall Meeting 2009. San Francisco, California
15. Earthquake Engineering Research Institute (2013) The November 7, 2012 M7.4 Guatemala earthquake and its implications for disaster reduction and mitigation. EERI, Oakland

Books and Reviews

Anon (1978) Joint conference on predicting and designing for natural and manmade hazards. In: Proceedings of ASCE-ICE-CSE. American Society of Civil Engineers, New York

Bolt AB (1998) Earthquakes. W. H. Freeman, New York

Dowrick DJ (1987) Earthquake persistent design. Wiley, Singapore

Elnashai AS, Sarno LD (2008) Fundamentals of earthquake engineering. Wiley, London

Isenberg J (1981) Proceedings of social and economic impact of earthquakes on utility lifelines: seismic consideration in lifelines, planning, siting and design. American Society of Civil Engineers, New York

Johnson RB, Degraff JV (1988) Principles of engineering geology. Wiley, New York

Kokusho T, Yoshimichi Y, Yoshimine M (2009) Performance based design in earthquake geotechnical engineering: from case history to practice. Taylor and Francis, Singapore

Smith K, Petley DN (2007) Environmental hazards. Routledge, New York

United Nations Educational Scientific and Cultural Organization (1978) The assessment and mitigation of earthquake risk. UNESCO, Paris

Dam Engineering and Its Environmental Aspects

Petar T. Milanović
Belgrade, Serbia

Article Outline

Glossary

Dam Civil structure planned, constructed, and operated to meet human needs in flood control, irrigation, supply of drinking water, electricity generation, recreation, and various other purposes.

Dam failure Collapse or movement of part of a dam or its foundation, so that the dam cannot retain water.

Guaranty ecological flow Required quantity and quality of flow to maintain the sustainability of the river ecosystem (ecological base flow).

Induced subsidence Collapse of the surface of the ground due to human activities, mostly reservoir operation and intensive pumping of groundwater.

Karst Terrain composed of highly soluble rocks (limestone, dolomite, gypsum, and salt), very risky environment for dams and reservoirs construction.

Large dams Dams having a height of 15 m from the foundation or, if the height is between 5 and 15 m, having a reservoir capacity of more than 3 million cubic meters.

Reservoir-triggered seismicity Seismic phenomena associated with impounding of reservoirs (reservoir-induced seismicity).

Definition of the Subject

Construction of dams and reservoirs involves considerable natural and anthropogenic impacts. These impacts are for the most part positive, but can have some negative influence on the environment. The main purposes of dam construction are focused on water regime improvement and consequently regional prosperity. Generally, the goal of dams and reservoirs is regional socioeconomic development by irrigation, flood control, power production, water supply, recreation purposes, reduction of deforestation, reduction of drought periods, fishing farms, mining purposes, navigation, to enhance landscape including development of new infrastructure, and to provide new possibilities for employment and many secondary benefits.

However, as a consequence of dam and reservoirs construction, a number of different and sometimes unpredictable negative environmental impacts and uncertainties cannot be avoided. Some common negative impacts are: the population migrates from inundated areas; the reservoirs cover arable land, settlements, and infrastructure; deep reservoirs provoke induced seismicity and induced collapses; water fluctuation provokes landslides along the reservoir banks; sedimentation of reservoirs; in some cases, important cultural and historical monuments are inundated; questionable impact on biodiversities, survival of wildlife,

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_308

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_308-3

and endemic species is endangered; tailings may contain dangerous chemicals; and regime of surface and underground water is considerably changed. In a number of cases, socioeconomic constrains related to migration from submerged regions are very pronounced.

The worst most negative and disastrous impact is dam failures. The karst environment is a particularly sensitive influence on dams and reservoirs, and a variety of positive and negative consequences are numerous in such terrain.

Introduction

Since time immemorial, people have considered how to tame the surface waters to prevent floods and to use water for different purposes, irrigation, water supply, and, much later, for electricity production. In many areas in the world, life was not determined by man but by the rivers. The people have had to cope with two kinds of misfortune: flood and drought. In many cases, the consequences were disastrous. One single flood of the Yangtze River in China (1931) devastated 3.3 million hectares of arable land and caused the suffering of 28 million people and the loss of more than 145,000 lives [1].

From very ancient times, dams appeared as the only effective structures to tame river waters. Construction of dams started a few 1000 years ago. Primary role of dams is to store or to divert waters. The oldest known are Jawa dams in Jordan (~3000 BC). In Egypt, the Kosheish Dam was constructed during the period 3000–2900 BC and Saad El-Kafra Dam about 2610 BC. The Anfantang reservoir in China was built in the sixth century BC, and a 30-m-high gabion dam was constructed around 240 BC in Shanxi Province. In Iran, dam constructions date from before 2000 years ago (Bahman Dam, Fig. 1); Shapour and Mizan dams were constructed during the reign of King Shapur I about 1700 years ago; the Tilkan and Sheshtarz Dam, 1000 years ago. The Amir Dam north of Shiraz, 1000 years old, still is operational [2]. In Spain, the Proserpina Dam, 22-m high, was built in the second century and still is operational.

Later in the twentieth century, individual dams were constructed to control the large hydro-systems, which consist of a number of dams and reservoirs, to change water regime at large catchment areas: Tennessee Valley Authority (29 dams), USA; dams along the Yangtze River in China; Southeast Anatolia Project (21 dams); the Volga-Kama cascades, Russia (11 dams); dams at Karun River catchment in Iran (16 dams); or dams in karst area like the Hydro-system Trebišnjica, Bosnia and Herzegovina (BiH), and Croatia (6 dams).

In many cases, dam projects were initially controversial and potentially disastrous. Failures of dams and potential environmental impacts are the main reasons for controversy and fear. However, with increasing demand on water resources and electric power, at issue is how to keep the balance between the necessity for development and preservation of the environment. At many locations, dams are still structures of great importance.

According to ICOLD World Register of Large Dams (1998), the main purpose of large dams is irrigation 37%, multipurpose use 22%, electricity generation 16%, water supply 12%, flood control 6%, recreation 3%, and other purposes 4%. Tailing dams are excluded from this survey.

In the USA, considering all dams (not large dams only), the main purpose is recreation 33.8%, flood control 15.6%, fire control 13.7%, irrigation 9.5%, water supply 9.4%, electricity production 2.9%, and the rest for many different purposes.

Environmental problems are coupled with political (transboundary) problems if dams are constructed at rivers bordering countries: for example, the Iron Gate Dam at the Danube River between Serbia and Romania; the Aswan Reservoir at the Nile River between Egypt and Sudan; dams on the Euphrates and Tigris rivers impacting Turkey, Syria, and Iraq; and the Itaipu reservoir along the border between Brazil and Paraguay.

Dam Engineering and Its Environmental Aspects, Fig. 1 Bahman Dam, Iran. Constructed approximately 2200 years ago

Flood Regulation: Regional Impact on Population

Historically, for thousands of years the primary role of dams was protection of settlements and arable land against flooding and for irrigation. Presently, flood control and irrigation are still of high importance, but usually dams and reservoirs are multifunctional, including power production, water supply, sediment control, landscape improvement, and recreation.

The Tennessee Valley Authority (Tennessee River, USA), founded in 1933, is one of the largest dam reservoir projects for flood control, navigation, power production, and irrigation [3]. Under natural conditions, the water regime was unfavorable for agriculture and life. Thirty percent of the population in the Tennessee Valley was affected by malaria. To construct 29 dams and reservoirs, more than 15,000 families were displaced. Electricity generation, flood control, and better organized water regimes have been of great benefit for the region. More than 1000 km of navigation channels have been constructed as part of this project. One important positive environmental impact is control of surface waters with regard to malaria prevention.

Large floods were a regular occurrence along the Nile River under natural conditions. At the same time, floodwater deposited about 4 million tons of nutrient-rich sediments per year. To prevent floods, the first modern dam at Aswan was built in 1889. The first dam was not high or effective enough to control flooding, and a new Aswan High Dam was constructed during 1960–1970 a few kilometers upstream. Over 60,000 Nubians were relocated from the reservoir area. The reservoir contains a volume of 169 billion cubic meters. About 17% of the reservoir is in Sudan. After dam construction, the annual floods are under control, and the navigation properties of the Nile River are considerably improved. However, artificial fertilizers have to be used instead of natural nutrients. Quality of the soil for farming has decreased. Negative impacts due to lack of rich sediments in the delta region is one negative environmental consequence.

Over the past two centuries, many people have died along the Yangtze River, China, due to catastrophic floods. In 1840, about 156,000 persons lost their lives during flood periods; in 1931, 145,000; in 1954 about 33,000; and during more recent flooding in 1998, over 1500 people died. Millions of hectares of arable land has been destroyed and is unusable. Numerous villages have completely disappeared. During the flood of 1954, 18 million were forced to evacuate from the area. For 3 months, Wuhan City with eight million people was covered with floodwater. After construction of the Three Gorges Dam (181 m high, 2335 m long), the frequency of major floods has been reduced to a minimum, and after project completion, large ships will be able to navigate from Shanghai 2400 km upstream. Due to project requirements, about 1.24 million people had to be relocated.

For construction of the Moses-Saunders Power Dam on the St. Lawrence River between Canada and the USA, 6500 people of Eastern Ontario were dislocated.

As result of construction of the Chira-Piura irrigation system in Peru, including the 9-km-long Poechos Dam, an area of 100,000 ha of uncultivated land has been turned into fertile farmland.

Construction of the Akosombo Dam in Ghana created one of the largest world reservoirs, Lake Volta; about 15,000 houses were inundated and 78,000 people were resettled.

The Gibe III Dam (240-m high at Omo River) in Ethiopia was controversial from its beginning. During the Omo River flood of 2006, at least 360 people and thousands of livestock were devastated. From an energy viewpoint, this project would provide great benefits for Ethiopia and Kenya. However, equally important is the projected negative impact on fisheries in adjoining Kenya's Lake Turkana, which threatens the livelihoods of about 0.5 million tribal people.

To construct the Ataturk Reservoir in Turkey, about 55,000 people were relocated, and from the Manantali Reservoir area (Mali), 15,000 people were displaced. It is estimated that the controversial Ilisu Dam project (Turkey) would displace at least 55,000 people. For construction of the Everkiiskaya Dam (Central Siberia, Russia), about 7000 local indigenous people would need to be relocated and a huge area flooded (9000 km^2).

Dam Failures

One of the worst negative environmental impacts of dams is the risk of failure. The first known dam failure, as reported by Herodotus, was the collapse of the Saad Ei-Kafra Dam in Egypt during flooding ~2500 BC. In modern history, one of the oldest reported failures was the Blackbrook I Dam (Great Britain, 1799). Official worldwide database and case histories of dam incidents and failures are not still complete. According to the ICOLD Bulletin 99, Dam Failures, Statistical Analysis (without China), a total of 5268 dams were built until 1950 (117 of them failed); 12,138 dams were built during the years 1951–1986 (59 of them failed).

Many incidents of dam failure occurred more than 50 years ago and involved older dams. In recent decades, the failure rate (particularly for dams more than 30-m high) has drastically decreased. At present, dams are constructed on the basis of thorough and detailed multidisciplinary investigations, including application of new technologies and detailed environmental studies, and operational stability is regularly monitored.

The world's recorded dam disaster occurred in 1975 in China (Banqiao Dam). Due to a strong hurricane and precipitation of 500 m over 3 days, the Banqiao Dam, together with Shimantan Dam and 62 small dams, was totally demolished. Water waves as high as 6–10 m and 12-km wide flooded more than a million hectares. More than 26,000 people were killed, and many more died afterward from resulting epidemics, for an estimated 150,000 total deaths. Eleven million people lost their homes.

Dam failures as a consequence of geology are sometimes catastrophic, causing loss of life or evacuation of thousands of people living downstream of the dams: for example, Malpasset (France), St. Francisco and St. Fernando (USA), Baldwin Hills (USA), and Teton (USA).

Total failure of the Malpasset concrete arch dam (66.5-m high) caused a huge flood on December 1959. Dam foundation consisted of gneissic rocks. Two sets of faults had crucial roles in the creation of a wedge failure. More than 325 persons lost their lives and a large area was devastated [4].

The worst American civil engineering failure of the twentieth century was the St. Francisco

Dam (California, USA), killing 450 people along the St. Francisco Canyon and St. Clara Valley. Foundation rock of the concrete gravity dam (60-m high) is conglomerate. During March 12–13, 1928, the St. Francisco Dam collapsed due to a paleo-landslide at the left abutment and strong uplift. According to Cooper and Calow (1998), the failure was partially attributed to gypsum dissolution [5].

The Teton Dam (USA) failure happened on June 5, 1976, killing 14 people; that embankment dam was 93-m high above riverbed. Dam foundation consists of basalt and rhyolite rock. Piping was identified as the most probable cause of failure [4].

Catastrophic failure of San Juan earth dam (Spain) occurred during the first filling of reservoir (2001). Due to intensive dissolution of gypsum, part of dam collapsed, provoking a huge flood downstream [6]. The flood caused by failure of the Lower San Fernando Dam (California, USA, 1971) due to a strong earthquake ($M = 6.6$) caused temporary evacuation of 80,000 people from the downstream area [4]. The Baldwin Hills Reservoir (USA) failed in 1963 causing enormous damage downstream; however, thousands of inhabitants were evacuated in time [4].

Some more dam failures at the USA are: East Fork Dam, Kentacky, 30-m high, collapsed during the night as consequence of piping through the karstified foundation rock (December 1978). Failure of the Fontenelle Dam, Wyoming, occurred in 1965 and 1982 when the large subsidence occurred. The Quail Creek Dam, Utah, collapsed in 1989, and failure of the Swift No. 2 Dam, Washington, occurred in 2002.

Thousands of inhabitants downstream from dams all over the world are permanently under psychological and mental pressure due to possible dam failure. Some of them seek relocation to safer places. If an area downstream from dam is populated, flood wave analysis in case of dam failure is very important. Analysis includes timing of water wave propagation, estimation of flood level, and installation of emergency alarm systems. Two important parameters are time of flood wave between dam site and urban areas and safe elevation for population evacuation – the elevation above which a disastrous wave is not effective.

In many cases, citizens living downstream from dams protest strongly against them, sometimes insisting on dam displacement. In the case of the Mulholland Dam (65-m-high concrete gravity dam) because of the disaster of St. Francisco, there was enormous protest from the citizens of Hollywood lining downstream from the dam. To solve this mostly psychological problem, the downstream dam face was covered in 1933 by a huge earthen mass, making it one of the most conservative dam structures in the world.

Mosul Dam (Iraq) is declared (2015) as the most dangerous dam in the world due to permanent seepage under the gypsiferous foundation rock. By measuring the dissolved materials in seepage water, an amount of 42 to 80 t/day was established (Guzina at all, 1986). According to analyses, it was estimated that the water wave in the case of failure will have a height of 54 m and a flow of approximately 551.000 m^3/s. This wave will reach the capital of Iraq – Baghdad – after 38 h. According to Al-Ansari et al. (2015), the expected wave height will be about 4 m in the urban area, and the potential number of casualties is estimated to be a half million. Dam structure still is under control; however, the risk for downstream settlements, including Mosul Town which is much closer to the dam, is not eliminated.

Strong earthquakes may affect a large area, and many dams may be subject to strong ground shaking and in extreme cases dam failures. Well-known dam examples affected by earthquakes over the past two decades are cited below.

Due to Maule earthquake in Chile, February 27, 2001 (M 8.8), several dams were damaged.

During the 2001 Bhuj earthquake in Gujart, India, 245 dams, mostly small embankment dams, have been rehabilitated. As consequence of the Wenchuan (China) earthquake, May 12, 2008, 1803 dams and reservoirs and 403 hydropower plants were damaged. However, only four of them had a height exceeding 100 m (Wieland 2014).

The $M = 9.0$ Tohoku Earthquake in Japan (March 11, 2011) resulted in failures of two dams: the 18.5-m-high Fujinuma earth dam and a smaller saddle dam. When the earthquake occurred, the reservoir level was almost full. Almost the entire length of the dam was overtopped approximately 25 min after the earthquake, destroying a small village and killing 8 people (Towhata at all 2011). Another 400 dams, subject to earthquake shaking, had to be inspected.

Reservoir Slope Instability

Reservoir slopes are exposed to different kinds of hazards. The most common is the potential for landslides to cause a wave which might overtop the dam crest, causing dam failure and a disastrous flood wave downstream. According to Schuster (2006), at least 254 large dams worldwide have been subjected to landslide activity [7]. The most common types of hazards are instability of slopes and deterioration of reservoir water quality due to solution processes if the slope consists of evaporates. The most frequent remedial measures to prevent instable rock masses from sliding are retention walls, prestress anchors, galleries and other drainage structures, and grouting and cutting of sliding planes.

Due to hydrodynamic force caused by reservoir fluctuation, the slopes are exposed to sliding, creeping, and toppling. This process can have catastrophic consequences. The difficulty is how to predict potential for a landslide on the basis of geological data and geological history of reservoir rims. An attempt has been presented by Moon (1997) in New Zealand [8]. He has established a magnitude-frequency curve for landslides based on geomorphological evidence of sliding in a valley and the geomorphological history of a valley.

In the case of Vaiont Dam (261-m-high dam in Italy), a huge landslide suddenly slid into the reservoir on October 9, 1963. The event lasted only 45 s, but a volume of about 300,000 million cubic meters plunged into the reservoir. The landslide length was 1.850 km and average thickness was 157 m. Maximal thickness of the slide rock mass was 330 m. A water wave was created which overtopped the 100-m-high dam crest. The catastrophic water wave completely demolished the small town of Longarone, 2 km downstream, and six more settlements. About 1700 inhabitants lost their lives and a number of industrial structures were completely destroyed [9]. The dam structure itself was not damaged at all.

In a number of cases, massive stabilization structures are constructed to eliminate hazards during reservoir operation. Induced slope stability is a problem along the 650 km of the Three Gorges Reservoir. The Lianziya potentially sliding rock mass, located about 25 km from the upper stream of the Sadoung dam site of the Three Gorges (China), has a volume of 2.26 million cubic meters. Deep prestress bolting, up to 3000 kN, is used to improve slope stability [10].

Tailings Dam Failures

Tailings dams are more vulnerable than other dam types. In the case of failure, environmental impact is catastrophic and long lasting. Tailings usually contain high concentrations of different chemicals. They represent a potential threat of environmental contamination, in some cases by extremely dangerous chemicals such as heavy metals or cyanides. One of the latest incidents (Baia Mare, Romania) occurred in January 2000 and released about 100,000 m^3 cyanide-contaminated water into the catchment area of Tisa River (tributary of Danube River), provoking great public concern in the huge and highly populated downstream area.

The last “Chronology of major tailings dam failure, 1960–2011,” is prepared on the basis of Bulletin 121, published by the United Nations Environmental Programme (UNDP), Division of Technology, Industry and Economics (DTIE), and International Commission on Large Dams (ICOLD) (Paris 2001, 144 p., updated March 2011) documenting 221 tailings dam incidents. Tailings dam failures occurred for many different reasons but mostly after heavy rainfall due to overtopping, seepage, foundation failure, or dam wall failure or liquefaction during earthquakes. Poor management and inadequate construction methods can also contribute to dam failure [11].

In many cases, failures are disastrous. Heavily polluted tailings flow can travel from a few 100 m up to more than 100 km. After failure of the tailing dam at Silverton, Colorado, USA (1975), tailings flow polluted 160 km of the Animas River; near Fort Mead, Florida, USA (1971), tailings flow traveled 120 km; and in the Inez failure in Martin County, Kentucky, USA (2000), polluted flow traveled 120 km; in the case of Huancavelica, Peru (2010), the Escalera and Opamayo rivers were contaminated 110 km downstream; and during dam failure of El Pocho, Bolivia (1996), 300 km of the Pilcomayo River were contaminated.

The worst impacts of tailings dam failure are the great number of people killed. From 1960 to 2010, 23 cases of tailings failures with fatalities were reported. Drastic examples are failure of the Sgorigrad, 1966 (Bulgaria), 488 people killed; the Stava, 1985 (Trento, Italy), 268 killed; Taoshi, 2008 (Shanxi Province, China), 254 killed; El Cordobe, 1965 (Chile), 200 killed; Aberfan, 1966 (Wels, UK), 144 killed; Buffalo Creek, 1972 (W. Virginia, USA), 125 killed; and Mufulira, 1970 (Zambia), 89 killed, Mount Polly 2014 (British Columbia).

One of the well-known recent tailings failures happened in Hungary (near Kolontar). The red mud with a pH of 13 flooded several settlements, killed 9, and injured 123 people.

A recent catastrophic failure occurred in the New Wales Plant, Florida 2016. A large sinkhole, 14-m wide, appeared in a phospogypsum stack, opening pathway for contaminated liquid into the underground. The liquid reached the Floridan aquifer, a major drinking water resource; 840,000 cubic meters of contaminated liquid were released.

Common impacts after tailing failures are demolition of homes, relocation of people, inundation of agricultural land, and catastrophic consequences for biodiversities in downstream areas, particularly for fish.

According to Rico et al. [12], for a group of 147 cases of worldwide tailings failures, 39% happened in the USA, 12% in Chile, 10% in the UK, and 4.8% in the Philippines [12]. Twenty-six occurred in Europe. With regard to tailings dam height, a greater percentage of failures occur if the height is not higher than 30 m. Tailings failures are frequently related to heavy precipitation or due to seismic liquefaction. More than 85% of failures occurred during mine operation, and only 15% of incidents were related to inactive tailings dams.

Reservoir-Triggered Seismicity: Induced Seismicity

From the very beginning, reservoir-triggered seismicity has been controversial. The first documented case was the case of the Hoover Dam (Lake Mead, USA). Presently, more than 60 cases of reservoirs are frequently cited to have experienced reservoir-triggered seismicity (Perman et al. 1983 and Gupta 1992). Magnitudes are greater in the case of greater depths of reservoirs: Koyna (India), depth 100 m, M = 6.3; Kremasta (Greece), 120 m, M = 6.3; Kariba (Zambia), 122 m, M = 6.25; Hsingfengiang (China), 105 m, M = 6.0; Srinagarind (Thailand), 133 m, M = 5.9; Oroville (USA), 204 m, M = 5.8; Aswan (Egypt), 90 m, M = 5.2; Hoover (USA), 191 m, M = 5.0; Kurobe (Japan), 186 m, M = 4.9; and Mratinje (Montenegro), 220 m, M = 4.1. However, except for a few cases (Koyna and Hsinfengkiang), human and material losses were negligible. The Hsinfengkiang Dam was considerably damaged, but so far no dam has collapsed due to the effect of induced seismicity [13].

According to experience in the Chinese karst regions, induced earthquakes are reported in the case of dams with height less than 100 m and in small reservoirs: Qinwo (Liaoning), 50 m, M = 4.8; Danjiangkou (Hubei), 97 m, M = 4.7; Qianjin (Hubei), 50 m, M = 3.0; Xindian (Sichuan), M = 4.2; Fengcun (Shaanxi), 30 m, M = 2.9; and Nanchong (Hunan), 45 m, M = 2.8 (Yuan D., 1991).

In general, triggered seismicity starts during the first impounding of the reservoir and increases with reservoir water levels; intensity of shaking sharply decreases with distance from the reservoir.

Certain earthquakes registered during reservoir impounding in karst indicate the role of karst in genesis of induced shocks. Those analyses indicate possible explosions of the compressed air during an abrupt reservoir impounding and simultaneous abrupt rising of the water table in the surrounding karst aquifer. Pressure of the air trapped in karst channels and siphons significantly increases. Trapped air pillows escape, creating strong explosions that are felt by inhabitants and recorded by seismographs at the surface. Environmental impact of this process is generally local and noisy but not harmful. Small damage is possible in the case of older village structures.

Dams and Heritage Protection

During dam construction, in some cases very important monuments and internationally recognized old-civilization heritage sites are threatened by inundations. Some are world heritage sites protected by UNESCO. In many cases, reservoirs may flood national parks, caves of archeological importance, or old necropoli, monasteries, graveyards, and ancient bridges.

The most famous are the Abu Simbel temples in Egypt built in the middle of the Nubian Desert (present reservoir) by Ramses II who ruled from 1290 to 1224 BC (19th dynasty). After construction of the Aswan Dam, both monuments of Ramses II and of his wife were sawn into 1036 blocks, 30 t each, plus 1110 blocks from surrounding rock. Monuments were reconstructed 90 m above the original level (Fig. 2).

Due to construction of dams along the Trebišnjica River (BiH), two ancient monasteries built in 1232 and during the first part of the fourteenth century, and one bridge constructed at the first part of the sixteenth century, nationally recognized cultural heritage sites have been relocated from the reservoir areas (Fig. 3).

Construction of the lower Gordon Dam in southwest Tasmania would have flooded a large karst area containing caves of great archeological importance. The project was abandoned for legal and environmental reasons in 1983 [14].

During construction of the Iron Gate Dam on the Danube River, a number of historical monuments, including ruins of the old bridge over the Danube River dating from 28 to 104 AD (time of Roman emperors Tiberius, Claudius, and Traianus) have been flooded. To preserve the monument "Tabula Traiana," heavy limestone blocks of 250 t were sawn and lifted 20 m to be above the reservoir level. The old prehistoric settlement (Lepenski Vir) dated between fifth and sixth millennia BC, which represents one of the oldest cultures in this part of Europe, was relocated above the Danube reservoir level.

Along the Three Gorges Reservoir (Yangtze River, China) about 1300 archeological sites, including 30 Stone Age localities, have been carefully investigated. About 1200 of them will be relocated to higher places. Some irreplaceable historical artifacts, however, have been permanently inundated.

The old Greek and Roman City of Zeugma, larger than Pompeii, founded in 300 BC, on the Euphrates River, was inundated after construction of the Birecik Dam, Southeast Anatolia Project, Turkey (1999). Zeugma mosaics have been declared one of the best preserved Roman mosaic collections in the world. This ancient city is submerged, but its famous mosaics were placed in the museum of Gaziantep.

The proposed Ilisu Dam (Southeast Anatolia Project, Turkey) on upper Tigris River, 65 km from the Syrian border, threatens to inundate Hasankeyf, an internationally recognized Roman, Byzantine, and Ottoman historical and cultural heritage site. Historical monuments include thousands of caves carved more than 2000 years ago. About 50 villages and 15 small towns along the Tigris valley would be displaced [15].

A large part of Munzur Valley National Park (Turkey) is to be flooded by construction of eight dams. Thousands of endemic species will become extinct after finalization of the project. Some of these dams are already constructed (Mercan and Uzuncayir dams).

A dam project in Coa Valley, Portugal, was canceled because of the important Ice Age rock art. Many irreplaceable artifacts of an ancient Mesopotamian city (2000 BC) are endangered by construction of the Makhal Dam in Iraq. In the case of construction of Sardar Sarovar and Narmada Sagar dams at Narmada River (India), 250,000 people will be displaced and 3000-year-old historical temples are potentially endangered.

Dam Engineering and Its Environmental Aspects, Fig. 2 Abu Simbel monument replaced from Aswan Reservoir, Egypt

Dam Engineering and Its Environmental Aspects, Fig. 3 Old monasteries at bottom of Bileća reservoir (BiH)

Dams and Ecosystems

After dam construction, it is not simple to keep ecosystem parameters upstream and downstream at the same levels as preconstruction conditions. Most frequently, temperature and flow regimes are disturbed, particularly if the purpose of the dam is power production, where a reservoir water body is not thermally and hydraulically homogenous. Magnitude and frequency of reservoir fluctuations are rapid and huge. Thermal stratification is quite pronounced. Intake structure position at the dam body is one of the important requirements to reduce the effect of complex processes in a reservoir. Water quality disturbances upstream from a dam are transferred to downstream flow. Enormous daily flow fluctuation and velocity due to hydroelectric power plant operation could have disastrous effects on flora, fauna, and physical properties of a riverbed. To minimize downstream negative environmental impacts, guaranty ecological flow is an essential requirement.

Guaranty ecological flow (or ecological base flow or environmental flow requirements) is usually one of the controversial requirements in dam construction. Different methods are proposed to define base flow (for instance, the Tennant method). Compared with frequently used sustainable flow or biological minimum flow, the guaranty ecological flow is accepted as the quantity of water flow which guaranties natural ecosystem sustainability [16]. Flora and fauna are the key parameters to establish balance between all ecological parameters, that is, to preserve ecosystem integrity. Adaptation processes may take place over many years before a new ecological balance is achieved. In the case of deep reservoirs for electricity production, with deep intake structures, the temperature of water downstream from a dam usually is much colder than under natural conditions. Daily, seasonal, and annual flows are quite different than under natural conditions. Consequently, the ecosystem is disturbed, and immediately below the dam structure, the water cannot be used for recreational purposes.

Dams reduce sediment load downstream. If a dam and a reservoir are constructed on a river close to a seacoast, the estuarial effect (saltwater intrusion) upstream is expected, that is, there can be a negative influence of brackish water on native biodiversity. Due to dam construction, the rate of deposition of sediments (fines, sand, and pebble) is reduced. If commercial excavation of sand and pebbles remains as before, the geometry of the riverbed drastically changes. If excavation of sand and pebbles occurred between the dam and the seacoast, the problem becomes more complicated. Excavated sediments cannot be naturally replaced; the river bottom becomes deeper and deeper. As a consequence, the saline water wedge penetrates upstream much faster. This effect can lead to declines in native aquatic vegetation, fish, and amphibian species. Over time, the natural balance is disturbed and brackish water species replace native species. In some cases, saltwater penetration endangers the quality of groundwater near riverbanks creating irrigation and water supply problem.

One of the most serious environmental consequences of dam construction is obstruction to fish migration. Dams are barriers for migratory fishes, such as salmon, trout, sturgeon, alewife, skate, eel, and many others. Construction of fish ways dates from about the seventeenth century in France. Presently, dams are widely equipped with several types of fish ways. These structures are not effective for all fishes or are only partially effective. In the case of downstream migration, mortality of fish passage through power plant turbines or over spillways is significant [17].

Particular problems appear in the case of reversible (pumping) power plants. When the power plant is in pumping regime, fishes can be sucked in by the turbines and transported through the pressure tube and headrace tunnel at the upper compensation reservoir.

Multiple dams along the river considerably worsen the situation for migratory fishes, but in the case of dams in the Glomma River system (Norway), efficient fish ways were constructed at eight dams along 122 km of river [18]. Four dams along the Peconic River (USA) have been equipped with fish ladders. Successful salmon migration has been studied in the case of Snake River dams, Lower Granite dams, and many other locations.

The Iron Gate Dam at the Danube River is not equipped with "fish ways." As a consequence, the caviar productive fish, sturgeon, and skate are unable to migrate from the Black Sea to the Danube River. The 11 dam cascades on the Volga and Kama rivers impede migration north from the Caspian Sea for several sturgeon migratory species including Beluga (Beluga caviar).

Construction of dams can negatively impact wetland ecosystems. Usually, wetlands are extremely rich in diversity of flora and fauna. Reduction of base flow which feeds wetland areas can lead to declines in aquatic vegetation, fish, and birds. Many wetlands are temporary recovery stations for migratory birds. Dam influence may cause reduction of some bird species; however, in some cases dam operation can be easily adapted to support, or even improve, wetland ecosystems.

Dams and Microclimate

As a consequence of damming, estimated total current reservoir surface is more than 400,000 km^2. Largest reservoirs are Lake Volta, Ghana, 8502 km^2; Aswan Reservoir, Egypt/Sudan, 5250 km^2; Itaipu Reservoir, Brazil/Paraguay, 1350 km^2; Ataturk Reservoir, Turkey, 817 km^2; and Keban Reservoir, Turkey, 675 km^2. Due to hot weather or strong winds, evaporation from reservoir surfaces is estimated at about 2 m^3/m^2 per year.

Exact climatological measurements and analyses related to reservoir impact are rare. Measurements taken at the reservoirs at Pournari and Mornos (Greece) and Bileća reservoir (BiH) indicate negligible increase of temperature close to the reservoir areas. For more precise conclusions, long-term monitoring of different parameters is necessary.

According to subjective impressions of local people, there is some climatological influence. In karst areas, intensity of deforestation decreased, and fogs are registered much frequently after dam and reservoir construction. Impact of the Krasnoyarsk Dam (Russia) on the Yenisei River reaches 200 km downstream and has an influence on local climate by increasing freezing fog.

Induced Subsidence (Collapse)

Origin of subsidence can be natural or induced. Induced subsidence is a consequence of human activities but mainly due to groundwater extraction, mining, and dam construction [19–22].

Subsidence development is a common process as a consequence of dam construction and reservoir operation. Induced subsidence is a series of spatially independent random events created by reservoir operation. Events such as these are unpredictable and practically instantaneous. Some prominent examples are the following reservoirs: Wolf Creek (USA), Hutovo (BiH), Slano and Vrtac (Montenegro), Tarbela (Pakistan), Mavrovo (FYR Macedonia), Perdika (Greece), Hammam Grouz (Algeria), Kamskaya (Russia), Lar (Iran), Keban (Turkey), North Dike (Florida, USA), Samanalawewa (Sri Lanka), Mosul and Haditha (Iraq), May (Turkey), Horsetooth Reservoir and Center Hill Dam (USA), La Loteta (Spain), and Huoshipo Reservoir (China). In some cases instantaneous seepage can be enormously high. In the case of the Keban Dam immediately after collapse occurred, seepage was 26 m^3/s.

Subsidence is induced by extensive water level fluctuation in reservoirs and results in extensive water leakage. In some cases, subsidence occurred after many years of successful operation. In the case of Hammam Grouz (Fig. 4), subsidence occurred after 17 years and in the case of Mavrovo, after 25 years.

In some cases, subsidence creates considerable environmental impact. For instance, the Mavrovo Reservoir collapse resulted in heavy damage to local roads and surrounding houses.

In the case of Kamskaye Reservoir (Russia), after dam construction the dissolution process in gypsum has intensified in the vicinity of the reservoir [23]. During the period 1956–1961, 11 collapses occurred. Prior to dam construction, in the same area, only two collapses were registered during the previous 50 years.

Spring Submergence Due to Dam Construction

Submergence of large springs by artificial reservoirs and the consequences on environment and reservoir integrity are frequently discussed. After construction of the 185-m-high Oymapinar Dam (Turkey), the large Dumanly Spring, Qmin = 35.6 m^3/s, was flooded by 120 m of water head at maximum storage level; the Trebišnjica Spring (BiH), Qav = 80 m^3/s, was flooded by 75 m of water column; the Neraidha Spring (Greece), Q = 10 m^3/s, was flooded by 40 m of Poliphiton Reservoir; the 220-m-high Piva Dam (Montenegro), the large Pivsko Oko Spring, Qav = 25.5 m^3/s, was flooded by 70 m; the Rama spring zone (BiH) was submerged by 40–60 m of water column; and the

Dam Engineering and Its Environmental Aspects, Fig. 4 Hammam Grouz Dam, Algeria. Subsidence occurred after 17 years of dam operation

Yarg Spring (Iran), Qav = 0.7 m^3/s, was flooded after construction of the Salman Farsi Dam (136 m) by 27 m; Oko Spring (BiH) was flooded after construction of 35-m-high Gorica Dam by 17 m of water column [24]. The Bel Spring (Sirvan River, Iran) is submerged more than 100 m by construction of the 230-m-high Darian Dam. Spring discharge varies between 0.2 m^3/s and 6 m^3/s. To keep two factories for water bottling and water supply of the local community operational, a very complex structure was constructed. This structure consists of a gallery and deep well behind the spring outlet. To prevent influence of reservoir water and deterioration of water quality, all natural channels between reservoir and tapping structure were plugged by concrete.

The most important questions to be answered are as follows: Is the water used for water supply or water bottling; if so, how are quantity and quality to be kept at acceptable levels? Are water losses from the reservoir possible? What is the possible submergence effect on the hydrologic regime in upstream areas? How could spring submergence affect induced seismicity in surrounding areas? In many analyzed cases, spring submergence does not increase considerable environmental consequences.

The Oko Spring is the only water supply source for the town of Trebinje (BiH, 20,000 people). After construction of the Gorica Dam, the tapping structure is above the reservoir level. Three large-diameter wells were drilled into the karst channel situated 25 m deeper than reservoir bottom. Impact of reservoir water to quality of potable water occurs only during extremely fast impounding of the reservoir.

Environmental Aspects of Dams in Karst

The complexity of karst presents a great variety of risk for dams and reservoirs in karst. The crucial role of dams and tunnels in karst is dewatering of temporarily flooded karst depressions for water transfer from one catchment (or political entity) to another and for electricity production. By dam construction in karst regions, some temporarily flooded depressions are changed to permanent reservoirs or, in other cases, to farmland areas. By applying different geotechnical measures, the karstified and pervious riverbeds are transformed into permanent river flows.

Construction of dams and reservoirs in general has considerable influence on regime and quality of surface and underground water downstream. In the case of karst, impacts are sometimes registered at remote springs at distances of 10–30 km, leading to local and transboundary environmental and political problems.

In many cases, the purpose of dam structures is rerouting and transferring of water from one catchment to another. These solutions create conflicts between owners of the dams (reservoirs) and users of the springs. This situation is especially delicate if the reservoir and springs are in different political regions. For example, by construction of Grančarevo and Gorica dams at BiH, the average yearly discharge of Ombla Spring (Croatia) has been reduced from Qav = 33.8 m^3/s to Qav = 24.4 m^3/s. A large part of the water is transferred through the headrace tunnel toward the power plant of Dubrovnik located in Croatia. There was no change in the minimal discharge of Ombla Spring. By construction of dams, the karst aquifer was starved of about 4 billion cubic meters of water annually as a result of rerouting through the tunnels and paved channels for power production and drainage of many swallow holes [24].

Karst underground is very rich with various fauna. Often, as a result of dam construction in karst, a large volume of caverns in the aeration zone are flooded, or temporarily flooded karst channels become permanently dry. In both cases, cave habitat for a number of rare and endemic species is endangered, for example, Normandy Dam (Tennessee, USA), Melones Dam (California, USA), Scrivener Dam (Australia), Grančarevo and Gorica dams (BiH), and Seymareh Dam (Iran).

A specific example is Popovo Polje in BiH (Fig. 5). Flooding under natural conditions before dam construction reached a height of 40 m in the lowest section of the polje; the polje was underwater an average of 253 days and was dry after only 112 days. During maximum flood, 7500 ha were under water. During dry periods the Trebišnjica River was dry also because of 65 m^3/s of seepage along the 65 km of riverbed [25]. The only variety of maize that grows here is called "hundred-day maize." In 100 days, it sprouts, grows, and bears fruit. After dam construction and the increase in impermeability of the riverbed, floods were almost eliminated and huge areas of arable land were created. New infrastructure (plantations, roads, irrigation canals, and settlements) changed complete environmental properties of the entire area.

Dewatering of the temporarily flooded Popovo Polje (BiH) has negative influence for aquatic organisms that inhabit temporary underground lakes during the dry period and ephemeral lakes during flood season. An example is the gaovica fish (*Paraphoxinus ghetaldi*) which spends dry months in numerous siphon lakes and pools of the underground karst. During flood periods, the fish leaves the underground through karst channels and openings of estavelles. For the duration of inundation, the fish lives in the intermittent lakes at the surface. For centuries, fishing at openings of estavelles was an important tradition and food source for inhabitants of Popovo Polje. By construction of two dams (Grančarevo and Gorica dams), the water regime has drastically changed; the gaovica fish lost connections with the surface in most locations and is now threatened with extinction. As a consequence of the same project, the large concentrations of endemic worm *Mariphugia cavatica*, mollusk *Kongeria*, and cave-dwelling aquatic endemic species *Proteus anguinus*, known as "human fish," are seriously endangered [26]. Dam and reservoir construction decreased the activity time and of discharge of a series of temporary and submarine springs along the seacoast. The operation of a commercial oyster and mollusk farm has been threatened because of reduced freshwater outflow through the submarine springs.

One of the richest caves with various fauna in the world is in the same area, Vjetrenica cave. Approximately 110 species have been identified in this cave. More than ten species are known only from this cave or the immediate vicinity [27].

From an environmental point of view, underground dams and reservoirs in karst have benefits. The main advantages of these unique structures are as follows: Arable land is not disturbed, water is not thermally stratified, water quality remains high and constant, catastrophic dam failure is not a possible outcome, and the landscape is not disturbed. In the karst regions of China, more than 20 underground reservoirs have been created with storage capacities between 1×10^5 and 1×10^7 m^3. Their purposes are water supply, irrigation, industry, and electricity production [28]. According to Yuan [29], in the Xiashi district (Guizhou Province) 16 underground dams have been constructed for irrigation of 3624 acres of farmland [29].

To control concentrated infiltration through the large ponors (swallow holes), a specific type of dam is constructed in karst areas – cylindrical dams. The purpose of cylindrical dams is to prevent natural plugging of large ponors, that is, to ensure fast dewatering of floodwater from farmlands. Some cylindrical dams constructed at Peloponnesus (Greece) in fourteenth century are still operational (Fig. 6).

Dam Engineering and Its Environmental Aspects, Fig. 5 Very pervious Trebišnjica riverbed covered by shotcrete. On the hillside is a visible line indicating flood level in natural conditions

Dam Engineering and Its Environmental Aspects, Fig. 6 Ancient dam at Peloponnesus, Greece, to protect natural plugging of large ponor (swallow hole)

In the Dinaric karst area (BiH), cylindrical dams were used from ancient time until the middle of the twentieth century for mills. These dams were constructed above the large ponors at the riverbanks. Water sinking into the ponor propels wooden turbines and millstones. Mills were equipped with simple intake structures and gates to control quantity of water. Recently, cylindrical dams were used to prevent leakage from reservoirs in the case of large ponors or estavelles in Chinese karst and at some other locations.

During filling of Salanfe reservoir (52-m-high dam, Switzerland), new thermal springs appear in the Val d'Illiez valley, at a distance of 8 km (1953). Springs are related to the leakage from Salanfe reservoir. The reservoir has never been filled completely since the first phase of impoundment [30].

The large number of failures and subsequent floods are registered in the case of dams situated in karstified rocks, particularly in evaporites (gypsum, salt). For instance the San Juan reservoir (Spain) failure occurred during test filling (2001). The flood wave of 3000.000 m^3 is transported fare downstream about 3350 tons of solid material (Gutierrez at all 2003).

Reservoirs in karst may fail to fill despite an extensive investigation program and sealing treatment. Dried reservoirs or reservoirs with unacceptable heavy leakage are common in many karst regions of the world: Hales Bar Dam (USA), Montejaque (Spain), Vrtac (Montenegro), Lar (Iran), May (Turkey), Perdika (Greece), Wolf Creek (USA), Apa Reservoir (Turkey), and many others. A distinctive example is the Montejaque Dam in Spain. The dam was abandoned because of huge leakage from reservoir, and cavernous system downstream from the dam is presently used as a training area for speleologists (Fig. 7).

Further Directions

Optimal strategies for water resources development are a key requirement for socioeconomic development. With increasing demands on energy, particularly in many underdeveloped countries, dams and reservoirs are still necessary structures. The environment has been modified by dam construction with possible detrimental impacts. In most instances, impacts are positive: flood control, irrigation, water supply, power production, infrastructure improvement, reduction of deforestation, recreation, fishing, and many secondary benefits. Some negative impacts cannot be avoided: population replacement, inundation of arable land,

Dam Engineering and Its Environmental Aspects, Fig. 7 Montejaque Dam, abandoned due to huge leakage historical and cultural monuments, influence on survival of endemic species and migratory fishes, deterioration of aquifers, the possibility of triggered seismicity, collapses, and similar events.

An important issue is how to keep the balance between the necessity for development and preservation of environment. Optimal environmental protection requires a multidisciplinary approach, including close cooperation of a wide spectrum of geologists, civil engineers, biologists, chemists, hydrologists, hydrogeologists, archeologists, sociologists, and many others. The ultimate aim is identification of crucial parameters that define causes and consequences between human activities (dam construction) and resulting impact on environment (cause-and-effect relations). Criteria for determining environmental protection, as well as regulatory procedures are important elements in the process.

Particularly, sensitive and complex is construction of large dams and reservoirs in highly developed karst because the majority of water flows through the underground karst conduits. Dam impacts on environment in karst may be unpredictable, occur rapidly, and may be unique. Similar situations are seldom, if ever, repeated. To "expect the unexpected" should be permanently kept in the mind as the basic philosophy of dam construction in karst. The major aims of proper planning of water resources systems in karst terrain are to minimize negative and to maximize positive environmental impacts by keeping water at surface level as much as possible.

Bibliography

Primary Literature

1. Liu J (1980) Changjiang (Yangtze River) China's largest river. Foreign Language Press, Beijing
2. Iranian National Committee on Large Dams (1993) A general view on Iranian dams, past-present-future. Ministry of Energy, Teheran
3. Tennessee Valley Authority (1949) Tennessee valley authority projects, geology, and foundation treatment. Technical report no 22, Knoxville, Tennessee
4. Flagg CHG (1979) Geological causes of dam incidents. Bull Int Assoc Eng Geol 20:196–201
5. Cooper AH, Calow RC (1998) Avoiding gypsum geohazards: guidance for planning and construction. British Geological Survey. Technical report WC/98/5
6. Gutierrez F, Desir G, Gutierrez M (2003) Causes of the catastrophic failure of an earth dam built on gypsiferous alluvium and dispersive clays, Altorricon, Huesca Province, NE Spain. Environ Geol 43:842–851
7. Schuster RL (2006) Interaction of dams and landslides. Case studies and mitigation. U.S. Geological survey professional paper 1723, p 107
8. Moon A (1997) Predicting low probability rapid landslides at Roxborough Gorge, New Zealand. In: Ctuden D, Fell R (eds) Landslides risk assessment. Balkema, Rotterdam, pp 272–284
9. Selli R, Trevisan L, Carloni GC, Mazzanti R, Ciabatti M (1946) La frana del Vaiont. Annali del museo geologico di Bolona, Bolongna
10. Lu Y (2001) Rational exploitation of resources and prevention of geohazards in karst regions. Acta geologica sinica (English Edition). J Geol Soc China 75(3):239–248
11. Chronology of major tailings dam failure, 1960–2011 (prepared on the basis of) Bulletin 121, Published by United Nation Environmental Programme Division of Technology, Industry and Economics and International Commission on Large Dams, Paris 2001, p 144, updated March 2011, (221 tailings dam incidents)
12. Rico M, Benito G, Salgueiro AR, Diez-Herrero A, Pereira HG (2008) Reported tailings dam failures. A review of the European incidents in the worldwide context. J Hazard Mater 152:845–852. Elsevier
13. Bozovic A, Wieland M (2004) Reservoirs and seismicity, state of knowledge. ICOLD committee on seismic aspects of dam design. Commission Internationale des Grands Barrages, Haussmann, Paris
14. Kiernian K (1988) Human impacts and management responses in the karsts of Tasmania. In: Proceedings of the international geographical union, study group Man's impact on karst, Sydney
15. Smith D (2000) Protest grow over plan for more Turkish dams. National Geographic News
16. Djordjević B, Dašić T (2007) Ecological guarantied discharge downstream from the hydropower plants. J Elektriprivreda 1. Belgrade, 5–13
17. Larinier M (2001) Environmental issues, dams and fish migration. In: Marmulla G (ed) Dams, fish and fisheries: opportunities, challenges and conflict resolution. FAO fisheries technical paper, 419. Food and Agriculture Organisation of the United Nations, Viale dell Terme di Caracalla, Rome
18. Linlokken A (1993) Efficiency of fishways and impact of dams on the migration of Grayling and Brown Trout in the Glomma river system, South-Eastern Norway. Regul Rivers: Res Manage 8:145–153. Wiley
19. Lamoreaux PE, Newton JG (1986) Catastrophic subsidences: an environmental hazard, Shelby County, Alabama. Environ Geol Water Sci 8(1/2):25
20. Hua SL (1987) Pumping subsidences of surface in some karst areas of China. In: Symposium on human influence on karst, Postojna

21. Tolmachev V, Ilyin A, Gantov B, Leonenko M, Khomenko V, Savarenski I (2003) The main results of engineering karstology research conducted in Dzerzinsk, Russia (1952–2002). In: Beck B (ed) Sinkholes and the engineering and environmental impacts of karst. Geotechnical special publication no 122. Alexander Bell Drive, Reston
22. Gutierrez F (2010) Hazards associated with karst. In: Alcantra-Ayala I, Goudie A (eds) Geomorphological hazards and disaster prevention. Cambridge University Press, Cambridge
23. Maximovich NG (2006) Safety of dams on soluble rock (The Kama hydroelectric power station as an example). On Russian. The Russian Federal Agency for Science and Innovations, Perm
24. Milanović P (2004) Water resources engineering in karst. CRC Press, Boca Raton
25. Milanović P (1981) Karst hydrogeology. Water Resources, Colorado
26. Milanović P (2002) The environmental impacts of human activities and engineering constructions in karst regions. Epizodes J Int Geosci 25(1):13–21
27. Sket B (2003) Vjetrenica Cave. In: Lučić I, Sket B (in Croatian), Zagreb-Ravno
28. Lu Y (1986) Some problems of subsurface reservoirs constructed in karst regions of China. Institute of Hydrogeology and Engineering Geology, Beijing
29. Yuan D (1990) The construction of underground dams on subterranean streams in South China karst. Institute of Karst Geology, Guilin, p 62
30. Roth P (1994) Reservoir induced earthquakes and thermal springs in valais. Proseis AG, Zurich

Books and Reviews

Abraham S (2008) Determining ecological baseflow needs for stream and springs in arid and semi-arid regions. Geol Soc Am, Abstracts with programs 40(6):87

Arthur HG (1977) Teton dam failure. In: The evaluation of dam safety: engineering foundation conference proceedings. ASCE, New York, pp 61–71

Al-Ansari N, Adamo N, Issa IE, Sissakian VK, Knutsson S (2015) Mystery of Mosul dam the most dangerous dam in the world: dam failure and its consequences. J Earth Sci Geotech Eng 5(3):95–111. Scienpress Ltd

Avakyan AB, Romashkov EG, Chestnaya (1979) Fishways and dams (Trans: Gidrotechnickoe Stroitelstvo, 1970). Power Techn Eng 4(11)

Bergkamp G, McCartney M, Dugan P, Mcneely J, Acreman M (2000) Dams, ecosystem functions and environmental restoration: thematic review II. I World Commission on Dams, Cape Town

Bizer JR (2000) Avoiding, minimizing, mitigating and compensating the environmental impacts of large dams. IUCN, Gland

Bozovic A (1974) Review and appraisal of case histories related to seismic effects of reservoir impounding. Eng Geol 8(1–2):9–27

Chandler RJ, Tosatti G (1995) The Stava tailings dam failure, Italy, July 1985. Proc Inst Civ Eng 113:67–79

Cogels FH, Coly A, Niang A (1977) Impact of dam construction on the hydrological regime and quality of a Sahelian Lake in the River Senegal Basin. Regul Rivers Res Manag 13(1):27–41

Collier M, Webb RH, Schmidt JC (1996) Dams and rivers: a primer on the downstream effects of dams. US Geological survey circular 1126. US Geological Survey, Tuscon

Djordjevic B (1990) Cybernetics in water resources management. WRP, Highlands Ranch

Dobry R, Alvarez L (1967) Seismic failures of Chilean tailings dams. J Soil Mec Found Div 93:237–260

Fell R, Bowles DS, Anderson LR, Bell G (2010) The status of methods for estimation of the probability of failure of dams for use in quantitative risk assessment. In: International commission on large dams, 20th congress, Beijing

Ford D, Williams P (2007) Karst hydrogeology and geomorphology. Wiley, Chichester

Fourie AB, Papageorfiou G, Blight GE (2000) Static liquefaction as an explanation for two catastrophic tailings dam failures in South Africa. In: Tailings and mine waste 2000, proceedings of the seventh international conference on tailings and mine waste 2000, Fort Collins, 23–26 Jan 2000, Balkema, Rotterdam, pp 149–158

Gupta HK (1992) Reservoir-induced earthquakes. Elsevier, Amsterdam

Gutierrez F, Desir G, Gutierrez M (2003) Causes of the catastrophic failure of an earth dam built on gypsiferous alluvium and dispersive clays, Altorricon, Huesca Province, NE Spain. Environ Geol 43:842–857

Guzina BJ, Sarić M, Petrović N (1991) Seepage and dissolution at foundations of dam during the first impounding of the reservoir. Commission Internationale Des Grandes Barrages, Vienne

Helms SW (1981) Jawa, lost city of the black desert. Methuen and Co Ltd

Holmgren R (2000) Experiences from the tailing dam failure at the boliden Atik and legal consequences, international report, County Administration of Norrbotten

ICOLD Bulletin 99 (1995) Dam failure, statistical analysis. Commission Internationale des Grands Barrages – 151, Paris

Jackson CD, Marmulla G (2001) The influence of dams on river fisheries. In: Marmulla G (ed) Dams, fish and fisheries: opportunities, challenges and conflict resolution. FAO, Fisheries technical paper, 419. Viale delle Terme di Caracalla, Roma

Jobin WR (1999) Dams and disease: ecological design and health impacts of large dams, canals, and irrigation systems. Taylor & Francis, London. ISBN: 0419223606 (2)

Keffer ML, Bjornn TC, Peery CA, Tolotti KR, Ringe RR, Stuehrenberg L (2003) Adult spring and summer Chinook salmon passage through fishways and transition pools at Bonneville, McNary, Ice. A report for project MPE-P-95-1. U.S. Geological Survey, Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho, Moscow, Idaho 83:844–1141

Lafitte R (2001) Ethics and dam engineers. Int J Hydropower Dams 4:58–59

Lamoreaux PE (1989) Water development for phosphate mining in a karst setting in Florida – a complex environmental problem. Environ Geol Water Sci 14(2): 117–153

Liu JK, ZT Y (1992) Water quality changes and effects on fish population in the Hanjiang River, China, following hydroelectric dam construction. Regul Rivers: Res Manage 7(4):359–368

McCartney M (2009) Living with dams: managing the environmental impacts. Water Policy 11(Supp 1): 121–139. IWA Publishing

McDonald MJ, Muldowny J (1982) TVA and the dispossessed: the resettlement of population in the Norris dam. University of Tennessee Press, Knoxville

McKartney MP (2007) Decision support systems for large dam planning and operation in Africa. Working paper 119. International Water Management Institute, Colombo

Nations Environmental Programme (UNDP) Division of Technology, Industry and Economics (DTIE) and International Commission on Large Dams (ICOLD) (2001) Risk of dangerous occurrences, lessons learnt from practical experiences. In: Bulletin, vol 121. Nations Environmental Programme (UNDP) Division of Technology, Industry and Economics (DTIE) and International Commission on Large Dams (ICOLD), Paris, p 144

Odum EP (1969) The strategy of ecosystem development. Science 164:262–270

Perman RC, Packer D, Coppersmith KJ, Kneupfer PL (1983) Collection of data bank on reservoir-induced seismicity. Woodward-Clyde Consultants, Walnut Creek. USGS contract no 14-08-0001-19132

Petrović P (2010) "The Kosheish Dam" has it really existed or not? Vodoprivreda no 246–248, Belgrade

Shen CK, Chen HC, Chang CH, Huang LS, Li TC, Yang CY, Wang TC, Lo HH (1973) Earthquakes induced by reservoir impounding and their effects on Hsinfengkiang dam. Sci Sinica 17(2):232–272. Beijing, China

Simpson DW (1988) Two types of reservoir induced seismicity. Bull Seismol Soc Am 78:2025–2040

Tanaka E (2007) Protecting one of the best roman mosaic collections in the world: ownership and protection in the case of the roman mosaics from Zeugma Turkey. Stanford J Archeol 5:183–202

Towhata I et al. (2011) On gigantic Tohoku Pacific earthquake in Japan. Earthquake News, Bulletin of International Society of Civil Engineering, 52

United Nations Environment Programme (1996) Environmental and safety incidents concerning tailings dams at mines: results of a survey for the years 1980–1996 by mining journal research services. United Nations Environment Programme, Industry and Environment, Paris, p 129

USCOLD (1997) Reservoir triggered seismicity. In: The sixteenth USCOLD annual meeting, LA. US Society of dams. Debver

U.S. Army Corps of Engineers (1996) Harbour, and lower granite dams. Technical report 2003–5. U.S. Army Corps of Engineers, Portland and Walla Walla District

U.S. Committee on Large Dams – USCOLD (1994) Tailings dam incidents. U.S. Committee on Large Dams – USCOLD, Denver, 82 p. 1-884575-03-X

Van Niekerk HJ, Viljoen MJ (2005) Causes and consequences of the Merriespruit and other tailings-dam failures. Land Degrad Dev 16:201–212

Vick SG (1990) Planning, design, and analysis of tailings dams. Bitech, Vancouver

Wieland M Dam safety aspects of reservoir-triggered seismicity. In: Wieland M, Ren Q, Tan SY, Taylor, Francis (eds) Book new developments in dam engineering. pp 95–100

Wieland M (2014) Seismic hazard and seismic design and safety aspects of large dam projects. In: Second European Conference on Earthquake Engineering and Seismology, Istanbul

Yuan D (1991) Karst in China. Geological Publishing House, Beijing

Dredging Practices and Environmental Considerations

Craig Vogt[1] and Greg Hartman[2]
[1]Craig Vogt Inc, Environmental Consultants, Hacks Neck, VA, USA
[2]Hartman Associates Inc, Waterway Engineering and Sediment Remediation, Redmond, OR, USA

Article Outline

Glossary

Beneficial use Placement or use of dredged material as resource materials in productive ways, which provide environmental, economic, or social benefits.

Confined disposal facility (CDF) An engineered structure for containment of dredged material consisting of dikes or other structures that enclose a disposal area above any adjacent water surface, isolating the dredged material from adjacent waters during placement. Other terms used for CDFs that appear in the literature include "confined disposal area," "confined disposal site," and "dredged material containment area."

Confined aquatic disposal (CAD) This is a form of dredged material disposal that involves controlled placement of dredged material into a subaqueous site with some form of lateral confinement. The lateral confinement may be provided by a bottom depression or by subaqueous berms. The contaminated material is then capped in most instances with clean sediment to physically separate it from the overlying environment (commonly called a CAD cell).

Contaminant A chemical or biological substance in a form that can be incorporated into, onto, or be ingested by or harm aquatic organisms, consumers of aquatic organisms, or users of the aquatic environment. Contaminated sediment or contaminated dredged material.

Contaminated sediments or contaminated dredged materials Those that may cause an unacceptable adverse effect on human health or the environment.

Dredged material Material excavated from fresh, estuarine, or ocean waters. The term "dredged

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_438

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_438-3

material" refers to material, which has been dredged from a water body and disposed in a disposal site. The term "sediment" refers to material on the bed of a water body prior to the dredging process.

Dredging Underwater excavation is called dredging. "Dredging" is the term given to removal by digging, gathering, or pulling out materials from the bed to deepen waterways and to create harbors, channels, and berths. Dredging is also conducted for construction purposes, for mining, and for environmental cleanup and enhancement.

Habitat The specific area or environment in which a particular type of plant or animal lives. An organism's habitat provides all of the basic requirements for the maintenance of life. Typical coastal habitats include beaches, marshes, rocky shores, bottom sediments, mudflats, and the water itself.

Open-water disposal Placement of dredged material for the purpose of disposal in rivers, lakes, estuaries, or oceans via pipeline or surface release from hopper dredges or barges, without confinement.

Placement of dredged material In this document, placement of dredged material is considered to be for another purpose other than disposal, such as placement on a beach for beach nourishment or placement on a wetland to enhance the ecological functioning.

Sediment Material, such as sand, silt, or clay, suspended in or settled on the bottom of a water body. Sediment input to a body of water comes from natural sources, such as erosion of soils and weathering of rock, or as the result of anthropogenic activities, such as forest or agricultural practices, or construction activities. The term "dredged material" refers to material which has been dredged from a water body, while the term "sediment" refers to material in a water body prior to the dredging process.

Suspended solids Organic or inorganic particles that are suspended in water. The term includes sand, silt, and clay particles as well as other solids, such as biological material, suspended in the water column.

Toxicity Level of mortality or other end point demonstrated by a group of organisms that have been affected by the properties of a substance, such as contaminated water, sediment, or dredged material.

Toxic pollutant Pollutants, or combinations of pollutants, including disease-causing agents, that after discharge and upon exposure, ingestion, inhalation, or assimilation into any organism, either directly from the environment or indirectly by ingestion through food chains, will cause death, disease, behavioral abnormalities, cancer, genetic mutations, physiological malfunctions, or physical deformations in such organisms or their offspring.

Turbidity An optical measure of the amount of material suspended in the water. Increasing the turbidity of the water decreases the amount of light that penetrates the water column. Very high levels of turbidity can be harmful to aquatic life.

Definition of Subject

Underwater excavation is called dredging. Dredging is the term given to removal by digging, gathering, or pulling out materials from the bed to deepen waterways and to create harbors, channels, and berths. Dredging is also conducted for construction purposes, for mining, and for environmental cleanup and enhancement. The complete dredging activity includes sediment excavation and removal from the bed and transport from the dredging site to a disposal area or placement site, which is located in either an open-water, nearshore, or upland location.

Operations that cause potential environmental impacts associated with the dredging process include (1) the sediment removal process from submerged excavation at the point of dredging and (2) the transport and placement for disposal of the dredged material. Environmental concerns relate to the location of the sediment removal by dredging and the disposal or placement site. General environmental considerations include:

- Physical and ecological impacts due to sedimentation and disposal (i.e., smothering)
- Turbidity in the water
- Resuspension of contaminants
- Acute and chronic toxicity due to chemicals in the dredged material
- Impacts to marine animals, including endangered species
- Impacts to shore-based wildlife from exposure pathways from shoreline and upland disposal
- Impacts to and loss of habitat
- Air pollutant emissions from the dredge, including greenhouse gases
- Other general construction issues, such as underwater and surface equipment noise
- Quality of life issues (e.g., noise, night lights)

Introduction

Dredges of various designs have been used for many years to create and maintain navigable waterways to move people, goods, and materials. It is theorized that thousands of years ago, blocks of stone that make up the pyramids in Egypt were barged from a distant quarry through a dredged canal. At that time, the canals were likely dredged using a barge with people using long-handled dipper shovels to raise solids out of a waterway and then place those solids on a haul barge deck for disposal elsewhere. Productivity gains likely came about when animal power was used to increase the digging power of early dredges. The late 1800s saw the development of electric and steam power units (Fig. 1), which enabled the construction of huge mechanical dredges with bucket ladders, backhoe dredges, and pipeline dredges with centrifugal pumps. Hydraulic technology made great advancements in the 1960s, with the result that hydraulic winches and hydraulic rotary cutter drives using suction of bottom sediments became a welcome replacement facilitating the removal of finer-grained sediment (compared to clunky and inefficient mechanical drives) [1].

Today, the dredge type can be hydraulic or mechanical and can be used for a multitude of purposes and projects. The primary purposes are navigation, environmental enhancement, and mining/construction [2].

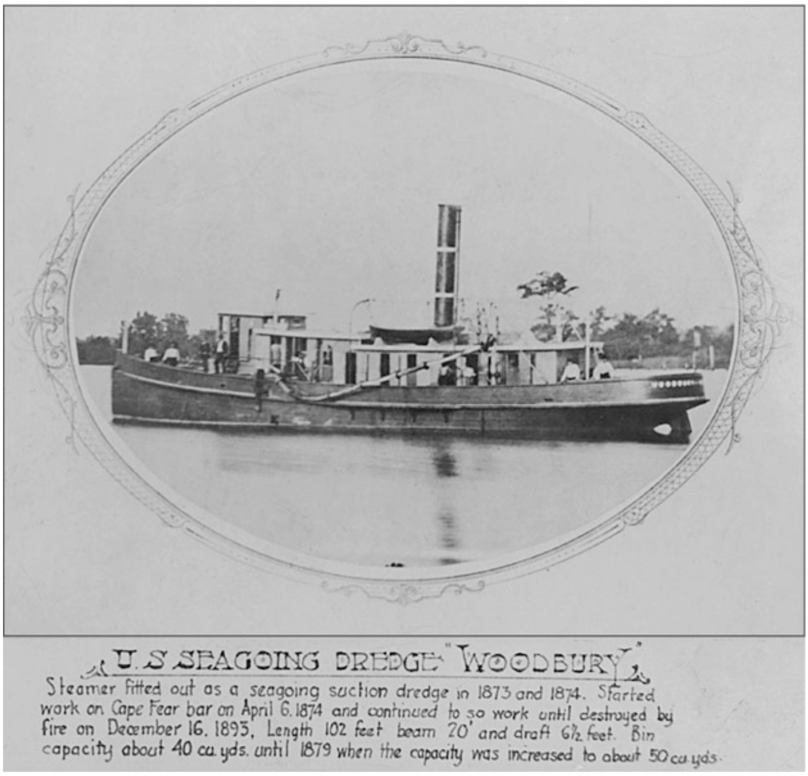
Navigation Dredging

Most coastal and river ports, harbors, and navigation channels are not naturally deep enough or wide enough to support safe passage of vessels. Navigation channels need to be dredged to create waterway channels with adequate channel area, depth, and access to port and harbor facilities. Nearly all the major ports in the world have at some time required dredging to deepen and widen the access channels, to provide turning basins, and to achieve appropriate water depths to and from shoreline facilities, such as cargo or cruise ship terminals.

Virtually, all of the navigation channels created in rivers and harbors have and continue to require maintenance dredging, i.e., the removal of sediments which naturally accumulate on the bottom of the dredged channel from upstream erosion. Navigation channel dredging can be categorized as two types. (1) New work dredging (also termed capital dredging) is the initial dredging conducted to excavate a channel with navigable depths greater than those that naturally exist. (2) Maintenance dredging is the dredging which removes accumulated sediments from dredged channels and ensures that the channel continues to provide adequate dimensions for vessels engaged in domestic and international commerce as well as for other types of vessels, such as recreational boating and commercial fishing.

Environmental Enhancement Dredging

In the last three decades, dredging has been successfully used to remove contaminated sediments from waterways, with the intention of improving water quality and restoring the health of aquatic ecosystems. Cleanup dredging (also termed environmental dredging) for removal of contaminants is used in waterways, lakes, ports, and harbors, usually in highly industrialized or urbanized areas that are suffering from past toxic waste and wastewater disposal practices. After removal from the bed, the contaminated sediments are transported and disposed under strict environmental controls (e.g., lined upland confined disposal facilities). In



Dredging Practices and Environmental Considerations, Fig. 1 US seagoing suction dredge "Woodbury" 1873 (Photo courtesy of Corps of Engineers)

some cases, the contaminated sediments may be treated, and some or all of the sediments are used for beneficial objectives. Under proper conditions, a viable alternative to removal is in situ isolation, i.e., the placement of a cap (i.e., a cover of clean material or amended materials to enhance isolation) over the contaminated sediments in their original location. In some situations, monitored natural recovery is an approach that allows natural forces to decrease exposure to contaminated sediments to acceptable levels.

Reclamation, Mining, and Construction Dredging

Dredging is an integral tool in many types of water-related construction projects, such as emplacement of pipelines or immersed tunnels, underwater foundations, creating new land (i.e., reclamation), and maintaining storage capacity in water supply and recreational reservoirs. In addition, dredging is important in mining activities, with a primary use to provide sand and gravel for construction and reclamation projects and a more recent role to remove seabed minerals from the deep sea, focused upon mining for gold, copper, manganese, nickel, lead, cobalt, lithium, titanium, platinum, and zinc.

Dredged aggregates have a wide range of uses, including:

- Land reclamation: Pressures arising from population growth have created a need to raise the elevation of low-lying areas for port and infrastructure development and/or to construct new land areas. Such pressures are likely to continue, but loss of marine habitat is an opposing issue when creating new land.
- Construction materials: An increasing quantity of aggregate mined from marine and freshwater borrow sites is used in concrete and fill construction.

Deep sea mining targets seabed polymetallic nodules, rich in concentrations of manganese, nickel, copper, and cobalt. They are found in abundance in the Clarion-Clipperton Zone (CCZ), a great abyssal plain as wide as the continental United States that lies 4,000 to 6,000 meters below the surface of the eastern Pacific Ocean. There, millions of the potato-sized nodules are scattered on top of or half-embedded within the muddy bottom of the CCZ. Their exploitation would probably involve scraping 5–10 cm (2–4 in.) off the top of the abyssal plain, separating the nodules from the mud, pumping the nodules to a surface ship by means of a giant tube, and returning the water and ne particles through another tube. Dredging or some modifications of dredging technologies are likely to be one of the approaches used in deep sea mining [3].

Dredging and Dredged Material Management

Dredging and dredged material management consists of the following three elements:

1. Excavation: The dislodgement and removal of sediments (clay, silt, sand, gravel, and rock) from the bed of the water body by a dredge, either mechanically, hydraulically, or by combination of the two dredging methods.
2. Transport: The transport of excavated material from the point of dredging to the final disposal site. This can be accomplished by haul barges separate from the dredge equipment, by a dredge equipped with hoppers, by pipeline from the dredge to the disposal or placement site, or, in the case of environmental dredging, by hauling by truck or rail.
3. Disposal or placement: The final disposal or placement of dredged material. Whether dredged material is disposed or placed (and reused for another purpose, such as creation of a wetland) is determined by a range of factors. The disposal site factors to be considered include sediment type to be dredged (e.g., grain size), location of the dredging project versus the disposal site or beneficial use site, future disposal site utilization, physical and chemical characteristics of the sediment (e.g., is it contaminated?), and available funding.

Dredged material is a resource having a wide number of beneficial use applications that must be considered in dredged material disposal management. This includes consideration for beach nourishment (e.g., replacing lost sand due to erosion to widen beaches), shoreline fill, and habitat creation or restoration. This concept as a resource is also an important environmental consideration for environmental dredging. Dredged material from environmental dredging is usually placed into well-controlled upland confined disposal facilities. Clean dredged material is often used to cap, cover, and isolate contaminated sediment in waterways.

Dredging

While dredging equipment varies widely in many sizes and types, dredging is actually accomplished basically by only two dredge types. They are mechanical dredges and hydraulic dredges. The type of dredge is derived from the method of sediment capture and removal from the bed.

Selection of dredging equipment and the methods used to perform the dredging depends on the following factors [4]:

- Physical characteristics of material to be dredged
- Quantities of material to be dredged
- Depth of material to be dredged
- Method of disposal or placement
- Distance to disposal or placement site
- Physical environment of the dredging area(s)
- Physical environment of the disposal area(s)
- Level of contamination of the material to be dredged
- Dredge production capability
- Type of dredges available
- Time, environmental, and economic limits of the project

Mechanical Dredges

Mechanical dredges remove bottom sediment through the direct application of mechanical force to dislodge and excavate the material at in situ densities. The mechanical dredges (Fig. 2) are well suited to removing hard-packed material or debris and to working in confined areas, such as in environmental cleanup dredging. Cohesive sediments that are mechanically dredged usually remain intact, with large pieces retaining their in situ density and structure through the dredging and placement process. Sediments excavated with a mechanical dredge are generally placed into a haul barge or scow for transportation from the dredging site to the disposal or placement site.

Mechanical dredges use some form of bucket to excavate and lift the dredged material from the bottom, then load it on to a haul barge or scow. When the haul barge is loaded, a tug or other attendant vessel will take the barge to the disposal or placement site. Mechanical dredges are classified by how the bucket is connected to the dredge. The three standard classifications include:

- Wire rope connected – barge-mounted crane (clamshell or cable bucket)
- Structurally connected – articulated fixed-arm dredge (e.g., backhoe, excavator)
- Chain and structurally connected – bucket ladder

A mechanical dredge is often labeled a clamshell dredge. The clamshell is actually a type of cable-connected bucket used as the digging part of the overall mechanical dredge and is the most common type of the mechanical dredges. Mechanical dredges that have the bucket with cable connection to the barge-mounted crane use a number of different bucket designs for differing sediment characteristics, such as mud, gravel, rock, or boulders.

An articulated fixed-arm mechanical dredge can be a back-acting (backhoe) excavating machine, an advance cut (dipper, excavator) excavating machine, or a bucket (grab) excavating machine.

A bucket ladder dredge is the oldest and the most common type of mechanical dredge. It is primarily used for mining applications. A bucket ladder dredge consists of a large number of buckets linked together in an endless chain, which is carried on a ladder that is raised and lowered by hoisting wires. The buckets dig into the face of the cut, and the sediment is carried up the ladder and dumped onto a conveyor belt, which conveys the dredged material to a waiting barge.

Dredging for environmental cleanup requires much greater precision than navigation dredging

Dredging Practices and Environmental Considerations, Fig. 2 Mechanical backhoe dredge, New York (Courtesy of Great Lakes Dredge & Dock Company) and can be accomplished using articulated fixed-arm mechanical dredges, which are similar to conventional upland excavators placed on a barge. The rigid arm, as compared to the cable-connected bucket, provides greater positioning control in placing the bucket on the bottom. Bucket dredges hanging from a crane on board a barge that are designed for a level cut and equipped to be enclosed after the cut are also effective in environmental dredging. These buckets (Fig. 3) minimize the leakage of water and contaminants during the excavation and placement of the contaminated material on the barge for transport [5].

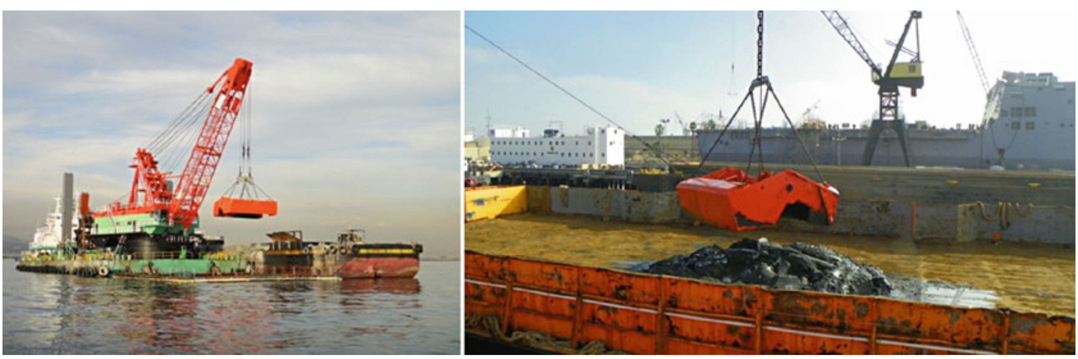
The dragline bucket dredge is similar to the clamshell dredge. They are effective in excavation of gravels, sand, and compact silts. The dragline bucket is not an enclosed bucket and can cause significant turbidity in the water column. Dragline buckets can be operated as a dredge from shore or mounted on a barge. The bucket does not load vertically. Instead, it is lowered to the bottom and then loaded by dragging toward the crane. A dipper dredge is a floating face shovel that digs forward into the face of the excavation and is mounted on a spud barge.

Hydraulic Dredges

Hydraulic dredges are identified by two primary types. They are the pipeline-cutterhead dredge and the trailing suction hopper dredge. The hydraulic dredge works by dislodging bed sediment and hydraulic removal of the sediment from the bed of the waterway by suction pipe.

The hydraulic pipeline dredge is generally comprised of the following equipment:

- Cutterhead. The pipeline dredge has an active cutterhead (Fig. 4) that rotates and dislodges the sediment from the bed. This allows the suction, created at the cutterhead by the suction pipe and pump, to be captured and pulled up the suction pipe.
- Suction pipe. The pipe that connects the cutterhead on the bed with the centrifugal pump in the dredge hull.
- Ladder. The ladder raises and lowers the suction pipe and the cutterhead.
- Barge hull. The pipeline dredge is a specialty barge that supports the machinery, pump cable winches, and other equipment in the hull with a lever room, anchor booms, and "walking" spuds on the hull.
- Discharge pipe. The discharge pipe transports the dredged material slurry from the pump to the disposal site.
- Dredge pump(s). The dredge pumps are large centrifugal pumps located on the dredge hull at the waterline and/or on the ladder near the suction mouth.
- Spuds/anchor wires. The spuds and anchor wires are used to anchor the dredge, swing the dredge across the cut, and move the dredge forward.



Dredging Practices and Environmental Considerations, Fig. 3 Mechanical dredges: environmental closed buckets (Courtesy of Cable Arm Company)

Dredging Practices and Environmental Considerations, Fig. 4 Typical hydraulic cutterhead dredge (Courtesy of Ellicott Dredges Company)

Dredging Practices and Environmental Considerations, Fig. 5 Cutterhead pipeline dredge (CSD), Texas (Courtesy: Great Lakes Dredge & Dock)

The pipeline dredge (Fig. 5) is not self-powered. It moves through the cut using the "walking" spud and then the working spud for dredging, thereby allowing the dredge to move forward as it swings the cutterhead from left to right and return.

The pipeline dredge size is classified by the size of the discharge pipe inner diameter. These dredges use hydraulic centrifugal pumps to provide the lifting force to capture and transport via pipeline the dredged material in a liquid and solid slurry, usually about 10–20% solids (dry weight). Pipeline dredges usually work well in loose, "unconsolidated" silts, sands, gravels, and soft clays. In dense consolidated sediments, the hydraulic pipeline dredge depends on an active cutterhead and/or waterjets that can break up the consolidated material at the mouth of the suction pipe.

The centrifugal pump on the hydraulic dredge operates at or near the water surface elevation to pull water into the suction mouth at the cutterhead. The combination of the vacuum and atmospheric weight acts to move the bed material

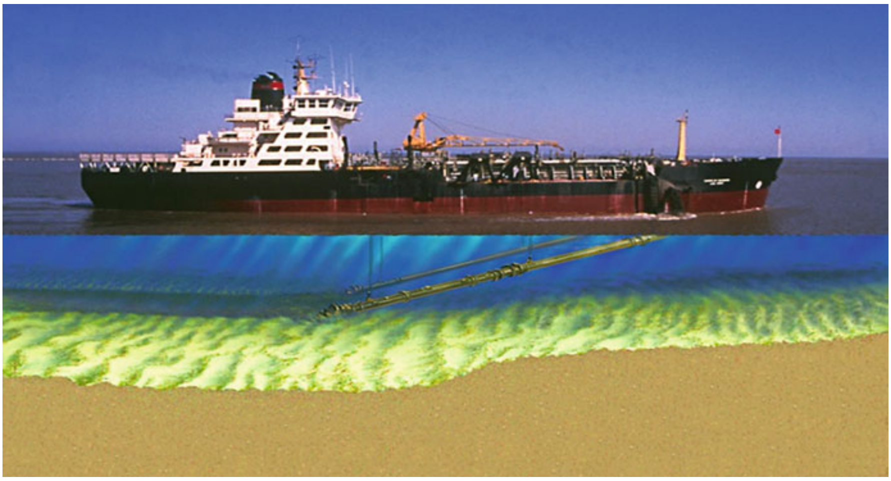
Dredging Practices and Environmental Considerations, Fig. 6 Pipeline booster pump (Courtesy: Great Lakes Dredge & Dock)

up through the suction pipe to the pump, and then the pump discharges the slurry into the discharge pipeline and directly to the disposal site. Because pipeline dredges pump directly to the disposal site, they operate continuously and can be very cost-efficient. A booster pump is used for long distances to the disposal site (Fig. 6).

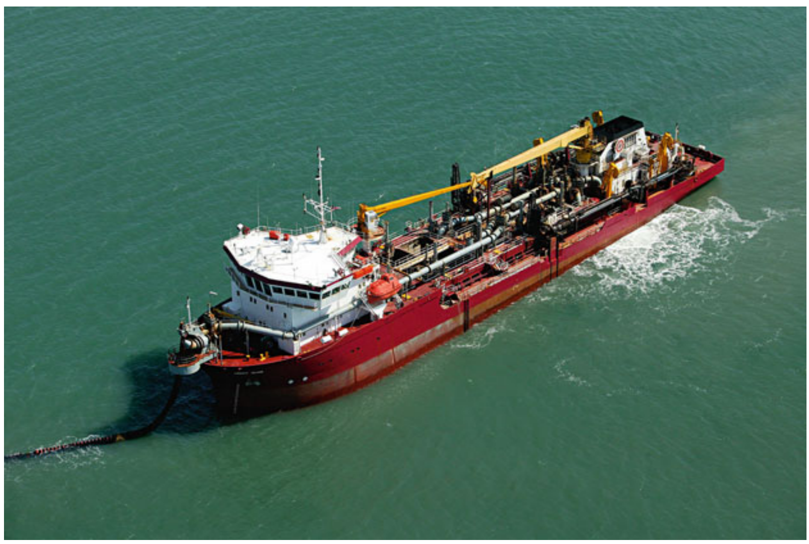
Cutterhead pipeline dredges work efficiently in large areas and in water depths up to 60–70 ft. The dredge production rate is the most efficient and creates minimum excess turbidity at the cutterhead when the loose sediment cut depth is equal or near the diameter of the dredge cutterhead. The dredged material slurry typically contains about 90% water and 10% solid volume, and the water must be contained in the disposal site until the solids settle out. The dewatering is accomplished typically by discharging surface water in the disposal site back into the waterway. If it is contaminated sediment, discharge controls are required to minimize impacts to the receiving waters.

The discharge pipe can be a floating pipe on the surface or a submerged discharge pipe. Typically, the discharge line is a floating line, and they are not well suited for work in rough seas, where lines can be broken apart, or in high-traffic areas, where the discharge pipeline can be an obstruction to navigation. The pipeline dredge is not an efficient dredge plant to work in an open estuary, entrance channel, or other open-water areas where significant wave conditions may occur. Rough waters with wave heights of significance will cause damage to the cutterhead, spuds, and the dredge hull, as well as the floating pipeline.

Hopper dredges are ships designed for dredging (Figs. 7 and 8). The trailing suction hopper dredge is a self-propelled seagoing ship equipped with a suction pipe, which trails over the side of the vessel or through a well in the hull. The sediment and water slurry is transported through the pumps just as the pipeline dredge, but when the sediment and water slurry passes through the pump to the discharge pipeline, it is discharged immediately into the hoppers of the dredge. The sediments settle out in the hopper, while excess water is separated and discharged over the weir structure. It can be advantageous to overflow hopper dredges to increase the amount of sediment in the hopper; however, this is not always acceptable due to water quality concerns near the dredging site. When the hoppers are full, the sediment and water slurry are transported by the ship to the disposal site.



Dredging Practices and Environmental Considerations, Fig. 7 Typical hopper dredge (Photo courtesy of Corps of Engineers)



Dredging Practices and Environmental Considerations, Fig. 8 Hopper dredge, Liberty Island (Photo courtesy of Great Lakes Dredge & Dock Company)

A hopper dredge can dispose of the dredged sediment in two manners. The split-hull hopper dredge design and the standard hopper dredge with bottom dump capability will discharge the dredged sediment to a submerged and acceptable disposal site. The alternative to bottom dumping is to pump out of the hoppers to an upland disposal site or to a shallow water area, for which the vessel draft prevents access.

Hopper dredges are well suited to dredging sandy sediments. They can maintain operations in relatively rough seas, and because they are mobile, they can be used in high-traffic areas. They are often used at ocean entrances where wave conditions prevent the use of pipeline dredges and limit mechanical dredging capability. They cannot be used efficiently in confined or shallow areas. Hopper dredges can move quickly to disposal sites under their own power. However, because the actual removal of sediment from the bed stops during the transit to and from the disposal area, the operation loses efficiency that can become critical when the haul distance to disposal is far.

There is a special hopper dredge class called side casters and pipeline dredges called dustpan dredges. Both of these dredges are unique within their family of dredges. They are specialty dredges used for unusual waterway conditions. They work best when dredging loosely compacted, coarse-grained, clean sediment with disposal in areas close to the dredging activity. They are not widely used.

Environmental Cleanup Dredges

Dredging of contaminated sediments is potentially very harmful to the local environment during dredging and disposal. Contaminants can be remobilized and/or released into the water column where they can detrimentally affect aquatic life and pose a risk to human health. Technological advances have fostered modification of existing dredge equipment, and creation of new dredging equipment to address the environmental issues. Contaminated sediment dredging focuses on minimizing suspension and release of problem sediments in the water column while increasing the precision of dredging to reduce overdredging. Examples of environmental cleanup dredges include the following:

- Encapsulated bucket lines for bucket chain dredges
- Closed buckets for backhoes
- Closed clamshells for grab dredges [5]
- Auger dredges, disk cutter, scoop dredges, and sweep dredges (all modified cutter dredges)

Selection of Dredging Equipment

Selection of dredging equipment and the methods used to perform the dredging depends on the following factors [6]:

- Physical characteristics of material to be dredged
- Quantities of material to be dredged
- Depth of material to be dredged
- Method of disposal or placement
- Distance to disposal site or staging area
- Physical environment of the dredging area(s)
- Physical environment of the disposal area(s)
- Level of chemical contamination of the material to be dredged
- Dredge production capacity

Transportation of Dredged Material

Transportation methods generally used to move clean and contaminated dredged materials are included in the three basic dredge types: pipelines, barges or scows, and hopper dredges.

- Pipeline transport is the method most commonly associated with cutterhead, dustpan, auger head, and other hydraulic dredges. Dredged material may be directly transported by hydraulic dredges through pipelines for distances of up to several miles, depending on a number of conditions. Longer pipeline pumping distances are feasible with the addition of booster pumps, but the cost of transport greatly increases proportionally with each booster pump added to the discharge line.
- Barges and scows, used in conjunction with mechanical dredges, have been one of the most widely applied methods of transporting large quantities of dredged material over long distances.
- Hopper dredges are capable of transporting the material for long distances in a self-contained hopper. Hopper dredges normally discharge the material from the bottom of the vessel hull by opening the hopper doors; however, most hopper dredges are equipped to pump out the material from the hopper and deliver the sediment much like a hydraulic pipeline dredge.

Dredged Material Disposal and Placement Alternatives

Evaluation and design of a proposed dredging project involves comprehensive assessment of alternatives for disposal or placement of the dredged material. Identification of the specific disposal site or beneficial use site involves a number of different considerations, including environmental, technical, and economic factors.

Three major disposal/placement alternatives are available:

- Open-water disposal at designated disposal sites in coastal waters, estuaries, rivers, or lakes, outside the navigation channel
- Confined disposal in open water (confined aquatic disposal (CAD)), confined disposal in a CDF along the shoreline, and confined disposal at an upland CDF
- Placement for environmental and beneficial use

Depending upon local factors, in certain cases of contaminated sediments and environmental dredging, treatment of the dredged material after temporary storage and before final disposal in an upland CDF may be a necessary approach.

Open-Water Disposal

Open-water disposal means that dredged material is placed at designated sites in oceans, estuaries, rivers, and lakes such that it is not isolated from the adjacent water. Clean dredged materials are the only acceptable dredged materials for disposal at open-water disposal sites. The determination that dredged material is "clean" is based upon a series of chemical and biological tests, the results of which must meet national environmental regulations. The disposal of contaminated material can be considered for open-water disposal but only with appropriate control measures, such as capping the contaminated sediments with clean materials to ensure long-term isolation from the surrounding environment.

The objective of capped in-water disposal is to isolate contaminated materials from the environment by covering the contaminated materials with clean materials, such as fine to coarse sand. The contaminated material is placed on a level bottom, in engineered deep constructed pits, or in bottom depressions. The cap of clean sediment that is placed on top must be designed to withstand erosion over time from bottom currents, waves, vessel movement, prop wash, and burrowing bottom creatures (Fig. 9). Caps should be monitored over time to ensure their integrity [7].

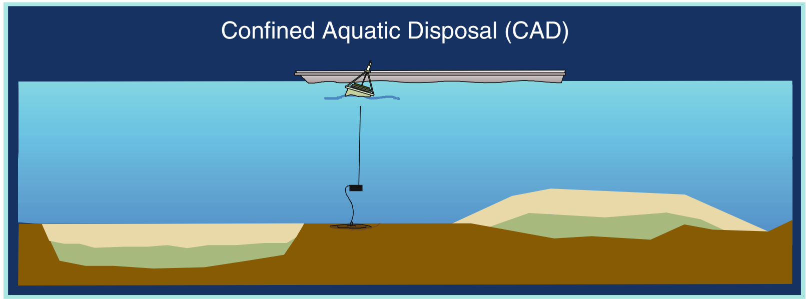
Confined Disposal Facilities

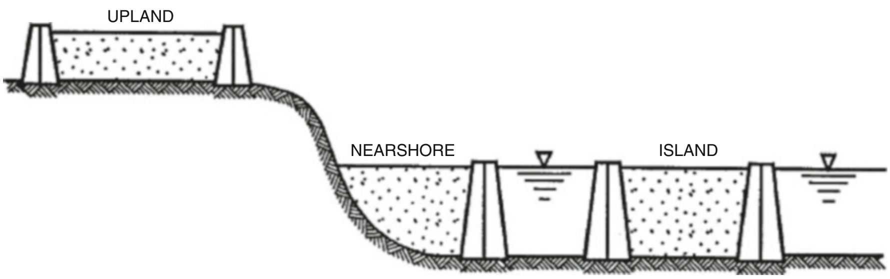
Confined disposal is placement of dredged material within engineered diked nearshore or upland confined disposal facilities (CDFs) via pipeline or barge delivery of sediments. CDFs may be constructed as upland sites, nearshore sites with one or more sides in water (sometimes called intertidal sites), as island containment areas, or as subaqueous contained capped cells (Figs. 10, 11, 12).

The primary objectives inherent in design and operation of CDFs are to provide for adequate storage capacity to meet dredged volume requirements, to maximize efficiency in retaining the solids, and to minimize exposure pathways to groundwater and wildlife. For facilities receiving contaminated material, an additional objective is to provide the efficient isolation of contaminants from the surrounding area. To achieve these objectives, depending on the degree of intended isolation, CDFs may be equipped with a complex system of control measures, such as surface covers and bottom liners, treatment of effluent, surface runoff and leachate monitoring, and management controls.

Hydraulic dredging adds several volumes of water for each volume of sediment removed. This excess water is normally discharged as effluent from the CDF during the filling operation, monitored for unacceptable levels of contaminants. The amount of water added depends on the design of the dredge, physical characteristics of the sediment, and operational factors, such as the pumping distance. When the dredged material is initially removed from the bed and deposited in the CDF, it can and will occupy several times its original volume. The settling process is a function of time, with sandy and gravelly sediment dewatering quickly and silts and clays dewatering very slowly. Silts and clays will eventually

Dredging Practices and Environmental Considerations, Fig. 9 Confined aquatic disposal (CAD) (Courtesy of Corps of Engineers)





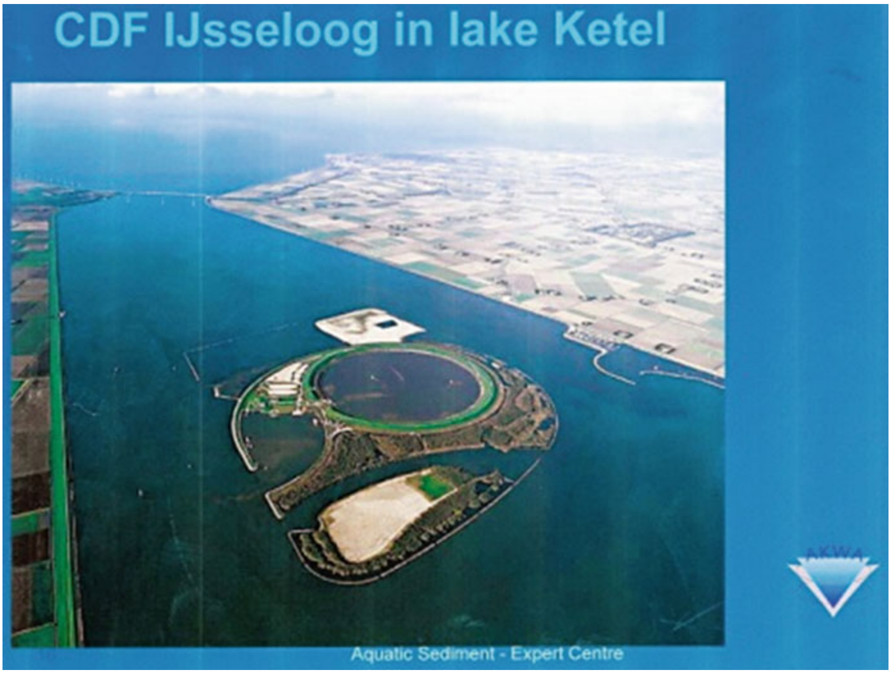
Dredging Practices and Environmental Considerations, Fig. 10 Types of confined disposal facilities (Courtesy of Corps of Engineers)

Dredging Practices and Environmental Considerations, Fig. 11 Nearshore CDF Huelva Estuary, Spain (Courtesy of Spain government)

consolidate to loose in situ volume or less if desiccation occurs. Adequate volume must be provided during the dredging operation to contain the total volume of sediment to be dredged, accounting for any volume changes during disposal.

Beneficial Use of Dredged Material

Dredged material is increasingly regarded as a resource rather than as a waste. More than 90–95% of sediments from navigation dredging



Dredging Practices and Environmental Considerations, Fig. 12 Island CDF at IJsseloog, the Netherlands (Courtesy of Dutch government)

comprise sediments acceptable for open-water disposal; in many cases, these are also considered acceptable for a wide range of environmentally and economically beneficial uses. The first step in examining dredged material management options is to consider possible beneficial uses of dredged material. Recent decades have seen increasing use of dredged materials for habitat creation, habitat restoration, beach nourishment, and coastal protection (Fig. 13).

Beneficial use is defined as "Utilizing dredged sediments as resource materials in productive ways, which provide environmental, economic, or social benefits" [8]. Broad categories of beneficial uses of dredged material, based on the functional use of the dredged material or site, include:

- Habitat development and restoration
- Parks and recreation
- Coastal protection
  - Beach and sand dunes nourishment
  - Riverbank and lakeshore protection
- Nearshore placement/littoral zone sediment management
- Construction and agricultural
  - Construction and industrial/commercial development (roads, dikes, levees, parking lots)
  - Land reclamation/remediation (brownfield restoration, strip mine reclamation)
  - Agriculture, forestry, horticulture, and aquaculture

Operational feasibility for these projects includes the availability of suitable material in the required amount at a particular time and within the project costs. These are crucial aspects of the most beneficial uses projects (Fig. 14).

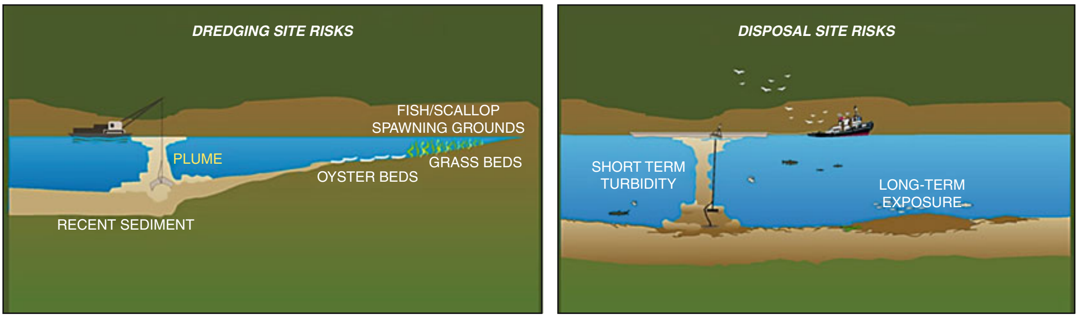
Treatment of Dredged Material

In certain cases of environmental cleanup by dredging, treatment of the dredged material may be necessary prior to confined disposal or reuse. A variety of treatment technologies are available to reduce the quantity or to reduce the contamination of the dredged material.

Dredging Practices and Environmental Considerations, Fig. 13 Deer Island Marsh Creation: Mississippi, USA (Courtesy of Corps of Engineers)

Dredging Practices and Environmental Considerations, Fig. 14 Beneficial use site: Evia Island bird nesting, Galveston Bay, Texas (Courtesy of Port of Houston Authority)

Treatment methods range from separation techniques, in which contaminated sediments are separated from relatively clean sand, to incineration. Some techniques are well developed to date, but others are still in the early stages of development. The products of the treatment methods have a wide array of potential beneficial uses (e.g., construction grade cement, lightweight aggregate, bricks, and manufactured soil). The problem is that treatment is usually very expensive, thereby limiting the feasibility of treatment or at least project scale. As a result, the treatment of small volumes of contaminated material is more likely than that of large volumes.



Dredging Practices and Environmental Considerations, Fig. 15 Environmental risks: at the dredging and disposal sites (Courtesy of Corps of Engineers)

Environmental Considerations

The potential environmental effects of navigation dredging and environmental cleanup dredging are the result of the actual dredging activity in the water and the disposal of the dredged material (Fig. 15).

During dredging, effects may arise due to the excavation of sediments causing resuspension in the water column, loss of material during transfer to the barge, overflow from the dredge while loading, and loss of material from the hopper dredge and/or pipelines during transport to disposal. Potential effects during disposal depend upon the physical, chemical, and toxicity characteristics of the dredged material and the selected disposal site (i.e., open water, nearshore, or upland).

During all dredging operations, the removal of material from the seabed also removes the surface-based (benthic) animals living on and in the sediments (benthic animals). With the exception of deep-burrowing animals or mobile surface animals that may survive a dredging event through avoidance, dredging can initially result in the complete removal of surface-dwelling biota from the dredging site. Where the channel or berth has been subjected to regular maintenance dredging over many years, it is very unlikely that well-developed benthic communities will occur in or around the dredged area. The recovery of disturbed habitats following dredging ultimately depends upon the nature of the new sediment at the dredge site, sources and types of recolonizing benthos, river width and bank line, and the extent of the disturbance [9].

The environmental issues associated with dredging and dredged material disposal include:

- Physical and ecological impacts due to turbidity and sedimentation
- Ecological and human health impacts: acute and chronic toxicity due to chemical contamination, e.g., PCBs, PAHs, dioxin, and metals, such as lead, cadmium, and mercury
- Loss of habitat due to dredging or placement of dredged material
- Impacts to endangered species (e.g., turtles) due to dredging
- Impacts to fish migration and spawning due to turbidity and exposure to toxic chemicals from dredging and disposal
- Impacts of noise from dredging operations upon aquatic living resources
- Others: emissions of air pollutants, quality-of-life issues (e.g., noise, night lights)
- Potential for toxicological impacts at CDFs or upland CDFs: pathways of exposure are numerous (e.g., groundwater and exposure to plants and animals at the CDFs)

All methods of dredging release suspended sediments into the water column during the excavation itself and during the overflow of dredging water from hoppers and barges. In many cases, the locally increased suspended sediments and turbidity associated with dredging and disposal are obvious from the turbidity "plumes," which may be seen trailing behind dredges and disposal sites [10]. When disposing of noncontaminated fine materials (e.g., silts, clays) in open-water disposal sites, the main environmental effects are associated with suspended sediments and increases in turbidity. Effects of dredging and disposal of clean dredged material include:

- Increases in suspended sediments and turbidity levels from dredging and disposal operations may under certain conditions have adverse effects on marine animals and plants by reducing light penetration into the water column and by physical disturbance.
- Increases in suspended sediments can impact filter-feeding organisms, such as shellfish, through clogging and damaging feeding and breathing equipment. Similarly, young fish can be damaged if suspended sediments become trapped in their gills, and increased fatalities of young fish have been observed in heavily turbid waters. Adult fish are likely to move away from or avoid areas of short-term high suspended solids, such as dredging sites, unless food supplies are increased as a result of increases in organic material.
- In important spawning or nursery areas for fish and other marine animals, dredging can result in smothering eggs and larvae. Shellfish are particularly susceptible during the spring when spatfall occurs.
- Increases in turbidity result in a decrease in the depth that light is able to penetrate the water column, which may affect submerged seaweeds and plants, such as eelgrass, *Zostera* species, by temporarily reducing productivity and growth rates.

The magnitude of sediment resuspension and resulting transport of contaminants during a dredging operation is influenced by many factors, including [11, 12]:

- Physical properties of the sediment (e.g., grain-size distribution, density, organic carbon content, acid-volatile sulfides (AVS) concentration)
- Vertical distribution of contaminants in the sediment
- Hydrodynamic regime in the dredging and disposal area (current direction and speed, mixing rate, tidal state)
- Type of dredge
- Methods of dredge operation
- Skill of operators
- Extent of controls on dispersion of resuspended sediment (e.g., silt curtains, sheet piling)
- Vertical profile of contaminant concentrations in sediment relative to the thickness of sediment to be removed
- Extent of debris
- Water salinity
- Extent of workboat/tugboat activity

Sediments dispersed during dredging and disposal may resettle over the seabed and the animals and plants that live on and within it. This blanketing can cause smothering of benthic animals and plants, may cause mortality, stress, and reduced rates of growth or reproduction, and, in the worst cases, the effects may be fatal. Generally, sediments settle within the vicinity of the dredged area, where they are likely to have little effect on the recently disturbed communities, particularly in areas where dredging is a well-established activity. However, in some cases, sediments are distributed more widely within the estuary or coastal area and may settle over adjacent subtidal or intertidal habitats possibly some distance from the dredged area.

At the disposal site, dredged materials will have a blanketing and smothering effect on benthic organisms in the immediate disposal site. The continual disposal of maintenance dredging at disposal sites may prevent the development of stable benthic communities and the partial or complete loss of benthic production and habitat. Recolonization is expected when disposal operations have been completed, depending on the characteristics of the dredged material and the changes to the hydrodynamic conditions at the disposal site.

A variety of harmful substances, including heavy metals, tributyltin (TBT), polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), dioxins/furans, and pesticides, are in the sediments in certain ports, harbors, and waterways. These contaminants are often of historic origin and from local or upstream sources. The highest levels of contaminants generally occur in industrialized estuaries. Dredging and disposal can release these contaminants into the water column, making them available to be taken up by animals and plants, with the potential to cause adverse acute and chronic toxicity. The risk of this occurring depends upon the type and degree of sediment contamination. If contaminants are released into the water column or are in the sediments at the open-water disposal site, they may bioaccumulate in marine animals and plants and transfer up the food chain to fish and sea mammals, with associated risks to human health.

Dredging can cause direct threats to endangered species, such as sea turtles and their nearshore marine habitats. Hopper dredges have been directly responsible for the incidental capture and the death of hundreds, if not thousands, of sea turtles in the United States. Development of specially designed hopper dredge drag heads and institution of best management practices in areas of turtle populations has helped alleviate the majority of the takings of turtles during dredging operations [13].

Nearshore or upland CDFs are the most commonly used disposal technique for contaminated dredged material. Strict testing requirements are used to determine the acceptability for open-water disposal. Confined aquatic disposal has also been successfully used around the world in many applications to isolate the contaminants from the surrounding environment.

Pathways for potential exposure to animals/plants and humans are similar for nearshore and upland CDFs. Potential pathways include the discharge into receiving waters (e.g., estuary or river) of the excess water from the dredged material, contamination of groundwater, and exposure of birds and animals to the dredged material in the CDF. Depending upon the level of contamination, controls can be used to minimize negative environmental impacts, such as using of impervious liners for disposal sites receiving dredged material from environmental dredging.

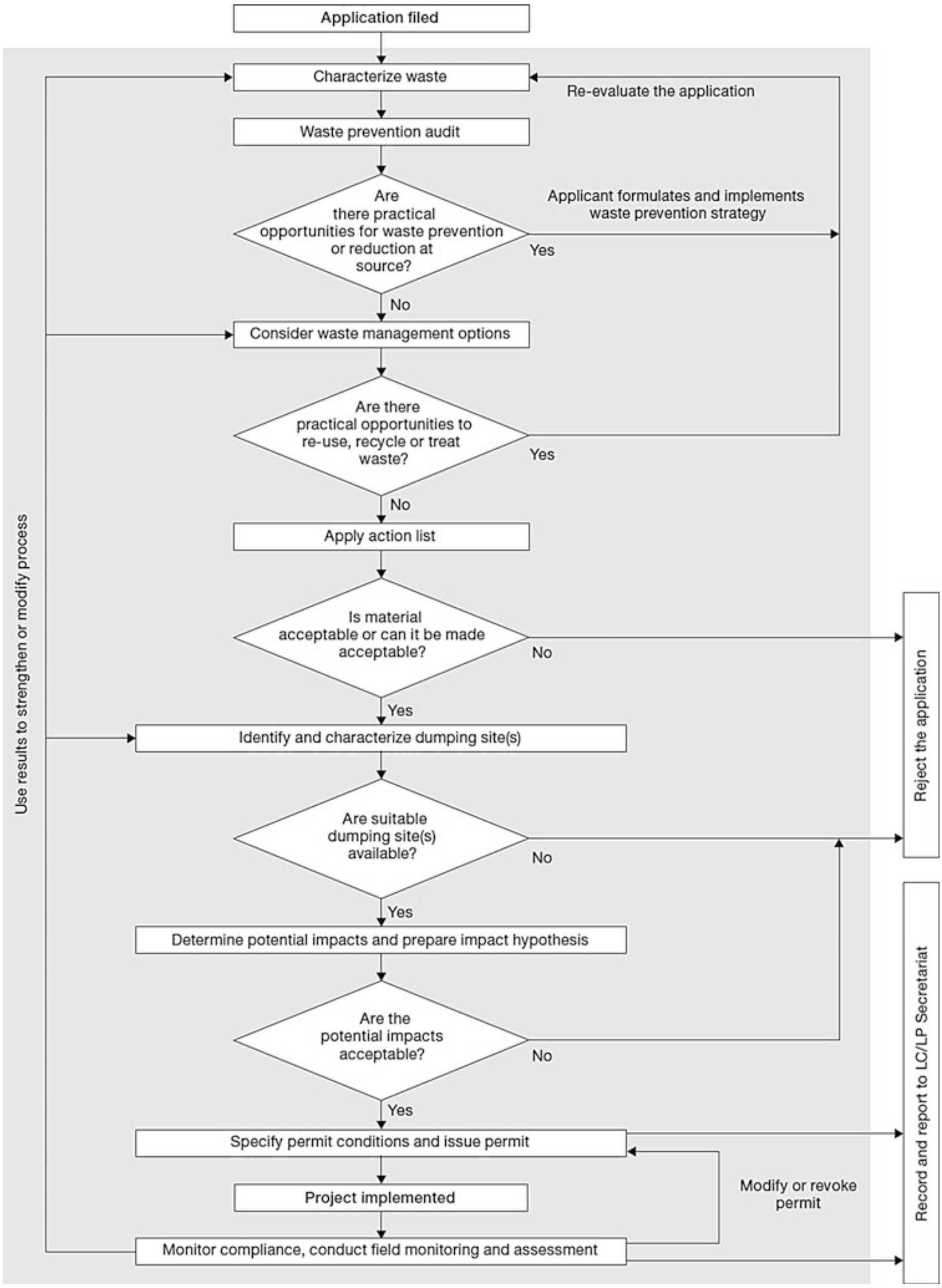
Environmental Regulation of Dredging and Dredged Material Disposal/Placement

The most widely applicable international regulatory instrument is the London Convention 1972 and London Protocol 1996 (LC/LP), which is part of the International Maritime Organization, an organization of the United Nations. The LC/LP regulates disposal of wastes into ocean waters, worldwide [14]. The LC/LP is an international treaty, which includes over 90 country signatories. Member countries are required to implement the conditions of the treaty including the waste assessment procedures noted below.

The LC/LP Waste Assessment Guidelines for Dredged Material allow disposal of dredged material into ocean waters, provided that strict environmentally protective criteria are met. A step-by-step process to evaluate a dredging project, criteria for selection of disposal sites or placement for beneficial use, and an action list for judging environmental acceptability of open-water disposal are specified in the guidelines [15]. Components of the international guidelines are shown in Fig. 16.

After an assessment of the need for dredging, major dredging or disposal projects should have studies carried out in order to ensure that any potential adverse effects are identified in advance and dealt with in an appropriate manner. Such investigations include characterization of the dredged material (physical, chemical, and toxicity), an examination of any sources of contamination and the potential to control those sources, an assessment of disposal or beneficial use placement alternatives, including identification and characteristics of the disposal site, and design of post disposal monitoring studies to determine whether any potential impacts are correctly predicted.

The environmental impact assessment should highlight both the positive and negative, short-



Dredging Practices and Environmental Considerations, Fig. 16 International guidelines for assessment of dredged material (Courtesy of London Convention)

and long-term impacts. Appropriate testing may be required to determine the physical behavior of the material at the disposal site. Also, testing and assessments of potential contaminants of concern may be required, depending upon existing knowledge of the dredging site and any potential contaminant pathways. Where potentially adverse effects are anticipated, management techniques should be implemented to reduce risks to acceptable levels. Possible controls for open-water alternatives include operational modifications, use of submerged discharges of dredged material, treatment, lateral containment, and capping or contained aquatic disposal. Possible controls for confined disposal facilities include operational modifications, treatment, and various site controls (e.g., covers and liners).

An important component in development of the environmental impact assessment and in identifying potential impacts and implementing acceptable measures is the involvement of interested groups and organizations, consulting with them, and reaching a consensus early in the process of determining the alternatives. It is in the best interests of the project sponsors and stakeholders that the decision-making process is transparent, stakeholders are involved, and the reasons for the selection of the preferred dredging and disposal or placement options are clearly understood.

Control of Upstream Sources of Sediments and Contaminants

Control of upstream sources of sediments and contaminants represents a significant part of the long-term solution to the continuing need for maintenance dredging of navigation channels and the continuing issues of contaminated dredged material. While control of upstream sources of sediments can contribute to lessening the need for dredging, sediments play a duel role in the watershed:

1. Sediments cause water quality problems and the need for more frequent dredging.
2. Sediments provide shoreline protection and maintaining habitats, such as marshes.

Sediments are increasingly recognized as scarce resources that can have costly impacts if not carefully managed, recycled, and conserved. Sediment overloading from land and stream erosion causes significant environmental and economic challenges, i.e., excessive needs for navigation dredging, and sediment in rivers, reservoirs, and estuaries may contribute to high turbidity, which can affect both habitat for aquatic life and human water uses, loss of flood storage capacity, and obstructions to waterway navigation and conveyance of floodwaters. Yet in other locations, a shortage of sediment can result in coastal and streambank erosion, erosion of beaches and shorelines losing developed properties and infrastructure, and loss of wetlands. Many water resource projects are designed to remedy local sediment problems, but they must be well planned and thought out as they can create even larger problems upstream, downstream, or downdrift along coastlines of the local project.

Achieving sediment management for source control usually involves substantial collaboration and cooperation between a number of organizations, agencies, and stakeholders. These groups need to be willing to identify the source and take measures to reduce or prevent further contamination of waterways with sediments and chemicals. Moving stakeholders and authorities together toward long-term successful outcomes of shoreline development and environmental protection is no easy task [16].

Future Directions: Sustainable Dredging and Dredged Material Management

The global economy and the dependence upon food and commodities via international trade require that vessels have sufficiently deep channels in ports, harbors, and waterways for safe passage. Other interests include national security, recreational opportunities, and reduced human and ecological exposure to contaminated sediments.

While sediment controls of upstream sources will help, the natural erosion process in rivers and estuaries will continue. Thus, navigation dredging will continue to be needed over the very long term. Environmental cleanup dredging will be needed for decades to come, even as improved controls are placed upon waste and wastewater sources. Legacy contaminants already in the sediments will continue to pose aquatic and human health risks until they are removed or isolated from the surrounding aquatic environment. Environmental considerations relate to the quality and quantity of the sediment to be dredged, the potential environmental risks from the dredging itself, and what to do with the dredged material.

Over time, dredging and dredged material management practices will move toward sustainability concepts.

Sediment Management and Sustainability

Dredging projects should be managed as part of the overall sediment system in the watershed and littoral cell, with dredged material being considered a resource. Opportunities for beneficial use of dredged material will increase as potential beneficial use projects (e.g., habitat restoration or creation, beach nourishment and coastal protection, restoration of brownfields, and construction purposes) and their sponsors are identified early in the dredging project planning process.

Dredging and other projects affecting or involving sediment should be undertaken with awareness of the littoral or fluvial sediment system and the effects that proposed projects and actions may have on other stakeholders. Regulatory authorities and stakeholders are becoming increasingly aware of the necessity to manage sediments on a system-wide basis, dredging being one part of the overall river, lake, or coastal system [17–19].

Principles of Sustainable Sediment Management [19]

- Recognize sediment as part of a system and as a valuable resource that is integral to economic and environmental vitality.
- Strive for balanced, economic, and environmentally sustainable solutions to sediment-related issues through integrated management of sediment from upland sources to estuaries and within coastal zones.

(*continued*)

- Involve external and internal partners and stakeholders to integrate and balance objectives and to leverage resources in implementation.
- In project decision-making, consider the sediment implications beyond the local site, including intended and potential effects, and over long time scales (decades or more), and be effective stewards of sediment and related resources.

Climate Change and Dredging

Climate change and sea level rise can cause more erosion in some places and less in others, with associated changes in the quantities of sediment needed to be removed by dredging for navigation purposes. Continued focus upon control of local and upstream sources of sediment and contaminants will begin to pay limited dividends by reducing the frequency of navigation dredging, but increased precipitation and storminess in some areas may negate the storm water and erosion control programs. Improved chemical and toxicological quality of those sediments will occur as additional environmental controls are put in place to control municipal and industrial discharges and storm water runoff from urban and rural areas, including farmlands. Instituting environmental controls on the disposal of hazardous waste has helped control runoff from hazardous waste disposal sites and over decades will reduce the need for environmental cleanup dredging. Environmental dredging, such as removal of PCB-contaminated sediments in the Hudson River, New York; Fox River, Wisconsin; and Passaic River, New Jersey, will contribute to improved sediment quality downstream.

Climate change and sea level rise will result in many more dredging projects, the purpose of which will be to create or enhance existing sand dunes and barrier islands. Sea level rise and increased storminess have the potential to cause serious flooding and erosion of shorelines with enormous economic, social, and environmental consequences. Thus, coastal communities will be seeking protection against waves and high water in building resilient shorelines.

Technological Innovations and Approaches

Driven by concerns about the potential impacts to aquatic life and human health, technology will continue to evolve in dredging hardware, treatment of contaminated dredged material, and use of dredged material in beneficial use applications. Innovations in dredging technology are focused upon dredge positioning for precision dredging and reduction in the disbursement of suspended solids and associated contaminants into the water column during dredging and disposal operations.

Further technological innovations in the types and efficiencies of treatment technologies will likely identify potential reuse opportunities for contaminated sediments, such as use as soil, fill, or aggregates. The barrier to the widespread use of these treatment technologies to create useful products will continue to be relative high costs.

The science and engineering of confined disposal facility design and use is well established and will continue. However, the likely trend is to increase the use of confined aquatic disposal cells, ensuring the isolation of contaminated sediment from the aquatic environment.

Other areas of dredging that are likely to see significant changes include:

1. Electric-powered dredges will contribute fewer diesel emissions and NO_x and SO_x in locations where compliance with air pollution standards is an issue.
2. Navigation channels will need to be dredged to accommodate larger ships (deeper channels), and at the same time, improved channel design will be necessary due to limited availability of funding (e.g., narrower channels and institution of vessel operational controls).
3. In management of contaminated sediments, instead of environmental dredging to attempt to remove contaminated sediments to acceptable levels in the waterway, the use of monitored natural recovery of contaminated sediments is likely to increase, as well as other approaches such as thin layer capping and enhanced natural recovery. Combinations of these approaches (i.e., removal or recovery) will be used depending upon local conditions [20].

Engineering/Building with Nature

The approaches captured in Engineering/Building with Nature enable more sustainable delivery of economic, social, and environmental benefits associated with water resources infrastructure. Engineering/building with nature includes the intentional alignment of natural and engineering processes to efficiently and sustainably deliver economic, environmental, and social benefits through collaborative processes.

Engineering/building with nature includes strategic placement of sediments, in combination with hydrodynamics, and natural transport processes, to build nearshore habitats. Key concepts [21, 22]:

- Engineering features to focus natural processes to minimize navigation channel infilling and to transport and focus sediments for positive benefits
- Cost-efficient engineering practices for enhancing the habitat value of infrastructure
- Natural systems, such as wetlands and other features, to reduce the effects of storm processes and sea level change on shorelines and coasts
- Science-based communication processes to improve stakeholder engagement and collaboration [23]

Implementation of Regulations

International guidelines (i.e., London Convention/London Protocol) are in place for protection of the environment from dredged material disposal in ocean waters. National and local regulations are in place in many countries that implement the London Convention/London Protocol guidelines to protect their coastal and ocean waters as well as for protection of internal country waters. These regulations are in various stages of implementation worldwide. Technical cooperation and assistance programs are ongoing to assist developing countries in their application. For example, low cost and low technology guidance for field monitoring of dredged material disposal is available on the London Convention and Protocol website [24].

One key aspect of the international (as well as national and local regulations) is the characterization of the dredged material prior to dredging. Updated procedures for testing sediments proposed to be dredged are providing better techniques to assess their acceptability for open-water disposal or for specific beneficial uses; these include improved bioassays, interpretive guidance, and the application of risk assessment in cases where high uncertainties exist [25]. New guidance on action levels for the acceptability of disposal of dredged material into the open water is available on the London Convention and Protocol website, and guidance on selection of open-water disposal sites is being developed by the London Convention and Protocol [26].

The final trend in the regulatory arena is the use of mitigation for unavoidable adverse impacts due to dredging and disposal. This will become more widespread as one of the tools for regulatory authorities.

Bibliography

Primary Literature

1. Willard (2009) *Willardsays*. . ..Dredge history done lite. http://www.willardsays.com/dredge_history.pdf. Accessed 25 May 2011
2. Cohen M (2005) Dredging: the environmental facts. Published by International Association of Dredging Companies; Typeset and printed by Opmeer Drukkerij bv, The Netherlands. ISBN: 90-75254-11-3. http://www.dredging.org/documents/ceda/downloads/publications-dredging_the_facts. Accessed 1 Apr 2011
3. Pew Charitable Trusts, A fact sheet on deep sea mining. March 2017. http://www.pewtrusts.org/~/media/assets/2017/04/sea_deep_sea_mining_the_basics.pdf
4. U.S. Army Corps of Engineers, U.S. Environmental Protection Agency (2004) Evaluating environmental effects of dredged material management alternatives – a technical framework. EPA842-B-92-008
5. Cable Arm website. 2017. http://www.cablearm.com
6. Bray RN, International Association of Dredging Contractors, Central Dredging Association (2008) Environmental aspects of dredging. Taylor and Francis/Balkema, Amsterdam
7. Otten MT, Hartman G (2002) Placing and capping fine-grained dredged material. In: ASCE conference proceedings of the third specialty conference on dredging and dredged material disposal, 5–8 May, Orlando, p 119
8. USEPA, Corps of Engineers (2007) Identifying, planning, and financing beneficial use projects using dredged material beneficial use planning manual. U.S. EPA, EPA842-B-07-001, Corps of Engineers, Washington, DC
9. International Council for the Exploration of the Sea (ICES) (1992) Report of the ICES working group on the effects of extraction of marine sediments on fisheries. ICES cooperative research report # 182, Copenhagen
10. UK Marine Protected Areas Centre (2001) UK marine SACs project. Dredging and disposal. http://www.ukmarinesac.org.uk/activities/ports/ph5.htm. Accessed 2 Apr 2011
11. U.S. EPA, Contaminated sediment remediation guidance for hazardous waste sites, EPA-540-R-05-12, OSWER 9355.0-85, Dec 2005
12. Bridges TS, Ells S, Hayes D, Mount D, Nadeau S, Palermo M, Patmont C, Schroeder P (2008) The four Rs of environmental dredging: resuspension, release, residual, and risk. Technical report ERDC/EL TR-08-4. U.S. Army Engineer Research and Development Center, Vicksburg
13. U.S. Corps of Engineers, Sea Turtle Data Warehouse. http://el.erdc.usace.army.mil/seaturtles/intro.cfm. Accessed 27 Sept 2011
14. IMO (2003) London convention 1972 and the 1996 protocol, 2nd edn. International Maritime Organization, London
15. IMO (2014) London convention and Protocol. Specific guidelines for assessment of dredged material. http://www.imo.org/en/OurWork/Environment/LCLP/Publications/wag/Pages/default.aspx
16. Vogt C (2007) Dredged material management in a watershed context – seeking integrated solutions. In: Proceedings of 26th WEDA conference, San Diego
17. U.S. Army Corps of Engineers (2004) Regional sediment management – primer. U.S. Army Corps of Engineers, Washington, DC
18. WODA, WODA Principles of Sustainable Dredging, 2013. https://www.dredging.org/documents/ceda/downloads/2013-wodaprinciples-signed.pdf
19. USACE Regional Sediment Management website: http://rsm.usace.army.mil/
20. Palermo M, Schroeder P, Estes T, Francingues N (2008) Technical guidelines for environmental dredging of contaminated sediments. Technical report ERDC/EL TR-08-29. U.S. Army Engineer Research and Development Center, Vicksburg
21. USACE, Engineering with Nature website: https://ewn.el.erdc.dren.mil

22. Building with Nature Presentation. Central Dredging Association. https://www.dredging.org/media/ceda/org/documents/presentations/ceda-nl/ceda_bwn_jan_2013_stefan_aarninkhof_intro.pdf
23. Bridges T (2013) Engineering with nature for sustainable water resources infrastructure, international coasts and ports conference, Sydney, 6–13 Sept 2013
24. London convention and protocol; Low cost, low technology field monitoring for dredged material. Prepared under contract by Craig Vogt Inc. 2016. http://www.imo.org/en/OurWork/Environment/LCLP/Pages/default.aspx
25. Moore DW, Bridges TS, Ruiz C, Cura J, Driscoll SK, Vorhees D, Peddicord D (1998) Environmental risk assessment and dredged material management: issues and application. In: Proceedings of workshop, San Diego
26. London Convention and Protocol website. http://www.imo.org/en/OurWork/Environment/LCLP/Pages/default.aspx

Books and Reviews

Adams JR, Herbich JB (1970) Gas removal systems. In: Proceeding world dredging conference, WODCON'70, Singapore

Addie GR, Whitlock L, Dresser J (2001) Dredge pump testing considerations. In: Proceedings of the Western Dredging Association 21st technical conference and 33rd Texas A & M Dredging seminar, Houston, 24–27 Jun 2001

Albar A (2001) Modeling of a bucket wheel dredge system for offshore sand and tin mining. PhD dissertation, Ocean engineering program, Texas A & M University, Ocean engineering program, Civil Engineering Department, College Station

Ancheta C, Santiago R, Sherbin G (1998) An innovative approach to harbour remediation, Thunder Bay. In: Proceedings of 15th world dredging congress, Las Vegas

Board M (1997) Contaminated sediments in ports and waterways. National Research Council, National Academy Press, Washington, DC

Bray RN, Bates AD, Land JM (1997) Dredging: a handbook for engineers, 2nd edn. Arnold, London

Brandon DL, Price RA (2007) Summary of available guidance and best practices for determining suitability of dredged material for beneficial uses, ERDC/EL TR-07-27. U.S. Army Engineer Research and Development Center, Vicksburg

Briot JE, Doody JP, Romagnoli R, Miller MM (1999) Environmental dredging case studies: a look behind the numbers. In: Proceedings of the western dredging association 19th technical conference and 31st Texas A & M Dredging seminar, Louisville, 15–18 May 1999

Cura JJ, Heiger-Bernays W, Bridges TS, Moore DW (1999) Ecological and human health risk assessment guidance for aquatic environments. U.S. Army Corps of Engineers, Engineer Research and Development Center, Vicksburg

Demars KR (1995) Dredging, remediation, and containment of contaminated sediments. ASTM committee D-18 on soil and rock, environment, Canada

Dickerson DD, Nelson DA, Dickerson CE Jr, Reine KJ (1993) Dredging related sea turtle studies along the Southeastern U.S., Coastal Zone 93, New Orleans

Eisma D (2005) Dredging in coastal waters. Taylor & Francis, London

Fredette TJ, Anderson G, Payne BS, Lunz JD (1986) Biological monitoring of open-water dredged material disposal sites. In: IEEE Oceans'86 conference proceedings, Washington, DC

Fredette TJ, Nelson DA, Clausner JE, Anders FJ (1990a) Guidelines for physical and biological monitoring of aquatic dredged material disposal sites. Technical report D-90-12. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Fredette TJ, Nelson DA, Miller-Way T, Adair JA, Sotler VA, Clausner JE, Hands EB, Anders FJ (1990b) Selected tools and techniques for physical and biological monitoring of aquatic dredged material disposal sites, Technical report D-90-11. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Graalum SJ, Randall RE, Edge BL (1999) Methodology for manufacturing topsoil using sediment dredged from the Texas Gulf intracoastal waterway. J Marine Environ Eng 1:1–38

Hayes DF (1986) Guide to selecting a dredge for minimizing resuspension of sediment, Environmental effects of dredging, Technical notes, EEDP-09-1. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Hayes DF (1999) Turbidity monitoring during Boston harbor bucket dredge comparison study. Project report, Submitted to USAE Waterways Experiment Station, Vicksburg

Hayes DF, Raymond GL, McLellan TN (1984) Sediment resuspension from dredging activities. In: Dredging. American Society of Civil Engineers, Clearwater, p 1984

Hayes DF, McLellan TN, Truitt CL (1988) Demonstration of innovative and conventional dredging equipment at Calumet Harbor, Illinois, MP EL-88-01. U. S. Army Engineer Waterways Experiment Station, Vicksburg

Herbich JB (1995) Removal of contaminated sediments: equipment and recent field studies. In: Demars KR et al (eds) Dredging, remediation and containment of contaminated sediments. ASTM Publication Code Number (PCN) 04-012930-38. American Society for Testing and Materials, Philadelphia

Herbich JB (2000a) Dredge pumps, Chapter 3. In: Herbich JB (ed) Handbook of dredging engineering. McGraw-Hill, New York, pp 3.36–3.58

Herbich JB (2000b) Removal of contaminated sediment by dredging. In: Dredging engineering short course notes, Center for Dredging Studies, Ocean Engineering Program, Civil Engineering Department, Texas A & M University, College Station

Herbich JB (2000c) Handbook of dredging engineering, 2nd edn. McGraw-Hill, New York

Herbich JB, Brahme SB (1991) A literature review and technical evaluation of sediment resuspension during dredging, Contract report HL-91-1. U.S. Army Engineer Waterways Experiment Station, Vicksburg

History Channel (2005) Modern marvels: dredging

Huston J (1970) Hydraulic dredging. Cornell Maritime Press, Cambridge

Iwasaki M, Kuioka K, Izumi S, Miyata N (1992) High density dredging and pneumatic conveying system. In: Proceedings of XIIIth world dredging congress, WODCON XIII, Bombay, 7–10 Apr 1992, pp 773–792

Jaglal K, McLaughlin DB (1999) Evaluation of large scale environmental dredging as a remedial alternative. In: Proceedings of the Western Dredging Association 19th technical conference and 31st Texas A & M Dredging seminar, Louisville, 15–18 May 1999

Kikeqawa K (1983) Sand overlying for bottom sediment improvement by sand spreader. In: Management of bottom sediments containing toxic substances: proceedings of the 7th US/Japan Experts meeting. Prepared for the U.S. Army Corps of Engineers Water Resources Support Center by the U.S. Army Engineer Waterways Experiment Station, Vicksburg, pp 79–103

Landin MC (2000) Beneficial uses of dredged material, Chapter 16. In: Herbich JB (ed) Handbook of dredging engineering. McGraw-Hill, New York

LaSalle MW, Clarke DG, Homziak J, Lunz JD, Fredette TJ (1991) A framework for assessing the need for seasonal restrictions on dredging and disposal operations. Technical report D-91-1. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Lee CR (1997) Manufactured soil from Toledo harbor dredged material and organic waste materials. In: International workshop on dredged material beneficial uses, Baltimore, 28 July–1 Aug 1997

Marine Board (1990) Managing troubled waters – the role of marine environmental monitoring. National Research Council, National Academy Press, Washington, DC

Marine Board (2001) A process for setting, managing and monitoring environmental windows for dredging projects. Transportation Research Board and Ocean Studies Board, National Research Council, National Academy Press, Washington, DC

Marine Board, National Research Council (1987) Sedimentation control to reduce maintenance dredging of navigation facilities in estuaries. National Academy Press, Washington, DC

Marine Board, National Research Council (1990) Managing coastal erosion. National Academy Press, Washington, DC

Marine Board, National Research Council (1997) Contaminated sediments in ports and waterways – cleanup strategies and technologies. National Academy Press, Washington, DC

Martin JL, McCutcheon SC (1992) Overview of processes affecting contaminant release from confined disposal facilities, Contract report D-92-1. U.S. Army Engineer Waterways Experiment Station, CE, Vicksburg, NTIS no AD A247 650

Matousek V (1997) Flow mechanism of sand-water mixtures in pipelines. PhD dissertation, Delft University of Technology, Delft

McLellan TN, Havis RN, Hayes DF, Raymond GL (1989) Field studies of sediment resuspension characteristics of selected dredges. Technical report HL-89-9. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Miertschin M, Randall RE (1998) A general cost estimation program for cutter suction dredges. In: Proceedings of 15th world dredging congress, Las Vegas, pp 1099–1115

Myers TE, Brannon JM, Tardy BA (1996) Leachate testing and evaluation for estuarine sediments. Technical report D-96-1. U.S. Army Engineer Waterways Experiment Station, Vicksburg, NTIS no AD A306 421

National Dredging Team (2003) Dredged material management: action agenda for the next decade, Jacksonville, 23–25 Jan 2003

Ocean Studies Board, National Research Council, Transportation Research Board (2001) A process for setting, managing and monitoring environmental windows for dredging projects; Special report 262

Otis MJ, Andon S, Bellmer R (1990) New Bedford harbor superfund pilot study: evaluation of dredging and dredged material disposal. US Army Engineer Division, New England, Waltham

Palermo MR, Maynard S, Miller J, Reible DD (1996) Guidance for in-situ subaqueous capping of contaminated sediments. Environmental Protection Agency, EPA 905-B96-004, Great Lakes National Program Office, Chicago

Palermo MR, Clausner JE, Rollings MP, Williams GL, Myers TE, Fredette TJ, Randall RE (1998) Guidance for subaqueous dredged material capping. Technical report DOER-1. U.S. Army Corps of Engineers, Washington, DC

Parchure TM (1996) Equipment for contaminated sediment dredging. Technical report HL-96-17. USAE Waterways Experiment Station, Vicksburg

Parchure TM, Sturdivant CN (1997) Development of a portable innovative contaminated sediment dredge. Final report, CPAR-CHL-97-2, Construction productivity advancement research (CPAR) program. USAE Waterways Experiment Station, Vicksburg

Pearson TH (1987) Benthic ecology in an accumulation sludge-disposal site. In: Capuzzo JM, Kester DR (eds) Ocean processes and marine pollution, vol 1, Biological processes and wastes in the ocean. Krieger Publishing, Malabar, pp 195–200

Pelletier JP (1992) Collingwood harbour sediment removal demonstration. Preliminary report on the water quality monitoring program, environmental protection, Environment Canada, Toronto

Pennekamp JGS, Quaak MP (1990) Impact on the environment of turbidity caused by dredging, Terra et Aqua,

42. International Association of Dredging Companies, The Hague

Pequegnat WE, Gallaway BJ, Wright TD (1990) Revised procedural guide for designation surveys of ocean dredged material disposal sites. Technical report D-90-8. U.S. Army Engineer Waterways Experiment station, Vicksburg

PIANC (1992) Beneficial uses of dredged material: a practical guide. PIANC, Brussels

Pianc, EnviCom (2009) Dredging management practices for the environment, Pianc report 100

Pilarczyk KW (1994) Novel systems in coastal engineering, geotextile systems and other methods, an overview. Rijkswaterstaat, Delft

Pilarczyk KW (1996) Geotextile systems in coastal engineering, an overview. In: Proceedings of the twenty-fifth international conference on coastal engineering, vol 2, pp 2114–2127

Randall RE (1992) Equipment used for dredging contaminated sediments. In: 1992 international environmental dredging symposium proceedings, Erie County Environmental Education Institute, Buffalo, 30 Sep–2 Oct 1992

Randall RE (1994) Dredging equipment alternatives for contaminated sediment removal. In: Conference on environmental dredging technology, East Brunswick, Sponsored by The Port Authority of New York/New Jersey, 21–22 Jun 1994

Randall RE (2000a) Estimating dredging costs. In: Herbich JB (ed) Handbook of dredging engineering, 2nd edn. McGraw-Hill, New York, Appendix 9

Randall RE (2000b) Pipeline transport of dredged material, Chapter 7, Section 2. In: Herbich JB (ed) Handbook of dredging engineering, 2nd edn. McGraw-Hill, New York

Randall RE (2004) Dredging, Chapter 11. In: Tsinker G (ed) Handbook of port engineering. McGraw-Hill, New York

Randall RE, Clausner JE, Johnson BH (1994) Modeling of Cap placement at the New York mud dump site. In: Dredging 94, Second international conference of dredging and dredged material placement, ASCE, Orlando, 13–16 Nov 1994

Randall R, Edge B, Basilotto J, Cobb D, Graalum S, He Q, Miertschin M (2000) Texas gulf intracoastal waterway (GIWW): beneficial uses, estimating costs, disposal analysis alternatives, and separation techniques, Texas Transportation Institute, Project report no 1733-S, Texas A&M University System, College Station, Sep 2000

Reine KJ, Dickerson DD, Clarke DG (1998) Environmental windows associated with dredging operations, DOER technical notes collection (TNDOER-E2) U.S. Army Engineer Research and Development Center, Vicksburg

Richardson M (2001) The dynamics of dredging. Placer Management Corp, Irvine

Sanders L, Killgore J (1989) Seasonal restrictions on dredging operations in freshwater systems, environmental effects of dredging. Technical note EEDP-01-16. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Sanderson WH, McKnight AL (1986) Survey of equipment and construction techniques for capping dredged material, Miscellaneous paper D-86-6. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Santiago R (2000) Contaminated sediment removal general management options and support technologies, Dredging engineering short course notes, Center for Dredging Studies, Ocean Engineering Program, Civil Engineering Department, Texas A&M University, College Station, Jan 2000

Schiller RE Jr (1992) Sediment transport in pipes, Chapter 6. In: Herbich JB (ed) Handbook of dredging engineering. McGraw-Hill, New York

Science Applications International Corporation (SAIC) (1995) Sediment capping of subaqueous dredged material disposal mounds: an overview of the New England experience, 1979–1993, DAMOS contribution #95, SAIC report no SAIC-90/7573&c84 prepared for the U.S. Army Engineer Division, New England, Waltham

Shook CA, Rocco MC (1991) Slurry flow: principles and practice. Butterworth-Heinemann, Boston

Steevens J, Kennedy A, Farrar D, McNemar C, Reiss MR, Kropp RK, Doi J, Bridges T (2008) Dredged Material analysis tools; performance of acute and chronic sediment toxicity methods, ERDC/EL TR-08-16. U.S. Army Engineer Research and Development Center, Vicksburg

Sumeri A (1989) Confined aquatic disposal and capping of contaminated bottom sediments in puget sound. In: Proceedings of the WODCON XII, Dredging: technology, environmental, mining, world dredging congress, Orlando, 2–5 May 1989

Togashi H (1983) Sand overlaying for sea bottom sediment improvement by conveyor barge. In: Management of bottom sediments containing toxic substances: proceedings of the 7th U.S./Japan experts meeting. Prepared for the U.S. Army Corps of Engineers Water Resources Support Center by the U.S. Army Engineer Waterways Experiment Station, Vicksburg, pp 59–78

Tsinker GP (2004) Port engineering: planning, construction, maintenance, and security. Wiley, Hoboken

Turner TM (1996) Fundamentals of hydraulic dredging, 2nd edn. ASCE Press, New York

USACE (1983) Dredging and dredged material disposal, Engineer manual, US Army Corps of Engineers, EM-1110-2-5025. US Government Printing Office, Washington, DC

USACE (1987a) Confined disposal of dredged material, Engineer manual, US Army Corps of Engineers, EM-1110-2-5027. US Government Printing Office, Washington, DC

USACE (1987b) Beneficial uses of dredged material, Engineer manual, US Army Corps of Engineers, EM-1110-2-5026. US Government Printing Office, Washington, DC

USEPA/USACE (1991) Evaluation of dredged material proposed for ocean disposal (Testing manual), EPA-503/

8-91/001. Office of Water, U.S. Environmental Protection Agency, Washington, DC

USEPA/USACE (1992) Evaluating environmental effects of dredged material management alternatives-A technical framework, EPA 842-B-92-008. US Government Printing Office, Washington, DC

USEPA/USACE (1996a) Guidance document for development of site management plans for ocean dredged material disposal sites. USEPA, Office of Water, Washington, DC

USEPA/USACE (1996b) Evaluation of dredged material proposed for discharge in inland and near-coastal waters – testing manual. Office of Water, U.S. Environmental Protection Agency, Washington, DC

USEPA/USACE (2007) The role of the federal standard in the beneficial use of dredged material from US Army Corps of Engineers new and maintenance navigation projects – beneficial uses of dredged materials. USEPA/USACE, Washington, DC

Van Drimmelen C, Schut T (1992) New and adapted small dredges for remedial dredging operations. In: Proceedings of WODCON XIII, Bombay, pp 156–169

Vorhees DJ, Driscoll SBK, Stackelberg K, Bridges TS (1998) Improving dredged material management decisions with uncertainty analysis. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Wilson KC, Addie GR, Sellgren A, Clift R (2006) Slurry transport using centrifugal pumps, 3rd edn. Blackie Academic, New York

Zappi PA, Hayes DF (1991) Innovative technologies for dredging contaminated sediments, Miscellaneous paper EL-91-20. U.S. Army Engineer Waterways Experiment Station, Vicksburg

Zeller RW, Wastler TA (1986) Tiered ocean disposal monitoring will minimize data requirements, Oceans'86, vol 3. Marine Technology Society, Washington, DC, pp 1004–1009

Mining and Its Environmental Impacts

Jörg Matschullat and Jens Gutzmer
TU Bergakademie Freiberg, Institute of Mineralogy and Helmholtz Institute Freiberg for Resource Technology, Freiberg, Germany

Article Outline

Glossary

Biota All life forms.

Decommissioning Removal of something from active status.

Eco-efficiency analysis Analysis of realizing the concept of creating goods and services with fewer resources and less waste and pollution.

Exploitation Act of using something (mineral resources) for any purpose.

Exploration Process of finding mineral resources for the purpose of mining.

Karst A geological feature in relatively soluble rocks, e.g., limestone, where sinkholes, caves, and similar hollows are formed above and below ground.

Lithosphere The outer rocky shell of planet Earth, comprising the oceanic and continental crust and part of the upper Earth mantle.

Long-term effect A change that will last or have an influence over a long period of time.

Nachhaltigkeit German for “sustainability,” first used in 1713 in Germany.

Open-pit excavation Process of extracting minerals from surface deposits.

Recultivation Making raw mineral soils (brownfields) fertile again through bioengineering and refertilization.

Rehabilitation Restoring land after some process has damaged it.

Remediation Removal of pollution or contaminants from the environment.

Sinkhole A natural depression at the Earth surface generated by subsurface erosion, particularly in karst areas.

Slag Partially glassy by-product from smelting ore, mostly consisting of a silicate matrix with metal oxides.

Sustainable mining Mining method that does not compromise environmental quality.

Definition of the Subject

The environmental impact of mining is the influence that mining activities have on the natural conditions and world in which humans and all biota live. The impact may involve diverse forms of environmental change or damage, from short- to long-term effects and from highly spatially restricted to long-distance consequences.

Just as any kind of human activity, mining has an inherent and partly unavoidable impact on the environment. From the first steps of exploration via exploitation and (ore) processing to the final stages of decommissioning and rehabilitation, environmental hazards and risks may be encountered and need to be addressed. The potential impacts and long-term aftermath of mining operations are manifold. Whether in fact and to which extent the impacts do lead to detrimental consequences in any one of the environmental compartments (atmosphere, hydrosphere, pedosphere, biosphere, cryosphere, and lithosphere) is difficult to predict. A thorough investigation of local conditions – both boundary conditions and operation-related conditions – is needed to answer

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_205

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC, 2012
https://doi.org/10.1007/978-1-4419-0851-3_205

that crucial question. Modern mines can be operated in a manner much less detrimental than was the standard up until only recently. Parallel to the mining industry's awareness of issues connected with the environment, however, new challenges appear: The exploitation of lower concentrations of the valuable constituents (e.g., minerals or metals) presents most demanding challenges that deal with difficult conditions and involve a larger footprint of mining and more complex approaches to beneficiation. Increasing challenges also apply to marine mining, be it nearshore or offshore; be it for diamonds and other placer deposits, oil, and gas; or be it for manganese nodules that are being procured in increasingly deeper marine environments [1].

Developing and implementing sustainable mining practices are tasks that have been high on the agenda of the international mining industry. These need to become the global standard to curtail the most long-lasting and detrimental impacts of mining. Although the necessary knowledge base is rapidly becoming available; there is still a need for basic research to further establish and foster sustainable solutions. Due to its exotic nature and less likelihood to disturb Earth's immediate biosphere and equilibrium, "extraterrestrial mining" is not dealt with here.

Introduction: Sustainable Mining – An Oxymoron?

Environmental impacts of mining appear to be most well known all over the world – almost beyond the necessity of further elucidation and questioning ([2, 3]; Table 1). The mining industry has recognized its impact on the environment and has identified the control and restriction of such impact as one of its key challenges [4, 47].

From a workable standpoint, mining encompasses a very large array of activities. Especially in highly industrialized nations, there is major resistance of societies against mining activities. For the building and construction industry, even aggregate materials from quarries, and sand and gravel from open-pit excavations are increasingly under scrutiny in densely populated areas. Generally, much larger open-pit or underground mines for rock salt, metalliferous ores, or precious stones are often no longer perceived as indicators of economic well-being and development, but rather symptomatic of visual, acoustic, and environmental perturbations with detrimental impact. The same is certainly true for energy resources, namely gas, oil, coal, tar sands, and uranium ores, whether or not these are being mined on land or in marine shelf areas. Not only the general public, but a considerable part of the decision makers in both industry and politics (at least in the western world) has developed a stance that mining is per se a dirty business and that its related activities can be left to (mostly) developing countries.

In the late twentieth century, the idea that "industry and the developed nations would always be able to buy the necessary commodities" prevailed. Since the advent of the twenty-first century, this position has been increasingly under scrutiny, simply because the growing world population demands increasing amounts of raw materials. To find acceptance and support in society, any future mining activity will demand a state-of-the-art environmental management and has to contribute to sustainable development [5]. Furthermore, the role of the mining industry is set to increase, as technological advancement demands rapidly increasing supplies of a rising number of raw materials that have never before found a significant industrial application, e.g., rare earth elements (REE) or lithium (Li).

To clarify current environmental issues connected to mining and to be able to develop alternatives to practices that are currently widely used, an understanding of the history of mining is needed, as well as an overview of the environmental effects of mining, differentiated by its relevant phases: exploration, exploitation and processing, decommissioning, and rehabilitation. Thereby, potentially negative impacts may be largely avoided or significantly abated if intelligent and foresighted precaution is taken. At the end of this contribution, future perspectives and the pathway to sustainable mining shall be evaluated.

A Concise Review of Mining History

Mining has been with mankind for much more than 40,000 years already (Paleolithic), when

Mining and Its Environmental Impacts, Table 1 Potential (and real) environmental impact of mining on environmental compartments

Compartment	Potential environmental impact and spatial extent
Atmosphere	Release of (toxic) gases (e.g., SO_2 emissions from sulfide ore roasting, CO_2 and CFC release from aluminum processing), dusts, and aerosols: very-short- (local) to long-range transport may contaminate vegetation cover and other biota, soil, and water. Often, the burning of fossil fuels (e.g., from energy generation) has more detrimental effects than the mining operation itself
	Alteration of local air humidity: effects on local microclimate and thus biota
Hydrosphere	*Surface water*: water level fluctuations, water losses, floods, direct contamination; accidental connection of surface water and soluble ore deposits with local to small regional effects
	Groundwater: direct contamination via seepage water from aboveground or directly from within the mine operation; lowering (due to new permeabilities) or rising of water table (due to ground softening and compression), vertical fluctuations of the water table, causing local underground erosion and loss of rock stability, hydraulic filling of underground cavities and aquifers after stopping the water pumping at mine closure – affects the aquifer extension
	Drainage and seepage water: saline water, acid water, and alkaline water, each with specific toxins – affects surface and groundwaters
	Coastal and marine waters: direct pollution by spills; placer deposit mining disrupts beach systems; use of deep-sea deposits threatens (rare) marine life; mostly local effects
Pedosphere	Soil loss (open-pit and underground operations) – competition for land use; large volumes of waste-rock heaps and tailings deposits; soil contamination (and water) by water spills and seepage of contaminated waters from slag and waste heaps, tailings deposits, improper operation, etc. – local effects
Biosphere	Disturbance of ecosystems, disruption of food chains, eviction of (key) species; silicosis and in general, inhalation of fibers (asbestos mining) as a health hazard to workers and high ambient dust concentrations in the vicinity of operations – local to small regional effects
Lithosphere	Mine structure (surface: collapse, subsidence; underground: pillar breaking, slab breaking), aquifer mine operation: surface overloading, surface vibrations and shaking (blasting), mine sludge, mine tailings, slag heaps – local effects
Anthroposphere	Damage to infrastructure (transport, buildings, etc.) due to surface movements and subsidence – local effects

commodities were procured from surface and even underground deposits from various places on several continents, in order to obtain flint stones for axes and arrowheads, clay and loam for pottery and construction (e.g., [6]), or iron oxide (hematite) for cosmetic purposes [7, 8]. The interest in metals developed later, as the related human activities in metal working became so widespread that entire epochs were named accordingly (Bronze Age, Iron Age, etc.) [9, 10]. With the advent of even more sophisticated technologies in the Chinese and Roman Empires, the spectrum of sought-after elements had expanded and included components such as silver (Ag), arsenic (As), gold (Au), copper (Cu), iron (Fe), mercury (Hg), lead (Pb), and tin (Sn) minerals [11–13]. There is no doubt that mining has been pursued by man millennia before the Industrial Revolution and on all continents (except for Antarctica), and independent of the global European influences that started with their conquest in the fifteenth century [14]. Following the Industrial Revolution in the late 18th to mid-19th centuries, this range of elements not only increased, but new orders of magnitude were reached in the demand for metals and other commodities, including fast-increasing amounts of nonrenewable energy resources (coal, oil, and gas – in that sequence). Today, the timing of Peak Oil is discussed in parallel and just as intensively as the possible shortage of REE and other metals that our modern industrial society technologically and economically depend upon. At the same time, possible causes for the collapse of historical civilizations are examined, and there is evidence that mining may have contributed to such human self-destruction in the past [15].

It was largely not before post-WWII economic recovery that people in the involved industrial countries started to look at the environmental impact of these mining activities. Using European history, the cradle of the Industrial Revolution, as an example, voices from almost two millennia back deserve mentioning.

Central Europe was mainly forest-covered until the advent of the Early Middle Ages. Mining had remained a rather small-scale business, even though during Roman times (ca. 300 AD), comparatively sophisticated mining and smelting technology is already known from various places (e.g., Harz Mountains, Germany [11]). The push towards the eastern frontier by Gallic and Germanic people under the guidance of Charlemagne (around AD 800) led to new settlements, a subsequent period of significant forest removal and a new period of mining exploration and exploitation. Cities like Annaberg and Freiberg, Erzgebirge in Saxony, Germany; Kutna Hora in Bohemia (Czech Republic); and many others may serve as examples. There, silver (Ag) was found, a key resource for coin making and luxury items – much like today. These successful mining cities developed, and production increased. Around the 1500s, most of the higher-elevation forests had been cut down already, leaving vast stretches of almost tree-barren landscape – much of which was used for agriculture.

In 1557, Georgius Agricola (Latinized version of his name Georg Bauer), a German medical doctor and early allround scientist from Chemnitz in Saxony, published the first and most comprehensive book (12 volumes) on mining and its implications, “De re metallica” (about the metal issues) [46]. His book did not only describe mineral exploration techniques (even touching the use of metallophilic plants), exploitation, and smelting techniques, but also explicitly introduced the reader to the detrimental side effects of mining. He reviewed the “bad smokes” and their effects on biota; he described the barren land where no plants would want to grow and the “dead waters” where fish would no longer live or spawn [16]. Indeed, one has to imagine such metalliferous provinces in Europe as being mainly forest-free areas after a few centuries of steadily increasing mining and smelting activities. The wood was needed and used both to fuel processes and to build support structures and equipment (water wheels, water ducts, etc.) in the mines.

As of the eighteenth century, such evidence on the effects of mining became even more prominent, and more publications related to these issues emerged. In 1713, the first book on sustainability was published to introduce the concept and coin the term “sustainability” (in German *Nachhaltigkeit*) [17, 18]. Hannß Carl von Carlowitz, a Freiberg mining engineer, was also responsible for the wood supply for the local mines and noticed the increasing depletion of this valuable natural resource. Von Carlowitz wrote that a sustainable forest management was then urgently necessary to avoid (and repair) the damage resulting from mining and smelting activities if this business was to continue. As a matter of fact, those days saw a significant decline in mine productivity for various reasons: an increasing lack of wood, a steadily increasing demand for more sophisticated technologies, and permanent wars between the small states and provinces in Central Europe. Until those days, mining and smelting were done as many centuries before, with minor technological advances. This now changed rather rapidly, with first the introduction of stipends for gifted young men (non-aristocrats) to receive a higher education in mining and in 1764 with the foundation of the world’s first mining academy (then named *Bergakademie*), today known as the Technical University (TU) Bergakademie Freiberg.

Soon, new technologies were introduced that increased the efficiency of mines by reducing the water and wood (fuel) demand per volume of ore. At the same time, forestry was developed as a scientific field, and in 1811–1816, the world’s first forest academy was founded in Tharandt, Saxony. Ever since, it serves to educate future foresters and forest scientists and is today part of the Technical University Dresden. Today, the Erzgebirge is largely forested again (i.e., similar to the Harz Mountains and the Black Forest in Germany, and many other historical mining areas) with a forest cover of about 30%.

Environmental Impacts of Mining

Throughout many historical mining districts, the less noticeable centuries-old legacy of mining is still perceptible to the trained eye. It yields many helpful lessons on avoiding further environmental damage and developing sustainable mining techniques.

In general, the environmental impact of mining takes place on many levels and may affect most environmental compartments – atmosphere, hydrosphere, pedosphere, biosphere, and lithosphere, and under certain aspects even the cryosphere. Some of the key "priority pollutants" are metals that are being liberated through mining and related activities (Table 2).

Mining requires exploration to identify the exploitation potential of a mineral deposit. Related investigations may include not only geophysical (electric, electromagnetic, gravity, and seismic investigations) and geochemical work at the surface (digging of pits, trenches, or rock cuts) but also drilling activities to verify obtained results. This enables and supports 3D modeling of the ore body, a basis for reducing technical and financial risks. Following a successful exploration phase, and depending on the decision for above-ground (open-pit) or underground mining, the required mining infrastructure will be developed and ensued by rather large surface excavation or the construction of shafts and tunnels. In most cases, extensive above-ground facilities are built concurrently, which encompass infrastructure for processing, workshops, storage, and a general infrastructure of offices, remote control rooms, transport access from helicopter ports and airfields, road and train access to the electrical and water supply, ore dressing and smelting facilities, and room for waste rocks and tailings deposits.

Mining and Its Environmental Impacts, Table 2 Priority pollutants: metals from natural and mining-related sources[a]

Ag	Native metal (Ag), chlorargyrite (AgCl), acanthite (Ag_2S); Cu, Pb, Zn ores	Mining	Metallic Ag, Ag–CN complexes, Ag halides, Ag thiosulfates
As	Metal arsenides and arsenates, complex sulfide ores (arsenopyrite, FeAsS), arsenolite (As_2O_3), volcanic gases, geothermal springs	Pyrometallurgical industry, soil heaps and tailings, smelting, mine drainage	As oxides (oxyanions), organo-metallic forms, methylarsinic acid ($H_2AsO_3CH_3$), dimethylarsinic acid ($(CH_3)_2AsO_2H$)
Cd	Zn sulfide ores	Mining and smelting, mine drainage	Cd^{2+} ion, Cd halides and oxides, Cd–CN complexes, $Cd(OH)_2$ sludges
Cr	Chromite ($FeCr_2O_4$)	Pyrometallurgical industry	Metallic Cr, Cr oxides (oxyanions), Cr^{3+} complexes with organic and inorganic ligands
Cu	Native metal (Cu), chalcocite (Cu_2S), chalcopyrite ($CuFeS_2$), bornite (Cu_5FeS_4)	Mining and smelting, pyrometallurgical industry, mine drainage	Metallic Cu, Cu oxides, Cu–humic complexes, alloys, Cu ions
Hg	Native metal (Hg), cinnabar (HgS), degassing from Earth's crust and oceans	Mining and smelting, mine drainage	Organo–Hg complexes, Hg halides and oxides, Hg^{2+}, $(Hg_2)^{2+}$, Hg^0
Ni	Pentlandite ($(Fe,Ni)_9S_8$), Ni hydroxy-silicate minerals	Mining and smelting	Metallic Ni, Ni^{2+} ions, Ni amines, alloys
Pb	Galena (PbS)	Mining and smelting, mine drainage	Metallic Pb, Pb oxides and carbonates, Pb-metal–oxyanion complexes
Sb	Stibnite (Sb_2S_3), geothermal springs	Pyrometallurgical industry, smelting, mine drainage	Sb^{3+} ions, Sb oxides and halides
Se	Polymetallic base metal sulfide ores	Smelting	Se oxides (oxyanions), Se–organic complexes
Tl	Polymetallic base metal sulfide ores	Pyrometallurgical industry	TI halides, TI–CN complexes
Zn	Sphalerite (ZnS)	Mining and smelting, pyrometallurgical industry, mine drainage	Metallic Zn, Zn^{2+} ions, Zn oxides and carbonates, alloys

[a]Table modified and focused on mining-related activities after Adriano [49] and Sparks [10]

Any one of these units must be seen as an integral part of the mining activity, each with a potential environmental imprint.

Exploration Phase

In the exploration phase, already and depending on the previous land use, land has to be cleared and roads and (minor) infrastructure constructed. Climatological and local conditions define the intensity and duration of exploration activities and thus play a role in the environmental impact. Largely, exhaust fumes and dust emissions may influence air quality during this operation [19]. In general, such works and the related noise emissions have a highly restricted local impact that will stop or rapidly decrease with the end of the exploration activities. Water resources can be impacted during exploration activities by improper handling of equipment and insufficient control of exploration drilling (spillage of drilling additives, oil losses, etc.). Primarily, temporary losses in aquatic biodiversity result; hence, the shorter the operation, the easier is the recovery. Yet, related impacts may remain evident for years and even decades. Soils have a much longer "memory" for human activities. The construction of drilling platforms (pressure and surface sealing) and the drill waste materials (including potentially toxic matter) may leave imprints for many decades or even centuries (arctic environments = potential impact on the cryosphere), albeit again, on a very local scale. Bio-spheric impact may be of critical importance since it is directly related to all other environmental compartments. Here, environmental impact assessment studies may be helpful prior to starting with the mining phase. Such assessments do not necessarily impede the progress of the exploration project and principally depend on the available ecosystem or site-specific knowledge of biologists or ecologists. These evaluations are in most cases restricted in time, and recovery is possible, provided that state-of-the-art operations and precaution are applied. The crucial and well-known risks related to immediate accidents (fatalities and injuries) and health problems during the mining process itself are not dealt with here. Impact on the lithosphere is restricted to excavations and boreholes themselves and may pose challenges mainly in unstable surface and in karst environments, e.g., triggering unwanted water pathways or rock-mechanical instabilities. In general, and particularly at locations with unsuccessful exploration activities, related legacies of failed prospecting and exploration may impact future land use much later due to non-documented activities in the mining phase that may compromise the free choice of subsequent land use.

While most environmental impacts are small-scale and short-term in the exploration phase, incognizant or careless practices can lead to serious consequences. Therefore, before start-up (and beyond closure) of each mining operation, responsible exploration activities should include a priori environmental assessment studies [20]. Related important work that involves the post-mining operations, such as reclamation and rehabilitation, is an essential source of information and a major support for all subsequent activities, including these post-mining operations. In many cases, it can be responsibly performed by trained personnel of the mining company, ideally jointly with local or regional NGOs and professionals from state agencies who will also accompany the subsequent phases.

Mining or Exploitation Phase

In principle, similar impacts as described above may occur with the establishment of a full mining operation, although these are a lot more extensive and persistent. In addition, a mining phase could result in a suite of considerably more hazardous and long-lasting impacts. For most environmental compartments, the impact duration is at least as important as the strength of the impact. Mines usually have an operating lifetime of at least 10 years to many decades, a period of direct impact. Such a lengthy span of impact has the potential to leave legacies for centuries or even millennia (see section "A Concise Review of Mining History").

Atmosphere. Both open-pit and underground mines generate exhausts and considerable amounts of dust, even when properly operated [19]. Dust is generated during aboveground and underground mining, drilling, blasting, and all processes involving transferring, dumping, discharging, crushing, hauling, and processing materials. Depending on local heat, humidity, and wind conditions (local climatology), the impact on the atmosphere may be comparatively large, covering substantial areas with mainly mineral dust and furnace residues (power plants) or even with fine metal aerosols (from smelter operations). Independent of their possible direct toxicity, the settling dusts and aerosols cover plants and soil surfaces, impeding plant respiration and altering the local soil chemistry. Although dusts, aerosols, and other exhausts may travel airborne for up to several thousands of kilometers away from the source, these usually remain within a limited "halo" around the operations. Apart from these, toxic gases may also be released, e.g., sulfur dioxide from mineral sulfides, a major precursor for long-range transport species of key aerosol components (e.g., ammonium sulfate). Very large operations are known to contribute to a great extent to hemispheric pollution, e.g., the Sudbury smelter in Ontario, Canada [21]; the Freiberg smelters in Saxony, Germany [22]; and the Nikel/Zapoljarnyi and Monchegorsk operations on Kola peninsula, Russia [48]. Apart from direct metal emissions, their SO_2 emissions contributed substantially to the atmospheric formation of acidic precipitation and are largely responsible for the related major air pollution with subsequent soil and water pollution in the last decades of the twentieth century [23]. Even the carbon dioxide balance of the operation comes under close scrutiny, since many countries use carbon-trading schemes in order to benefit from implementing smart technology and to penalize big energy wasters. By their very nature, mining activities and equipment can emit high levels of noise and vibration. This usually appears to adversely affect people, including workers, more than most animal species, while no known related impact has been determined on plants.

Hydrosphere. Most mining operations demand comparatively large amounts of energy and water. For some high-volume, high-mass operations, such as coal mining, entire power plants are needed to meet the energy demand of the operations. In addition to cooling water, a water resource is needed in very many parts of the operational stages. The cooling water and its evaporation in cooling ponds or towers may influence the local microclimatology, which involves largely uncritical humidity increases. Open-pit mines may use very large amounts of water for the mining process itself (e.g., high-power water jets, air stripping, machine cooling), and to safeguard infrastructure (e.g., "constant" water spraying to suppress dust generated on haul roads and stockpiles), and particularly for the ore dressing and smelting process (milling, classification and transport as slurries, flotation processes). Additional high water consumption derives from leaching and bioleaching operations. For this reason, water demand itself can pose a major challenge, particularly in dry or semidry environments. The required lowering of groundwater levels around the mining operation (to keep the mine dry, safe, and operable) is another direct imminent impact within the area that is being dewatered throughout the era of active mining. Competition for this water supply with resident people and with terrestrial and aquatic ecosystems can be a contentious issue. In consequence, the water balance at mining sites is altered, and a persistent lowering of water tables or even diminishment of aquifers is often encountered. The described applications lead not only to water losses but potentially to hazardous water contamination. The input of polluted waste-water may directly impact biota – and indirectly, the human body. Acidic mine waters are another big issue, again mostly restricted in their spatial impact – and potentially easy to control. For decades following the 1970s, considerable attention on and inquiry into acid mine drainage (AMD) was triggered by lasting operational and environmental concerns, making it a hot topic [24, 25]. As a consequence of this acidic outflow of water, surface waters (rivers and lakes) and ground-waters may be seriously affected. Sediment pollution needs to be considered [9, 26, 27] since ample examples exist of non-retained mining materials traveling (and contaminating) hundreds of kilometers downstream (e.g., Ok Tedi mine, Papua New Guinea) [28] or of dam failures and subsequent accidents (e.g., Tysa river, Hungary) [29]. Access to clean drinking water remains a challenge in many parts of the world. This issue will remain with us for a long time to come, with a growing world human population exceeding the seven billion humans mark in 2011, and nine billion around the year 2050. While many waterborne pollutants may have a rather limited lifetime (e.g., cyanide from gold mining), persistent organic pollutants (POPs) from drilling operations, ore dressing, and energy conversion and potentially toxic metal species that may reside in aquatic systems for very long periods of time (decades to centuries) pose a lasting challenge. Such pollutants require particular attention and necessitate safeguarding against any kind of spill, leakage, and loss [30].

Pedosphere. While all mining operations require an initial removal of the natural unconsolidated land surface material (overburden), this is particularly true for open-pit operations. It is common practice to remove and store the nutrient-rich and potentially fertile top-soil separately. The deeper mineral soil material also is removed to free the deposit for active mining and stored separately. This avoids disposal of this subsoil overburden (depending on the operation), which may consist of millions of tons of rock material (usually soft and permeable). In underground mining, the equivalent to this requiring storage is the waste rock from the mining operation. Ore dressing and smelting operations produce partly extensive amounts of tailings, slags, and similar materials that need to be disposed of. Valley filling still appears to be the most sought-after option. As a result and independent of surface or underground mining, comparatively large areas that far exceed the immediate area of the mining facility may become part of the mining operation and of its environmental footprint. Valleys filled or soils covered with such "waste" materials can no longer provide their useful ecological services. Their former habitat function has ended too. While the new morphology and material will attract new life and new ecological equilibria may form (over extended periods of time), the previous ecosystem is no longer functional, and thus, profoundly and permanently impacted. If in addition, the deposited materials contain toxic components, both from the mined raw materials and from chemicals added during the beneficiation processes, these may again further enhance longer-term environmental degradation. Gold-mining legacies with related arsenic toxicity serve as an example [31]. "White mining," the mining of rock salt, further illustrates the challenges: large amounts of impure salt rock debris are being deposited on spoil tips that will persist for centuries. If not covered and not equipped with drain controls (effluent treatment), the easily dissolvable material will deliver excessive amounts of salt into adjacent soils, groundwater, and surface waters. The detrimental effects of excessive amounts of simple mineral salts on plants and many other biota are well known.

Biosphere. With the discussion on ecosystem functions in both the hydrosphere and the pedosphere, it is obvious that the biosphere is strongly impacted too. The first – and often key issue – is habitat loss. This is most certainly the most crucial and critical element of potentially very long-lasting detrimental consequences of mining operations. Although life can re-establish itself even in the most hostile and apparently devastated environments, previous ecosystems may never reestablish. Such consequences could be tolerated if it did not happen at very many places worldwide and if refuge areas did not become increasingly scarcer. Options to protect the biosphere from detrimental impacts are available but often disregarded or considered excessively expensive or demanding. In detail, again a very large array of developments and consequences emerges, depending on biome and local ecosystems. Even if individual species are being extinguished at a specific location, this loss may lead to a domino effect on the web of organisms on all levels – from microbial life via all levels of plants and insects to molluscs, amphibians, fish, reptiles, birds, and mammals. Nutrient supply may become limited due to the mostly fresh rock and overburden materials; the absence of fine materials may further inhibit the growth of higher plants (resettlement). Without further management options, recovery of such sites may take centuries.

Cryosphere. Permanent ice cover and permafrost environments yield potentially attractive mineral resources. These do not only occur at very high latitudes on both hemispheres but also at higher alpine elevations (e.g., Bolivia, Peru).

With ongoing global warming, so far mostly inaccessible areas mostly in North America, Siberia, and Greenland as well as in Argentina and Chile become potentially available and feasible for exploration and exploitation. Such environments are extremely sensitive to impacts and will remain sensitive. Their slow biogeo-chemical cycles retain negative imprints for very long periods of time, and recovery is accordingly extremely slow. Although ice or frozen ground may be compromised, mining will impact exceedingly on the water cycle, the soils (generally very shallow), and the low biodiversity (this low abundance characterizes the rather extreme vulnerability of such environments).

Lithosphere. Even the lithosphere itself can experience a lasting impact, detrimental to future use. Mining subsidence, sinkholes, and drying-up of aquifers are among the most prevalent potential environmental impacts of mining. It is well known that sinkholes may form at the surface over former underground mining operations and that mining subsidence can affect areas of hundreds of square kilometers in size. A notable example (also for a major impact on regional aquifers) is the very densely populated Ruhr area (Ruhrgebiet) in western Germany, where deep coal mining leaves its legacy [32].

Decommissioning and Recultivation Phase

In most modern mines, recultivation commences long before production ceases and the mine is abandoned. An intelligent long-term advance planning may even turn environmental legislation demands into profits. Planning recultivation and handling of environmental issues are key prior to any action. Impressive positive examples can be taken from lignite open-pit mines in the Lusatian basin in Germany (e.g., [33, 34]) and various other places. However, there are still regions where recultivation starts only after decommissioning – if it starts at all. Ever so often, mining companies claim bankruptcy at the end of the operation to save the necessary costs related to recultivation. As leading mining companies, joined in the Global Mining Initiative (GMI; [35]), actively demonstrate their responsibility, a certain fraction of the global mining enterprises still follows a different route – and contributes to the above-mentioned notion of mining being a dirty business. In most countries with a well-developed mining sector, mining companies are forced to put aside funds (usually into trust funds managed by government regulators) that will suffice for recultivation and clean-up of facilities during decommissioning, so that future land use is not compromised. Once the active mining has stopped, all facilities and infrastructures need to be dismantled, removed, and, wherever possible, recycled. Theoretically, the landscape should be returned to its original state prior to the mining-related activities. Water, soils, and biota should be able to recover rapidly. In this phase, however, disturbances are unavoidable, albeit moderate as compared with the active mining phase. The slightly suboptimal reality should be countered by a discussion of some important aspects, namely on the dimensions of scale. An unusual and generally very positive example can be taken from the German superfund site of the Wismut operations, which was a "secret" Russian uranium mining and processing operation in former East Germany. This was one of the world's largest mine closures and remediation projects, "*including five underground mines, and more than 3,700 ha of contaminated areas with ca. 500 million* m^3 *of solid, radioactively contaminated material*" [36].

Mining enterprises and related activities range from spatially highly restricted small-scale (or artisanal) mining, usually run by local people and often without appropriate training, to very large projects, mostly run by national and international companies with access to highly sophisticated equipment and technology. Such variety cannot be discussed on the same level. Sometimes, small-scale mining may be considerably less environmentally friendly by unit, but if the enterprise remains highly localized, this size restriction at the same time reduces the ecological footprint. At the other end of the scale, a very large operation that manages the site with state-of-the-art techniques may still be making an unsustainably large footprint, simply because of its sheer size. For this reason, it also appears obvious that the boundary conditions of any mine's location play a crucial role in realistically assessing the true impacts of the operation.

The Bottom Line

Mining per se must not be a devastating enterprise as it is ever so often perceived – and undoubtedly often with reason. Only if a comprehensive and open-minded environmental impact assessment is professionally performed from the very beginning and if related recommendations are followed, then the mining operation and its surrounding related activities could be regarded as a rather sustainable enterprise. A skeptic will immediately point out the related assessment costs that may well suppress any entrepreneurial activity and increase financial risks beyond feasibility. That would be a valid argument only if certain boundary conditions are not seen and met.

First and foremost, it has to be acknowledged that the twenty-first century marks the very first human generation that is capable of "seeing the global consequences" of its own activities. Prior to the development and employment of remote-sensing technologies, this awareness was outright impossible. Still, most people only perceive their immediate habitat and often make far-reaching decisions based on that limited worldview. In 2011, the world human population is the largest ever and is predicted to reach nine billion by 2050. This population increase is but one of the many global change challenges: climate change, soil and biodiversity loss, water scarcity, etc., mark a few other hotspots.

Without mining, however, the growing human population would neither be able to improve its standard of living nor maintain its well-being due to the shortage of primary raw materials essential for developing technology and building houses and infrastructure of any kind. One has only to consider the technological demands of modern medicine. The need for new materials emerges only with scientific and technological advancement. Mining will remain a necessity, since even the very best recycling rates cannot provide the amounts needed of various commodities. To avoid or at least drastically curb the damaging side effects of mining and related activities, a different approach deems necessary and paramount – the approach of sustainable mining.

Future Directions – Sustainable Mining

What Is Sustainable Mining?

The strictly regulated mining industry worldwide may strive for but can never attain a completely sustainable mining scenario. Still, when looking at related publications from the mining industry and authorities (e.g., [20, 37–40]), the notion of sustainable mining has taken a stronghold and increasingly focuses on the social and environmental issues. Sustainability has been clearly defined as having the social, the economic, and the environmental perspective in view [41]. Yet, there is a need for a strong practical bias towards environmental issues when thinking about sustainable mining. Without a "healthy" environment, there would be a rather grim future for both social and economic issues. Rajaram et al. [42] provide a helpful discussion in this respect while shying away from a distinct definition.

Simply spoken, sustainable mining is the kind of mining activity that does not compromise the future long-term well-being of people on or near sites of earlier mining. There may be a discussion on what constitutes "well-being." Hence, a pragmatic, less philosophical approach is suggested by defining "well-being" as the state of a human being where basic social and health needs are met. Since these demand a healthy environment, this three-part perspective is essential to sustainability. How difficult this may be in detail, however, has been addressed by Marker et al. [43] with various multi-scale examples, particularly from the developing world. They argue for a concept of an "*ideal sustainability model as one that minimizes negative environmental impact and maximizes benefits to society, the economy and regional/national development*." They also acknowledge the long-term character of such an approach, if taken seriously and if broad acceptance is to be achieved.

As a result, the near future will most likely see both conventional mining and also emerging new methods and technologies. These may include phyto-mining and the use of microbial assemblies to access and bring forth desired commodities without large rock and material movement. It will include in situ leaching and in situ processing of ores and will avoid the buildup of waste-rock piles and tailings deposits. It will also see a changing approach from the focus on a single or limited commodity to a broader and more long-term view that avoids producing “wastes” and rather safeguards and leaves future options open. At the same time, however, more surface operations that exploit increasingly lower concentrations of the commodities – with all potential risks involved – are seen. Both underground and surface (open-pit) mining can be done without compromising the environment for future generations. To understand such a claim, an even more complex vision needs to be developed.

The Complete Budget

The term “waste” is purposely accentuated here in quotation marks in order that those unneeded materials not be regarded as waste in the literal sense but rather as a potential future resource. It may be equally necessary not only to look at the entire mining business as an enterprise that will deliver commodities but also to address the issue in a much broader context.

Just like a water reservoir can be seen as a constructed body for the provision of drinking water or water for industrial purposes alone, it can also be seen as a multifunctional construct that potentially provides hydro energy, flood protection, fish-farming, recreation, and more opportunities and services. Obviously, these additional services may deliver a significant benefit to society. Can mining be seen and interpreted in a similar fashion? It can, although such a perspective demands a rather radical redefinition of the role of mining.

The paradigm change needed demands a complex and holistic long-term view, where a mining company plays a role as a service provider for society at large and not just as an independent private business. As a consequence, a much closer and partnership-based relation would be developed between all stakeholders: the company, the government (local, regional, or national), and the regional populace. A strategy developed by the chemical industry that serves as an example (however, which would need to be adapted to the mining sector) is the concept of eco-efficiency analysis [44, 45]. Adapted for use by the mining industry, a company would benefit from delivering additional services up to the decommissioning and possibly the rehabilitation phases. It would earn its money not only through selling a commodity to the global market but also through the complex added values, hence improving the socioeconomic situation in the region (which is often done already). It might develop post-enterprise industrial activities to ensure the subsequent benefit for the region, and could plan and establish the rehabilitation activities, based not on the minimum but the maximum possible requirements. This includes looking at mining wastes as a potential future commodity that needs to be safeguarded for easy, energy-efficient, and safe retrieval at a later time. All of these added activities generate additional costs, although if done properly, these may save a lot of future costs that are paid by the tax payer and easily excel the monetary benefit of the mining operation itself. One visionary example further illustrates this point where a back-end approach is taken. It is the complex knowledge of an ore body or reserve and its setting that drives the planning for exploiting the mine. The planning is not driven by momentary market prices (that contradict maximum resource efficiency) but by the objective and longer-term necessities and requirements for an efficient, safe, and complete utilization of all commodities in that deposit. “Waste” could be used as construction and building materials, and all toxic components could be extracted as by-products, recycled, or stored in a safe manner to serve future generations as a secondary resource.

Thus, the aim is to establish a long-term partnership and win-win situation for the benefit of all – the company, the employers and residents, and the environment – and to further the development of the region. It basically turns from a single business economy approach to a long-term perspective of political economics. With the most likely future political developments (long-term perspective) in mind, this will translate to international political economy rather than national economy.

The downside, at least as it may be perceived by a company, clearly means a much longer planning phase, the demand for early and truly open communication with all stakeholders (including risk communication), and the necessity of a much more transparent operation throughout as compared to the prevalent current standards. There are quite a few "walls to surmount" and even more prejudice and traditional concepts to overcome. Particularly, the mining industry is still largely characterized by a rather conservative approach.

The benefits are obvious: mining companies and related enterprises can no longer be perceived as obscure omnipotent malevolent entities, interested in basically nothing but the provision of industry with commodities, but as badly needed and responsible partners. Sustainable mining operations will be involved not only in the necessary acquisition and refinement of raw materials but also in the recuperation and delivery of the exploited area to future generations without compromising that future.

Such a vision is nothing short of revolutionary. Yet, it may need truly revolutionary attempts to successfully face the global challenges and to support a still growing human population – without waging wars and without turning a blind eye to extreme socioeco-nomic disparity.

Bibliography

Primary Literature

1. Schneider J (1998) Environmental impact of marine mining. N Jahrb Geol Paläont Abh 208:397–412
2. Chamley H (2003) Geosciences, environment and man. In: Chamley H (ed) Developments in earth and environmental sciences, 1. Elsevier, Amsterdam, 527 p
3. Ellis D (1989) Environments at risk. Case histories of impact assessment. Springer, Berlin/New York, 329 p
4. IRMA (2011) Documents. The initiative for responsible mining assurance. http://www.responsiblemining.net/documents.html. Accessed 8 Sept 2011
5. Kausch P, Ruhrmann G (2001) Environmental management. Environmental impact assessment of mining operations. Logabok, Köln, 133 p
6. Bednarik RG (1992) Early subterranean chert mining. Artefact 15:11–24
7. Dart RA (1967) The antiquity of mining in Southern Africa. S Afr J Sci 63(6):264–267
8. Dart RA, Beaumont PB (1968) Ratification and retrocession of earlier Swaziland iron ore mining radiocarbon datings. S Afr J Sci 64(6):241–246
9. Matschullat J, Ellminger F, Agdemir N, Cramer S, Liessmann W, Niehoff N (1997) Overbank sediment profiles – evidence of early mining and smelting activities in the Harz mountains, Germany. Appl Geochem 12:105–114
10. Sparks DL (2005) Toxic metals in the environmental: the role of surfaces. Elements 1(4):193–197
11. Klappauf L, Linke FA, Brockner W, Heimbruch W, Koerfer S (1990) Early mining and smelting in the Harz region. In: Pernicka E, Wagner GA (eds) Archaeometry, vol 90. Birkhäuser Verlag, Basel, pp 77–86
12. Rebrik BM (1987) Geologie und Bergbau in der Antike. Deutscher Verlag für Grundstoffindustrie, Leipzig, 183 p
13. Rosman KJR, Chisholm W, Hong S, Candelone JP, Boutron CF (1997) Lead from Carthagian and Roman Spanish mines isotopi-cally identified in Greenland ice dated from 600 B.C. to 300 A.D. Environ Sci Technol 31:3413–3416
14. MHN (1997) The mining history network. http://projects.exe-ter.ac.uk/mhn/. Accessed 8 Sept 2011
15. Diamond J (2005) Collapse. How societies choose to fail or survive. Penguin, London, 575 p
16. Down CG, Stocks J (1977) Environmental impact of mining. Applied Science, London, 380 p
17. von Carlowitz HC (1713) Sylvicultura oeconomica. Anweisung zur wilden Baum-Zucht. Reprint of the 1713 ed Leipzig, Braun, revised by Klaus Irmer and Angela Kießling, TU Bergakademie Freiberg and Akademische Buchhandlung, Freiberg 2000, ISBN 3-86012-115-4; Reprint of the 2nd ed from 1732, Verlag Kessel, ISBN: 978-3-941300-19-4
18. Grober U (2010) Die Entdeckung der Nachhaltigkeit. Kulturgeschichte eines Begriffs. Kunstmann Antje GmbH, 300 p
19. Plumlee GS, Ziegler TL (2005) The medical geochemistry of dusts, soils and other Earth materials. In: Sherwood Lollar B (ed) Environmental geochemistry. In: Holland HD, Turekian KK (ser eds) Treatise on geochemistry, vol 9, issue 7, pp 263–310
20. PDAC (2009) e3plus – a framework for responsible exploration, 34 p. http://www.pdac.ca/e3plus/. Accessed 8 Sept 2011
21. Gunn JM (ed) (1995) Restoration and recovery of an industrial region, Environmental management. Springer, New York, 358 p
22. Ilgen G, Fiedler HJ (1990) Smelter smoke damage at Freiberg in the 19th century, and its study by Professors Reich (Freiberg) and Stöckhardt (Tharandt) II Explaining the causes of damage by agricultural chemistry methods. Wiss Z TU Dresden 29(6):115–118
23. Last FT, Watling R (1991) Acidic deposition – its nature and impacts. Proc Royal Soc Edinburgh B Biol Sci 97:343
24. Blowes DW, Ptacek CJ, Jambor JL, Weisener CG (2005) The geochemistry of acid mine drainage. In:

Sherwood Lollar B (ed) Environmental geochemistry. In: Holland HD, Turekian KK (ser eds) Treatise on geochemistry, vol 9, issue 5, pp 149–204
25. Singer PC, Stumm W (1970) Acidic mine drainage: the rate-determining step. Science 167(3921):1121–1123
26. Knittel U, Klemm W, Greif A (2005) Heavy metal pollution downstream of old mining camps as a result of flood events: an example from the Mulde river system, eastern part of Germany. Terr Atmos Ocean Sci 16(4):919–931
27. Ridgway J, Flight DMA, Martiny B, Gomez-Caballero A, Macias-Romo C, Greally K (1995) Overbank sediments from central Mexico: an evaluation of their use in regional geochemical mapping and in studies of contamination from modern and historical mining. Appl Geochem 10:97–109
28. Pernetta JC (1988) Potential impacts of mining on the Fly river, UNEP 99, 191 p
29. Hum L, Matschullat J (2003) Gold kann schmutzig sein. Welche längerfristigen Auswirkungen hatte das Unglück bei Baia Mare auf die Theiss? In: Unland G, Herzog P (eds) Der Bergbaubezirk Baia Mare, Rumänien. Eine komplexe Betrachtung der Lagerstätte, des Bergbaus, der Aufbereitung sowie der Umweltfolgen. TU Bergakademie Freiberg, Freiberg
30. Goudie A (2006) The human impact on the natural environment, 6th edn. Blackwell, Oxford, 357 p
31. Deschamps E, Matschullat J (2011) Arsenic: natural and anthropogenic. In: Bundschuh J, Bhattacharya P (ser eds) Arsenic in the environment, vol 4. CRC Press, Balkema, 209 p
32. Bell FG, Stacey TR, Genske DD (2000) Mining subsidence and its effects on the environment: some differing examples. Environ Geol 40(1–2):135–152
33. Hüttl RF (1998) Ecology of post-mining landscapes in the Lusatian lignite mining district, Germany. In: Fox HR, Morre HM, McIntosh AD (eds) Fourth International conference of the internat affiliation of land reclamationists. Balkema, Nottingham
34. Krümmelbein J, Horn R, Raab T, Bens O, Hüttl RF (2010) Soil physical parameters of a recently established agricultural recultivation site after brown coal mining in East Germany. Soil Tillage Res 111(1):19–25
35. Littlewood G (2000) The global mining initiative. Address to Mining 2000, Melbourne September 20. www.icmm.com/document/104. Accessed 8 Sept 2011
36. Paul M, Mann S (2010) Environmental clean-up of the East German uranium mining legacy: discussion of some key experiences made under the Wismut remediation program. In: Lam E, Rowson J, Ozberk E (eds) Uranium 2010 – Proc 3 rd Internatational conference uranium, vol II, 15–18 Aug, Saskatoon, pp 481–493
37. AusIMM (2011) Australasian institute of mining and metallurgy. http://www.ausimm.com.au/. Accessed 8 Sept 2011
38. CSIRO (2011) Sustainability. Commonwealth Scientific and Industrial Research Organisation. http://www.csiro.au/sci-ence/Sustainability.html. Accessed 8 Sept 2011
39. Mining Association of Canada (2011) Towards sustainable mining. http://www.mining.ca/www/Towards_Sustaining_Mining/index.php. Accessed 8 Sept 2011
40. PDAC (2007) Prospectors and developers association of Canada. http://www.pdac.com.br/2007/english/index.htm. Accessed 8 Sept 2011
41. United Nation (1987) Report of the World commission on environment and development: our common future. http://www.un-documents.net/wced-ocf.htm (Brundtland Commission)
42. Rajaram V, Dutta S, Parameswaran K (2005) Sustainable mining practices: a global perspective. CRC Press, Baco Raton, 370 p
43. Marker BR, Petterson MG, McEvoy F, Stephenson MH (eds) (2005) Sustainable minerals operation in the developing world. Geological Society Special Publication 250, 249 p
44. Saling P, Kicherer A, Dittrich-Krämer B, Wittlinger R, Zombik W, Schmidt I, Schrott W, Schmidt S (2002) Eco-efficiency analysis by BASF: the method. Int J Life Cycle Assess 7(4):203–218
45. Shonnard DR, Kicherer A, Saling P (2003) Industrial applications using BASF eco-efficiency analysis: perspectives on green engineering principles. Environ Sci Technol 37(23):5340–5348
46. Agricola G (1556) De re metallica. Libri XII. English language version from 1912 by Hoover H and Hoover LH; ISBN 0-486-60006-8; 650 p
47. IIED (2002) Breaking new ground: mining, minerals and sustainable development. 462 p. http://www.iied.org/sustainable-markets/key-issues/business-and-sustainable-development/ mmsd-final-report. Accessed 8 Sept 2011
48. Reimann C, Äyräs M, Chekushin V, Bogatyrev I, Boyd R, de Caritat P, Dutter R, Finne TE, Halleraker JH, Jæger Ø, Kashulina G, Lehto O, Niskavaara H, Pavlov V, Räisänen ML, Strand T, Volden T (1998) Environmental geochemical atlas of the Central Barents Region. NGU-GTK-CKE Special Publication, Geological Survey Norway, Trondheim, 745 p. http://www.schweizerbart.de/publications/detail/isbn/9783510652631. Accessed 8 Sept 2011
49. Adriano DC (2001) Trace elements in terrestrial environments. Biogeochemistry, bioavailability and risks of metals. 2nd ed. Springer, New York, 867 p

Books and Reviews

Abdelouas A (2006) Uranium mill tailings: geochemistry, mineralogy, and environmental impact. Elements 2(6):335–341

Breitkreuz C, Drebenstedt C (eds) (2009) Sustainable mining and environment – a German – Latin American perspective. TU Bergakademie Freiberg, Freiberg

Einaudi MT (2000) Mineral resources: assets and liabilities. In: Ernst WG (ed) Earth systems: processes and issues, 23. Cambridge University Press, Cambridge, pp 346–372

Figueiredo BR (2000) Mine'rios e ambiente. Editora da Unicamp, Coleção Livro-Texto, 401 p

Fubini B, Fenoglio I (2007) Toxic potential of mineral dusts. Elements 3(6):407–414

Hüttl RF, Heinkele T, Wisniewski J (1996) Minesite recultivation. Springer, New York, 172 p

Maskall J, Whitehead K, Thornton I (1995) Heavy metal migration in soils and rocks at historical mining sites. Environ Geochem Health 17:127–138

Mining, People and the Environment (online magazine) http://www.mpe-magazine.com/. Accessed 8 Sept 2011

Morin G, Calas G (2006) Arsenic in soils, mine tailings, and former industrial sites. Elements 2(2):97–102

Sharma AK (no year) Scientific and sustainable mining. www.fedmin.com/html/goapaper.pdf. Accessed 8 Sept 2011

Woodward J, Place C, Arbeit K (2000) Energy resources and the environment. In: Ernst WG (ed) Earth systems: processes and issues, vol 24. Cambridge University Press, Cambridge, pp 373–401

Part VI

Earthquakes and Volcanoes

Earthquake Faulting: Ground Motions and Deformations

Ömer Aydan
Department of Civil Engineering, University of the Ryukyus, Nishihara, Okinawa, Japan

Article Outline

Glossary

Earthquake fault The fault which produces an earthquake

Ground motions The movement of ground induced by earthquakes and they involve displacement, velocity and acceleration of ground

Engineering structures Structures built by engineers for a given purpose and they generally involve buildings, civil engineering structures such as bridges, tunnels, dams, pylons etc

Damage physical harm impairing the normal function of a given object

Definition of the Subject and Its Importance

Ground motion characteristics, deformation, and surface breaks of earthquakes depend upon the causative faults. Their effects on the seismic design of engineering structures are not considered in the present codes of design although there are attempts to include in some countries (i.e., the USA, Japan, Taiwan). This chapter describes ground motions and the effect of surface ruptures associated with earthquake faulting on response and stability of engineering structures.

Introduction

Earthquakes are known to be one of the natural disasters resulting in the huge losses of human lives and properties as experienced in the recent earthquakes. Since there is no way to prevent the occurrence of earthquakes in earthquake-prone countries such as Turkey, Japan, the USA, and Taiwan, the design of structures and residential and industrial developments must be done according to possible types of earthquakes. It is well known that ground motion characteristics, deformation, and surface breaks of earthquakes depend upon the causative faults [1–13]. While many large earthquakes occur along the subduction zones, which are far from the land, and their effects appear as severe shaking, the large inland earthquakes may occur just beneath or nearby urban and industrial zones as observed in the recent great earthquakes (Fig. 1).

The seismic design of engineering structures is generally carried out by considering the possible shaking characteristics of the ground during earthquakes in a given region. It is a fact that the residual (permanent) relative displacement of the ground is not considered in any seismic code all over the world, except for very long linear structures such as pipelines. This problem is currently considered to be beyond the capability of seismic design concept for structures in the earthquake engineering, although it must be dealt with somewhat [1, 2].

In this chapter, the author first describes ground motions and deformations in view of laboratory experiments on rock samples and recordings in earthquake having different faulting mechanism. Then the effects of surface ruptures

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_906

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_906-3

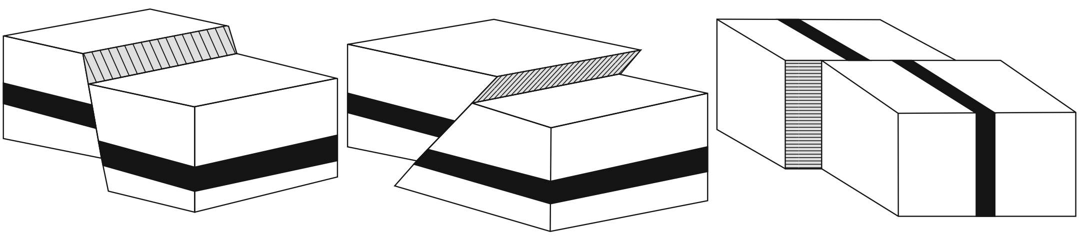
and deformations due to earthquake faulting on the response and stability engineering structures through observations in recent great earthquakes are presented and discussed.

Ground Motions

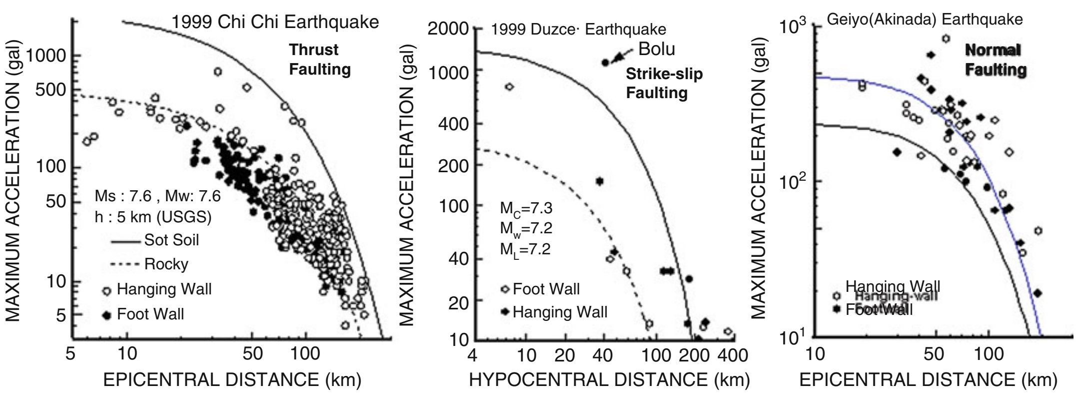
It is observationally known that the ground motions induced by earthquakes could be much higher in the hanging-wall block or mobile side of the causative fault as observed in the recent earthquakes such as the 1999 Kocaeli earthquake (strike-slip faulting), the 1999 Chi-chi earthquake (thrust faulting), the 2004 Chuetsu earthquake (blind thrust faulting), and the 2000 Shizuoka earthquake and L'Aquila earthquake (normal faulting) [7, 10, 12–15]. Figure 2 illustrates the effect of hanging-wall effect on the maximum ground accelerations observed in 1999 Chi-chi earthquake (Taiwan), 1999 Düzce earthquake (Turkey), and 2001 Geiyo earthquake (Japan) with different faulting mechanisms [9].

The recent advances in measuring, monitoring, and logging technologies enable us to measure and to monitor the dynamic responses of geomaterials during fracturing and slippage. Therefore, the studies concerning the dynamic responses of geomaterials during fracturing and slippage can now be easily undertaken as compared with that in the past. Such studies have been recently undertaken [1, 6–9, 16–19]. The experiments have been performed on geomaterials ranging from very soft materials to hard rocks such as siliceous sandstone by using different loading schemes and loading frames and specially designed stick-slip experimental device.

Aydan and Ohta [3] and Ohta and Aydan [9] reported some tests on samples of various rocks such as Ryukyu limestone, tuff, granite, porphyrite, andesite, and sandstone. Two examples are Fuji-TV No.1 and Mitake Sandstone 107 MS2 (Fig. 3).



Earthquake Faulting: Ground Motions and Deformations, Fig. 1 Fault types [1]



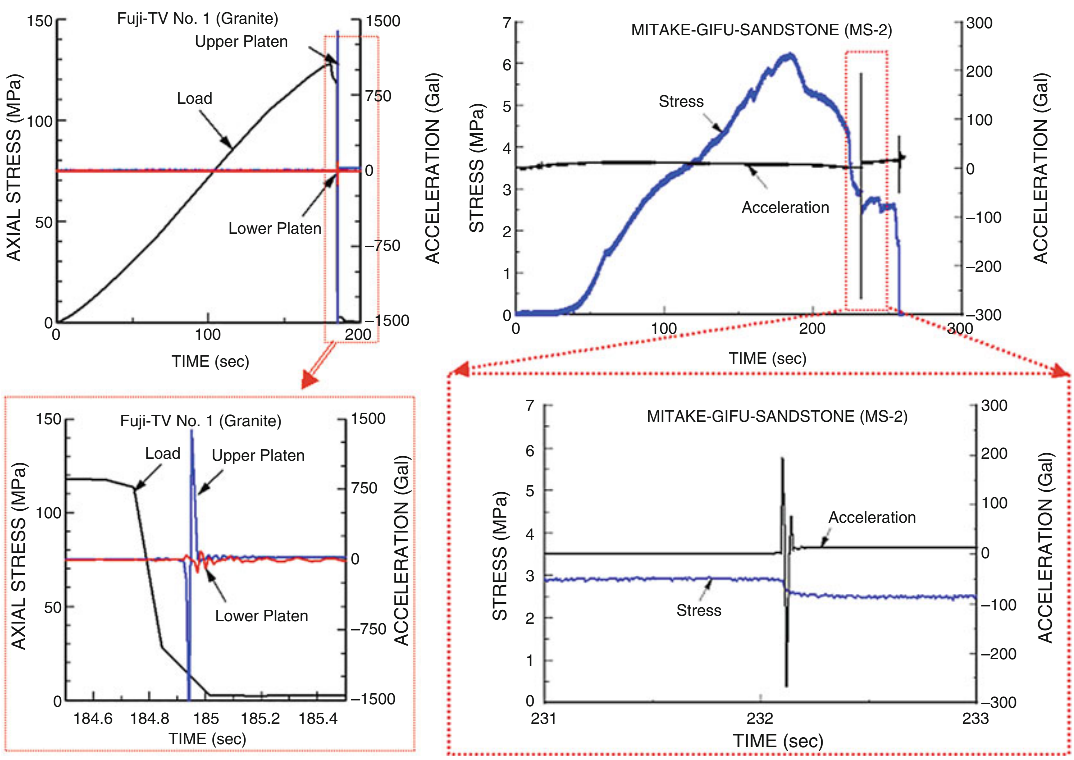
Earthquake Faulting: Ground Motions and Deformations, Fig. 2 Attenuation of maximum ground accelerations for some earthquakes

Fuji-TV No.1 granite is a prismatic sample 108 (100x100x200mm) and is quoted here. The acceleration responses start to develop when the applied stress exceeds the peak strength and it attains the largest value just before the residual state is achieved as seen in Fig. 3. Another important aspect is that the acceleration of the mobile platen is much larger than that of the stationary platen. This is a common feature in all experiments. In another words, the amplitude of accelerations of the mobile part of the loading system is higher than that of the stationary part.

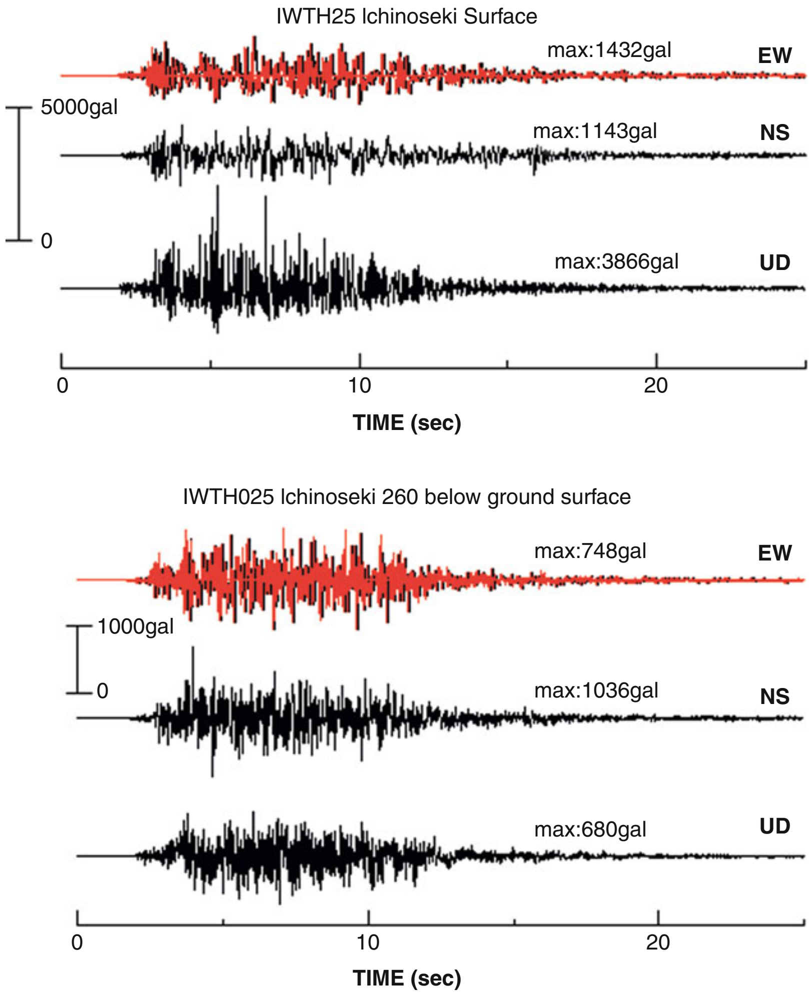
It is well known that the ground motions are generally smaller than those at ground surface. Nasu [20] carried out first instrumental studies on tunnels during the aftershock activity following 1924 Izu earthquake with a 2.4 m offset. Kanai and Tanaka [21] measured ground acceleration in underground caverns and at the ground surface. These measurements indicated that the surface acceleration was generally twice or greater than twice of that at depth as expected theoretically.

Figure 4 shows the records of accelerations at the ground surface and at bedrock 260 m below at Ichinoseki strong motion station (IWTH25) of KIK-NET [22] strong motion network of Japan measured during the 2008 Iwate-Miyagi earthquake. The strong motion station was located on the hanging-wall side of the fault, and it was very close to the surface rupture. As noted from the figure, the ground acceleration of the UD component was amplified by 5.67 times that at the bedrock. This record is also the highest strong motion recorded in the world so far.

Figure 5 shows the acceleration records measured on the ground surface (GSA) and underground gallery (GSG) during the 2009 Mw 6.3 L'Aquila earthquake [14, 23]. The GSA station is at Assergi, and the GSG station is located in an underground gallery of Gran Sasso Underground Physics Laboratory of Italy. Both stations are



Earthquake Faulting: Ground Motions and Deformations, Fig. 3 Acceleration and axial response of sandstone and granite samples



Earthquake Faulting: Ground Motions and Deformations, Fig. 4 Acceleration records at ground surface and bedrock at Ichinoseki strong motion station IWTH25 of KIK-NET in Iwate-Miyagi earthquake

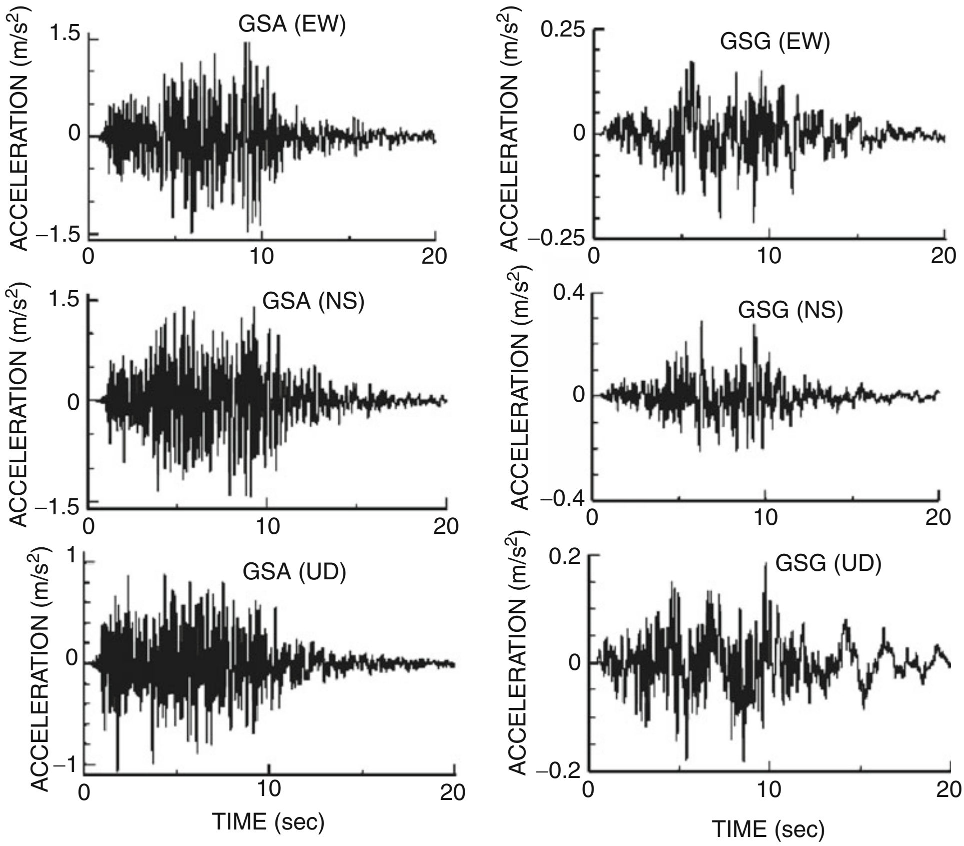
founded on Eocene limestone with a shear wave velocity of 1 km/s. Although the epicentral distances and ground conditions are almost the same, the acceleration at ground surface is amplified almost 6.4 times that in the underground gallery.

There are many empirical attenuation relations for estimating ground motions [24–28]. Including the next generation attenuation (NGA) relations, all these equations are essentially spherical or cylindrical attenuation relations, and they cannot take into account the directivity effects. As it is shown in the beginning of this section, ground motions such as maximum ground acceleration (AMAX) and maximum ground velocity (VMAX) have strong directivity effects in relation to fault orientation. Furthermore, these relations are generally far below the maximum ground acceleration, and they are incapable of obtaining the maximum ground acceleration (AMAX) or the preferred term "peak ground acceleration (PGA)."

Aydan and Ohta [4] proposed an attenuation relation by combining their previous proposals [5, 25, 26] together with the consideration of the inclination and length of earthquake fault using the following functional form (Fig. 6):

$$\alpha_{max} = F_1(V_s)^* F_2(R, \theta, \phi, L^*)^* F_3(M) \qquad (1)$$

where V_s, θ , ϕ , L^*, and M are the shear velocity of ground and the angle of the location from the



Earthquake Faulting: Ground Motions and Deformations, Fig. 5 Acceleration records at GSA and GSG strong motion stations [14]

strike and dip of the fault (measured anti-clockwise with the consideration of the mobile side of the fault) and earthquake magnitude. L^* (in km) is a parameter related to the half of the fault length.

The following specific forms of functions in Eq. (1) were put forward as:

$$F_1(V_s) = Ae^{-V_s/B} \tag{2a}$$

$$F_2(R, \theta, \phi, L^*) = e^{-R\left(1-D\sin\theta+E\sin^2\theta\right)(1+F\cos\phi)/L^*} \tag{2b}$$

$$F_3(M) = e^{M/G} - 1 \tag{2c}$$

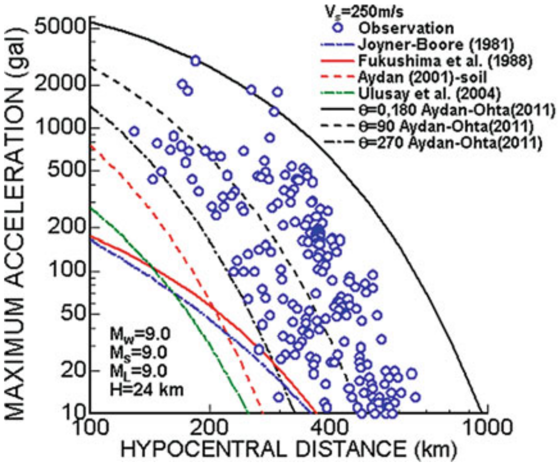
The same form can be also used for estimating the maximum ground velocity (V_{max}). The constants of the functions for maximum ground acceleration and velocity for intraplate earthquakes without consideration of faulting sense are given in Table 1.

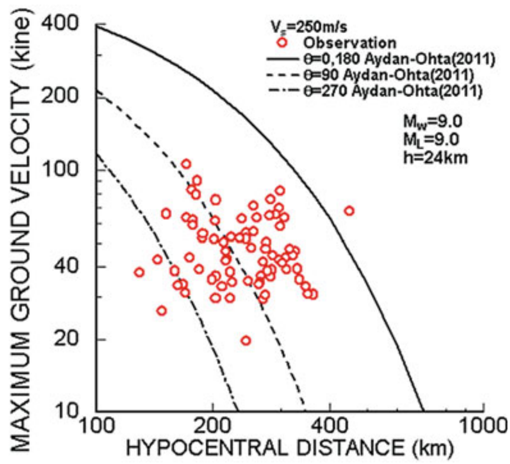
For interplate and slab earthquakes, the value of constant G is only different, and it has the value of 1.05, while other parameters remain the same.

L^*, which is approximately associated with the half of the fault length, is related to the moment magnitude in the following form:

$$L^* = a + be^{cM_w} \tag{3}$$

This parameter easily accounts the effect of fault length, and the similar form was initially proposed by Aydan et al. [18] for the effects of earthquake faults on underground structures. However, it should be noted that the attenuation characteristics of the Earth's crust varies depending upon the location and the sense of





Earthquake Faulting: Ground Motions and Deformations, Fig. 6 Comparison of estimated attenuation of maximum ground acceleration and ground velocity with observations for the 2011 Great East Japan earthquake [4]

Earthquake Faulting: Ground Motions and Deformations, Table 1 Values of constants in Eq. (5) for intraplate earthquakes

	A	B(m/s)	D	E	F	G(M) (Ms)	G(M) (Mw)
Amax	2.8	1000	0.5	1.5	0.5	1.11	1.16
Vmax	0.4	1000	0.5	1.5	0.5	1.11	1.16

Earthquake Faulting: Ground Motions and Deformations, Table 2 Values of constants in Eq. (3)

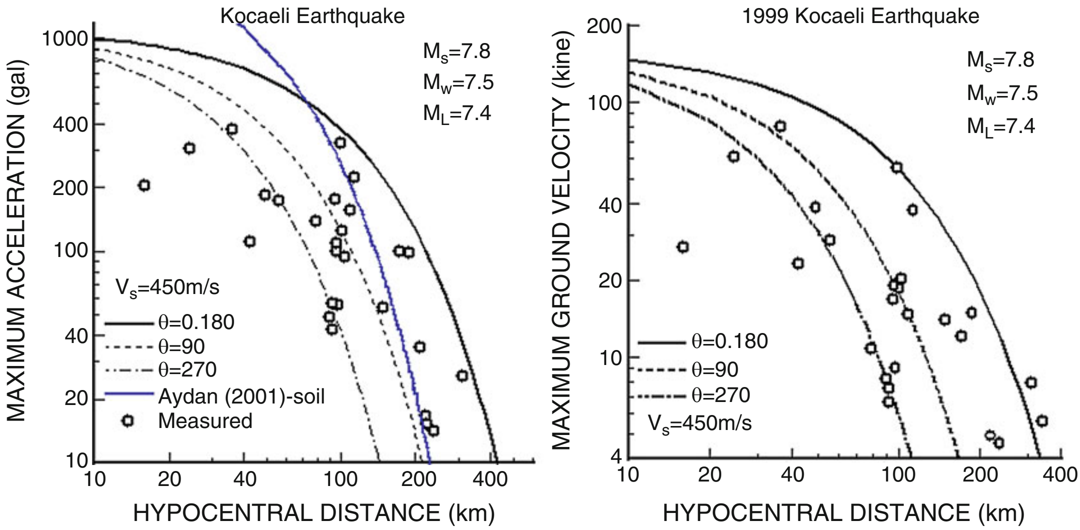
Faulting type	*a*	*b*	*c*
Normal faulting	30	0.002	1.35
Strike-slip faulting	20	0.002	1.40
Thrust faulting	30	0.003	1.45

faulting. The values of coefficients of Eq. (3) are given in Table 2.

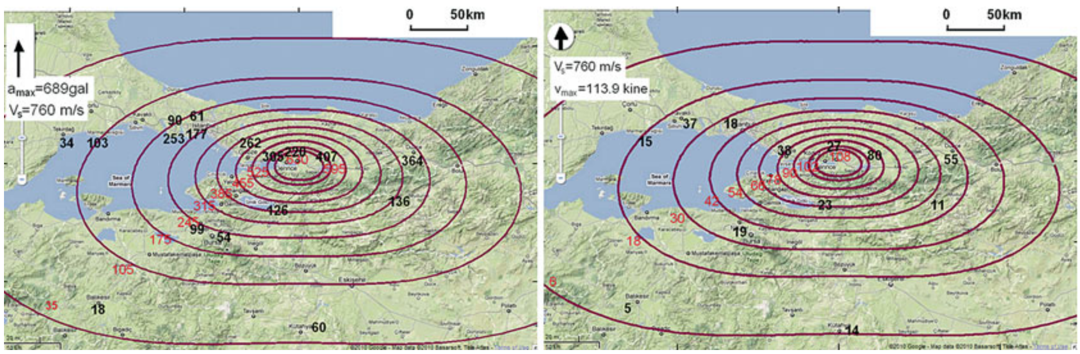
Equation (1) takes into account the location of observation points with respect to earthquake fault and ground conditions, which are fundamentally the main causes of scattering of observational data seen in spherical models for attenuation of ground motions including NGA relations [24–28].

The attenuation relation given by Eq. (1) was used to evaluate the maximum ground acceleration and ground velocity of 2011 Great East Japan Earthquake (GEJE) and 1999 Kocaeli earthquake and compared with actual observation data in Figs. 6 and 7. The same equation is used to evaluate the areal distribution of maximum ground acceleration and velocity for the Kocaeli earthquake and compared with observational data in Fig. 8. For large earthquakes, the application of Eq. (1), the estimations based the segmentation of faults, may be more appropriate. Aydan and Ohta [4] compared the estimated maximum ground accelerations for the single and double source models for the 2008 Wenchuan earthquake as shown in Fig. 9. They concluded that it would be more appropriate to utilize multiple source models based on segmentation of the causative fault for better estimations.

Strong ground motions including permanent ground deformations may also be obtained through numerical methods incorporating the fault and its rupturing process. Such an analysis was reported by Iwata et al. [29], who investigated the strong motions induced by the 2014 Nagano-Hokubu earthquake. The model is based on 3D FEM version. Figure 10a shows the fault parameters, and Fig. 10b shows the 3D mesh of the earthquake fault and its vicinity. Figure 11a shows the time histories of surface acceleration at distances of 1 and 2 km from the surface rupture in 3D FEM model. Rupture time is about 7–8 s. The maximum acceleration is higher in the east



Earthquake Faulting: Ground Motions and Deformations, Fig. 7 Comparison of estimated attenuation of maximum ground acceleration and ground velocity with observations for the 1999 Kocaeli earthquake [4]

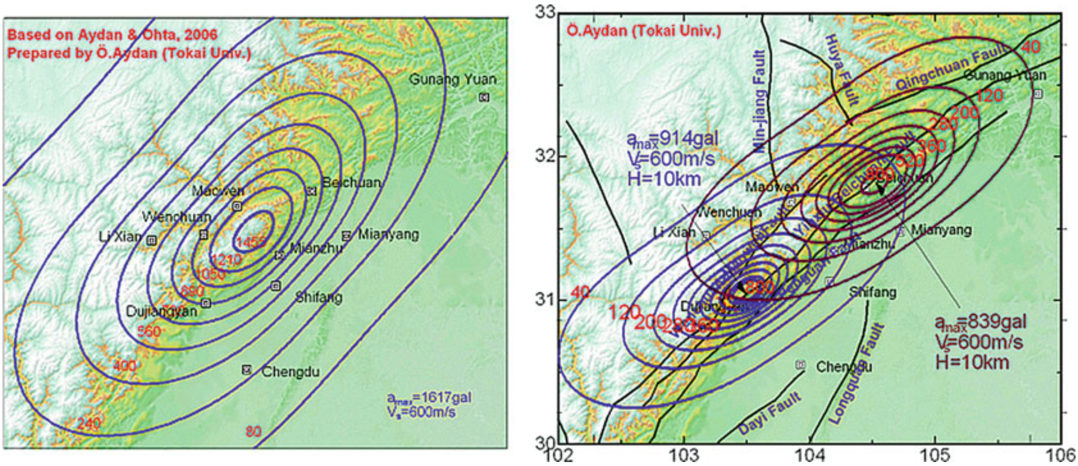


Earthquake Faulting: Ground Motions and Deformations, Fig. 8 Comparison of estimated contours of maximum ground acceleration and ground velocity with observations for the 1999 Kocaeli earthquake [4]

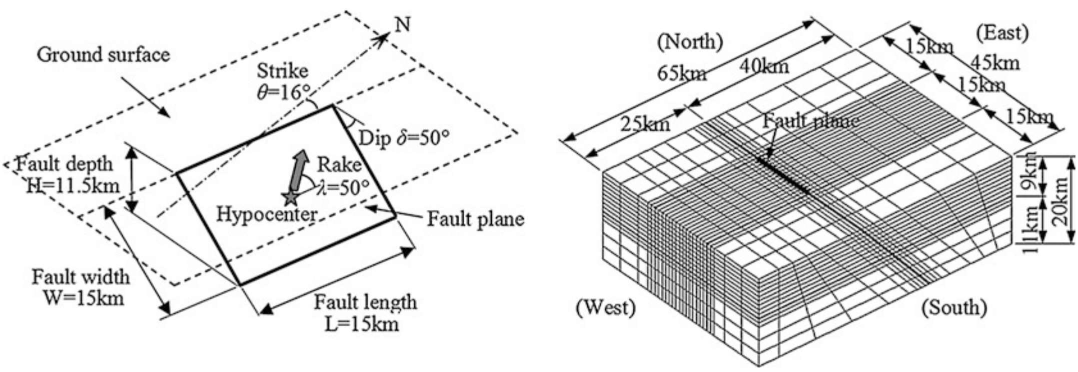
side (hanging wall) than that in the west side (footwall), which is close to the general trend observed in strong motion records. Nevertheless, the computed accelerations was less than the measured accelerations. Figure 11b shows the time histories of surface displacement at distances of 1 and 2 km from the surface rupture. The east side of the fault moves upward with respect to the footwall together with movement to the north direction, and the vertical displacement of the east side is larger than that of the footwall, and the computed results are close to the observations. However, the utilization of finer meshes would result in better simulations of ground accelerations, which would definitely require the use of the supercomputers.

Ground Deformations

The fault is geologically defined as a discontinuity in geological medium along which a relative displacement took place. Faults are broadly classified into three big groups, namely, normal faults, thrust faults, and strike-slip faults, as seen in Figs. 1 and 12. A fault is geologically defined as active if a



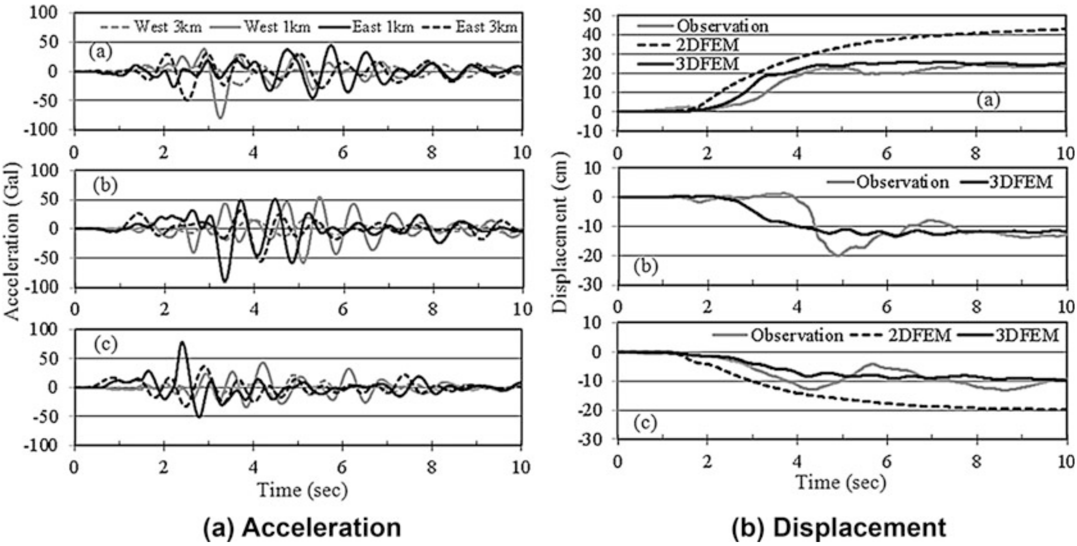
Earthquake Faulting: Ground Motions and Deformations, Fig. 9 Comparison of single and double source models for maximum ground acceleration for the 2008 Wenchuan earthquake [4]



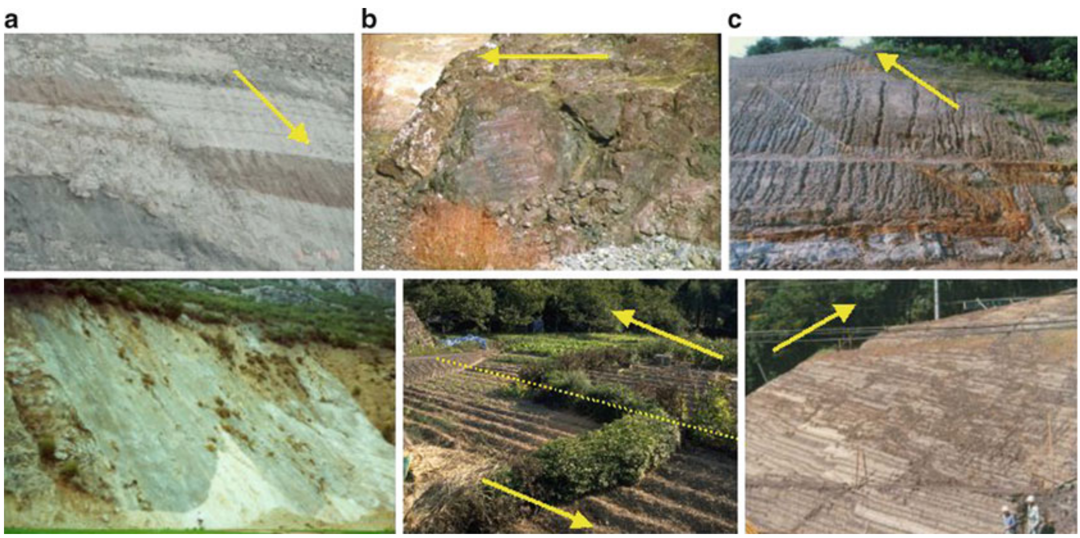
Earthquake Faulting: Ground Motions and Deformations, Fig. 10 Fault model and 3D FEM mesh [29]

relative movement took place in a period less than 2 million years. It is well known that a fault zone may involve various kinds of fractures and it is a zone having a finite volume [30]. In other words, it is not a single plane. Furthermore, the faults may have a negative or positive flower structure as a result of their trans-tensional or transpressional nature and the reduction of vertical stress near the earth surface [2]. For example, even a fault having a narrow thickness at depth may cause a broad rupture zones and numerous fractures on the ground surface during earthquakes. The appearance of ground breaks is closely related to geological structure, characteristics of sedimentary deposits, their geometry, the magnitude of earthquakes, and fault movements.

It is well known that the earth's crust is ruptured and contains numerous faults and various kinds of discontinuities and it is almost impossible to find a piece of land without faults. During the construction of structures such as tunnels, dams, power plants, roadways, railways, power transmission lines, bridges, elevated expressways, etc., it is almost impossible not to cross a fault or faults. Therefore, one of the most important items is how to identify which fault segments observed on ground surface will move or rupture during an earthquake. It is well known that a fault zone may involve various kinds of fractures as illustrated in Fig. 13 and it is a zone having a finite volume [30]. In other words, it is not a single plane. Furthermore, the faults may have a negative or



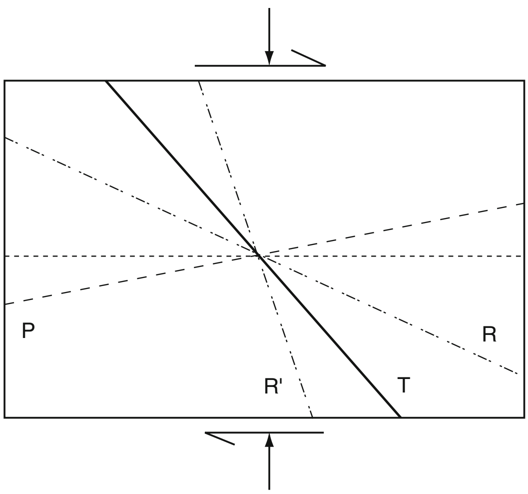
Earthquake Faulting: Ground Motions and Deformations, Fig. 11 Computed acceleration (**a**) and displacement (**b**) responses [29]



Earthquake Faulting: Ground Motions and Deformations, Fig. 12 Some examples of faulting. (**a**) Normal faulting (**b**) Strike-slip faulting (**c**) Thrust faulting

positive flower structure as a result of their trans-tensional or transpressional nature and the reduction of vertical stress near the earth surface as shown in Fig. 14. For example, even a fault having a narrow thickness at depth may cause a quite broad rupture zones and numerous fractures on the ground surface during earthquakes. Furthermore, the movements of a fault zone may be diluted if a thick alluvial deposit is found on the top of the fault (i.e., 1992 Erzincan earthquake [31]).

The recent global positioning system (GPS) also showed that permanent deformations of the ground surface occur after each earthquake (Fig. 2). The permanent ground deformation may result from different causes such as faulting, slope failure, liquefaction, and plastic deformation induced by ground shaking [18]. This type of ground deformations will have limited effect on small structures as long as the surface breaks do not pass beneath those structures. However, such deformations may cause tremendous forces on long and/or large structures. The ground deformation may induce large tensile or compression forces as well as bending stresses in structures depending upon the character of permanent ground deformations. As an example, the ground deformations reported by Reilinger et al. [32] are shown in Fig. 15, which were caused by a strike-slip fault during the 1999 Kocaeli earthquake in Turkey. Blind faults and folding processes may also induce some peculiar ground deformations and associated folding of soft overlaying sedimentary layers. Such deformations caused tremendous damage on tunnels during the 2004 Chuetsu earthquake although no distinct rupturing took place.



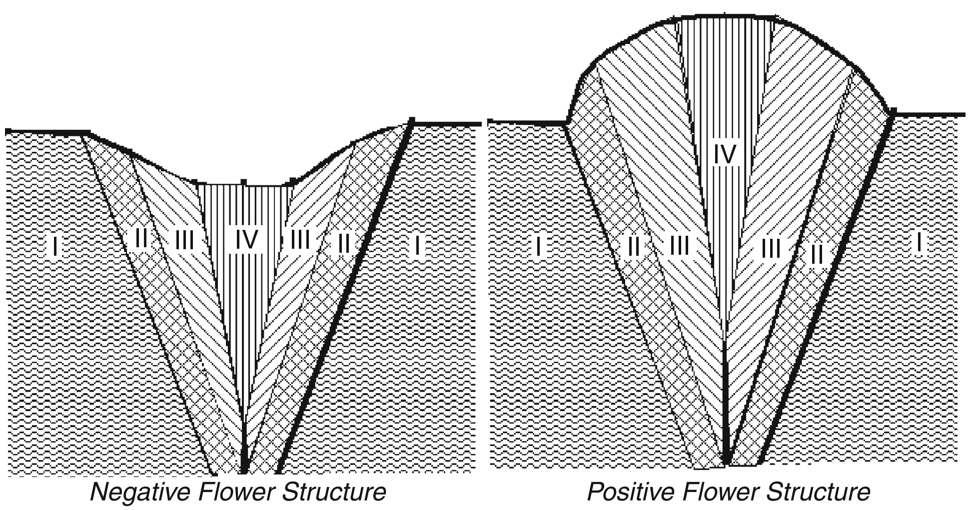
Earthquake Faulting: Ground Motions and Deformations, Fig. 13 Fractures in a shear zone or fault [30]

Effects of Surface Ruptures on Structures

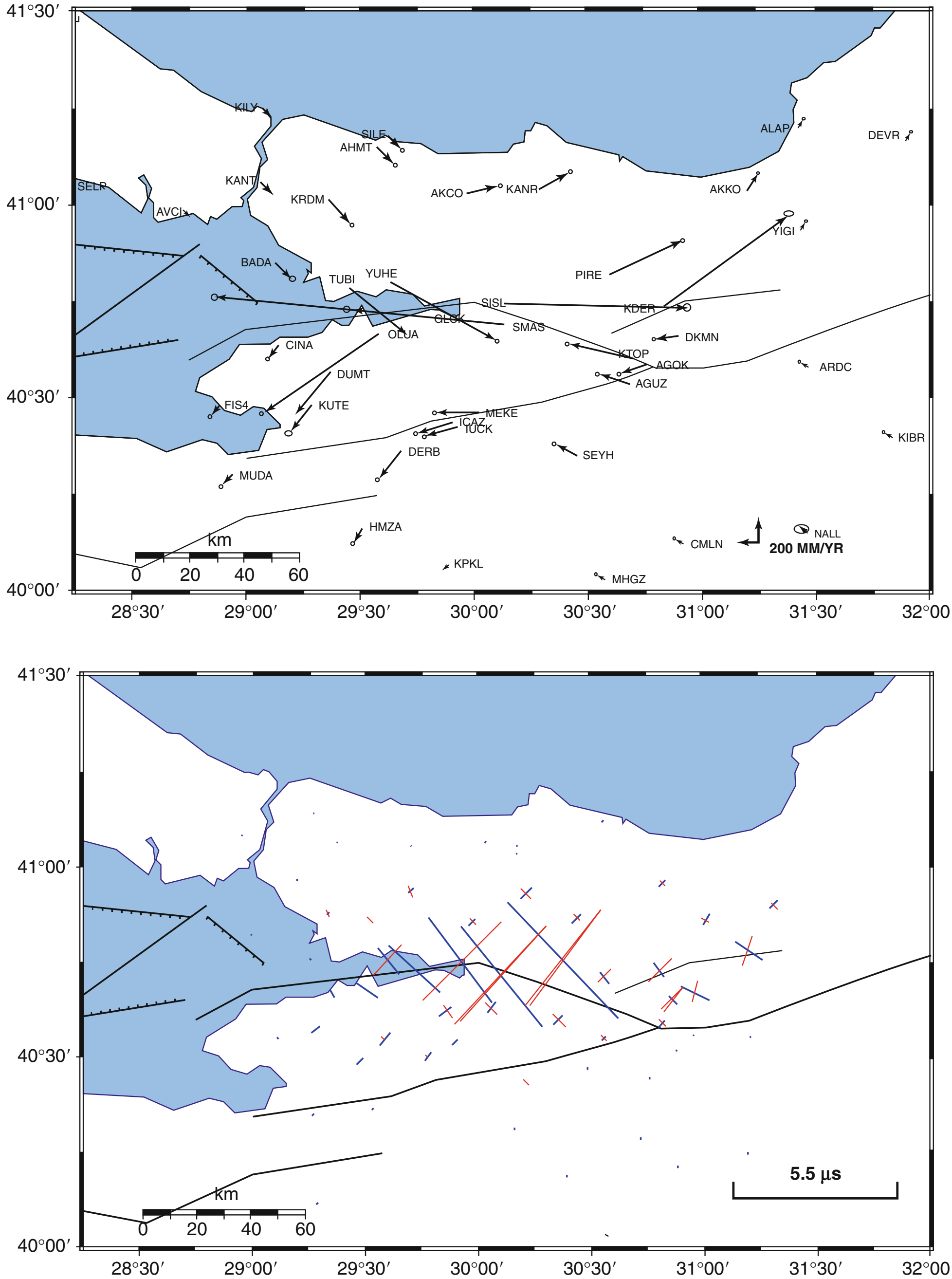
In this section, typical examples of damage to various structures induced by the fault breaks observed in recent large earthquakes since 1995 are presented, and details can be found in the quoted references [1, 2, 11, 33–51].

Roadways and Railways

The Trans-European Motorway (TEM) was damaged at three different locations by the earthquake fault caused by the 1999 Kocaeli earthquake. The motorway with east and west bounds having three



Earthquake Faulting: Ground Motions and Deformations, Fig. 14 Negative and positive flower structures due to trans-tension or transpression faulting (Modified from [2, 11, 30])

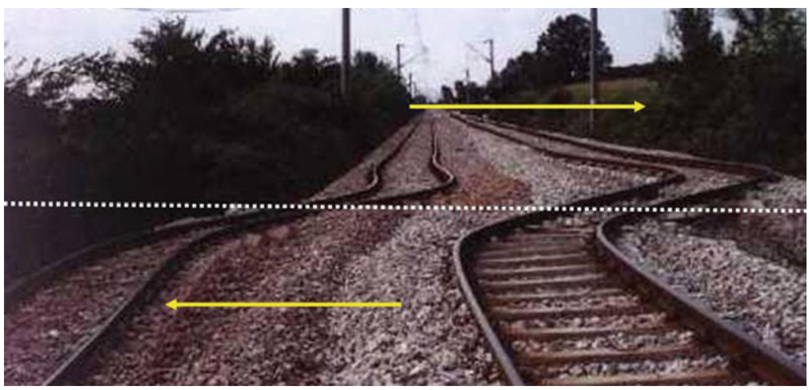


Earthquake Faulting: Ground Motions and Deformations, Fig. 15 Permanent ground deformations and associated straining induced by the 1999 Kocaeli earthquake

lanes each was slightly elevated through embankments in the earthquake-affected region. The surfacing of the motorway was damaged by rupturing and buckling as seen in Fig. 16. The railways were also built on the existing ground surface. The railways were buckled near Tepetarla station where the earthquake fault crossed the railways at an angle of $50–55^0$ with well-known "S" shape (Fig. 17).

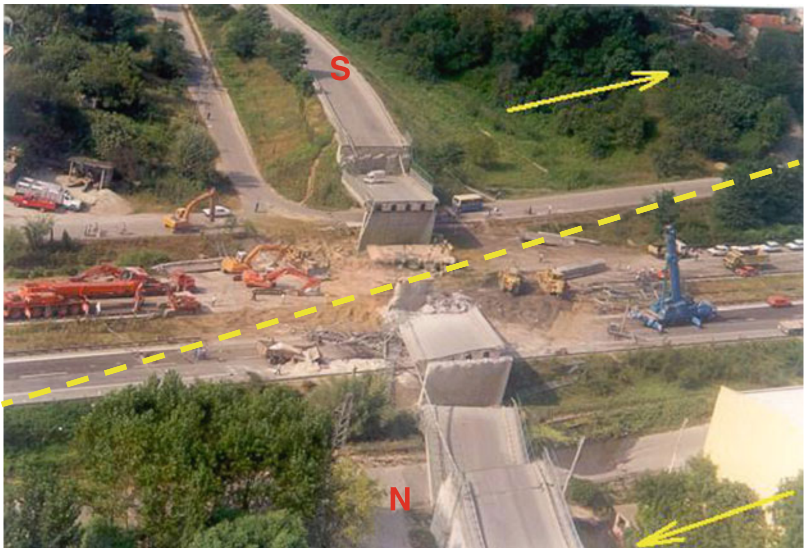
Earthquake Faulting: Ground Motions and Deformations, Fig. 16 Buckling of roadway surfacing

Earthquake Faulting: Ground Motions and Deformations, Fig. 17 Buckled railways

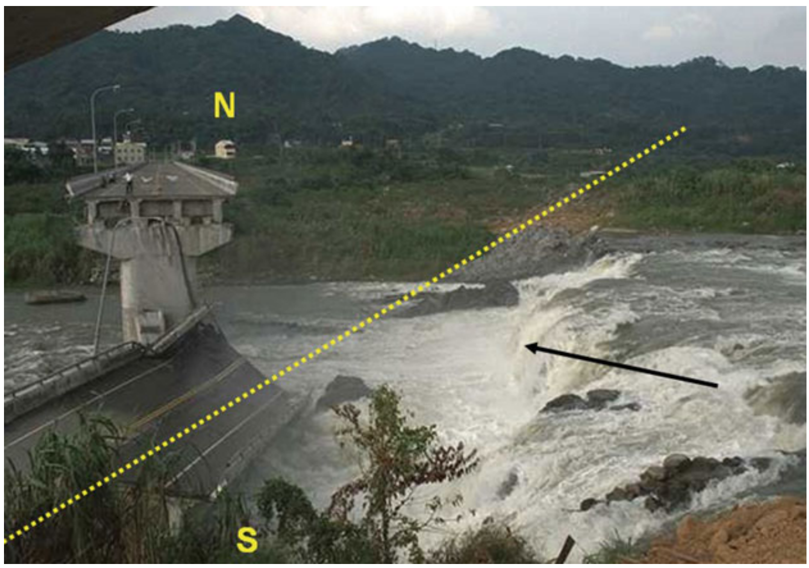


Bridges and Viaducts

Along the damaged section of the TEM motorway mentioned above, there were several overpass bridges. Among them, a four-span overpass bridge at Arifiye junction collapsed as a result of faulting (Fig. 18). The fault rupture passed between the northern abutment and the adjacent pier. The overpass was designed as a simply supported structure according to the modified AASHTO standards, and girders had elastromeric bearings. However, the girders were connected to each other through prestressed cables. The angle between the motorway and the strike of the earthquake fault was approximately 15°, while the angle between the axis of the overpass bridge and the strike of the fault was 65°. The measurements of the relative displacement in the vicinity of the fault range between 330 and 450 cm. Therefore, an average value of 390 cm could be assumed for the relative displacement between the pier and the abutment of the bridge. The Pefong Bridge collapsed due to thrust faulting in



Earthquake Faulting: Ground Motions and Deformations, Fig. 18 The collapse of the overpass



Earthquake Faulting: Ground Motions and Deformations, Fig. 19 Collapse of Pefong Bridge at Arifiye (Note the uplifted ground on RHS)

the 1999 Chi-chi earthquake, which passed between the piers near its southern abutment as seen in Fig. 19.

Dams

The Shihkang dam, which is a concrete gravity dam with a height of 25 m, was ruptured by the thrust type faulting during the 1999 Chi-chi earthquake (Fig. 20). The relative displacement between the uplifted part of the dam was more than 980 cm. Liyutan rockfill dam with a height of 90 m and a crest width of 210 m, which was on the overhanging block of Chelungpu fault, was not damaged even the acceleration records at this dam showed that the acceleration was amplified 4.5 times of that at the base of the dam (105 Gal). The deformation zone of faulting during the 2008 Wenchuan earthquake caused some damage at Zipingpu dam with concrete facing.

Earthquake Faulting: Ground Motions and Deformations, Fig. 20 Failure of Shihkang dam due to thrust faulting

Earthquake Faulting: Ground Motions and Deformations, Fig. 21 Examples of damaged portals of tunnels

Tunnels

The past experience on the performance of tunnels through active fault zone during earthquakes indicates that the damage is restricted into certain locations [18, 42–44]. Portals and the locations where the tunnel crosses the fault may be damaged as it occurred in the 2004 Chuetsu, 2005 Kashmir, and 2008 Wenchuan earthquakes (Fig. 21). A section nearby Elmalık portal of Bolu Tunnel collapsed (Fig. 22). This section of the tunnel was excavated under very heavy squeezing conditions [52]. The well-known examples of damage to tunnels at locations, where the fault rupture crossed the tunnel, are mainly observed in Japan. The Tanna fault ruptured during 1930 Kita-Izu earthquake caused damage to a railway tunnel, and the relative displacement was about 100 cm [53, 54]. The 1978 Izu-Oshima-Kinkai earthquake induced damage to Inatori railway tunnel [55, 56]. Similar type of damage with a small amount of relative displacements due to motions of Rokko, Egeyama, and Koyo faults to the tunnels of Shinkansen and subway lines through Rokko mountains were also observed [34]. During the 1999 Chi-chi earthquake, the portal of the water intake tunnels was ruptured for a distance of 10 m as a result of thrust faulting. Except this section, the tunnel was undamaged for its entire length.

Jiujiaya Tunnel is a 2282-m-long double-lane tunnel. It is 226.6 km away from the earthquake epicenter, and it is about 3–5 km away from the earthquake fault of Wenchuan earthquake. The tunnel face was 983 m from the south portal at the time of the earthquake. The concrete lining follows the tunnel face at a distance of approximately 30 m. Thirty workers were working at the tunnel face, and one worker was killed by the flying pieces of rock bolts, shotcrete, and bearing plates caused by intense deformation of the tunnel face during the earthquake [17, 18]. The concrete lining was ruptured and fallen down at several

Earthquake Faulting: Ground Motions and Deformations, Fig. 22 Collapse of Bolu Tunnel during the 1999 Düzce earthquake

Earthquake Faulting: Ground Motions and Deformations, Fig. 23 Earthquake damage at Jiujiaya Tunnel due to permanent deformations [18]

section (Fig. 23). However, the effect of the unreinforced lining rupturing was quite large and intense in the vicinity of the tunnel face. The rupturing of the concrete lining generally occurred at the crown sections although there was rupturing along the shoulders of the tunnel at several places. Furthermore, the invert was uplifted due to buckling at the middle sections.

The Kumamoto earthquake on April 16, 2016, caused heavy damage to several tunnels in the vicinity of Tateno and Minamiaso Villages. Damages to Tawarayama Roadway Tunnel and Aso Railway Tunnel and Minamiaso Tunnel were publicized [57]. The damage to Tawarayama Tunnel occurred at two locations (Fig. 24). The first damage occurred approximately 50–60 m from the west portal of the tunnel, and the concrete lining was displaced by about 30 cm almost perpendicular to tunnel axis. The heaviest damage occurred for a length of 10 m about 1600 m away from the west portal and about 460 m from the east portal. The angle between the relative movement and tunnel axis was about 20°–30°. At this location, the non-reinforced concrete lining collapsed for a length of about 5 m. Although the tunnel is located about 2 km away from the main fault, the tunnel was damaged by secondary faults associated with the trans-tension nature of the earthquake fault.

Subways

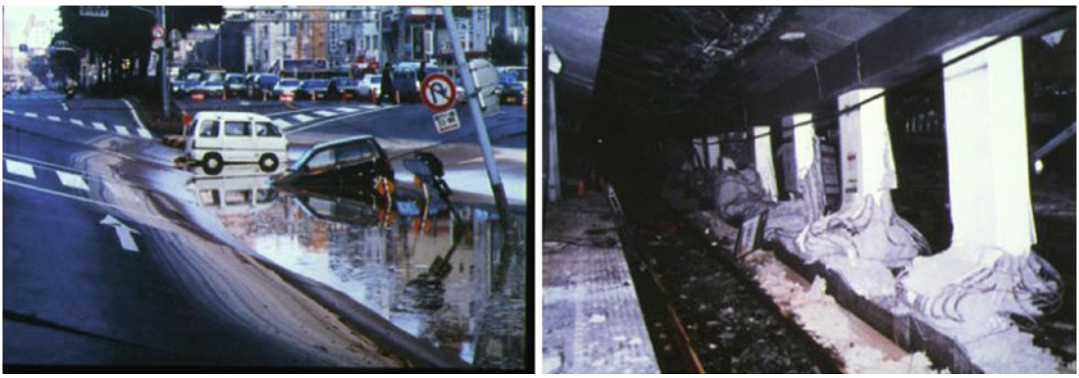
The behavior of subways in active fault zones is basically quite similar to that of tunnels. The Daikai station of the subway line in Kobe was caused by the lateral strike-slip movement of Egeyama fault just beneath this station although some tried to associate the collapse of the station with the intensity of shaking (Fig. 25). The investigation of the collapse of this station by the first author showed that the collapse was not due to shaking as the central columns of the station were subjected to torsional failure due to the permanent ground displacement, which was consistent with the lateral strike-slip movement of Egeyama fault [34].

Earthquake Faulting: Ground Motions and Deformations, Fig. 24 Views of damage and their locations at Tawarayama Tunnel [57]

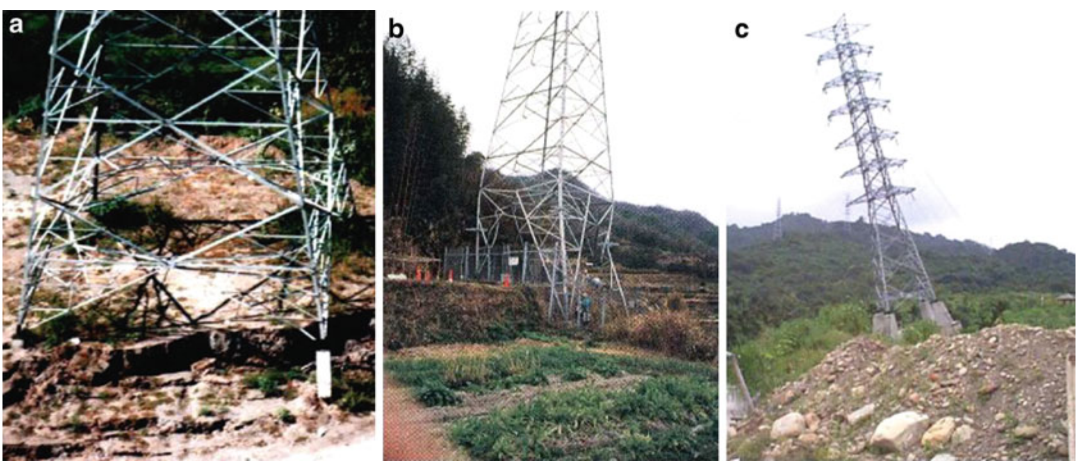
Power Transmission Lines

Power transmission lines generally consist of pylons and power transmission cables. The design of pylons and cables is generally based on the wind loads resulting from typhoons or hurricanes. The cables do not fail during earthquakes unless the pylons are toppled due to faulting, shaking, or slope failure. During the 1999 Kocaeli earthquake, only one pylon was damaged nearby Ford Otosan automobile factory at Kavaklı district of Gölcük town. At this site, a normal fault, which is a secondary fault to the main lateral strike-slip faulting event, crossed through the foundations of the pylon and its vertical throw was about 240 cm. One of the foundations of the pylon was pulled out of the ground and was exposed as seen in Fig. 26a. Some of its truss elements were slightly buckled. Similar types of damage to pylons straddling the Nojima fault break in the 1995 Kobe earthquake (Fig. 26b) and the Chelungpu fault during the 1999 Chi-chi earthquake were observed as shown in Fig. 26c.

Pylons are generally earthquake resistant due to their flexibility against ground shaking. Nevertheless, many pylons were damaged by the Kumamoto earthquake on April 2016 due to permanent deformation of their foundations as a result of ground movements induced by faulting and slope failures. Figure 27 shows several examples of damage to pylons. Although the earthquake did



Earthquake Faulting: Ground Motions and Deformations, Fig. 25 The collapse of Daikai station



Earthquake Faulting: Ground Motions and Deformations, Fig. 26 Damage to pylons due to faulting. (**a**) 1999 Kocaeli earthquake (**b**) 1995 Kobe Earthquake (**c**) 1999 Chi-Chi Earthquake

Earthquake Faulting: Ground Motions and Deformations, Fig. 27 Views of damage to pylons [57]

not cause the toppling of the pylons, they were either tilted or deformed, and some segments became loose due to rupture of the bolts.

Line-Like and Tubular Structures

Tubular structures may be specifically designated as petrol and gas pipelines, water pipes, and sewage systems. They can be also classified as line-like structures. These structures may fail either by buckling or separation during a faulting event. Five such incidents were observed during the 1999 Kocaeli earthquake (Fig. 28). One of the incidents involved the separation of a ductile iron pipe as a result of faulting near the collapsed overpass bridge. The second incident took place at the pumping facility of the Seka papermill plant at Sapanca Lake. The third incident occurred near Tepetarla village where the railways were buckled. The fourth and fifth incidents took place at Arifiye and nearby Başiskele. The fifth incident was quite important since the fault caused a heavy damage to the main water pipe having a diameter of 2 m. Similar type of failures took place in the sewage pipe networks whenever faulting breaks were observed. The natural gas pipeline lines crossing the İzmit Gulf between Yalova and Pendik were undamaged. Some brittle asbestos water pipes were also damaged in Kaynaşlı and Fındıklı due the fault rupture of the 1999 Düzce earthquake.

Buildings

Many buildings along the earthquake faults of the 1999 Kocaeli, Düzce, and Chi-chi and 2008 Wenchuan earthquakes behaved in various manner. During the site investigation of the 1999 Kocaeli earthquake, one could see either totally collapsed, severely damaged, or intact buildings just on or next to the traces of the fault breaks. The examples are many and it is quite difficult to quote all of them. Two typical examples are given and briefly discussed. Nevertheless, this topic deserves more thorough investigations. The first example is a single story reinforced concrete house with a raft foundation in Fındıklı village. The fault passed just underneath the building (Fig. 29a). The relative displacement of the fault break was about 200 cm with a 100 cm downthrow. The building was tilted, but no damage to this building was observed. A very peculiar behavior of an apartment complex consisting of eight five-story apartment blocks was observed in Kullar village as shown in Fig. 29b. Seven apartment blocks failed in a pancake mode, while one apartment block remained self-standing. One of the failed apartment blocks just crossed by the fault break which has a relative horizontal displacement of 240 and 20–25 cm vertical throw (north side down). The ground surface was sloping to the north. One of the two apartment blocks

Earthquake Faulting: Ground Motions and Deformations, Fig. 28 Damage to pipes in the 1999 Kocaeli and 1999 Düzce earthquakes

on the southern side was damaged, while the other one collapsed toward the east in a pancake mode in accordance with the movement of its foundation. The five blocks on the northern side completely collapsed in a pancake mode toward the west in accordance with the direction of shaking. Except for the apartment block over the fault break, the failure of five blocks on the northern side of the fault break may be considered to be purely due to shaking. The Bailu secondary school-reinforced concrete building survived the Wenchuan earthquake even though the surface rupture of the thrust fault passed underneath the building (Fig. 29c). Although the intensity of shaking on the southern side of the fault break should be the same, the damaged self-standing apartment block should deserve some special consideration. Whatever the reason is, it is of great interest that the most vulnerable buildings may also survive within a distance of 5–6 m to the fault break during the inland earthquakes.

Landslides and Rockfalls

The recent large earthquakes caused mega-scale slope failures and rockfalls particularly along the surface ruptures on the hanging-wall side of the fault. The slope failure induced in Beichuan town during the 2008 Wenchuan earthquake [41, 51] is of great interest. In association with the sliding motion of the earthquake fault, NW- or SE-facing slopes failed during this earthquake. There were two large-scale slope failures (landslides) in Beichuan town, which destroyed numerous buildings and facilities. The NW-facing landslide (Jingjiashan) involved mainly limestone, while the SE-facing landslide (Wangjiaya) involved phyllite (it is mudstone according to some) rock unit (Fig. 30). Limestone layers dip toward the

Earthquake Faulting: Ground Motions and Deformations, Fig. 29 The behavior of reinforced concrete buildings over surface ruptures. (**a**) Düzce earthquake (**b**) Kocaeli earthquake (**c**) Wenchuan earthquake

Earthquake Faulting: Ground Motions and Deformations, Fig. 30 Views of landslides in Beichuan

valley side with an inclination of about 30°. Furthermore, there are several faults dipping parallel to the failure surface within the rock mass. The angles of the lower and upper parts of the failed slope are 60° and 30–35°, respectively. The existence of several faults dipping parallel to the slope with an inclination of about 60–65° creates a stepped failure surface.

The SE-facing slope (Wangjiaya landslide) may involve a slippage along the steeply dipping bedding plane (fault plane?) and shearing through the layered rock mass. In other words, it may be classified as a combined sliding and shearing sliding [58]. The angles of the lower and upper parts of the failed slope are 40–45° and 30–35°, respectively. The layers dip at angle of 40° toward the valley and shearing plane are inclined at an angle of 20°.

Future Directions

The characteristics of ground motions and deformations associated with earthquake faulting are presented in this chapter. Furthermore, the response and damage to structures due to surface ruptures are illustrated through actual observations in recent great earthquakes.

The advanced strong motion attenuation relations by Aydan and Ohta [4] seem to explain the wide range of observational data, which are considered to be scattering when spherical or cylindrical symmetric attenuation relations are used. Multi-segment or multisource model may be used if surface ruptures due to earthquakes are relatively long.

Numerical models incorporating the earthquake faults and capable of simulating the rupture process can estimate both transient and permanent ground deformation as a natural output of computations. Although the displacement response could be more easily simulated, acceleration response simulation requires fine meshes, which undoubtedly require the use of supercomputers. It should be also noted that FEM models could easily simulate both permanent displacements in addition to strong ground motions.

Case histories including those of the 2016 Kumamoto earthquakes indicated that the damage to rock engineering structures might be classified as shaking-induced damage, and permanent ground deformations induced damage. Permanent ground deformation-induced damage is generally caused either by faulting or slope movements. However, the present seismic design only considers the effect of shaking, and there is an urgent necessity to include the effect of permanent straining in the seismic design.

The GPS measurements of ground deformations during earthquakes ($M > 6$) clearly indicated that permanent ground deformations do occur. The overall deformation response involves both coseismic and time-dependent components. The permanent ground deformation may result from different causes such as faulting, slope failure, liquefaction, and plastic deformation induced by ground shaking. They may cause tremendous forces on long and/or large structure underground structures such as tunnels, powerhouses, and underground storage facilities of oil, gas, and nuclear wastes. Therefore, the current seismic design codes must be revised to include these effects.

As a conclusion, *it is almost impossible for mankind to prevent the damage to structures if a fault break happens to be just passing underneath the structures*. Nevertheless, the adverse effects of fault breaks on structures can be minimized to a negligible level in view of findings from the actual examples presented in this chapter. The following suggestions may be put forward for developing the seismic codes for structures in active fault zones:

1. Roadways and railways should not be elevated, and they should be constructed on the existing ground surface in active fault zones. In doing so, the recovery of roadways and railways to be damaged by fault breaks in the event of the earthquakes would very rapid. As a result, the effective rescue and rehabilitation works would become possible.
2. Bridges and viaducts should be constructed as redundant structures and simply supported structures should be avoided. Although prestressing can be used to increase the confinement of girders along the longitudinal axis of the bridge, the structures should not be wholly prestressed in order to prevent domino action in the event of failure. Furthermore, T-type piers must not be used as they constitute top-heavy situations which are undesirable in case of faulting as well as shaking caused by earthquakes.
3. Dams should never be built on active faults defined in engineering sense as they have severe impacts on structures and residential areas in downstream. For some reasons, if they have to be built, their height should be kept to a minimum, and the type of dams should be rockfill.
4. Observations and studies on various underground openings showed that they are strong against shaking. Nevertheless, the existence of discontinuities makes them vulnerable to collapses particularly in case of shallow underground openings. This may have some important implications on areas where shallow abandoned mines, underground shelters, and old tunnels exist. As the damage to tunnel linings due to fault breaks would be localized, there is no need for extra precautions at those locations. Underground openings, crossing faults, and fracture zones may be enlarged to accommodate relative slips along faults and fracture zones. The lining of the openings should be ductile to accommodate permanent

ground deformations at such zones. Furthermore, the brittle linings of the existing underground structures should be lined with ductile thin plates or fiber-reinforced polymers together with rock bolts at fracture and fault zones, where permanent ground deformations may occur.

5. Tubular structures should be designed with the use of flexible joints to accommodate relative displacements with the consideration of likely relative displacement of the ground.
6. Buildings should be designed as rigid boxlike structures with the use of shear walls. Walls with fragile hollow bricks having no structural strength must not be allowed to be used in building construction. If the brick walls are to be used, the walls should be constructed before the construction of columns and beams in order to attain the structural integrity of the structure, and they should be light, resistant, and ductile.
7. Vertical accelerations of inland earthquakes could be 0.65–0.75 times the maximum horizontal accelerations. Sometimes, they may exceed even the maximum horizontal acceleration. The effect of vertical accelerations must be considered in structural design.
8. The structures of great importance should be sited on the rock foundations rather than on loose alluvial ground since the amplification of the loose ground may be several times that of the rock foundation. Therefore, the failure risk of structures during earthquakes could be drastically reduced by such a selection. The *nuclear power plants* and even arch dams built on rock foundation performed very well without any damage during the recent great earthquakes.
9. Global, regional, and local geological and geotechnical maps of cities and towns for residential and industrial developments in seismically active fault zones should be prepared. Furthermore, the detailed geological and geotechnical investigations should be required before any new development so that it may be possible to reduce the seismic risks to structures during earthquakes in active fault zones.

Bibliography

1. Aydan Ö (2003) Actual observations and numerical simulations of surface fault ruptures and their effects engineering structures. The Eight U.S.-Japan Workshop on Earthquake Resistant Design of Lifeline Facilities and Countermeasures Against Liquefaction. Technical Report, MCEER-03-0003, pp 227–237
2. Aydan Ö, Ulusay R, Hasgür Z, Hamada M (1999) The behaviour of structures built in active fault zones in view of actual examples from the 1999 Kocaeli and Chi-chi Earthquakes. ITU-IAHS International Conference on the Kocaeli Earthquake 17 August 1999: A Scientific Assessment and Recommendations for Re-building, Istanbul, pp 131–142
3. Aydan Ö, Ohta Y (2006) The characteristics of ground motions in the neighbourhood of earthquake faults and their evaluation. Symposium on the Records and Issues of Recent Great Earthquakes in Japan and Overseas, EEC-JSCE, Tokyo, 114–120
4. Aydan Ö, Ohta Y (2011) A new proposal for strong ground motion estimations with the consideration of characteristics of earthquake fault. Seventh National Conference on Earthquake Engineering, Istanbul
5. Aydan Ö (2007) Inference of seismic characteristics of possible earthquakes and liquefaction and landslide risks from active faults (in Turkish). The 6th National Conference on Earthquake Engineering of Turkey, Istanbul, vol 1, pp 563–574
6. Ohta Y, Aydan Ö (2004) An experimental study on ground motions and permanent deformation nearby faults. J Sch Mar Sci Technol 2(3):1–12
7. Ohta Y (2011) A fundamental research on the effects of ground motions and permanent ground deformations nearborhoud earthquake faults on civil engineering structures (in Japanese). Doctorate Thesis, Graduate School of Science and Technology, Tokai University, 272 pp
8. Ohta Y, Aydan Ö (2009) An experimental and theoretical study on stick-slip phenomenon with some considerations from scientific and engineering viewpoints of earthquakes. J Sch Mar Sci Technol 8(3): 53–67
9. Ohta Y, Aydan Ö (2010) The dynamic responses of geomaterials during fracturing and slippage. Rock Mech Rock Eng 43(6):727–740
10. Chang T-Y, Cotton F, Tsai Y-B, Angelier J (2004) Quantification of Hanging-Wall Effects on Ground Motion: Some Insights from the 1999 Chi-chi Earthquake. Bull Seismol Soc Am 94(6):2186–2197
11. Ulusay R, Aydan Ö, Hamada M (2002) The Behavior of structures built on active fault zones: examples from the recent earthquakes of Turkey. Struct Eng Earthq Eng JSCE 19(2):149–167
12. Somerville PG, Smith NF, Graves RW, Abrahamson NA (1997) Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity. Seismol Res Lett 68:199–222

13. Tsai YB, Huang MW 2000 Strong ground motion characteristics of the Chi-chi Taiwan earthquake of September 21, 1999. 2000 NCHU-Waseda Joint Seminar on Earthquake Engineering, 17–18 July 2000, Taichung, vol 1, pp 1–32
14. Aydan Ö, Kumsar H, Toprak S, Barla G (2009) Characteristics of 2009 l'Aquila earthquake with an emphasis on earthquake prediction and geotechnical damage. J Mar Sci Technol Tokai Univ 9(3):23–51
15. Abrahamson NA, Somerville PG (1996) Effects of the hanging wall and footwall on ground motions recorded during the Northridge earthquake. Bull Seismol Soc Am 86(1B):593–599
16. Aydan Ö (2003) An experimental study on the dynamic responses of geomaterials during fracturing. J Sch Mar Sci Technol Tokai Univ 1(2):1–7
17. Aydan Ö, Ohta Y, Geniş M, Tokashiki N, Ohkubo K (2010) Response and Earthquake induced Damage of Underground Structures in Rock Mass. J Rock Mech Tunnel Technol 16(1):19–45
18. Aydan Ö, Ohta Y, Geniş M, Tokashiki N, Ohkubo K (2010) Response and stability of underground structures in rock mass during earthquakes. Rock Mech Rock Eng 43(6):857–875
19. Aydan Ö, Daido M, Tokashiki N, Bilgin A, Kawamoto T (2007) Acceleration response of rocks during fracturing and its implications in earthquake engineering. 11th ISRM Congress, Lisbon, vol 2, pp 1095–1100
20. Nasu N (1931) Comparative studies of earthquake motions above ground and in a tunnel. Bull Earthq Res Inst Tokyo Univ 9:454–472
21. Kanai K, Tanaka T (1951) Observations of earthquake motion at different depths of the earth. Bull Earthq Res Inst Tokyo Univ 28:107–113
22. K-NET and KiK-Net (2007, 2008) Digital acceleration records of earthquakes since 1998. http://www.k-net.bosai.go.jp/ and http://www.kik.bosai.go.jp/
23. Kawashima K, Aydan Ö, Aoki T, Kishimoto I, Konagai K, Matsui T, Sakuta J, Takahashi N, Teodori SP, Yashima A (2010) Reconnaissance Investigation on the Damage of the 2009 L'Aquila, Central Italy Earthquake. J Earthq Eng 14:817–841
24. Joyner WB, Boore DM (1981) Peak horizontal acceleration and velocity from strong motion records from the 1979 Imperial Valley California Earthquake. Bull Seismol Soc Am 71(6):2011–2038
25. Aydan Ö, Sezaki M, Yarar R (1996) The seismic characteristics of Turkish Earthquakes. The 11th world conference on earthquake Engineering, CD-2, Paper No 1270
26. Aydan Ö (2001) Comparison of suitability of submerged tunnel and shield tunnel for subsea passage of Bosphorus (in Turkish). Geol Eng J 25(1):1–17
27. Campbell KW (1981) Near source attenuation of peak horizontal acceleration. Bull Seismol Soc Am 71(6): 2039–2070
28. Ambraseys NN (1988) Engineering Seismology. Earthq Eng Struct Dyn 17:1–105
29. Iwata N, Adachi K, Takahashi Y, Aydan Ö, Tokashiki N, Miura F (2016) Fault rupture simulation of the 2014 Kamishiro Fault Nagano Prefecture Earthquake using 2D and 3D–FEM. EUROCK2016, Ürgüp, pp 803–808
30. Aydan Ö, Shimizu Y, Akagi T, Kawamoto T (1997) Development of fracture zones in rock masses. International symposium on deformation and progressive failure in geomechanics, IS-NAGOYA, pp 533–538
31. Hamada M, Aydan Ö (1992) A report on the site investigation of the March 13 Earthquake of Erzincan, Turkey. ADEP, Association for Development of Earthquake Prediction, 86 pp
32. Reilinger RE, Ergintav S, Burgmann R, McClusky S, Lenk O, Barka A, Gürkan O, Hearn L, Feigl KL, Çakmak R, Aktug B, Özener H, Toksöz MN (2000) Coseismic and postseismic fault slip for the 17 August 1999, M = 7.5, Izmit, Turkey Earthquake. Science 289
33. Asakura T, Sato Y (1998) Mountain Tunnels Damage in the 1995 in Hyogo-ken Nanbu Earthquake. 39(1), Railway Technical Research Institute (RTRI), pp 9–16
34. Aydan Ö (1996) Faulting and characteristics of earthquake waves in Hyogo-ken Nanbu Earthquake of January 17, 1995 (in Turkish). Jeoloji Mühendisliği 48:63–77
35. Aydan Ö, Kawamoto T (2004) The damage to abandoned lignite mines caused by the 2003 Miyagi-Hokubu earthquake and some considerations on its causes. 3rd Asian Rock Mechanics Symposium, Kyoto, pp 525–530
36. Aydan Ö, Kumsar H (2010) An Experimental and Theoretical Approach on the Modeling of Sliding Response of Rock Wedges under Dynamic Loading. Rock Mech Rock Eng 43(6):821–830
37. Aydan Ö, Hamada M (2006) Damage to Civil Engineering Structures by Oct. 8, 2005 Kashmir Earthquake and Recommendations for Recovery and Reconstruction. J Disaster Res 1(3):1–9
38. Aydan Ö, Hamada M, Suzuki Y (2005) Some observations and considerations on the damage induced by the tsunami of the 2004 Sumatra earthquake on structures and coast. J Sch Mar Sci Technol 3(1):79–94
39. Aydan Ö, Miwa S, Kodama H, Suzuki T (2005) The Characteristics of M8.7 Nias Earthquake of March 28, 2005 and Induced Tsunami and Structural Damages. J Sch Mar Sci Technol Tokai Univ 3(2):66–83
40. Aydan Ö, Ohta Y, Hamada M (2009) Geotechnical evaluation of slope and ground failures during the 8 October 2005 Muzaffarabad earthquake in Pakistan. J Seismol 13(3):399–413
41. Aydan Ö, Hamada M, Itoh J, Ohkubo K (2009b) Damage to civil engineering structures with an emphasis on rock slope failures and tunnel damage induced by the 2008 Wenchuan earthquake. J Disaster Res 4(2): 153–164
42. Hashimoto S, Miwa K, Ohashi M, Fuse K (1999) Surface soil deformation and tunnel deformation caused by the September 3, 1998, Mid-North Iwate

Earthquake. 7th Tohoku Regional Convention, Japan Society of Engineering Geology
43. Yashiro K, Kojima Y, Shimizu M (2007) Historical earthquake damage to tunnels in Japan and case studies of railways tunnels in the 2004 Niigata-ken Chuetsu earthquake. QR of RTRI 48(3):136–141
44. Asakura T, Shiba Y, Sato Y, Iwatate T (1996) Mountain tunnels performance in 1995 Hyogo-ken Nanbu Earthquake. Special Report of the 1995 Hyogo-ken nanbu Earthquake, Committee of Earthquake Engineering, JSCE
45. Aydan Ö, Kumsar H (1997) A site investigation of Oct. 1, 1995 Dinar Earthquake. Turkish Earthquake Foundation,TDV/DR 97–003
46. Aydan Ö, Ulusay R, Kumsar H, Sönmez H, Tuncay E (1998) A site investigation of Adana-Ceyhan Earthquake of June 27, 1998. Turkish Earthquake Foundation, TDV/DR 006–30, 131 pp
47. Aydan Ö, Ulusay R, Hasgür Z, Taşkın B 1999 A site investigation of Kocaeli Earthquake of August 17, 1999. Turkish Earthquake Foundation, TDV/DR 08–49, 180 pp
48. Aydan Ö, Ulusay R, Kumsar H, Tuncay E 2000a Site investigation and engineering evaluation of the Düzce-Bolu Earthquake of November 12, 1999. Turkish Earthquake Foundation, TDV/DR 095–51, 307pp
49. Ulusay R, Aydan Ö, Erken E, Kumsar H, Tuncay E, Kaya Z (2003) Site Investigation and Engineering Evaluation of the Cay-Eber Earthquake of February 3, 2002. Turkish Earthquake Foundation, TDV/DR 012–79. 213 pp
50. Wang WL, Wang TT, JJ S, Lin CH, Seng CR, Huang TH (2001) Assessment of damage in mountain tunnels due to the Taiwan Chi-chi earthquake. Tunn Undergr Space Technol 16:133–150
51. Aydan Ö, Ohta Y, Hamada M, Ito J, Ohkubo K (2009) The response and damage of structures along the fault rupture traces of the 2008 Wenchuan Earthquake. International Conference on Earthquake Engineering: the 1st Anniversary of Wenchuan Earthquake, Chengdu, pp 625–633
52. Aydan Ö, Dalgıç S, Kawamoto T (2000) Prediction of squeezing potential of rocks in tunnelling through a combination of an analytical method and rock mass classifications. Ital Geotech J 34(1):41–45
53. Kuno H (1935) The geologic section along the Tanna Tunnel. Bull Earthq Res Inst Univ Tokyo 14:92–103
54. Sakurai T (1999) A report on the earthquake fault appearing in the Tanna tunnel under construction by North-Izu Earthquake 1930 (in Japanese). J Japan Soc Eng Geol 39(6):540–544
55. Kawakami H (1984) Evaluation of deformation of tunnel structure due to Izu-Oshima Kinkai earthquake of 1978. Earthq Eng Struct Dyn 12(3):369–383
56. Tsuneishi Y, Ito T, Kano K (1978) Surface faulting associated with the 1978 Izu-Oshima- Kinkai earthquake. Bull Earthq Res Inst, Univ Tokyo 53:649–674
57. Aydan Ö, Tomiyama J, Matsubara H, Tokashiki N, Iwata N (2017) The characteristics of damage to rock engineering structures induced by the 2016 Kumamoto earthquakes. Proceedings of the 14th Japan Rock Mechanics Symposium, Japan Society for Rock Mechanics, January 2017, IRMS044, 6 pp
58. Aydan, Ö., Y. Shimizu and T. Kawamoto (1992). The stability of rock slopes against combined shearing and sliding failures and their stabilisation. Int. Symp. on Rock Slopes, New Delhi, 203–210.

Books and Reviews

Fowler CMR (1990) The solid earth – an introduction to global geophysics. Cambridge University Press, Cambridge

Okamoto S (1973) Introduction to earthquake engineering. University of Tokyo Press, Tokyo

Yeats RS, Sieh K, Allen CR (1997) The geology of earthquakes. Oxford University Press, Oxford, NY

Induced Seismicity

Caitlin Barnes[1] and Todd Halihan[2]
[1]Educational Research, Oklahoma State University, Stillwater, OK, USA
[2]Boone Pickens School of Geology, Oklahoma State University, Stillwater, OK, USA

Article Outline

Glossary

Basement Igneous or metamorphic rocks, often Precambrian, that unconformably underlie unmetamorphosed sedimentary strata [75]

Baseline pressure A starting point of pressure for comparisons [37]

Bottomhole The lowest point in a well [37]

Fault A fracture along which rocks have been displaced in a horizontal, vertical, or oblique sense [75]

Formation A body of rock strata that consists of a certain lithology or combination of lithologies; a lithological unit [73]

Hypocenter The point within the earth where an earthquake rupture starts. Also commonly termed the focus [38]

Injection interval A formation or aquifer "used for the injection of fluids for any purpose, including artificial recharge and waste disposal" [73]

Reservoir A highly porous and permeable mass of rock that is able to hold or transmit fluids [75]

Storativity The volume of water an aquifer releases from or takes into storage per unit surface area of the aquifer per unit change in head. It is equal to the product of specific storage and aquifer thickness [18]

Definition of the Subject

Induced seismicity due to fluid injection is a phenomenon observed scientifically beginning in the 1960s. Since 2009, the USA has experienced a significant increase in seismicity in the mid-continent region, also known as the Central and Eastern United States (CEUS). The increase in earthquakes is associated with recent changes in hydrocarbon production methods. Fluid-induced earthquakes are generated when a fault is exposed to an increase in pressure from fluids, typically wastewaters, which are injected into a nearby disposal well. Research is continuing to determine the best management practices to avoid these events. This topic does not include seismicity generated by other anthropogenic sources nor injection seismicity due to geothermal or carbon sequestration sources.

Introduction

A combination of advancements in the hydrocarbon recovery process and an increase in magnitude of industrial fluid injection during the early twenty-first century has played a role in the increase in US seismic activity, primarily within the Central and Eastern United States (CEUS) zone [28, 67]. To address the concerns of increased seismicity within these areas, it is first important to describe the origin and effects of industrial fluid processes and how they relate to induced seismicity.

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_976

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_976-1

Hydraulic Fracturing

Hydraulic fracturing, also referred to as fracking, is a technique used to enhance the production of natural resources, predominately natural gas. Hydraulic fracturing was first introduced as a new technique in the USA in the 1940s [35]. Hydraulic fracturing is a method used to stimulate or treat production wells in unconventional reservoirs [27, 35, 37]. Traditionally, oil and natural gas are trapped in large pockets of sedimentary rock layers, where the resource accumulates in "pools" within the reservoir. In unconventional reservoirs, the oil and natural gas are trapped within the pore spaces of sedimentary rocks, such as shale and porous limestone. Hydraulic fracturing is a nontraditional method allowing the extraction of natural gas otherwise inaccessible by traditional drilling methods [27].

The hydraulic fracturing process begins with drilling a borehole into the subsurface and stopping when a natural gas-bearing rock formation is reached, which can be anywhere from 2 to 4 km (1 to 2.5 miles) beneath the surface on average. After drilling, the borehole is lined with a casing typically consisting of cement and steel piping. Using a perforation device with explosive charges, small holes and fractures are created through the casing and the targeted rock formation [27]. This completes the well establishment process.

Large volumes of pressurized fluids are pumped into the completed well to generate fractures to yield hydrocarbons in the formation. The fluids used in the process contain a mixture of water, granular solids (typically sand), and chemical additives. The numerical percentage of the fluid, that is, water, sand, or chemical additive varies depending on the type and depth of the well and the type of rock formation in which the hydraulic fracturing is occurring [85]. The water works as a transport mechanism to get the sand into the newly created fractures. Also, the water is pressurized, so it can extend into both the newly created fractures and naturally occurring weaknesses in the formation. The sand is referred to as a proppant or propping material, which is critical to the hydraulic fracturing process as it props open the newly formed fractures allowing natural gas to flow out of the rock once the fluid is removed [36]. The purposes of chemical additives within the water are extensive, including but not limited to corrosion inhibitors to protect the casing, acids to dissolve minerals, polymers to maintain viscosity as temperature increases, gels to thicken the water and suspend the proppant, and surfactant to reduce surface tension and improve fluid recovery [27]. The operators pump a percentage of the water back to the surface before the natural gas is collected, referred to as flow back.

Although hydraulic fracturing began in the 1940s, recent advances in technology have significantly increased the number of well sites using the technique in the last 20 years [28]. In a process known as horizontal drilling, the borehole is extended horizontally into the targeted formation for an additional 1–2 km (1 mile). Horizontal drilling, using advanced GPS and drilling technologies, allows industries to pinpoint the exact location within the formation most conducive to production, increasing productivity. Industries drilling for natural gas are able to reduce the number of well sites by drilling horizontally within the formation, which lowers the impact at the surface. More than 82,000 hydraulically fractured wells are currently operating in the USA [66].

Industrial Wastewater

There are many terms used in reference to the waters used in the hydraulic fracturing process. Hydraulic fracturing fluid is water treated with the proppant and chemical additives and injected into the well. Produced wastewater is the fluid mixed with formation minerals after injection, which returns to the surface through the well after the production of natural gas. Produced wastewater contains chemical and metallic contaminants from the formation, which may be harmful to the environment [28]. Contaminated water is produced through drilling site preparation, drilling itself, operation, and the use of the hydraulic fracturing technique [28]. Commonly, the original water used in the process is fresh water retrieved from lakes, streams, or municipalities [86]. The work to utilize saline fluids as an alternative is referred to as slick water fracking. The selection between fluids is often determined based on availability and economics.

A maximum of 75% of the injected fluid is retrieved in the hydraulic fracturing process [27]. Flewelling and Sharma [19] found physical constraints prevent the upward migration of hydraulic fracturing fluid; therefore, the remaining 25% of original fracturing fluids are trapped in the fractured formation. After well operators retrieve wastewaters from the well, it must either be treated or disposed.

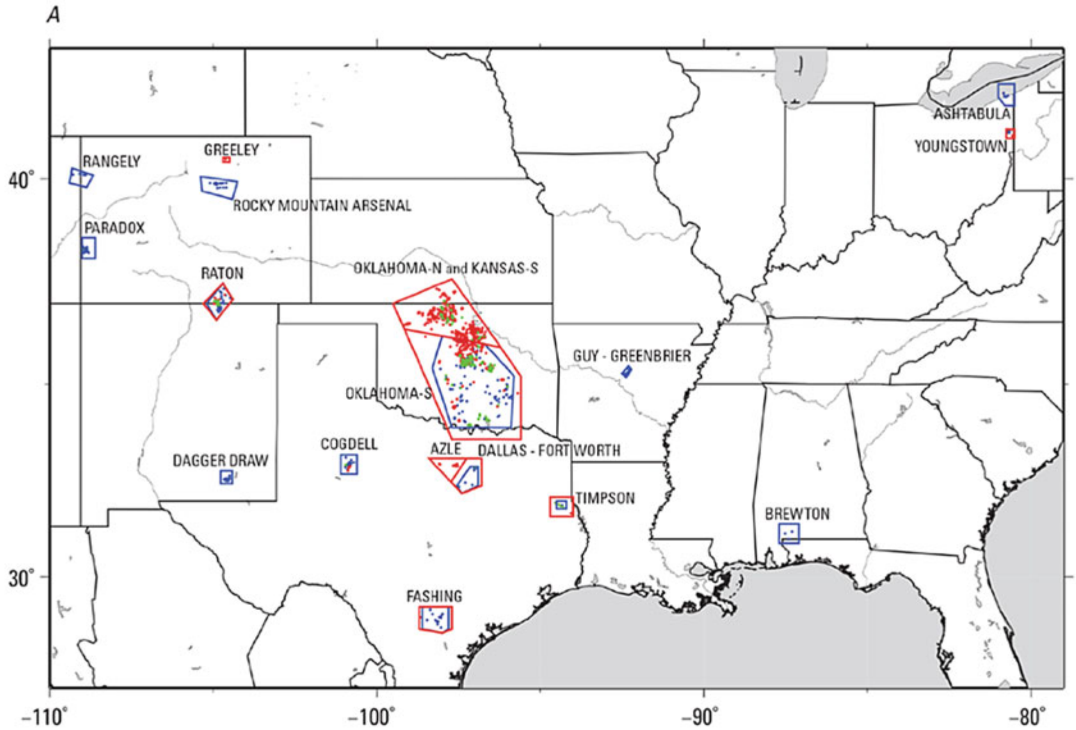
Energy companies have implemented several strategies to address produced wastewater. According to Hammer and VanBriesen [28], there are five basic strategies for managing the chemically treated wastewater retrieved in the process: (1) minimizing the produced wastewater, (2) recycling, (3) treatment, (4) beneficial reuse, and (5) disposal. Minimization of wastewater production is implemented at the well site. Advanced technologies and mechanical blocking devices are minimization methods, but these methods are not as popular with oil and gas companies because the technology is still being developed and the effects are uncertain [28]. Wastewater can be reused outside of the hydraulic fracturing process as water for livestock, vegetable cultures, irrigation, and fire control [1]. Treatment and recycling methods are more commonly used for managing wastewater. Within recent years, some oil and natural gas companies have begun creating facilities and/or management procedures regarding the treatment or recycling of wastewater. Chesapeake Energy Corporation has designed water filtration processes for eight different formations across the country to help reduce the amount of contaminants [77]. Devon Energy has constructed a water recycling facility to treat and reuse wastewater within the hydraulic fracturing process. The recycled fluids will be reused until the level of chlorides reaches 30,000 parts per million. Wastewater with this level of chlorides may clog the well, which makes it hazardous to the process, and it must be disposed [78]. When industrial wastewater is too contaminated or too costly to treat and reuse, it must be disposed. According to the US Environmental Protection Agency's (EPA) Underground Injection Control (UIC) section of the Safe Drinking Water Act (SDWA) of 1974, using a Class II injection disposal well is the safest option for disposing wastewater or storing hydrocarbons produced from the oil and natural gas production process in order to protect the environment and public drinking water [14].

Injection Seismicity

The US National Research Council (NRC) [53] identified 48 locations within the USA with observed occurrences of human-influenced or human-induced seismicity. The seismicity within these locations is linked to industrial fluid processes, such as hydrocarbon storage, retrieval, or wastewater disposal [53]. The US Geological Survey (USGS) announced in 2014 the locations of 14 areas that may be linked to these industrial fluid processes [60]. Three additional areas were added to the USGS list in 2015, making a total of 17 seismic areas "at risk" for industrially induced seismicity [61]. The USGS-identified potentially induced seismic areas (PISAs) stretch across eight states. The names and locations of these specified potentially induced seismic areas, or PISAs, are shown in Fig. 1.

In all 17 PISAs, peer-reviewed literature indicates the increase in seismicity is likely associated with subsurface injection disposal wells. Sixteen of the 17 locations have peer-reviewed literature attempting to determine the origin of seismic onset (research within the remaining unstudied location is ongoing). The radial movement of injection fluid from the center location of a disposal well, as it crosses naturally occurring faults in the subsurface, can cause earthquakes [2, 29–31, 34, 49, 54]. Research in some of the USGS-identified PISAs has confirmed induced seismic events, such as the earthquakes in Rangely and Rocky Mountain Arsenal in Colorado [29, 65].

Not all research has been successful with linking injection wells to local seismic activity with high levels of scientific certainty. Keranen, Savage, Abers, and Cochran [40] explained hydraulic data such as reservoir pressure, formation permeability, and injection fluid volume are needed to conclusively link Oklahoma's seismic activity to subsurface fluid injection. Similarly, results in the Dagger Draw Oil Field, New Mexico, indicated a lack of data regarding surface pressures, fluid injection, and earthquake data to be able to conclusively verify



Induced Seismicity, Fig. 1 The locations of the 17 areas removed from the USGS National Seismic Hazard Map are indicated with red (nondeclustered) and blue (declustered) polygons. Red (2014), green (2013), and blue (2012) dots indicate earthquake epicenters. (Image from Petersen et al. [61])

induced seismicity [70]. Researchers investigating increased seismicity in Brewton, Alabama, explained the difficulty identifying induced seismicity without knowing the fluid properties, movement of subsurface pressure, and material properties [25]. The lack of hydrogeologic data remains a consistent theme among these induced seismicity investigations. Ellsworth [13] emphasizes the lack of hydrogeologic data, stating more subsurface formation properties are needed to enhance induced seismicity research. McGarr et al. [50] further support this claim, adding the importance of locating faults and identifying subsurface formation properties as crucial to managing injection-induced earthquakes. Barnes and Halihan [4] provide a systematic review of the uncertainty in the links between injection well activity and seismicity.

The 17 USGS-identified PISAs were excluded from the 2014 US National Seismic Hazard Map for the final analysis due to the anthropogenic influence (Fig. 1). In 2016, the USGS attempted to quantify the seismic hazard within these areas by evaluating earthquake patterns, rates, and ground motion data. However, the USGS does not attempt to address the cause of seismicity and acknowledges there are many gaps in the scientific research, such as fault locations and orientations, hydrogeologic characteristics, and incomplete records of injection and formation pressures, rates, and volumes [62, 87].

Injection Disposal Wells

The purpose of an injection disposal well is to prevent the upward migration of contaminants into groundwater or surface water; therefore, the injection well targets a porous, permeable rock formation for which the fluids are confined [14]. Several classes of disposal wells exist, but in the USA, a Class II well exclusively targets wastewater and hydrocarbons resulting from oil and natural gas production [14]. The management and regulation of these wells is generally passed from the federal to state government in the US state organizations such as the Texas Railroad Commission and the Oklahoma Corporation Commission regulate and monitor each Class II injection well within their region. Injection disposal wells are created in a process similar to traditional oil and natural gas wells, except the goal is to drill into a porous, permeable rock formation where the injected wastewater will be contained [14]. In some cases, the wastewater can by injected back into an existing production well after the hydrocarbon resources are exhausted.

There are over 11,000 active and inactive injection wells currently in the US state of Oklahoma, which contains the two largest PISAs in surface area [56]. Over 55,000 injection wells are located in the state of Texas, which includes five of the 17 PISAs [64]. As of 2012, there were over 150,800 Class II injection disposal wells in the USA [44]. Not all regions of the USA are conducive to injection disposal wells. For example, there are only eight disposal wells in the entire state of Pennsylvania [46], even though hydraulically fractured wells are common in this state. This is because Pennsylvania does not have significant confined permeable injection intervals capable of preventing the vertical migration of wastewater fluids. When using a Class II injection disposal well, the injected wastewater extends in all directions throughout the rock formation with no artificial barriers containing the fluids. Although natural gas retrieval companies must follow state and federal standards for disposing wastewater, there is always the threat of contamination or other harmful environmental impacts when dealing with chemically treated water, whether by accident or ignorance.

Environmental Balance of Wastewater Disposal

Wastewater disposal is a balance between economics and societies desire for fossil fuels and environmental impacts to water quality or inducing seismicity. This section looks at these competing issues around wastewater disposal.

Economic Impact

Natural gas retrieved from the hydraulic fracturing process (i.e., shale gas) is one of the fastest-growing energy sources in the USA, accounting for over 60% of the US gas supply [83]. Americans consumed 22,467 billion cubic feet of natural gas in 2011 [82]. Natural gas consumption increases each year, with a total of 27,474 billion cubic feet consumed in 2015 [82]. Because Americans are increasing their consumption of natural gas, there is a demand to maintain a national supply. According to an IMPLAN model developed by Miller and Blair [51] for the state of Pennsylvania, the shale gas industry in 2008 was responsible for "$2.2 billion in economic activity, the creation of 29,284 jobs, and the payment of $238.5 million in state and local taxes within the commonwealth of Pennsylvania" [42]. Oklahoma's oil and gas industry supports 364,300 jobs, employing a quarter of the state's population, and the oil and gas industry contributes $50 billion of Oklahoma's $150 billion economy [58]. The oil and gas industry in Texas generates over 315,000 jobs and established a "Rainy Day Fund" of over $2.2 billion, which helps statewide shortfalls in education, health insurance, child protective services, and disaster recovery programs just to name a few [59]. In 2011, the nationwide employment in the oil and gas industry consisted of over 5% of the total employment. The oil and gas industry generated over $550 billion dollars, which was 8% of the US total economy [3]. Despite the economic value, resistance to hydraulic fracturing and wastewater disposal exists in the USA due to health and safety concerns.

Societal Impact

If natural gas is unattainable by traditional methods in the quantities required, hydraulic fracturing is critical to maintaining stores of natural gas for US consumption. However, there is public opposition and confusion regarding hydraulic fracturing. A survey of Americans in 2012 found only 26% of Americans were well-informed about the hydraulic fracturing process, 35% had heard nothing at all, and for those who had heard of it, 35% were opposed to its use [63]. Boudet, Clarke, Bugden, Maibach, Roser-Renouf, and Leiserowitz [8] found 39% of Americans had heard nothing at all regarding the hydraulic fracturing process and only 9% of Americans were well-informed. Additionally, 22% of Americans were strongly opposed to hydraulic fracturing, and 20% of the population supported it, regardless of how well-informed they were about the process [8]. Boudet et al. [8] found those who opposed tended to be more informed about the process and referenced environmental impacts of hydraulic fracturing. Several organizations maintain websites advocating against the process of "fracking" for oil and natural gas extraction: americansagainstfracking.org, nyagainstfracking.org, artistsagainstfracking.org, dangersoffracking.com, californiansagainstfracking.org, dontfrackwithus.org, nationalgrassrootscoalition.org, and many more. According to these anti-fracking websites, social anxiety over fracking stems from the environmental issues surrounding the hydraulic fracturing process or, more generally, hydrocarbon production techniques.

Little to no peer-reviewed research exists examining perceptions of induced seismicity. Misconceptions and inaccuracies regarding induced seismicity are reported through media outlets, which exacerbate public confusion [67]. Common misconceptions include all earthquakes are caused by hydraulic fracturing (only a small percentage), there would be no wastewater disposal without hydraulic fracturing techniques (nearly all production wells produce wastewater), and all injection wells create earthquakes (most do not [67]). Despite addressing common misconceptions among the public, there remain genuine concerns regarding impacts to the environment.

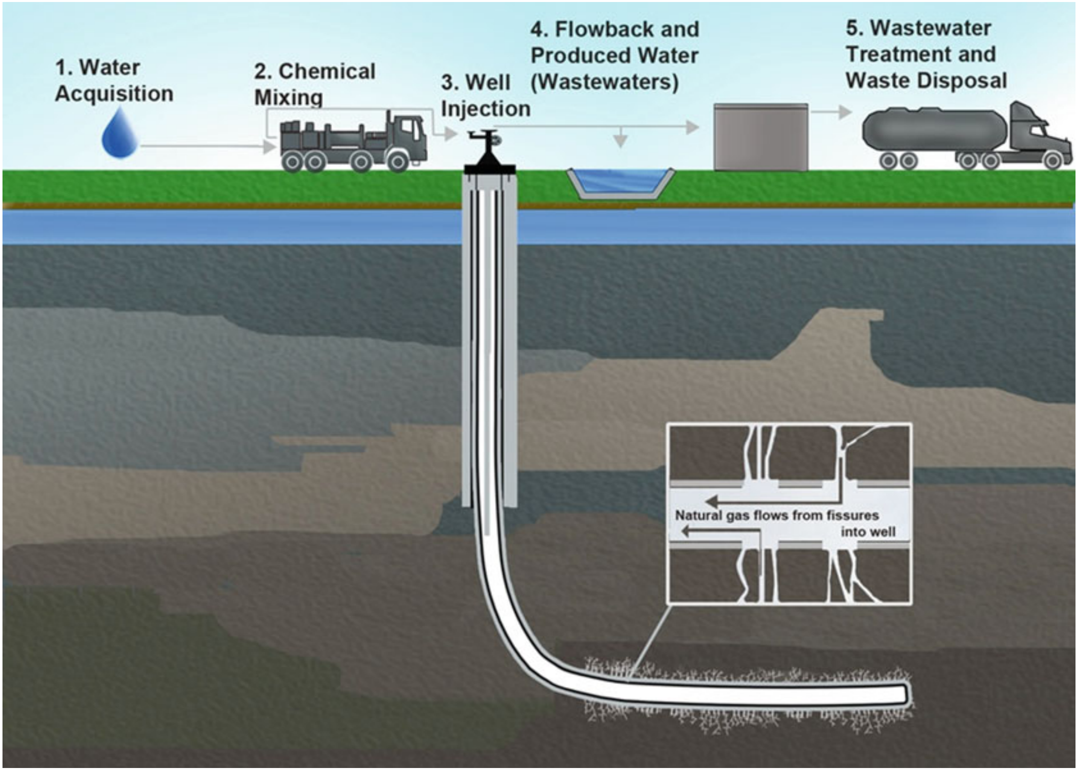
Environmental Impact

A major environmental concern with hydraulic fracturing is the large volume of water required for the process. For example, approximately 3,800,000 gallons of water per well is needed to complete the hydraulic fracturing process while drilling in the Marcellus Shale, which is a large oil and natural gas retrieval location in northeastern USA [86]. The use of millions of gallons of fresh water for hydraulic fracturing could be a concern in areas with water scarcity [86]. For reference, the amount of fresh water used in the hydraulic fracturing process includes approximately 4% of the total estimated uses of US fresh water [45]. Other uses of fresh water include public supply at 12%, irrigation for agriculture at 33%, and thermoelectric power at 45% (Fig. 2 [45]).

The multiple ways of creating wastewater also provide multiple opportunities for contamination of the surrounding environment. Hydraulic fracturing can require hundreds of semi-truckloads to transfer the millions of gallons of water to each well site and then to treatment facilities or injection well locations, which can increase carbon emissions. Well operators and truck drivers may accidentally spill the wastewater onto the land surface at the well site, in transit to disposal wells, or in transit to other destinations such as water treatment facilities [28]. Any locations where wastewater is transferred at the surface can become a source area.

Groundwater contamination can occur when the induced fractures are hydraulically connected to a fresh water reservoir or through improperly plugged wells [26]. Air pollution is another environmental concern, as hydraulic fracturing releases dust, diesel fumes, methane, and other particulate matter into the atmosphere [26]. Other contamination possibilities arise through the hydraulic fracturing process: secondary pollution due to transferring wastewater from production site to storage facility, loss of land use due to extreme salt contamination, impacts of water withdrawals, and improper sealing of abandoned wells [28].

According to the EPA [15], the oil and natural gas industry is exempt from the SDWA of 1974. Diesel and the disposal of wastewater through injection disposal wells are the only aspects of hydraulic fracturing held accountable by the SDWA and the EPA [15]. This means tracing contaminants found in local water sources back to hydraulic fracturing sites would be difficult, since the industry does not have to disclose any chemicals (except for diesel) used in the injected



Induced Seismicity, Fig. 2 Artist depiction of a typical hydraulic fracturing well implementing horizontal drilling. This image also shows natural gas flowing through the man-made fractures. (Image source: EPA [16])

fluids. However, the Resource Conservation and Recovery Act (RCRA) does give individual states the authority to require disclosure of harmful chemicals. There are currently 23 US states disclosing their industrial chemicals on FracFocus.org. Oklahoma passed fracking disclosure rules forcing natural gas industries to post all of the chemicals used in their hydraulic fracturing fluids [47]. The Oklahoma Department of Environmental Quality and Oklahoma Corporation Commission list the regulations regarding monitoring hydraulic fracturing sites and well water on their websites. Air and water quality monitoring at well sites consists of sampling 2–3 days apart or averaging within a 24-hour period, which tests for a broad spectrum of common environmental concerns [55, 57]. Some of the chemicals used in the hydraulic fracturing fluid are not found naturally. This means water quality regulators could trace harmful chemicals found in drinking water sources back to hydraulic fracturing sites if pathways exist.

The connection between wastewater disposal and induced seismicity from fluid injection led researchers to investigate closely the relationship between these two components. Holland [32] correlated the intense pressure of fluid injected into a subsurface fault during the process of hydraulic fracturing to shallow earthquakes with magnitudes of 0.6–2.9, which are sometimes felt at the surface but not strong enough to cause damage. Although Holland linked earthquakes directly to hydraulic fracturing, the low risk factor encourages scientists to focus on induced earthquake sources capable of causing seismic damage. As stated earlier, increased seismicity within the CEUS has been linked to injection disposal wells.

The disposal of wastewater through injection disposal wells has likely induced seismicity in all of the 17 PISAs, with each location containing various degrees of scientific certainty based on evidential support [4]. Earthquakes have increased in Alabama, Arkansas, Colorado, Kansas, Ohio, New Mexico, Oklahoma, Texas, and Virginia, where the epicenters, or earthquake location at the surface, have all been at or near injection disposal well sites [13].

Damage to structural property and residents has been reported in Oklahoma and Texas [23, 40]. To begin investigating induced seismicity, it is important to discuss the mechanism of earthquakes and how earthquakes are located and measured.

Induced Seismicity

Natural vs. Induced Earthquakes

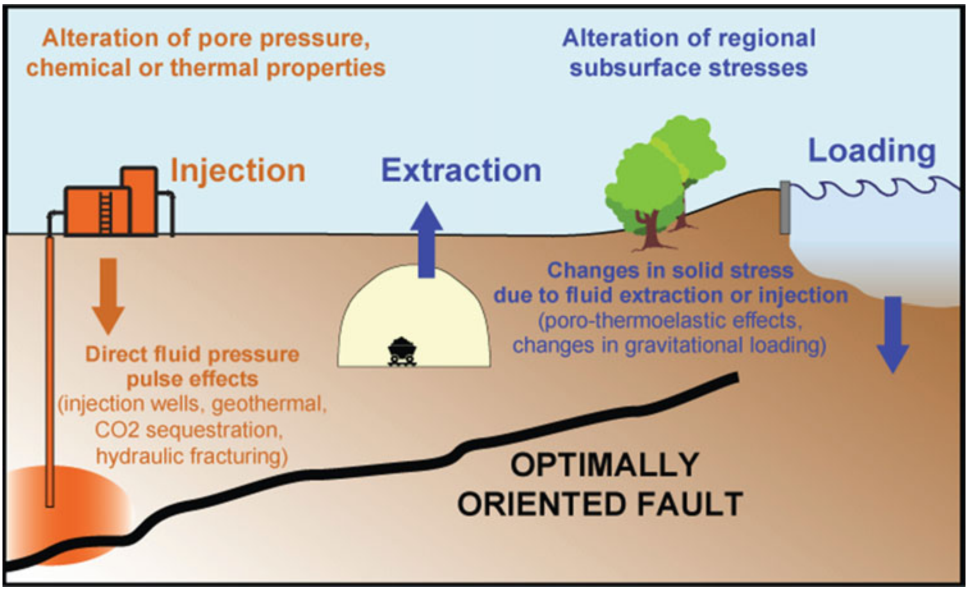
Natural earthquakes typically occur along faults in regions classified as tectonically active. An earthquake is a sudden release of slowly accumulated stress at a fault [5]. Seismologists measure earthquakes with three basic scales: Richter, moment magnitude, and Mercalli Intensity [84]. The Richter scale measures earthquake magnitude, which is a combination of the amplitude of earthquake waves and duration of event. The Mercalli scale uses human observations and surface destruction to categorize intensity. The Mercalli scale only considers the human impact, and the Richter scale does not measure strong earthquakes accurately [84]. The USGS often uses the Modified Mercalli Intensity Scale to depict the strength of earthquakes to the general public. However, the most common measurement method used by seismologists for earthquakes is the moment magnitude scale. The moment magnitude scale combines the Richter and Mercalli scales to measure earthquake magnitude and intensity. Calibrated seismometers placed in several locations throughout a region can triangulate moment magnitude. The moment magnitude is a logarithmic scale, increasing by 10 each time the magnitude is increased by one. Most seismologists, geologists, and other scientists studying seismicity use the moment magnitude as the scale for peer-reviewed research [84]. Throughout this chapter, magnitude values are assumed as moment magnitude unless otherwise denoted.

Although earthquakes with moment magnitudes of 3.0 or higher typically occur along tectonic plate boundaries, earthquakes are possible along faults in the continental interior due to high shear stress [60, 81]. Areas with high shear stress are at or near the strength limit of the crust. This strength limit means any distress, large or small, applied to a critically stressed fault can trigger an earthquake [49, 54, 60]. In the USA, naturally occurring earthquakes with moment magnitudes of 5.0 or greater are rare east of the Rocky Mountains. However, from 2008 to 2011, the annual number of earthquakes in the Oklahoma region was 11 times greater than the annual average number of earthquakes from 1976 to 2007 [40]. Rubenstein, Ellsworth, McGarr, and Benz [68] claim a 40-fold increase in earthquakes in the CEUS since 2001. At peak seismic activity in 2015, Oklahoma experienced 904 felt earthquakes, with a background rate of 1.6 per year.

Distinguishing between natural events and induced seismicity is difficult. There are some differences between the two, which recently have been revealed due to the ongoing research in induced seismicity. Induced seismic earthquakes may tend to have lesser magnitudes than naturally occurring earthquakes [48]. Induced earthquakes often occur in swarms and at shallower depths than natural earthquakes [25, 72]. Ground-shaking patterns are often more intense with induced seismicity due to the shallow hypocenter locations, but additional research is required to confirm these conclusions [62]. Ellsworth [13] visualizes the changes in the number of earthquakes within the CEUS with magnitude of 3 or higher, by comparing the projected seismic rate versus the actual seismic rate from 1967 to 2012.

Induced Seismicity

There are four known anthropogenic processes capable of inducing a felt earthquake: reservoir loading, mining, geothermal activity, and fluid injection [13, 74]. Reservoir loading is the addition or removal of a large volume of water, which changes the ground stress levels quickly and drastically (Fig. 3). The underground excavation in mines leads to the removal of large masses of rocks from beneath the surface and is often accompanied by removal of groundwater mass,



Induced Seismicity, Fig. 3 Depiction of three induced seismicity mechanisms. Note in the fluid injection depiction the fluid does not have to reach the fault to trigger an earthquake [52]. (Image after portions of Ellsworth [13])

thus weakening the formation integrity. Geothermal energy extraction induces earthquakes by removing large volumes of fluids from beneath the surface. This extraction process has similar affects as fluid injection with a different measurement system. To correlate geothermal energy extraction and seismic events, one must quantify the net volume of produced fluid, as opposed to only quantifying the volume of injected fluid [13]. Fluid and gas injection into subsurface rock increases the pore pressure within the formation, thus decreasing effective stress and increasing the formation pressure. The increase in overall formation pressure affects physical structures (i.e., fractures and faults) within the injected region [74].

Certain conditions may make faults more susceptible for an induced seismic event, such as high shear stress or increasing pore pressure [54]. The Mohr-Coulomb failure equation is used to determine the critical stress most likely to trigger a seismic event. Mohr-Coulomb failure is expressed by the following equation:

$$\tau_{crit} = \tau_\upsilon + \mu\sigma_\eta$$

"where τ_{crit} is the critical shear stress required to cause slip on a fault, τ_υ is the frictional stress on the plane of slip, μ is the coefficient of friction, and σ_η is the normal stress acting across the fault" [54]. An increase in fluid pressure on a fault can trigger Mohr-Coulomb failure by reducing the critical threshold within the surrounding rock structures. Nicholson and Wesson [54] were among the first researchers to supply suggestions for deep fluid injections outlining considerations such as site location, distance from faults, stress estimate, and the natural seismicity of the region before the establishment of an injection disposal well. Current researchers reference Nicholson and Wesson [54] frequently to validate induced seismicity findings. For example, Ake et al. [2] correlated the seismic events at Paradox Valley, Colorado, to the fluid injection rates at a nearby injection disposal well through criteria supplied by Nicholson and Wesson [54].

Davis and Frohlich [11] published criteria for rationally assessing whether an event is a natural event or induced seismicity. The article provided a starting point for determining whether an event is natural or induced. The authors describe seven questions to ask after a specific seismic event occurs. These questions help evaluate the likelihood of induced seismicity. The authors provide examples of earthquake events, which they submitted through their questionnaire to see if the questions were valid. For example, the authors ran both the Rangely and Rocky Mountain Arsenal cases (both locations were established as seismically induced) through the questionnaire, and the results indicated induced seismicity [12]. The most important question in the questionnaire is: Do earthquakes occur naturally in the region? Other questions within this research include parameters such as location of the earthquake epicenters, fluid pressures, and correlation of seismic event to fluid injection [11]. The authors explained how these questions, if results indicate induced seismicity, were not an absolute indicator of induced seismicity. Although many of the established induced seismic events in literature align with the questionnaire results, further research at each site is necessary to provide evidentiary support. These questions can guide seismologists and other researchers to pursue induced seismicity research or find alternative sources for the onset of an earthquake.

Ellsworth [13] encourages constant seismic monitoring around injection disposal wells to better understand the hazards of induced seismicity. Current monitoring regulations for Class II injection disposal wells only cover fracture pressure, total injection volume, and average injection pressure [13]. Ellsworth [13] compares the magnitude of natural earthquakes to induced earthquakes. Hazards for major seismic events include liquefaction, landslides, surface rupturing, and tsunamis if located in or near an ocean. Bird and Bommer [6] explain these hazards may occur with any ground-shaking event, specifically a magnitude of 3 or greater. Ellsworth [13] concludes induced seismic events can have magnitudes as high as 6. Keranen et al. [40] claim this number should be increased due to the 5.7, which hit Prague, Oklahoma, in 2011. An earthquake in San Salvador, El Salvador, with the same magnitude as the Prague earthquake in 2011 affected over 110,000 residents, including the deaths of over 1500 people [13]. Although the likelihood of human death or injury in this range is low in areas with modern buildings, there are areas in the USA where some buildings are not constructed with earthquake standards and could cause significant damage if subjected to an earthquake with a moment magnitude of 6.

Classifying earthquakes and discovering the triggering mechanism for each event are part of the seismological aspect of induced seismicity. Structural geology helps identify critically stressed faults and their orientations, which is important for locating and determining potential hazards of induced seismicity. Locating hypocenters, determining the stress load of nearby critical faults, and mapping distances to injection disposal well locations are all common practices in induced seismicity literature. Hydrogeology of the injection interval has not been as prevalently addressed in the current literature as the structural and seismological components of induced seismicity. Researchers recognize their claims need further hydrogeological data to support their findings beyond a reasonable doubt [25, 40].

Hydrogeologic Parameters

Increasing pore pressure near a critically stressed fault can produce earthquakes [12], and fluid injection increases subsurface pore pressure [65]. Earthquakes can be induced by the increase in pressure alone, meaning the fluid itself does not have to reach the fault [52]. In order to calculate the distance pressure will travel over time, the hydrogeologic characteristics of the injection interval are needed.

When investigating an area of potentially induced seismicity, it is critical to correlate the injection rate to the location of earthquake hypocenters [88]. Injection rate is important because it indicates how fast the fluid is being pushed into the injection interval. How quickly the fluids move through the injection interval is dependent on the hydrogeology, which is comprised of rock properties (grain size, orientation, porosity, and permeability) and fluid characteristics (density and viscosity [18]). Hydraulic conductivity is the ability of rocks to transmit water, also known as the coefficient of permeability, and is affected by a number of parameters. The property of the rock that is governed by the size of openings available for fluid movement is the intrinsic permeability. Fluid density and viscosity also affect the migration [18]. Fluid viscosity changes with temperature: viscosity increases as temperature decreases. Fluid density is altered with pressure, temperature, or added minerals. For example, salt water has a different fluid density than fresh water. The intrinsic permeability of a rock is dependent on primary openings formed as the rock was formed and secondary openings formed after rock formation. Typically shales have low hydraulic conductivity (10^{-9}–10^{-13} m/s) and are often used to line solid waste disposal sites due to the difficulty of fluids to move through [18]. Chemically precipitated rocks, such as limestone or dolomite, can have high hydraulic conductivity (10^{-3}–10^{-5} m/s). These types of rocks are often the target for injection disposal wells, as long as they have a confining layer, such as a shale bed, above and below. Additionally, limestones and dolomites are susceptible to secondary openings caused by dissolution. Crystalline rocks, such as igneous basement rock, typically have very low hydraulic conductivity (10^{-9}–10^{-13} m/s). Secondary openings within these rocks can increase fluid flow by orders of magnitude [18]. The majority of injection disposal wells within the PISAs investigated inject into crystalline basement or limestones and dolomites lying directly above basement rock [2, 22, 29, 33, 34, 40, 41, 70, 72, 88]. This is why it is crucial to know the extent of fractures and fluid pathways in these systems to discover how quickly and over what distance fluid pressure pulse can migrate.

The most rational approach for discovering hydraulic pathways to a critically stressed fault is to calculate the rate of pressure migration based on injection rate and hydrogeologic characteristics of the injection interval. Davis and Pennington [12] used the Theis equation to determine fluid migration rates. With slight adjustments to accommodate site-specific conditions, the Theis equation has been used to calculate the propagation of pressure waves in the subsurface [10, 17, 43, 71]. The Theis equation is used for transient or unsteady-state conditions [79]. For steady-state conditions, the Thiem equation is appropriate [80]. Both equations assume a uniform interval with an infinite extent and equal hydraulic properties in all directions being injected at a constant rate. These equations can be used in a variety of circumstances and across a wide range of geologic settings. By using the Theis or Thiem equation as a basis to estimate the pressure migration radiating from injection wells, it is critical to obtain the following parameters: injection interval thickness, hydraulic conductivity, storativity, and injection rates. Additionally, pressure measurements should be obtained throughout the injected reservoir. Davis and Pennington [12] explain although bottomhole pressures at the injection site are important, these measurements are not typically helpful when used in isolation. The pressure analysis for the injection interval could be inaccurate. Useful pressure measurements include data further from the well and throughout the injected formation to gain a better understanding of pressure migration.

Review of Potentially Induced Seismic Areas

A thorough investigation into a PISA would include data sets from structural geology, seismology, and hydrogeology. These three disciplines are not always equally represented within induced seismicity investigations. The following section provides an overview of the 17 PISAs, including descriptions of the location and significant findings.

Rocky Mountain Arsenal, Colorado

Healy et al. [29] published one of the first investigations of induced seismicity directly linking fluid injection to earthquake events. In 1961, the Rocky Mountain Arsenal installed an injection disposal well for chemical waste, and fluid injection began in 1962. There were two seismometers located in the Denver, Colorado metropolitan area measuring earthquakes in 1962. The authors, in collaboration with the USGS, installed several more to record the increased seismic activity. The increase in earthquakes began 7 weeks after fluid injection began. All of the earthquakes classified within the Denver earthquake sequence originated within a radius of 65 km (40 miles) of the injection disposal well. The authors explained the probability of a natural earthquake sequence originating in the same area as the injection disposal well and occurring simultaneously with the onset of fluid injection was 1/2,500,000. In other words, it is highly unlikely the earthquake sequence in Denver from 1962 to 1967 was a naturally occurring seismic event. Earthquakes continued in the Denver metropolitan area even after the termination of the disposal well in 1966. The largest earthquake of the series occurred over a year after the well's termination in 1967, with a magnitude of 5.0. The researchers explained the wastewater beneath the surface continued to radiate outward from the well and increase stress on the surrounding faults, long after the injection ceased [29].

Rangely, Colorado

The Denver, Colorado, earthquake sequence (correlating injection fluid to earthquakes) sparked interest in researchers to conduct further testing. If fluid injection causes earthquakes, and humans control fluid injection, can earthquake sequences be controlled? Raleigh et al. [65] tested this question. Experiments could not be conducted on the closed injection disposal well from Rocky Mountain Arsenal. The researchers needed to maintain control of stress factors and fluid pressure within the area, have the ability to locate the origin of earthquakes, and minimize the risk of inducing a damaging earthquake. The researchers collaborated with the Chevron Oil Company to conduct research with four injection disposal wells in the Rangely Oil Field, which had no record of seismicity prior to the experiment. In 1969, the researchers began pumping fluid into the designated injection disposal wells at Rangely and measured the seismicity. One year later, the researchers stopped pumping and began a period of flow back (fluid retrieval) for 6 months. The researchers recorded over 900 earthquake events from 1969 to 1970, with over one-third of the earthquakes originating 1.0 km (0.6 miles) away from one of the four designated injection disposal wells. During the 6-month flow back period, the pressures measured in the formation dropped significantly, and the earthquakes averaged approximately one earthquake per month. In 1972, the oil company turned the wells back on, and pressure within the formation began to increase. From 1972 to 1973, there was an average of 26 earthquakes per month. The company shut the wells down in 1973, and earthquakes in the region have now ceased [65]. From this research, the authors discovered a critical piece of information: seismic activity dramatically increased after disposal wells reached an injection pressure of 3727 psi (25.7 MPa [65]). All subsequent research targets this injection pressure (also referred to as bottomhole pressure) as a factor for induced seismicity.

Paradox Valley, Colorado

The injection disposal well in Paradox Valley was used to dispose excess salt water from the Colorado River. Recognizing that Colorado had experienced earthquakes previously due to subsurface injection, this well was a government project to see if they could economically and environmentally dispose of the excess salt water while minimizing the earthquake hazard. Seismicity began 4 days after initial injection occurred in 1991 [2]. A seismic swarm occurred in 2013, which was within 6–8 km (4 to 5 miles) of the injection well and 4.1 km (2.5 miles) in depth [7]. With little to no background seismicity, the increase in seismicity at Paradox Valley was determined as induced [2, 7]. The authors used the locations of earthquakes to discover the pressure migration from injection wells. They concluded calculating the pressure pulse provides a reliable estimate of how fast and far the pressure will travel, to help prevent the pressure from reaching a critical fault. Additionally, the pressure propagation could identify the maximum rate to inject within a location [2].

Greeley, Colorado

As of May 2016, there is no peer-reviewed research in the area of Greeley, Colorado, regarding induced seismicity. Yeck et al. [89] presented preliminary findings at the American Geophysical Union in 2014 and described a recent increase in seismicity within this location. With the occurrence of a 3.2 magnitude earthquake in proximity to injection disposal wells, the researchers deployed additional seismic stations to begin consistent monitoring of Greeley. Additional aftershocks including a 2.6 magnitude earthquake occurred 3 weeks after the seismic stations were deployed. The Colorado Oil and Gas Corporation Commission (COGCC) recommended an immediate cease of injection for the first time in its history [89]. Additional research is necessary in this area to support scientific certainty of induced seismicity.

Raton Basin, Southern Colorado, and Northern New Mexico

The Raton Basin lies across the borders of Colorado and New Mexico. The Oil and Conservation Division of the Energy, Minerals, and Natural Resources Department of New Mexico is responsible for regulating the injection disposal wells in New Mexico. No injection data prior to 2006 are available through this entity. Therefore, researchers in this area relied heavily on available data provided through the COGCC. There are several injection disposal wells within the Raton Basin, with over 20 on the Colorado section alone. Rubinstein et al. [68] completed a statistical analysis regarding the increase of seismicity in the Raton Basin since 2001. They calculated a 3% probability earthquakes would happen naturally in this area. They were able to make spatial and temporal correlations between earthquakes and wells with high injection rates and volumes. The authors attempted to associate high rate and volume injection wells to the increase in seismicity. Unfortunately, they could not determine whether rate or volume was more important in determining if or when earthquakes will occur [68].

Dagger Draw, New Mexico

As previously stated, there are no injection records prior to 2006 in the state of New Mexico. Sanford et al. [70] claimed there were no surface pressures and fluid injection data and lacked earthquake numbers and strengths to make suitable correlations to injection activities. Dagger Draw has experienced previous seismicity due to a large magma body at approximately 19 km (12 miles) depth near this location. With a lack of data and a history of seismicity, the authors could not establish a clear connection to induced seismicity. There was a temporal correlation between onset of injection and onset of earthquakes. From 1999 through 2004, the seismicity activity in New Mexico almost doubled [70]. With additional data, it is possible clearer correlations can be made in the Dagger Draw PISA.

Brewton, Alabama

Brewton, Alabama, experienced a 4.9 magnitude earthquake in 1997. Gomberg and Wolf [25] described two injection wells within 5.0 km (3 miles) of the main shock having a focal depth of approximately 4.5 km (2.8 miles). The injection wells reached a depth of 2.1 km (1.3 miles) into a sandstone and shale formation. The volume of extraction and injection of fluids was used to determine if correlations exist between industrial processes and seismicity. The authors could not find a spatial-temporal relationship between volumes, pressures, and earthquake occurrence. However, injection rates were not evaluated in this study. The authors recognize this relationship is difficult to quantify without fluid properties, an idea of pressure migration, and injection interval material properties.

Ashtabula, Ohio

In the 1980s, earthquakes began shaking the town of Ashtabula, Ohio. Ashtabula is not a seismically active area. Similar to the induced seismic events in the Rocky Mountain Arsenal near Denver, researchers directly correlated earthquakes in Ashtabula to an injection disposal well. Seeber et al. [72] discovered the earthquakes originated from one of two major faults, which lie within 5.0 km (3 miles) of an injection disposal well. The researchers concluded the wastewater injected into the well acted as a lubricant to facilitate fault movement and trigger earthquakes. The largest earthquake in the Ashtabula, Ohio sequence occurred after the termination of the injection disposal well [72].

Youngstown, Ohio

Research in Youngstown, Ohio, began when over 100 earthquakes affected this area with no natural background seismicity. Kim [41] explained Ohio has a natural earthquake zone called the Anna

Seismic Zone. However, in the area of Youngstown, there is no record of earthquakes prior to 2011. After 2011 (onset of fluid injection), earthquakes began occurring near the injection disposal well site and radiated outward over time. The researchers clearly mapped a radial pattern of earthquakes with the disposal well in the center of the circle. Kim [41] concluded the increase in pore pressure in the subsurface is due to wastewater injection spread outward from the disposal well, inducing earthquakes in the affected region. Again, the earthquakes did not stop entirely after the disposal well shut down: seismicity continued to decrease steadily over the following months [41].

Azle, Texas

In 2013, seismic activity increased near Azle, Texas. Multiple injection wells were used to dispose wastewater into the Ellenberger formation, which is comprised of dolomite lying directly above the crystalline basement rock. Hornbach et al. [33] emphasize the importance of obtaining baseline pressure (bottomhole), and permeability values, while constantly monitoring fluid rates and volumes throughout the injection process. The authors explain the Texas Railroad Commission does not keep a record of bottomhole pressures; therefore, they attempted to model the changes in subsurface pressure in the Azle area. The authors validated pressure increases within the area capable of producing an increase in seismicity through subsurface pressure modeling.

Fashing, Texas

Frohlich and Brunt [21] published research on earthquakes originating in the Eagle Ford Shale in Texas near the city of Fashing. The Eagle Ford Shale is an unconventional reservoir containing large amounts of oil and natural gas; thus, well operators implement fracking to release the resources. The authors examined 14 hypocenters of a specific earthquake sequence and found a correlation between 10 of the hypocenters and injection disposal wells. However, there was no correlation between four of the hypocenters and injection disposal wells. This led to mixed results. The authors concluded a possible correlation between earthquakes and wastewater injection exists, but the four uncorrelated earthquakes caused doubt [21]. The authors concluded the earthquakes in Fashing were induced but through fluid extraction and not injection. It is possible the injection process contributed to the overall increase in seismicity, but the subsurface pressure reduction from the extraction process exceeded injection pressure [21]. The authors recognized a detailed analysis of subsurface hydrogeologic parameters would help support their findings.

Cogdell, Texas

Saltwater disposal has occurred in the Cogdell Oil Field since 1956 [12]. The seismicity near Snyder, Texas, and within the Cogdell Oil Field was determined induced, even though seismic swarms did not occur until 20 years after injection commenced [12]. In 1971, industries started injecting CO2 into the same Canyon Reef limestone used for saltwater disposal. Gan and Frohlich [24] found gas injection could be responsible for increased seismicity for the first time. These researchers suggest extensive modeling of subsurface stress and hydrogeology would help explain why the Cogdell Oil Field experiences earthquakes while surrounding regions are not. Gan and Frohlich [24] researched injection disposal wells with rates equal to those correlated to seismic events that did not have earthquake hypocenter within 5.0 km (3 miles) of the well. Gan and Frohlich concluded these wells were not in the near vicinity of a fault. However, the author did not have enough data of the subsurface structure to be able to definitively support this hypothesis [24].

Dallas/Fort Worth, Texas

Frohlich [20] conducted a correlation study in the Barnett Shale, an unconventional oil and gas retrieval location in Texas. In the Barnett Shale study, Frohlich [20] examined 24 hypocenters occurring near the Dallas/Fort Worth International Airport. This study found all 24 were within 3.5 km (2 miles) of injection disposal wells. From this research, Frohlich [20] determined wastewater injection into the Barnett Shale caused the 24 seismic events. Additionally, a critical injection rate of 150,000 B.P. (24,000 m2/month) was correlated to each major seismic swarm. Frohlich [20] acknowledges the critical rate will depend on site-specific subsurface properties.

Timpson, Texas

A 4.8 magnitude earthquake occurred near Timpson, Texas, in 2014. The earthquake caused damage to several houses and woke up residents as far as 50 km (31 miles) away from the epicenter. With two high volume injection wells within 3.0 km (2 miles) of the earthquake swarm, researchers began investigating the cause of the Timpson seismicity. They found although sufficient evidence exists to correlate injection to the increase in seismicity, they could not rule out natural causes within this location [23]. The authors recognize a more complete understanding of subsurface properties such as hydrogeology and stress conditions would provide a better understanding of the connection between seismic events and fluid injection [23].

Central Oklahoma

Keranen et al. [40] closely examined a specific earthquake sequence from 2011 in Oklahoma. The sequence included three earthquakes with moment magnitudes of 5.0 or greater, the largest being a 5.7. The largest event caused structural damage at the epicentral region and two human injuries. At least 17 states felt the earthquake. The 5.7 earthquake near Prague, Oklahoma, is the largest induced seismic event in US history (as of March, 2016). Oklahoma has experienced a 200-fold increase in earthquakes since 2009 [87]. Most of the research for this event was retroactive, meaning the authors used aftershocks to locate the focus of each major earthquake. The researchers deployed seismometers 24 hours after the initial 5.7 earthquake affected the area. Data from over 1000 aftershocks provided the location for the earthquake hypocenters within the Wilzetta fault zone. The Wilzetta fault zone lies within the Wilzetta Oil Field, and the three major earthquakes originated within 5 kilometers of an injection disposal well. The wastewater injection into the injection disposal well began in 1993, which is 17 years before the first noted earthquake occurred in Oklahoma. Using the Davis and Frohlich [11] criteria for induced seismicity, the authors concluded the cause of the 2011 major seismic event in Oklahoma was most likely, but not definitively, wastewater injection into Class II injection disposal wells [40].

An opposition statement from the state of Oklahoma exists regarding the Prague, Oklahoma, earthquake sequence. Keller and Holland [39], as representatives of the Oklahoma Geological Survey and in collaboration with the Oklahoma Corporation Commission, issued a brief statement on the Prague earthquakes. The authors explain how the Wilzetta fault zone has a history of earthquakes: swarms of earthquakes, such as the Prague event, are natural to the area. The authors declare fluid injection began in 1955 preventing the correlation of fluid injection to the earthquake swarm. Keller and Holland [39] conclude the earthquake swarm in Prague was no more than a naturally occurring event, but continued monitoring would provide more insight into this event.

Sumy, Cochran, Keranen, Wei, and Abers [76] conducted an additional study on the Oklahoma earthquakes. The researchers looked at the intensity of the earthquakes and the Coulomb failure criteria of the Prague, Oklahoma, earthquake sequence. The team concluded the three earthquakes with magnitudes of 5.0 and higher resulted from Mohr-Coulomb stress failure within the adjacent rock structures and triggered several additional earthquakes. Results from the team's research imply the seismic hazard, or risk assessment, for induced seismicity may be greater than previous estimates.

Northern Oklahoma and Southern Kansas

Research with the region of Northern Oklahoma and Southern Kansas is still in progress. The Mississippi Limestone (directly above crystalline basement) is the target injection interval for this region, which lies across the Kansas and Oklahoma border. Kansas started experiencing earthquakes in 2013. Rubinstein, Ellsworth, Llenos, and Walter [69] suggested the earthquakes in Kansas might not be natural in origin. The USGS and

Kansas Geological Survey have deployed additional seismic stations and are currently locating earthquake hypocenters [9]. Monthly saltwater injection data are not available for Kansas [87]. Walsh and Zoback [87] provided injection rate and earthquake correlations among several locations in Northern Oklahoma. They concluded earthquakes in this region are likely associated with industrial activities, and preexisting geological conditions may be more indicative of predicting seismic magnitude than pore pressure [87].

Guy-Greenbrier, Arkansas

In Arkansas, Horton [34] published research findings from yet another recent earthquake sequence. Unlike the other research articles, Arkansas has a strong history of earthquakes. The name of this active region is the New Madrid Seismic Zone (NMSZ). The NMSZ is the most seismically active region east of the Rocky Mountains. Evidence of ancient liquefied soil and sediments called paleoliquefaction in the region indicates earthquakes with magnitudes of 7 or higher occurred within the last 1100 years. After a 98% increase in earthquakes appeared within 6.0 km (3.7 miles) of disposal wells, the researchers began investigating the possibility of induced seismicity. The researchers measured seismicity before and after the installation of new disposal wells in the area in order to determine whether earthquakes were natural or induced. Due to known injection formation properties, such as rock type and fault lines, they knew the formations were directly connected to basement rock. The Arkansas Oil and Gas Commission ordered an emergency shutdown of the observed disposal wells after hundreds of small magnitude earthquakes appeared within 28 days of initial injection. The researchers verified induced seismicity by correlating initial injection with earthquakes and by the significant reduction of earthquakes after the emergency shutdown. Earthquakes continued to shake the region even after the shutdown, mimicking the research from the Rocky Mountain Arsenal and Rangely Oil Field [34].

Future Directions

Each of the PISAs included within this research has experienced an increase in seismicity. Some of these locations are confirmed induced seismic areas, while others lack enough data to make scientific conclusions. The process of hydraulic fracturing includes a large volume of produced wastewater, which must be treated or disposed for public health and safety. Since subsurface injection is one of the most inexpensive options and is accepted by the EPA, it is the most often used disposal method by the industry. Hydraulic fracturing and wastewater disposal can cause several environmental issues, which are troublesome among concerned citizens. However, these industrial processes are beneficial to producing states and to the overall US economy. If the USA is to continue supplying the national demand for natural resources such as natural gas, then it is imperative to find an environmentally and economically feasible solution the public can accept. Regulatory agencies and other monitoring entities must find a way to mitigate or even prevent seismic activity in regions of injection disposal. Additional data are required in structural, seismological, and hydrogeologic disciplines in order to make realistic decisions on how to manage seismic hazard.

Bibliography

1. Adebambo O (2011) Evaluation of the beneficial re–use of produced water: a review of relevant guidelines and produced water toxicity. Doctoral dissertation, Duke University
2. Ake J, Mahrer K, O'Connell D, Block L (2005) Deep-injection and closely monitored induced seismicity at Paradox Valley, Colorado. Bull Seismol Soc Am 95(2):664–683
3. American Petroleum Institute (API) (2013) Economic impacts of the oil and natural gas industry on the U.-S. economy in 2011. Retrieved 16 May 2016, from American Petroleum Institute. http://www.api.org/~/media/files/policy/jobs/economic_impacts_ong_2011.pdf
4. Barnes C, Halihan T (2017) The availability of hydrogeologic data associated with areas identified by the US geological survey as experiencing potentially induced seismicity resulting from subsurface injection. Hydrogeol J. https://doi.org/10.1007/s10040-017-1699-5

5. Bates RL, Jackson JA (1987) Glossary of geology. American Geological Institute, Alexandria, p 788
6. Bird JF, Bommer JJ (2004) Earthquake losses due to ground failure. Eng Geol 75(2):147–179
7. Block LV, Wood CK, Yeck WL, King VM (2014) The 24 January 2013 ML 4.4 earthquake near Paradox, Colorado, and its relation to deep well injection. Seismol Res Lett 85(3):609–624
8. Boudet H, Clarke C, Bugden D, Maibach E, Roser-Renouf C, Leiserowitz A (2014) "Fracking" controversy and communication: using national survey data to understand public perceptions of hydraulic fracturing. Energy Policy 65:57. Retrieved from http://search.proquest.com/docview/1470981315?accountid=4117
9. Buchanan RC (2015) Increased seismicity in Kansas. Lead Edge 34(6):614–617
10. Cihan A, Zhou Q, Birkholzer JT (2011) Analytical solutions for pressure perturbation and fluid leakage through aquitards and wells in multilayered-aquifer systems. Water Resour Res 47(10):W10504
11. Davis SD, Frohlich C (1993) Did (or will) fluid injection cause earthquakes?-criteria for a rational assessment. Seismol Res Lett 64(3–4):207–224
12. Davis SD, Pennington WD (1989) Induced seismic deformation in the Cogdell oil field of west Texas. Bull Seismol Soc Am 79(5):1477–1495
13. Ellsworth WL (2013) Injection-induced earthquakes. Science 341(6142):1225942
14. Environmental Protection Agency (EPA) (2012) Basic information about injection wells. Retrieved 8 Dec 2014, from EPA. http://water.epa.gov/type/groundwater/uic/basicinformation.cfm
15. Environmental Protection Agency (EPA) (2014) Regulation of hydraulic fracturing under the safe water drinking act. Water.epa.gov. Retrieved 25 Apr 2014 from http://water.epa.gov/type/groundwater/uic/class2/hydraulicfracturing/wells_hydroreg.cfm
16. Environmental Protection Agency (EPA) (2016). The hydraulic fracturing water cycle. Retrieved 25 May 2016, from Environmental Protection Agency. https://www.epa.gov/hfstudy/hydraulic-fracturing-water-cycle
17. Ferris JG (1952) Cyclic fluctuations of water level as a basis for determining aquifer transmissibility. US Department of the Interior, Geological Survey, Water Resources Division, Ground Water Branch, Washington, DC
18. Fetter CW (1994) Applied Hydrogeology. Prentice Hall, Upper Saddle River, p 691
19. Flewelling SA, Sharma M (2014) Constraints on upward migration of hydraulic fracturing fluid and brine. Ground Water 52:9–19. https://doi.org/10.1111/gwat.12095
20. Frohlich C (2012) Two-year survey comparing earthquake activity and injection-well locations in the Barnett shale, Texas. Proc Natl Acad Sci USA 109(35):13934–13938. https://doi.org/10.1073/pnas.1207728109
21. Frohlich C, Brunt M (2013) Two-year survey of earthquakes and injection/production wells in the eagle ford shale, Texas, prior to the Mw4. 8 20 October 2011 earthquake. Earth Planet Sci Lett 379:56–63
22. Frohlich C, Potter E, Hayward C, Stump B (2010) Dallas-fort worth earthquakes coincident with activity associated with natural gas production. The Leading Edge, Tulsa
23. Frohlich C, Ellsworth W, Brown WA, Brunt M, Luetgert J, MacDonald T, Walter S (2014) The 17 May 2012 M4. 8 earthquake near Timpson, East Texas: an event possibly triggered by fluid injection. J Geophys Res Solid Earth 119(1):581–593
24. Gan W, Frohlich C (2013) Gas injection may have triggered earthquakes in the Cogdell oil field, Texas. Proc Natl Acad Sci 110(47):18786–18791
25. Gomberg J, Wolf L (1999) Possible cause for an improbable earthquake: the 1997 mw 4.9 Southern Alabama earthquake and hydrocarbon recovery. Geology 27(4):367–370
26. Groat CG, Grimshaw TW (2012) Fact-based regulation for environmental protection in shale gas development. The energy institute at the University of Texas at Austin: Flawn Academic Center. Retrieved from http://www.scribd.com/doc/82147814/Fact-Based-Regulation-for-Environmental-Protection-in-Shale-Gas-Development-by-The-Energy-Institute-at-the-University-of-Texas-at-Austin-February-201
27. Groundwater Protection Council (GWPC) and Interstate Oil and Gas Compact Commission (IOGCC) (2014) Why chemicals are used. Retrieved 8 Dec 2014, from Frac Focus Chemical Disclosure Registry. http://fracfocus.org/chemical-use/why-chemicals-are-used
28. Hammer R, VanBriesen J (2012) In fracking's wake: new rules are needed to protect our health and environment from contaminated wastewater. National Resources Defence Council, New York
29. Healy JH, Rubey WW, Griggs DT, Raleigh CB (1968) The Denver earthquakes. Science 161(3848):1301–1310. https://doi.org/10.1126/science.161.3848.1301
30. Herrmann RB, Park S-K, Wang C-Y (1981) The Denver earthquakes of 1967–1968. Bull Seismol Soc Am 71(3):731–745
31. Herrmann R, Benz H, Ammon C (2011) Monitoring the earthquake source process in North America. Bull Seismol Soc Am 101(6):2609–2625
32. Holland AA (2013) Earthquakes triggered by hydraulic fracturing in south-central Oklahoma. Bull Seismol Soc Am 103(3):1784–1792
33. Hornbach MJ, DeShon HR, Ellsworth WL, Stump BW, Hayward C, Frohlich C, ..., Luetgert JH (2015) Causal factors for seismicity near Azle, Texas. Nat Commun 6:6728
34. Horton S (2012) Disposal of hydrofracking waste fluid by injection into subsurface aquifers triggers earthquake swarm in central Arkansas with potential for damaging earthquake. Seismol Res Lett 83(2):250–260
35. Hubbert MK, Willis DG (1972) Mechanics of hydraulic fracturing. Pet Eng AIME 210:153–168

36. Huitt JL, Glenshaw, McGlothlin, BB (1964) U.S. Patent No. 3,121,464. U.S. Patent and Trademark Office, Washington, DC
37. Hyne N (2014) Dictionary of petroleum exploration, drilling & production. PennWell Corporation, Tulsa
38. Hypocenter (2016) In U.S. geological survey earthquake glossary online. Retrieved from http://earthquake.usgs.gov/learn/glossary/?term=hypocenter
39. Keller G, Holland A (2013) Oklahoma geological survey evaluation of the Prague earthquake sequence of 2011. Oklahoma Geological Survey, Norman
40. Keranen KM, Savage HM, Abers GA, Cochran ES (2013) Potentially induced earthquakes in Oklahoma, USA: links between wastewater injection and the 2011 mw 5.7 earthquake sequence. Geology 41(6):699–702
41. Kim W-Y (2013) Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio. J Geophys Res Solid Earth 118(7):3506–3518. https://doi.org/10.1002/jgrb.50247
42. Kinnaman TC (2011) The economic impact of shale gas extraction: a review of existing studies. Ecol Econ 70(7):1243–1249
43. Lee MK, Wolf LW (1998) Analysis of fluid pressure propagation in heterogeneous rocks: implications for hydrologically-induced earthquakes. Geophys Res Lett 25(13):2329–2332
44. Lustgarten A, Schmidt KK (2012) State-by-state: underground injection wells. Retrieved 18th May 2016, from ProPublica: Journalism in the Public Interest. http://projects.propublica.org/graphics/underground-injection-wells
45. Maupin MA, Kenny JF, Hutson SS, Lovelace JK, Barber NL, Linsey KS (2014) Estimated use of water in the United States in 2010 (No. 1405). US Geological Survey. https://pubs.usgs.gov/circ/1405/pdf/circ1405.pdf
46. McCurdy R (2011) Underground injection wells for produced water disposal. In: Proceedings of the technical workshops for the hydraulic fracturing study: water resources management. EPA
47. McFeeley M (2012) State hydraulic fracturing disclosure rules and enforcement: a comparison. (Issue brief 16–06-A). Natural Resources Defense Council. Retrieved from http://www.nrdc.org/energy/files/Fracking-Disclosure-IB.pdf
48. McGarr A (2014) Maximum magnitude earthquakes induced by fluid injection. J Geophys Res Solid Earth 119(2):1008–1019
49. McGarr A, Simpson D, Seeber L (2002) Case histories of induced and triggered seismicity. Int Handb of Earthq Eng Seismol 8(A):647–664
50. McGarr A, Bekins B, Burkardt N, Dewey J, Earle P, Ellsworth W, ..., Sheehan A (2015) Coping with earthquakes induced by fluid injection. Science 347(6224): 830–831
51. Miller RE, Blair PD (2009) Input-output analysis: foundations and extensions. Cambridge University Press, Cambridge
52. Mulargia F, Bizzarri A (2014) Anthropogenic triggering of large earthquakes. Sci Rep 4
53. National Research Council (US), Committee on Induced Seismicity Potential in Energy Technologies (NRC) (2013) Induced seismicity potential in energy technologies. National Academies Press, Washington, DC
54. Nicholson C, Wesson RL (1990) Earthquake hazard associated with deep well injection. U.S. Government Printing Office, Washington, DC
55. Oklahoma Corporation Commission (OCC) (2016) Permanent rules and approved emergency rules. Occeweb.com. Retrieved 25 Apr 2016 from http://www.occeweb.com/rules/rulestxt.htm
56. Oklahoma Corporation Commission (OCC) (2016). Oil and gas data files. Retrieved 15 May 2016, from Oklahoma Corporation Commission Oil and Gas Division. http://www.occeweb.com/og/ogdatafiles2.htm
57. Oklahoma Department of Environmental Quality (DEQ) (2016) Oklahoma DEQ rules and regulations/legal information. Deq.state.ok.us. Retrieved 25 Apr 2016 from http://www.deq.state.ok.us/mainlinks/deqrules.htm
58. Oklahoma Oil and Gas Administration (OKOGA) (2016) Economic impact. Retrieved 16 Apr 2016, from Oklahoma Oil and Gas Administration. http://okoga.com/economic-impact/
59. Permian Basin Petroleum Association (PBPA) (2014) Economic impact of oil and gas in Texas. Retrieved 18 May 2016, from Permian Basin Petroleum Association. http://pbpa.info/education/permian-basin-industry-statistics/economic-impact-of-oil-and-gas-in-texas/
60. Petersen MD, Mueller CS, Moschetti MP, Hoover SM, Rubinstein JL, Llenos AL, Michael AJ, Ellsworth WL, McGarr AF, Holland AA, Anderson JG. Incorporating induced seismicity in the 2014 United States National Seismic Hazard Model: Results of 2014 workshop and sensitivity studies. US Department of the Interior, US Geological Survey; 2015
61. Petersen MD, Mueller CS, Moschetti MP, Hoover SM, Rubinstein JL, Llenos AL, Michael AJ, Ellsworth WL, McGarr AF, Holland AA, Anderson JG (2015) Incorporating induced seismicity in the 2014 United States national seismic hazard model—results of 2014 workshop and sensitivity studies: U.S. Geological survey open-file report 2015–1070, 69 p. https://doi.org/10.3133/ofr20151070
62. Petersen MD, Mueller CS, Moschetti MP, Hoover SM, Llenos AL, Ellsworth WL, ..., Rukstales KS (2016) 2016 One-year seismic hazard forecast for the central and Eastern United States from induced and natural Earthquakes (No. 2016–1035). US Geological Survey
63. Pew Research Center for the People and the Press (2012) AsGasPricesPinch, support for oil and gas production grows [WWWDocument]. PewRes.Cent. People Press. Retrieved from https://www.people-press.org/2012/03/19/as-gas-prices-pinch-support-for-oil-and-gas-production-grows/
64. Railroad Commission of Texas (RRC) (2015) Injection and disposal wells. Retrieved 23 May 2016, from

Railroad Commission of Texas. http://www.rrc.state.tx.us/about-us/resource-center/faqs/oil-gas-faqs/faq-injection-and-disposal-wells/
65. Raleigh C, Healy J, Bredehoeft J (1976) An experiment in earthquake control at Rangely, Colorado. Science 191(4233):1230–1237
66. Riddlington E, Rumpler J (2013) Fracking by the numbers. Environment America Research & Policy Center. Retrieved from http://www.environmentamerica.org/sites/environment/files/reports/EA_FrackingNumbers_scrn.pdf
67. Rubinstein JL, Mahani AB (2015) Myths and facts on wastewater injection, hydraulic fracturing, enhanced oil recovery, and induced seismicity. Seismol Res Lett 86(4):1060–1067
68. Rubinstein JL, Ellsworth WL, McGarr A, Benz HM (2014) The 2001–present induced earthquake sequence in the Raton Basin of Northern New Mexico and Southern Colorado. Bull Seismol Soc Am 104(5):2162–2181. https://doi.org/10.1785/0120140009
69. Rubinstein JL, Ellsworth WL, Llenos AL, Walter SR. Is the recent increase in seismicity in southern Kansas natural? Abstract S53E-08 presented at 2014 Fall Meeting, AGU, San Francisco, Calif., 15–19 Dec
70. Sanford AR, Mayeau TM, Schlue JW, Aster RC, Jaksha LH (2006) Earthquake catalogs for New Mexico and bordering areas II: 1999–2004. New Mexico Geol 28:99–109
71. Saucier A, Frappier C, Chapuis RP (2010) Sinusoidal oscillations radiating from a cylindrical source in thermal conduction or groundwater flow: closed-form solutions. Int J Numer Anal Methods Geomech 34(16):1743–1765
72. Seeber L, Armbruster JG, Kim WY (2004) A fluid-injection-triggered earthquake sequence in Ashtabula, Ohio: implications for seismogenesis in stable continental regions. Bull Seismol Soc Am 94(1):76–87
73. Sharp JM (2003) A glossary of hydrogeological terms. Department of Geological Sciences, The University of Texas, Austin
74. Simpson D (1986) Triggered earthquakes. Annu Rev Earth Planet Sci 14:21
75. Smith DG (1981) The Cambridge encyclopedia of earth sciences. In: Smith, DG (ed) The Cambridge encyclopedia of earth sciences, Crown and Cambridge University Press, Cambridge, 496 p, 1
76. Sumy DF, Cochran ES, Keranen KM, Wei M, Abers GA (2014) Observations of static coulomb stress triggering of the November 2011 M5. 7 Oklahoma earthquake sequence. J Geophys Res Solid Earth 119(3):1904–1923
77. Terry-Cobo S (2013) Adviser for OKC-based Chesapeake energy details wastewater reuse. J Rec. Retrieved from http://search.proquest.com/docview/1446101145?accountid=4117
78. Terry-Cobo, S. (2013). Conservation in drilling: OKC-based Devon reuses water in gas production process. J Rec. Retrieved from http://search.proquest.com/docview/1283189143?accountid=4117
79. Theis CV (1935) Relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using ground water storage. Trans Am Geophys Union 2:519–524
80. Thiem G (1906) Hydrologische methoden. JM Gebhardt's verlag, Leipzig
81. Townend J, Zoback MD (2000) How faulting keeps the crust strong. Geology 28(5):399–402
82. U.S. Energy Administration (2016) Summary of natural gas supply and disposition in the United States, 2014–2016. Retrieved 18 May 2016, from U.S. Energy Information Administration. http://www.eia.gov/naturalgas/monthly/pdf/table_01.pdf
83. U.S. Geological Survey (2014) Introduction to hydraulic fracturing. Retrieved 8 Dec 2014, from USGS: Science for a Changing World. http://www.usgs.gov/hydraulic_fracturing/
84. U.S. Geological Survey (2014) Measuring the size of an earthquake. Retrieved 8 Dec 2014, from USGS. http://earthquake.usgs.gov/learn/topics/measure.php
85. Veatch RW (1983) Overview of current hydraulic fracturing design and treatment technology-part 2. Soc Pet Eng. https://doi.org/10.2118/11922-PA
86. Veil JA (2010) Water management technologies used by Marcellus shale gas producers. U.S. Department of Energy. Argonne National Laboratory, Argonne
87. Walsh FR, Zoback MD (2015) Oklahoma's recent earthquakes and saltwater disposal. Sci Adv 1(5):e1500195
88. Weingarten M, Ge S, Godt JW, Bekins BA, Rubinstein JL (2015) High-rate injection is associated with the increase in US mid-continent seismicity. Science 348(6241):1336–1340
89. Yeck WL, Sheehan AF, Weingarten M, Nakai J, Ge S. The 2014 Greeley, Colorado earthquakes: science, industry, regulation, and media. Abstract PA23C-05 presented at 2014 Fall Meeting, AGU, San Francisco, Calif., 15–19 Dec

Infrared Thermographic Imaging in Geoengineering and Geoscience

Ömer Aydan
Department of Civil Engineering, University of the Ryukyus, Nishihara, Okinawa, Japan

Article Outline

Glossary

Earthquake prediction It is a branch of earthquake science concerned with the prediction of the time, location, and magnitude of earthquakes

Geoengineering It fundamentally refers to engineering on/in the earth crust and it is a hybrid discipline of civil engineering, mining engineering, petroleum engineering, and earth science

Geoscience It is a branch of science concerning the mechanism of formation of geological processes, earthquakes, mountain ranges and ocean basins and the movements of continental plates of the earth

Infrared imaging It is an imaging technique utilizing thermographic cameras to detect radiation in the long-infrared range of the electromagnetic spectrum and to produce images of that radiation, called thermograms

Rockburst It is a spontaneous, violent fracture and detachment of rock. It occurs in underground excavations subjected to high stress condition and it is also reported to occur on surface excavations such as high slopes and foundation of dams

Definition of the Subject and its Importance

The prediction of earthquakes in geoscience and the prediction of the failure of structures in geoengineering are the long-desired goals of geoengineers and geoscientists. Furthermore, infrared thermographic imaging technique could be a tool for the maintenance purpose of structures in geoengineering. The entry explores the possibility of the utilization of infrared thermographic imaging technique for such purposes.

Introduction

Earthquakes are known to be one of the most destructive natural disasters resulting in the huge losses of human lives and properties as experienced in the past and recent earthquakes. As there is no way to prevent the occurrence of earthquakes, the prediction of earthquakes is of great importance for mankind.

Rockbursts under high stress environment in mining and civil engineering are of prime concerns during excavations of underground as well surface structures. Therefore, the prediction of real-time stability of underground excavations in geoengineering is also of great importance.

The maintenance of various structures is also of prime importance for engineers. When materials constituting structures deteriorate, their ther-

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_1025

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_1025-1

mal properties and heat absorption characteristics change, which may be important elements for determining their deterioration state.

The utilization of the infrared thermographic imaging (IRTIT) was explored by the author through an experimental and monitoring program for assessing the real-time stability of underground excavations against rockburst in underground excavations, thermomechanical responses of discontinuities, and fault gouges in relation to heat release during shearing process, which may be also of great value in earthquake prediction [1–3]. This entry explores the possibility of applying the infrared thermographic imaging technique in geoengineering and geoscience. First, the principles of thermographic imaging by infrared camera are briefly explained. Then, the experimental results of a series of laboratory tests on rock specimens and large rock blocks having a circular hole are tested under compression environment in relation to the deformation and rupture processes; discontinuities subjected to dynamic shearing are presented. Later, some observations that were carried out at the actual structures are presented. Finally, the possibility of applying the infrared imaging in geoengineering and geosciences is presented and discussed.

Principles of Thermographic Imaging by Infrared Camera

Temperature variations can be measured through thermographic imaging technique utilizing infrared cameras [4–8]. The wavelength of infrared rays used in the infrared cameras ranges from 0.7 μm to 14 μm, and the focal plane arrays were one-dimensional initially, and they become two-dimensional. Infrared thermography technology has been recently advanced, and it is possible to measure temperature variations with a sensitivity of 0.1 mK at a sampling rate of 380 Hz. With a lower resolution, the sampling interval can be much smaller. The infrared thermographic imaging technique utilizes long-wave infrared rays (LWIR), medium-wave infrared rays (MWIR), and short-wave infrared rays (SWIR). Infrared cameras can be of cooled or non-cooled type.

It is also well known that temperature variations do occur during loading of geomaterials and their amplitude depends upon the rate of loading. Weber [9] was first to recognize the temperature variations during the loading of solid materials. Kelvin was the first to formulate the effect of temperature in thermoelasticity [10]. The fundamental reasoning to those variations can be simply explained through the energy conservation law of the continuum mechanics, which is given below (see, e.g., [11]):

$$\rho c \frac{\partial T}{\partial t} = -\nabla \cdot \mathbf{q} + \boldsymbol{\sigma} \cdot \dot{\boldsymbol{\varepsilon}} \pm Q \qquad (1)$$

where $\rho, c, T, \mathbf{q}, \boldsymbol{\sigma}, \dot{\boldsymbol{\varepsilon}}, Q$ are density, specific heat, temperature, directional derivative operator, heat flux, stress, strain rate, and heat production (or sink), respectively. The second term on the right-hand side of Eq. 1 corresponds to mechanical work done during deformation process per unit time. This term corresponds to the energy variation during deformation. During plastic deformation of rocks or slippage along discontinuities, some part of dissipated mechanical energy would also be transformed to heat, which would induce temperature increase on rocks. These temperature variations may be monitored using the infrared cameras, and this type of monitoring may be designated as "passive monitoring." When heat environment is almost constant, the variations of temperature would be almost very small so that the thermographic images could not be distinct to observe the deteriorated state of surrounding materials for maintenance monitoring purposes. It should be also noted that if the body is externally excited by mechanical actions, for example dynamic forces such as cyclic loading, impacts etc., one should also expect some temperature variations.

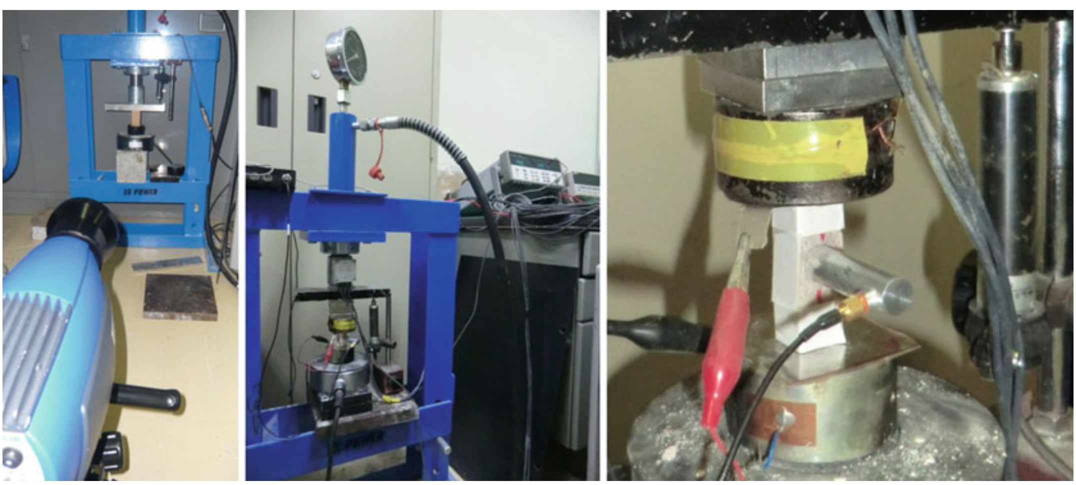
When Eq. 1 is reconsidered, steady-state or transient temperature variations may be observed depending upon the characteristics of heat source or heat fluxes and dynamic passage of waves. These variations would depend upon heat absorption or desorption characteristics of constituting materials in view of their density and specific heat constants. Thermographic imaging by infrared camera using external excitation sources, such as electrical heating (induction, conduction), optical heating (intense light, lasers), mechanical loading, vibrations (high-power ultrasound), or convective heating, may be designated as "active monitoring." In this entry, two types of infrared thermographic imaging are experimentally investigated in the first part, and then their possible applications in geoengineering and geosciences are presented in the remaining part.

Uniaxial Compression and Brazilian Experiments on Minerals, Rocks, and Fault Gouges

Experimental Setup, Devices, and Materials

The experiments were first carried out at the rock mechanics laboratory of Aydan, when he was associated with Tokai University (TU) until the end of October 2013 [1, 2]. The loading machine used by the author is a low-stiffness machine with a loading capacity up to 100 kN. The loading was manually performed so that the environmental electrical and electronic noises due to the loading system were absent as seen in Fig. 1. The applied load and induced displacement were automatically measured and stored on the hard disk of a laptop computer. The acceleration responses of the samples during fracturing were measured by Yokogawa WE7000 measurement station using the AR-10TF-type accelerometers of TOKYO SOKKI, which can measure three components of accelerations up to 10G with a frequency range of 0–160 Hz. The accelerometers were either attached at top and bottom platens. However, if the sample was large enough, the accelerometers were attached directly onto samples. The accelerations were measured in the direction of loading and in two mutually perpendicular directions to the loading direction.

As expected from the energy conservation law of continuum mechanics given by Eq. 1, the thermal response of geomaterials would be observed as some of the mechanical energy is transformed into heat during deformation and fracturing. The author and his group have first attempted to measure temperature variations using contact-type temperature measurement setups as well as infrared thermographic imaging technique in the past (i.e., [12]). Recently, the infrared thermographic imaging technique is greatly improved, and the author reused this technique to observe the thermal response of minerals and rocks in compression and Brazilian tensile experiments and compression of quartz fault gouge. The infrared cooled camera used in experiments is produced by FLIR, and its type is SC5500 (Fig. 1). It has an image resolution of 320 x 256 pixels for a frame rate of 380 Hz. It has InSb focal plane area, and the wavelength of



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 1 Views of experimental devices and setups infrared rays ranges between 3 and 5 μm. In the experiments the first 12 tests were conducted with a resolution of 160 x 128 pixels and a frame rate 200 of Hz. The rest of the five experiments were performed using a resolution of 160 x 128 pixels and a frame rate 500 of Hz.

Experiments on Minerals

Several crystals such as quartz, orthoclase, tourmaline, calcite, aragonite, and gypsum were selected, and experiments were conducted. During experiments, high strength minerals such as quartz and tourmaline failed in a mode of explosion, while gypsum, rocksalt, and aragonite failed in a very ductile manner. In this section, the experiments on quartz and gypsum are only presented.

Quartz

Responses of several measurable parameters such as stress, strain, electric potential, and cumulative AE are shown in Fig. 2 together with views of quartz crystal before and after the failure. The failure mode of the quartz crystal was like an explosion, and the remains of the sample after the experiment were powderized as seen in the same figure. The maximum acceleration was 13 times the gravitational acceleration. Distinct variations of various measurable parameters such as electric potential and acoustic emission besides load and displacement were observed during deformation and fracturing processes.

An infrared thermographic image and temperature response along the selected line of the quartz crystal sample during the initiation of failure is shown in Fig. 3. The maximum temperature was observed almost at the center of the top surface of the crystal, and the temperature difference was about 16 °C from the ambient temperature. Furthermore, one can easily notice the ejection trajectories of some fragments from the failing quartz sample.

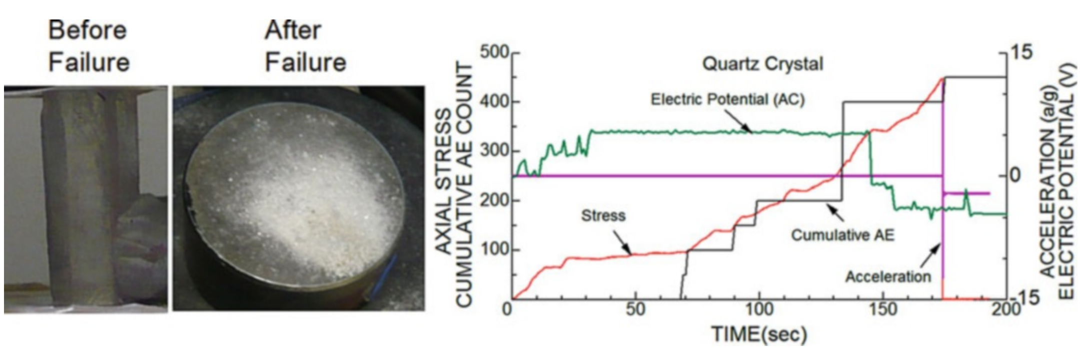
Gypsum

Gypsum crystal failed in a very ductile manner. Figure 4 shows responses of several measurable parameters such as stress, strain, and cumulative AE that are shown together with views of gypsum crystal before and after the failure. As noted from Fig. 4, the gypsum crystal failed in the mode of buckling existing cleavage system in the crystal, which cannot be distinguished from the visual observation. The maximum acceleration was only 16 gals, and it is 0.0163 g times the gravitational acceleration, which is extremely small as compared with that observed during the failure of quartz mineral.

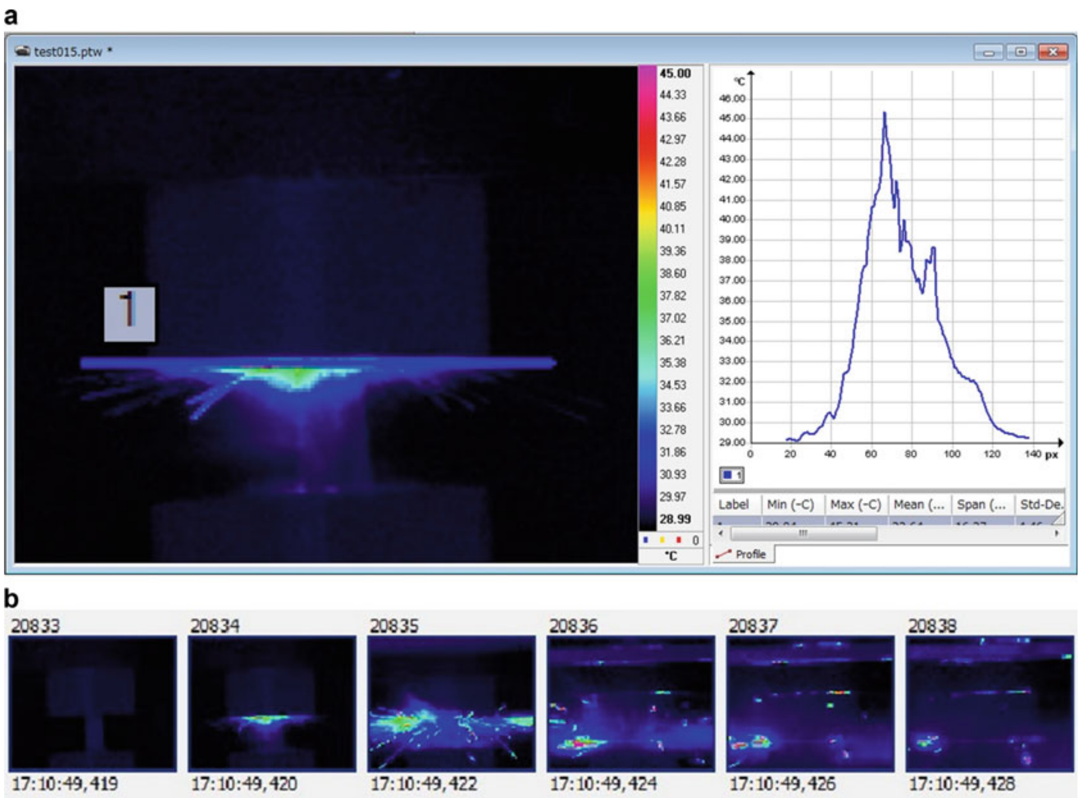
An infrared thermograph image and temperature response along the selected line in the observed outer surface of gypsum crystal sample during the initiation of failure is shown in Fig. 5. The maximum temperature rise of the sample was only 0.85 °C, which is very low compared with that observed during the experiment on the quartz crystal.

Rocks

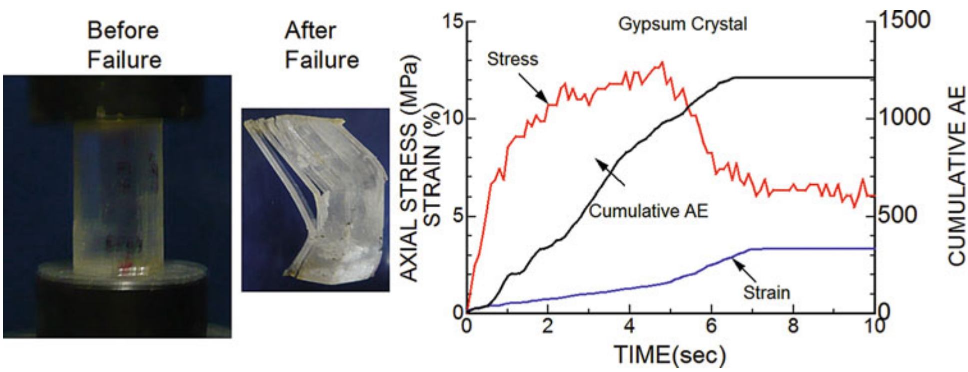
Observations on real-time infrared thermographic responses of some typical rocks ranging from soft mudstone to quartzite were carried during the deformation and fracturing processes. Rocks



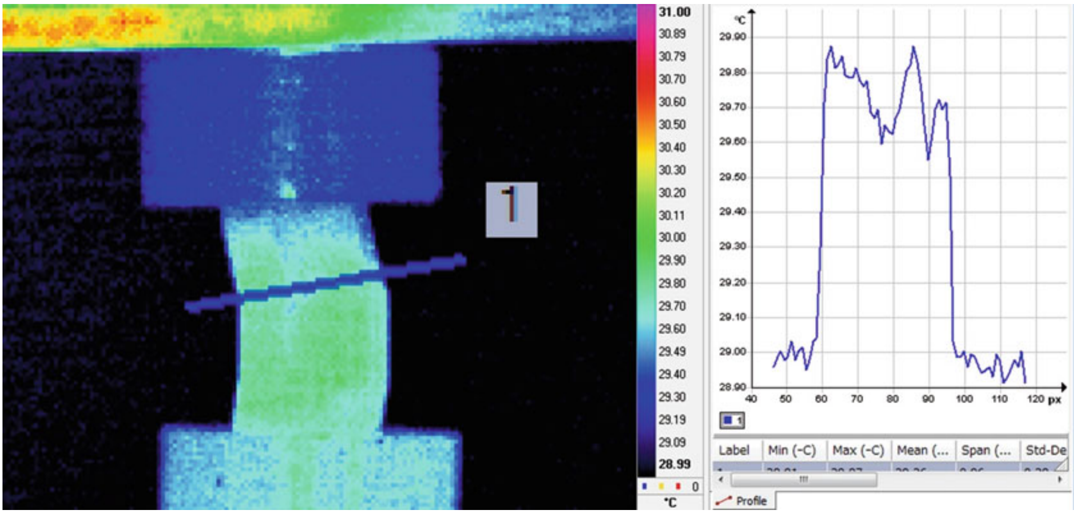
Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 2 Multiparameter responses of quartz crystal sample and its view before and after failure were mudstone from Seyitömer (Turkey), shale from Donghekou (China), tuff and andesite from Sinop (Turkey), chert from Balakot (Kashmir), sandstone from Ehime (Japan), marbles from Kuşini (Selçuk, Turkey) and Muğla (Turkey), porous basalt from Mt. Fuji (Japan), granite from Inada (Ibaraki, Japan) and Kaore (Gifu, Japan), and rhyolite from Kaore (Gifu, Japan). Most of



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 3 (**a**) An infrared thermograph image of the quartz crystal during failure and temperature rise along the selected line, (**b**) sequential infrared thermographic images during fracture process



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 4 Multiparameter responses of gypsum crystal sample and its view before and after failure



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 5 Infrared thermograph image of the gypsum crystal during failure and temperature rise along the selected line

experiments were carried out under uniaxial compression loading condition, and two Brazilian tests were carried out on granite and rhyolite samples. In this section, we present results of some of these experiments and discuss their implications.

Uniaxial Compression Tests

Low Strength Rocks

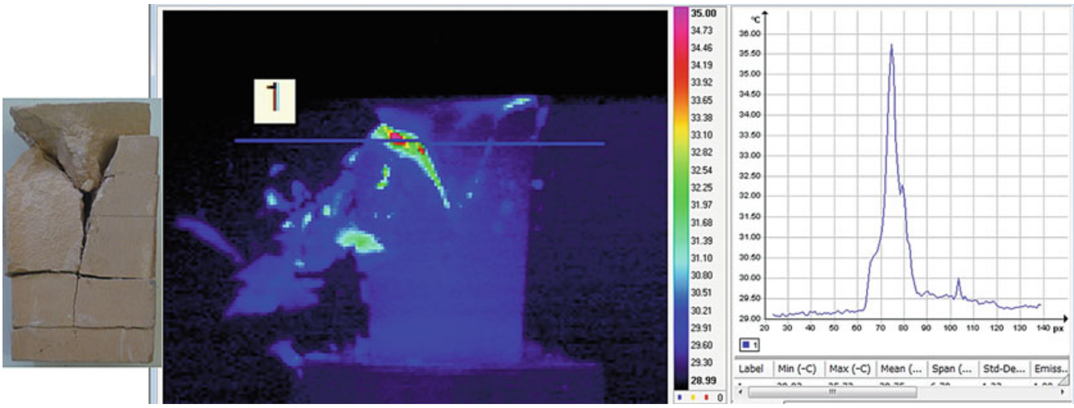
Experimental results on mudstone sample from Seyitömer open-pit lignite mine and tuff sample from Sinop Peninsula of Turkey are presented herein. Figure 6 shows an example of infrared thermographic image of Seyitömer mudstone and temperature distribution along the line indicated in the image. It is also interesting to note that a thermal band, which coincides with fracture zone, is observed. This may be of great value for identifying the possibility of rockburst and their locations in rock engineering and earthquake faults in geoscience.

Figure 7 shows the multiparameter response of the mudstone sample from Seyitömer. It is of great interest that acceleration waves with lower amplitude occurred when the macroscopic crack initiation started to occur at the peak strength while the maximum acceleration observed during the failure state as noted in previous experiments reported by Aydan et al. [13] and Ohta and Aydan [14]. Furthermore, the maximum acceleration is about 120 gals.

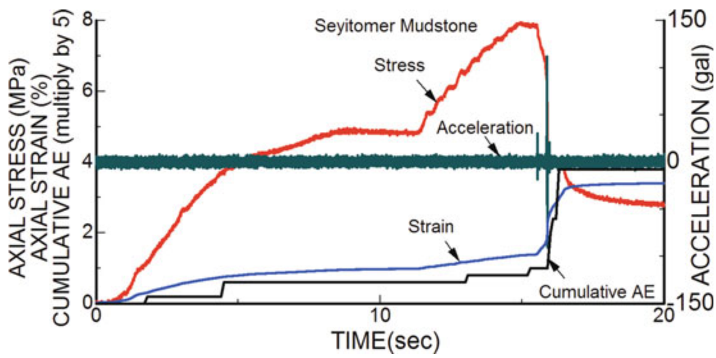
The next example of rock is a tuff sample from Sinop Peninsula. Figure 8 shows a view of the sample after experiment, infrared thermal image of the sample during fracturing, and its associated temperature distribution along the fracture plane and stress-strain relation. It is of great interest that the temperature distribution is higher along the fracture plane. The first sudden drop in strain-stress relation is associated during the splitting crack in the middle of the sample. Then, the sample again continued to sustain a higher load.

Medium Strength Rocks

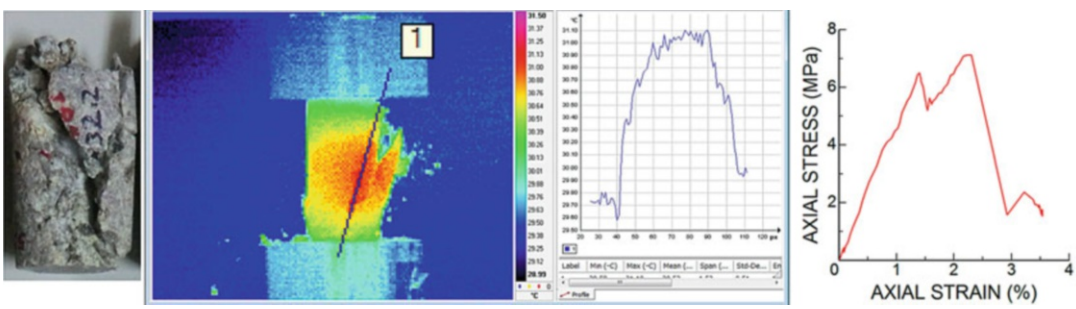
Although a number of experiments were carried out on medium strength rocks, experiments on two examples are described herein. The first example is a marble sample collected from Kuşini antique underground quarry (Selçuk, Turkey) [15]. The antique underground marble quarry, which was, at least, exploited about 2000 years ago, was used in building the structures and monuments in Phrygian Efes (Ephesus) antique city. Figure 9 shows an example of infrared thermographic image of Kuşini marble and temperature



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 6 Actual view and infrared thermograph image of Seyitömer mudstone together with temperature distribution along the chosen line



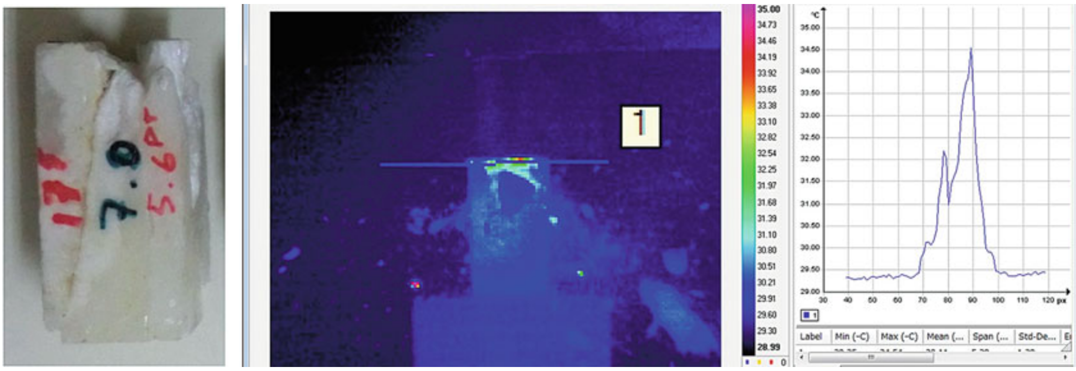
Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 7 Multiparameter responses of Seyitömer mudstone



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 8 Actual view and infrared thermograph image of Sinop tuff together with temperature distribution along the chosen line and strain-stress response during the experiment

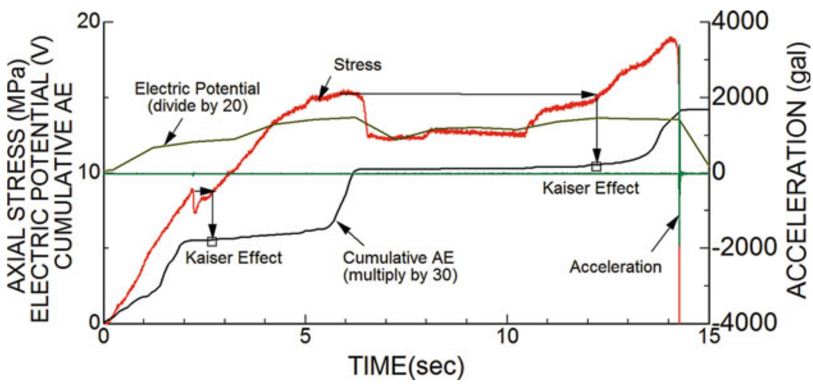
distribution along the line indicated in the image during fracturing. The temperature inferred reached about 34.5 °C compared to the ambient temperature of about 29–30 °C.

Figure 10 shows the multiparameter response of the marble sample from Kuşini antique underground quarry. There are several important observations in the figure. As the load was manual,



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 9 Actual view and infrared thermograph image of Kuşini marble together with temperature distribution along the chosen line

Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 10 Multiparameter responses of Kuşini marble

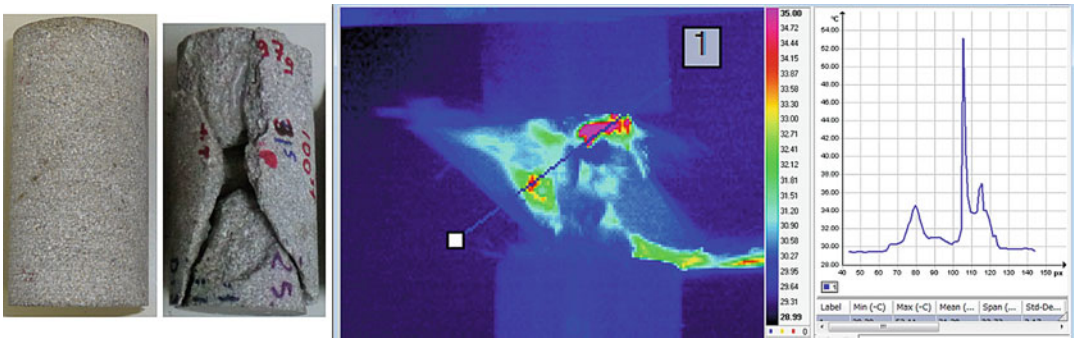


some load reduction occurred during the loading steps. In particular, it is interesting to notice that Kaiser effects known in the acoustic emission technique for in situ stress inference are clearly observed at two stress levels as indicated in Fig. 10. It is also interesting to note that electric potential decreases during relaxation or reduction of load, which was also noted in previous studies by Aydan et al. [12, 16]. In this example, it is also of great interest that the maximum acceleration is observed during the failure state as noted in previous experiments reported by Aydan [17] and Ohta and Aydan [14]. As rock becomes stronger, the maximum acceleration was about 3391 gals and the wave form was asymmetric with respect to time axis as noted previously by Aydan [17].

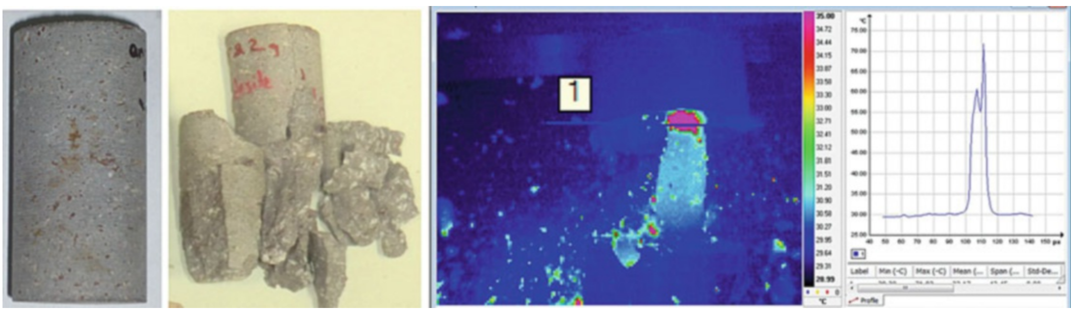
A sandstone sample from Ehime Prefecture (Japan) with a uniaxial compressive strength of 62 MPa was tested. The views of the sample before and after experiment and an infrared thermograph image together with temperature distribution along the selected line are shown in Fig. 11. The sample failed along two conical conjugate fracture planes as seen in the figure. The temperature difference obtained from the infrared observation was about 33 °C. As noted from the figure, temperatures are much higher in the vicinity of fracture planes.

Hard Rocks

An andesite sample from Sinop Peninsula (Turkey) was tested. Views of the sample before and after experiment and an infrared thermographic image together with temperature distribution along the selected line are shown in Fig. 12. The sample failed in a violent manner with too many fracture planes as seen in the figure. The temperature difference obtained from the infrared observation was about 42 °C. As noted from the figure, temperatures are much higher in the sample just below the



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 11 Views of the sample before and after the experiment and infrared thermograph image of Ehime sandstone together with temperature distribution along the chosen line



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 12 Views of the sample before and after the experiment and infrared thermograph image of Sinop andesite together with temperature distribution along the chosen line

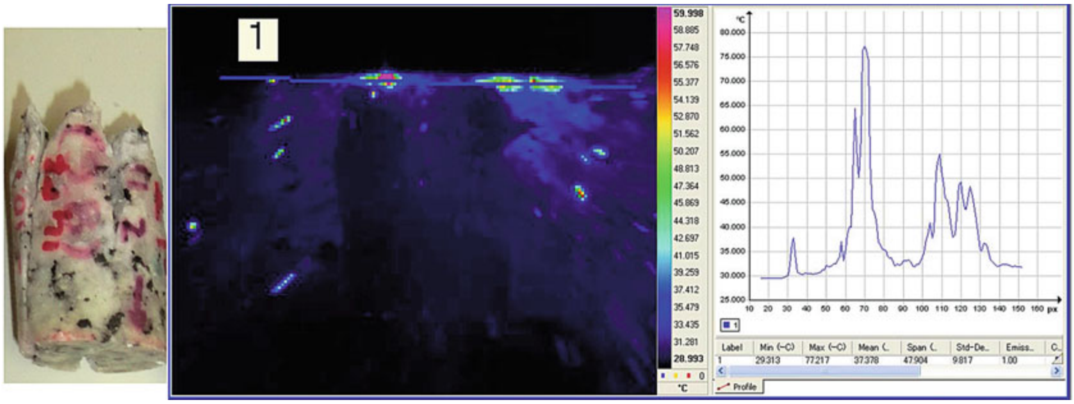
mobile loading platen. The ejection trajectories of rock fragments in space were also observed.

Inada granite from Ibaraki Prefecture of Japan was selected as another example of hard rocks. Figure 13 shows a view of the failed sample, and an infrared thermograph image together with temperature distribution along the selected line is shown in the figure. The temperature difference was more than 47 °C at the time of initiation measured by the infrared thermal camera. The temperatures were higher at the corners where crack initiation started. The sample failed in a violent manner, and the trajectories of fragments of the rock sample together with a powder cloud are easily noticed from the image. From the images it is also possible to obtain ejection velocities, which may be of great importance for assessing the possible ejection distance as well as destruction potential of the rockbursts in actual constructions.

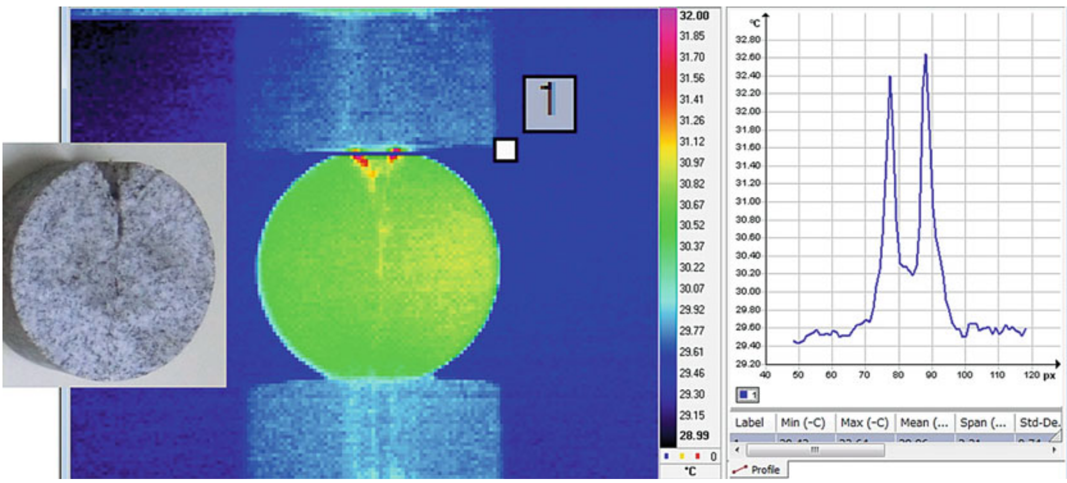
Brazilian Tests

Observations by infrared thermal camera were done in Brazilian tensile strength experiments on Inada granite and Kaore rhyolite samples. Figure 14 shows the sample after experiment and the infrared thermograph image together with temperature distribution along the chosen line shown in the same figure. As noted from the figure, high-temperature bands coincide with the fracture planes. Furthermore, the crack initiation locations coincide with highest temperature spots.

A view of the rhyolite sample from Kaore underground powerhouse (Gifu, Japan) after experiment and its infrared thermograph image together with temperature distribution along the chosen line are shown in Fig. 15. As noted in previous experiments, high-temperature bands appear along some zones before rupture and these high-temperature bands eventually constitute the



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 13 View after the experiment and infrared thermograph image of Inada granite together with temperature distribution along the chosen line

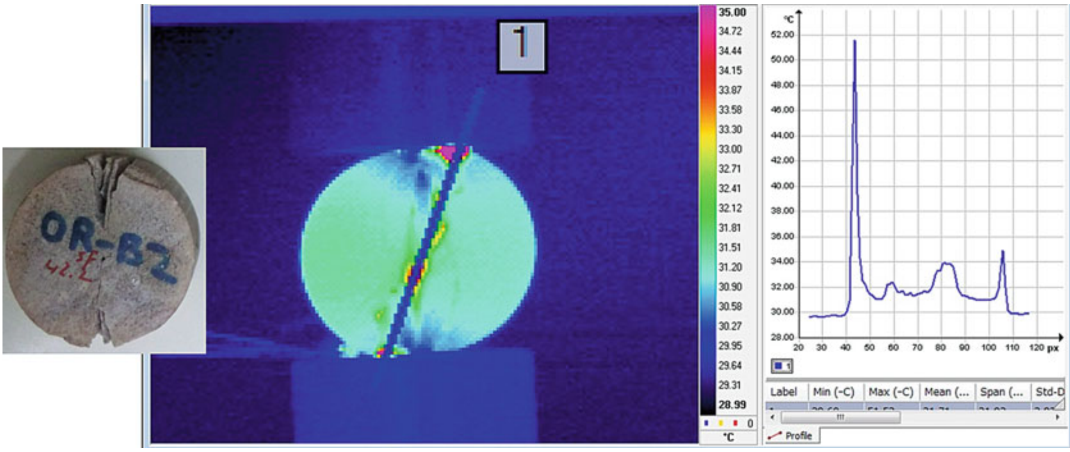


Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 14 Views of the granite sample after the experiment and its infrared thermograph image and temperature distribution along the chosen line

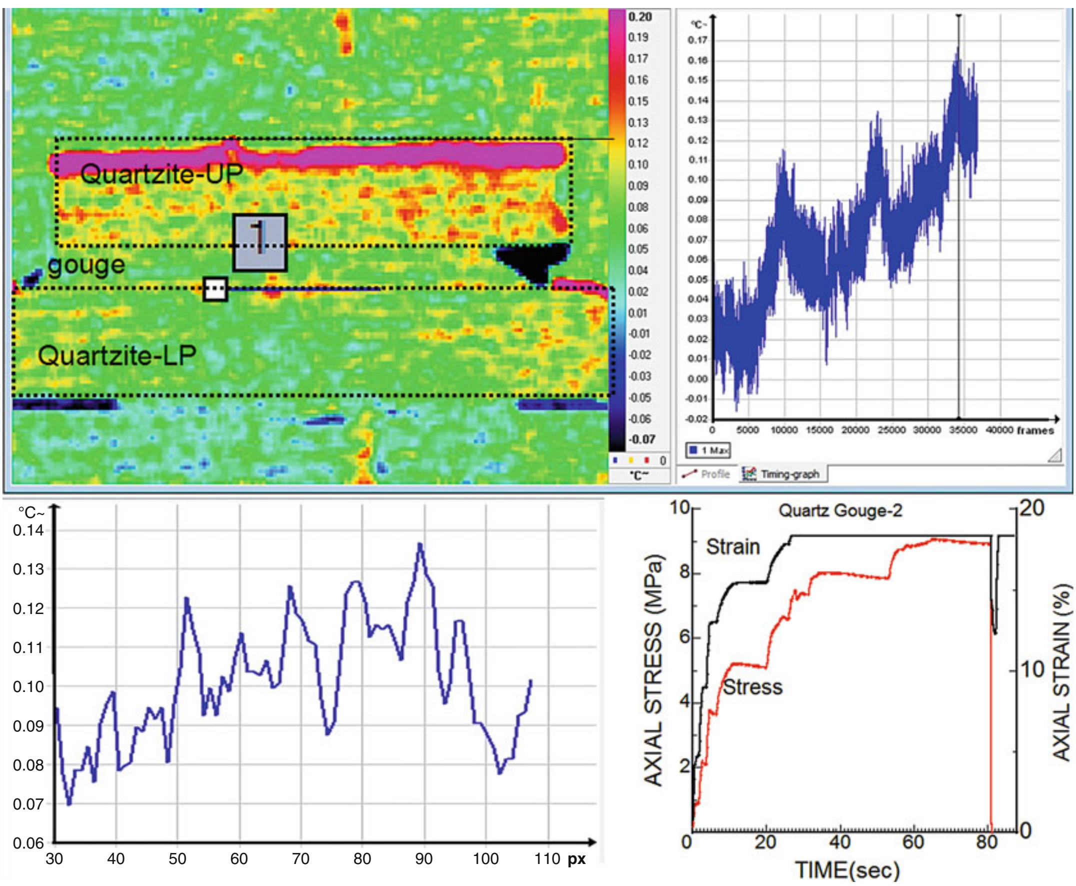
major fracture zones. The temperature rise of the rhyolite sample was higher than that of the granite sample. In other words, temperature is higher if rock becomes harder and fails in a brittle manner.

Gouge

Two experiments were carried out to see the thermal response of gouge material (granular quartzite) sandwiched between plates of steel (GE-S) and quartzite (GE-Q) during compression. The main purpose was to study possible temperature variations along shear zones and fracture/fault zone in actual circumstances. In this section, the observations on the experiment denoted QE-Q (Gouge Experiment 2) are reported herein. Figure 16 shows the infrared thermograph image, temperature distribution along the chosen line, temperature variation with time, and strain-stress response during experiments. Although temperature variations were somewhat less than those expected, very interesting observations were done. As gouge material was granular quartz, temperature distribution fluctuates along the chosen line. Lower temperature is expected at those grains a little behind from those where temperature is higher. Another important observation



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 15 Views of the rhyolite sample after the experiment and its infrared thermograph image and temperature distribution along the chosen line



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 16 Infrared thermograph image of quartz gouge together with temperature distribution along the chosen line and strain-stress response during the experiment is the temperature rise in relation to the imposed load on the gouge material. When stress and strain increase, temperature increases. However, if the load is kept constant, temperature tends to decrease. All these responses can be directly explained if the energy conservation law is taken into account.

Experiments on Rock Discontinuities during Dynamic Shearing

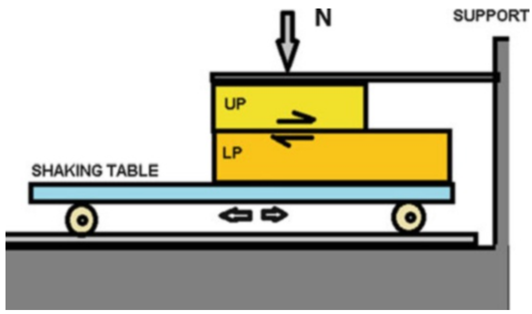
The first set of experiments was carried out on the schistosity plane of metamorphic quartzite in 2014, and similar experiments are recently repeated on discontinuities in andesite, green schists, diorite, granite, gabbro, basalt, limestone, and quartzite. The results are fundamentally similar. In this part, the experiments on the schistosity plane of quartzite are presented.

Experimental Setup

The experimental setup consists of a shaking table, Testo 885 or 890–2 infrared cameras, accelerometers, non-contact displacement measurement acquisition system, and a specimen. In addition a contact-type temperature sensor is attached next to the discontinuity plane on the upper block. The shaking table is utilized to induce cyclic shearing on the discontinuity planes. The stroke of the shaking table is 30 mm with a maximum speed of 240 RPM, and its speed can be adjusted at any desired level.

The shearing motions are monitored using accelerometers and non-contact laser transducers produced by KEYENCE. A stand-alone-type accelerometer QV3-OAM-SYC or TOKYO SOKKI is attached to the shaking table to measure imposed accelerations, which may also be used to evaluate the other dynamic motion parameters.

The specimen consists of lower and upper blocks. The lower block is attached to the movable shaking table, while the upper block is attached to the unmovable support as illustrated in Fig. 17. The reason to attach the lower block to the shaking table was to prevent rotational movements during shearing. Figure 18 shows a view of a typical experimental setup.



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 17 An illustration of experimental setup

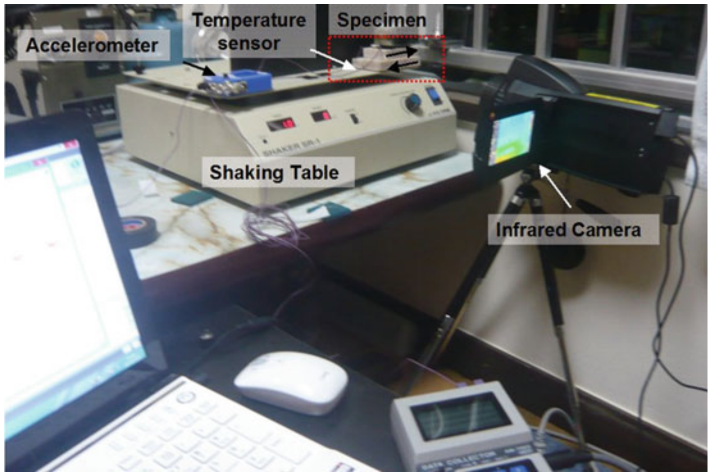
A contact-type temperature sensor was also fixed near the discontinuity plane in the first experiment. As the measured responses by contact gauges were not as accurate as those from the infrared camera, it was decided not to use it in the rest of experiments.

Some experiments were also carried out using Testo 880 IR camera. However, the sensitivity of the camera and the resolution of IR images were much lower than Testo 885 or 890–2 IR cameras. Furthermore, the continuous recording was not possible due to limitations of the Testo 880 IR camera. Therefore, the experiments done using the Testo 880 IR camera are not reported herein.

Experiments

Geometrical dimensions and mechanical and frictional properties of blocks are given in Tables 1, 2, and 3. Schistosity planes are chosen as shearing planes. Schistosity planes also include some micaceous minerals such as muscovite.

The number of cyclic shear tests was five, and the normal stress was varied in each test. Table 4 gives the normal stress levels in a respective cyclic shear test. Although the level of normal stress is lower compared with those used in the conventional shearing test, these normal stress levels should be appropriate if distinct heating response is achieved. There is no doubt that the heating would be higher if normal stress level and/or shearing velocity becomes higher. Although the normal stress level is comparatively small, the maximum shearing velocity was about 1 mm/s in all experiments. Furthermore, the duration of



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 18 A view of experimental setup and instrumentation

Infrared Thermographic Imaging in Geoengineering and Geoscience, Table 1 Geometry and weight of blocks

Block	Length (mm)	Width (mm)	Height (mm)	Weight (gf)
UB	73.1	49.7	14.8	134
LB	90.4	60.8	14.2	197

Infrared Thermographic Imaging in Geoengineering and Geoscience, Table 2 Material properties

Unit weight (kN/m^3)	UCS (MPa)	Elastic modulus (GPa)	P-wave velocity (km/s)	S-wave velocity (km/s)
24.9–25.2	194	196	4.86	2.47

Infrared Thermographic Imaging in Geoengineering and Geoscience, Table 3 Friction properties

State	Fresh surface	Worn surface
Friction angle	23–26	21–23

Infrared Thermographic Imaging in Geoengineering and Geoscience, Table 4 Test names and normal stresses

Test name	Normal stress (kPa)
CST-1	0.37
CST-2	2.05
CST-3	3.73
CST-4	5.87
CST-5	11.37

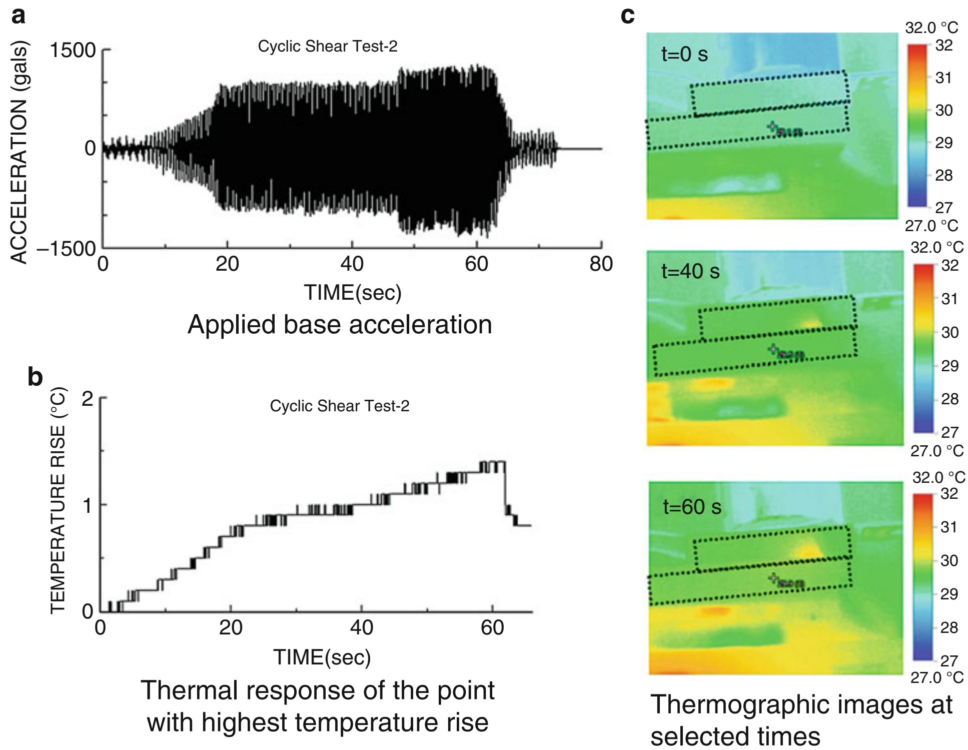
shearing is another factor influencing the thermal responses. The responses and results of experiments are numbered CST-2, CST-3, and CST-5.

Cyclic Shear Test 2 (CST-2)

Figure 19a, 19b shows the acceleration applied to the shaking table and temperature response of the point with the highest temperature rise during the experiment. The maximum surface temperature rise was 1.4 °C at the point where the highest temperature was observed. The maximum amplitude of the acceleration was 1346 gals. Surface temperature of the sample during three selected time steps (namely, 0 s, 30 s, and 60 s) is shown in Fig. 19c. Although the discontinuity plane is apparently almost planar, surface temperature distribution of the specimen was not uniform. The highest temperature rises apparently occur at contact areas over the discontinuity plane. In other words, the temperature rise or heat release would be higher at asperities of rock discontinuities during the shearing process. It is also interesting to note that temperature rise is much steeper in relation to the increase of amplitude of acceleration.

Cyclic Shear Test 3 (CST-3)

Figure 20a, 20b shows the applied acceleration to the shaking table and temperature response of the point with the highest temperature rise during the experiment. The maximum amplitude of the acceleration was almost the same as that of the previous experiment. The maximum surface temperature rise



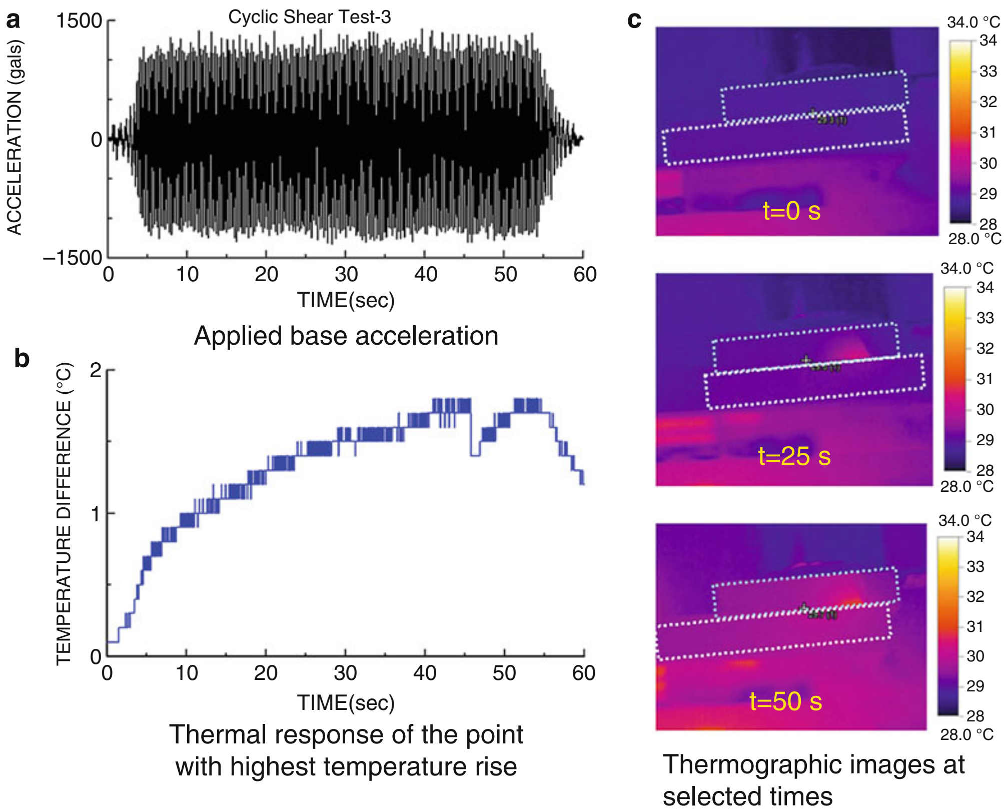
Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 19 Applied acceleration, thermal response, and thermographic images

was 1.8 °C at the point where the highest temperature was observed. The increase in the temperature rise in this experiment compared to that in the previous experiment is directly associated with the increase in normal stress level.

Surface temperature of the sample during three selected time steps (namely, 0 s, 25 s, and 55 s) is shown in Fig. 20c. Surface temperature distribution of the specimen was not uniform, and the highest temperature rises apparently occurred at contact areas (namely, asperities) over the discontinuity plane. It is also interesting to note that temperature rise is also much steeper, while the amplitude of acceleration is increasing. On the other hand, if the shearing speed is constant, the heat release or temperature rise increases at a constant rate. A drop in time-temperature rise response is probably due to external causes related to temperature fluctuations in the vicinity of the testing environment.

Cyclic Shear Test 5

The normal stress level increased and it was almost twice that of the experiment denoted cyclic shear test 5. Figure 21a, 21b shows the applied acceleration to the shaking table and temperature response of the point with the highest temperature rise during the experiment. The maximum amplitude of the acceleration was slightly lower than that in other experiments, and its maximum amplitude was 1186 gals. The maximum surface temperature rise was 3.6 °C with a fluctuation range of ± 0.3 °C. The initially selected point moves in space during shaking, while the selected point in the thermographic image remains the same. Nevertheless, the maximum temperature rise is twice that of the experiment denoted cyclic shear test 3. The increase in the temperature rise in this experiment compared to that in the previous experiment is directly associated with the increase in normal stress level. Regarding the temperature



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 20 Applied acceleration, thermal response, and thermographic images

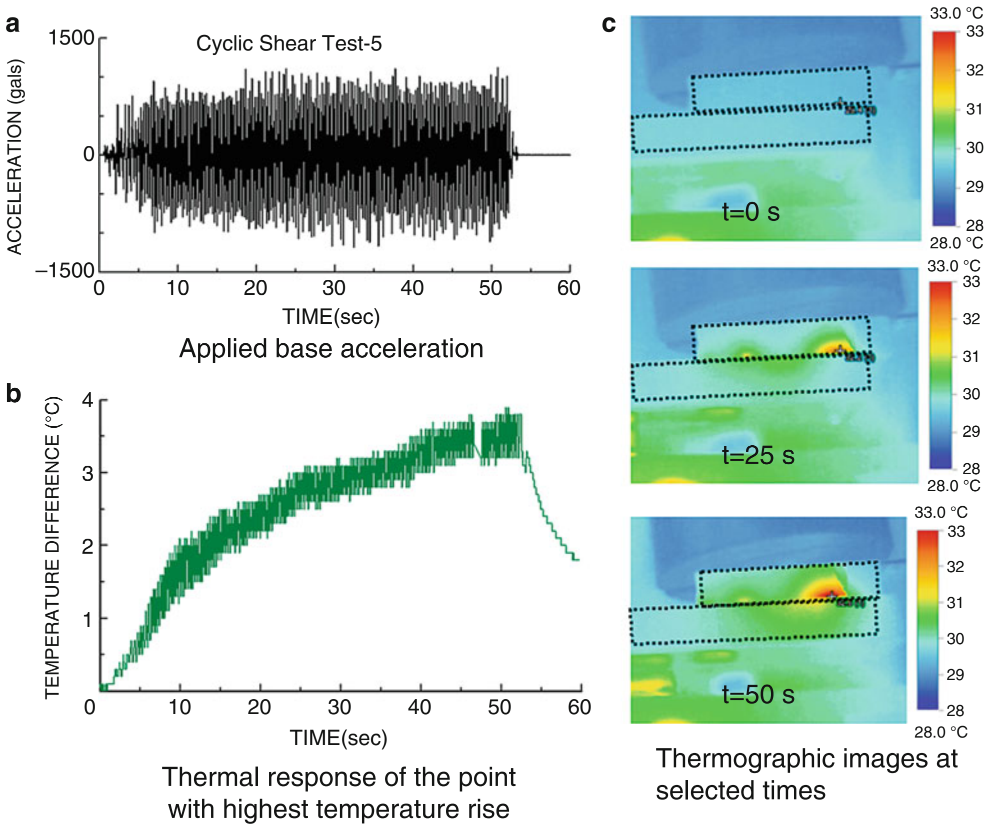
rise, it is again worth to notice that temperature rise is much steeper, while the amplitude of acceleration is increasing. On the other hand, if the shearing speed is constant, the heat release or temperature rise increases at a constant rate.

Surface temperature of the sample during three selected time steps (namely, 0 s, 25 s, and 50 s) is shown in Fig. 21c. It is also noted that the surface temperature distribution of the specimen was not uniform, and the highest temperature rises apparently occurred at contact areas (namely, asperities) over the discontinuity plane. Compared to the contact areas in the previous experiments, the number of contact areas increased due to higher normal stress applied in the experiment. After the experiments, it was noted that a thin powder layer accumulated on the surface of discontinuity surfaces as seen in Fig. 22. In other words, asperities were partially worn out and the damage to asperities of the lower mobile block was higher than that of those at the stationary upper block.

The responses of temperature rises for all experiments are plotted in Fig. 23. As noted from the figure, the temperature rise is highest for the experiment denoted cyclic shear test 5, while it is lowest for the experiment denoted cyclic shear test 1. The temperature rise is higher during the increase of shearing rate, and temperature increase becomes linear as the shearing rate becomes constant. The temperature rise is also related to the duration of shearing. The temperature becomes higher as the duration of shearing increases.

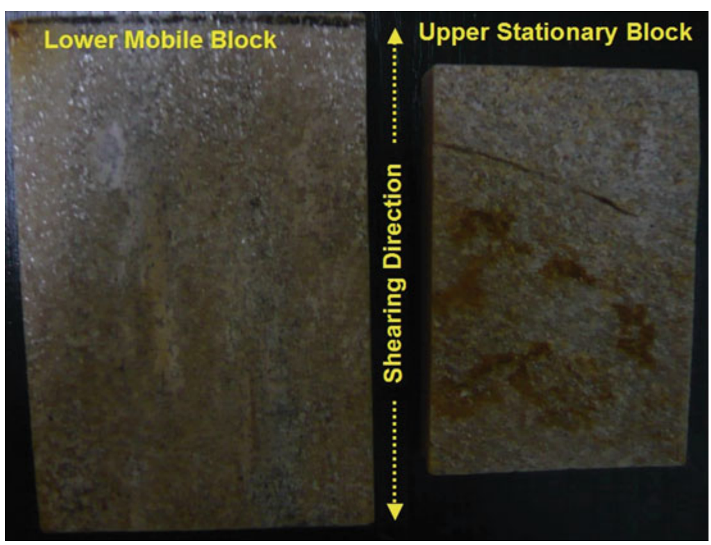
Laboratory Experiments on Rockburst Phenomenon by Infrared Imaging

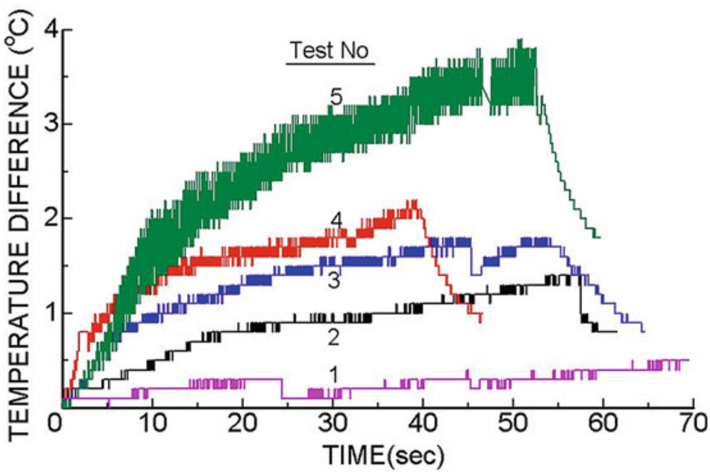
An experimental setup which consists of a compression testing device with a capacity of 2000 kN



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 21 Applied acceleration, thermal response, and thermographic images

Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 22 Views of sheared surfaces





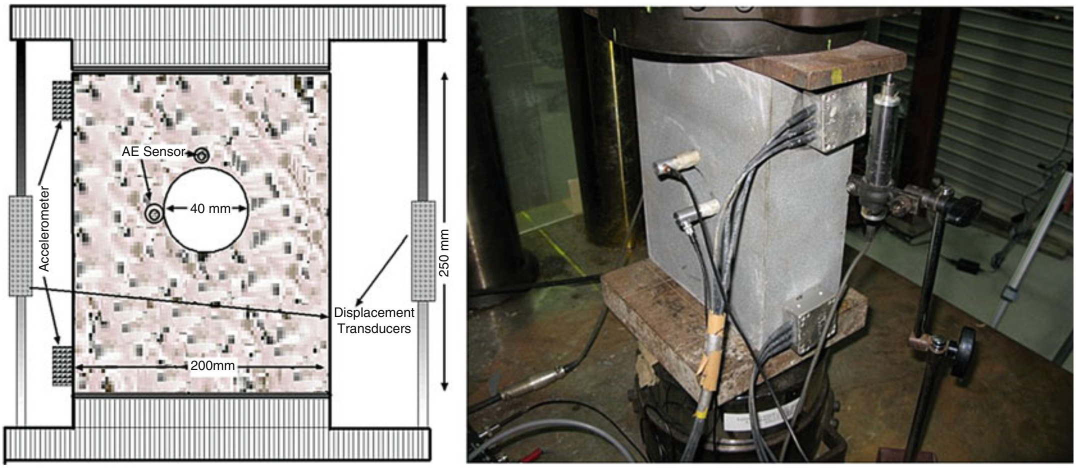
Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 23 Comparison of responses of temperature rises for all experiments

was used and three prismatic rock blocks with a size of 200 x 250 x 90 mm having a circular opening with a diameter of 40 mm subjected to uniaxial compression. Rock samples were obtained from Taru-Toge Tunnel, which is being constructed using drilling-blasting technique as a part of the expressway project connecting Shin-Tomei Expressway and Chuo Expressway at the boundary of Shizuoka and Yamanashi Prefectures in Central Japan [18]. The cores obtained from the rock blocks were tests for obtaining their uniaxial strength. The uniaxial strength of rocks ranged from 26 MPa to 110 MPa, implying that the rock strength could be quite anisotropic with respect to bedding plane orientation (BP). During experiments, the rock blocks were attached with accelerometers and acoustic emission sensors in addition to displacement transducers and load cell to measure the dynamic response of rock blocks during deformation and fracturing processes (Fig. 24).

An infrared camera X8400sc/X6500sc produced by FLIR was used to observe infrared thermographic imaging. Frame rate was 100/s during observations (Fig. 25). Figures 26, 27, 28, 29, 30, and 31 show the multiparameter responses of rock blocks and associated infrared images together with visual images at several time steps for each block. The temperature of infrared thermographic images ranges between 30 °C and 33 °C for all experiments except the test numbered TGT2–2; the blocks failed in a violent manner. Although the appearance of tensile fractures at the roof, which is expected from the stress state induced by the applied loading and boundary conditions, was not clearly observed, spalling started to occur at side walls as expected. While visible spalling is observed at 80–85% of the total load level, AE occurrences started at 35–40% of the total load level, which roughly corresponds to the anticipated yielding stress level at the sidewalls.

Before the macroscopic failure of samples, the ejection of fragments from the side wall of the openings induces vibrations. These vibrations are clearly noted at acceleration responses. When rock blocks fail, the amplitude of accelerations reaches the level of 3 to 4.5 times the gravitational acceleration. The acceleration response is also not symmetric with respect to time axis as noted previously by Aydan [17], Ohta and Aydan [14], and Aydan et al. [19], previously. If the experiment terminated before the total failure, the acceleration level is much less and it is about 240 gals. Nevertheless, it should be noted that the measured accelerations are measured at the top and bottom of the sample near loading platens. In other words, accelerations at the points where the ejection of fragments occurred should be much greater than those measured at the top and bottom of the sample.

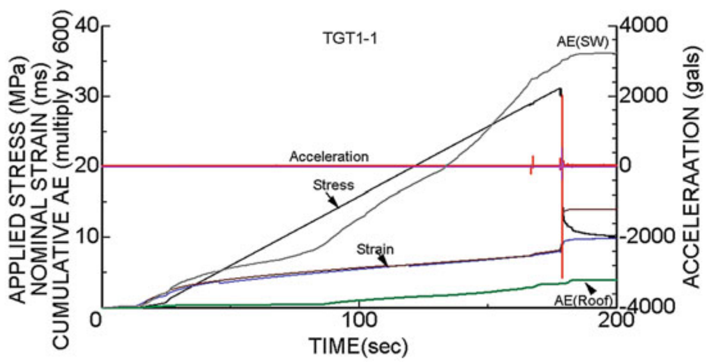
As noted from infrared thermographic images shown in Figs. 27, 29, and 31, high-temperature bands are observed at locations where plastic straining and cracks occurred. These high-temperature spots grow, and ejected fragments appear as high-temperature spots. Once large



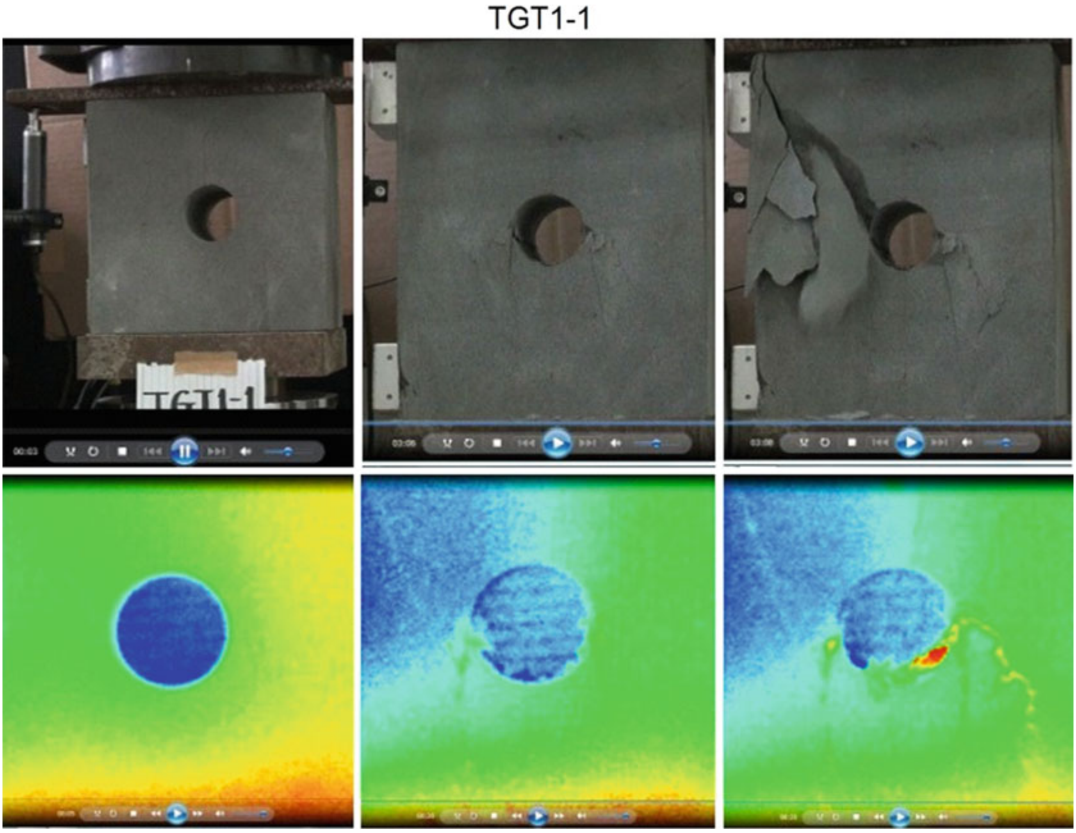
Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 24 Illustration of instrumentation of a rock block sample and its view

Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 25 Views of an infrared camera used in experiments and a typical experiment

Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 26 Multiparameter responses of TGT1–1 rock sample with circular opening

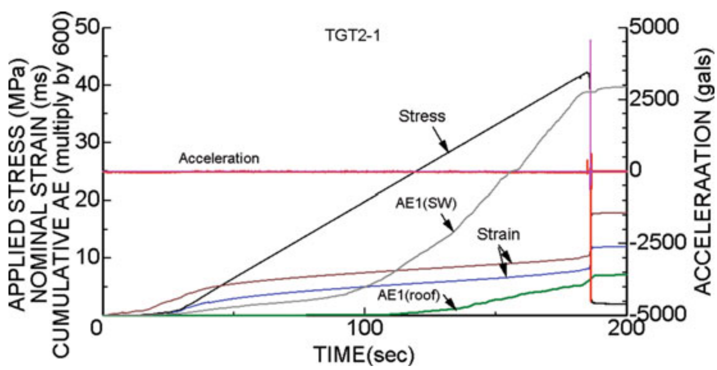


cracks appear, then those high-temperature bands cool down. These observations clearly show that infrared thermographic imaging technique could be a great tool for engineers to identify high-temperature zones as an indicator of likely locations of rock failure in the vicinity of excavation surfaces for assessing the real-time safety of underground excavations.



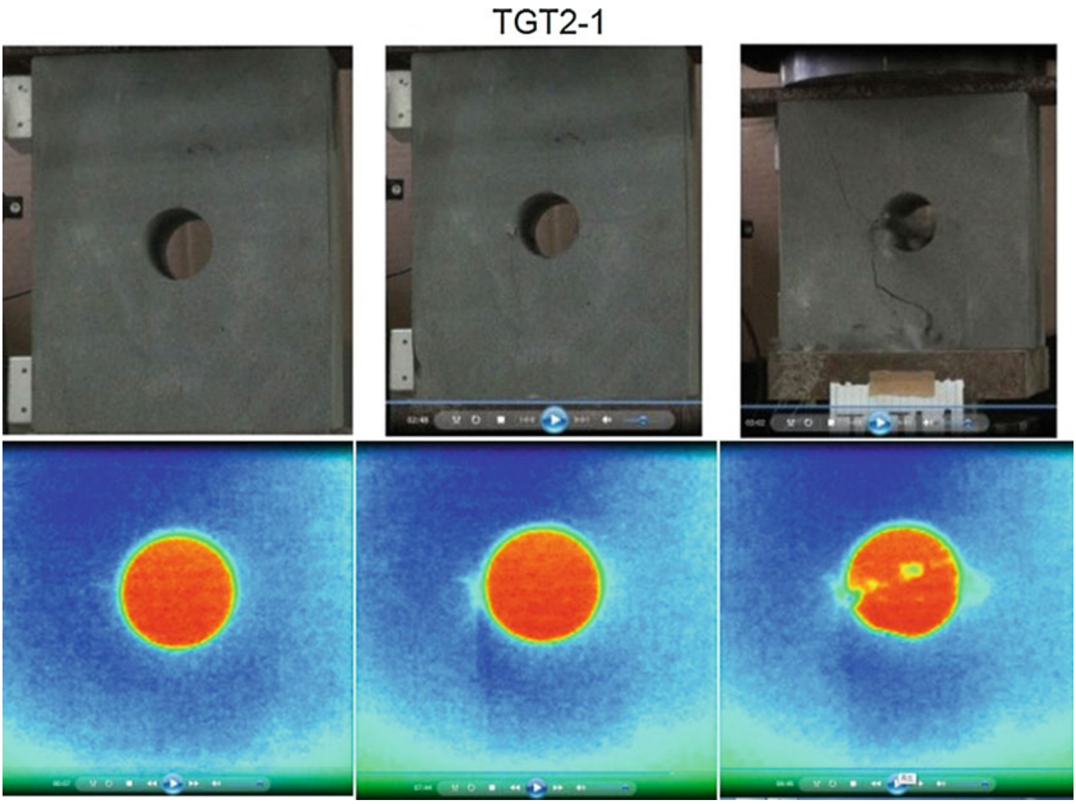
Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 27 Visual and infrared thermographic images of rock block sample TGT1–1

Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 28 Multiparameter responses of TGT2–1 rock sample with circular opening



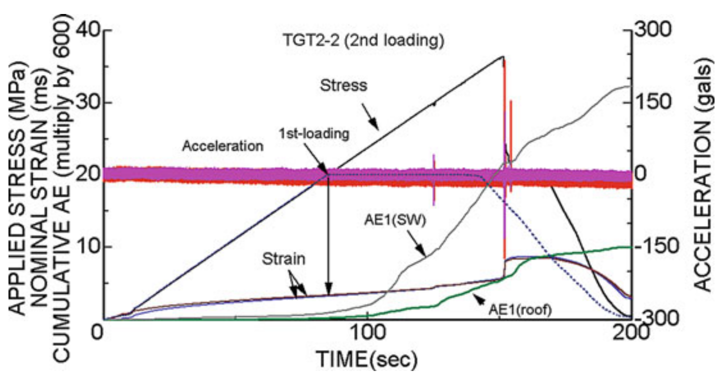
Applications to Underground Excavations

As an application of the infrared thermographic imaging technique to real underground excavations, some observations during the excavation of Taru-Toge Tunnel (Shizuoka Prefecture, Japan), from which rock blocks were obtained, were made. The tunnel is excavated by drilling and blasting technique, which involves the blasting, mucking, shotcreting, installation of rock bolts and steel ribs, and drilling of holes for



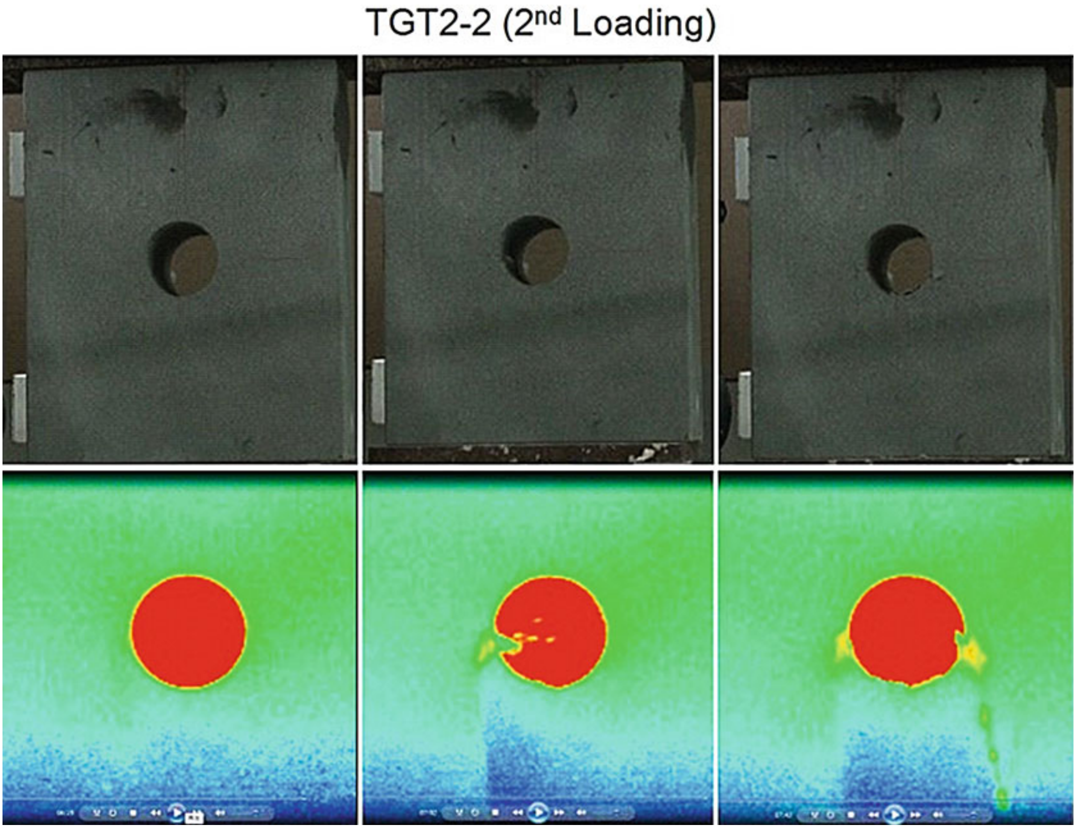
Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 29 Visual and infrared thermographic images of rock block sample TGT2–1

Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 30 Multiparameter responses of TGT2–2 rock sample with circular opening

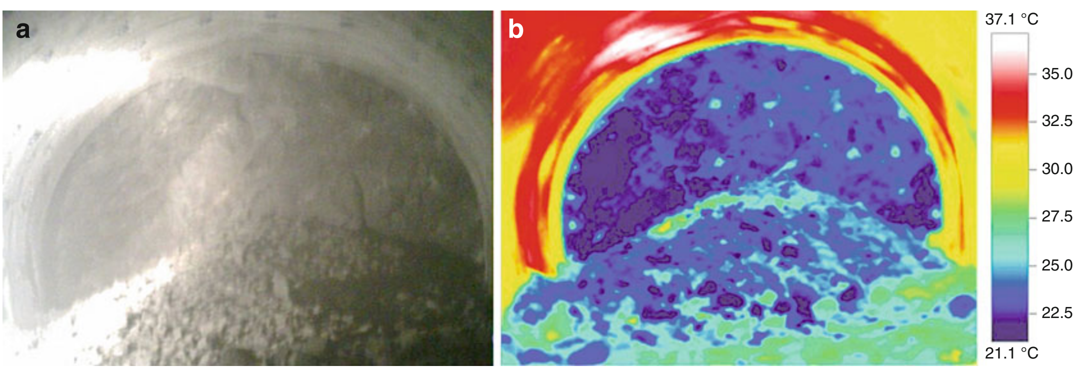


the next blasting round. Forced ventilation is imposed soon after the blasting operation to clear the dust and cloud from the blasted tunnel face. In addition, high-temperature zone next to the blasted tunnel face is the shotcrete layer, which was undergoing hydration process. This observation may also be of great value to assess the quality of shotcrete and its hardening process (Fig. 32b).

Figure 32 shows visible and infrared thermographic images of the evacuation tunnel soon before and after the blasting. An interesting observation is that the infrared thermographic image clearly illustrates the tunnel face condition, while

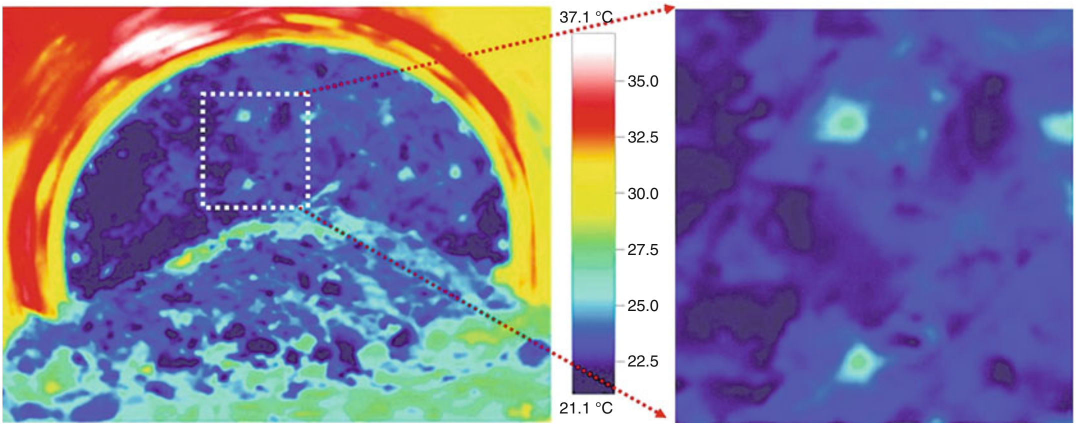


Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 31 Visual and infrared thermographic images of rock block sample TGT2–2

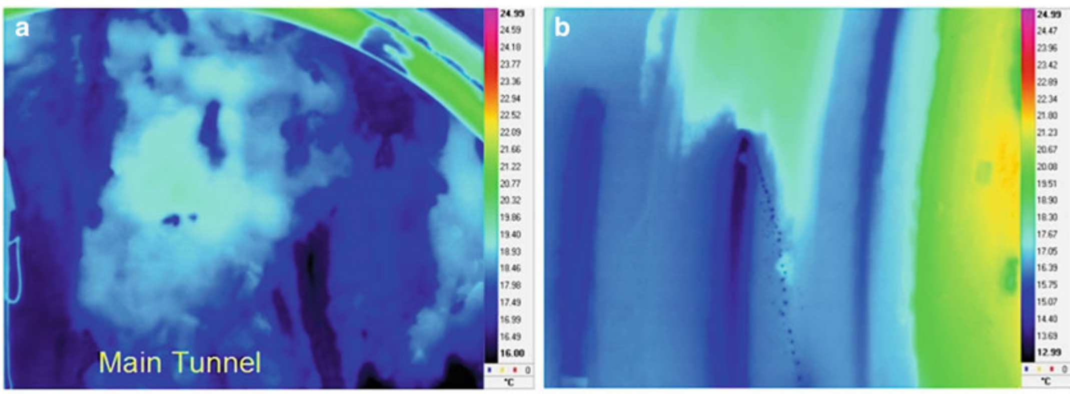


Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 32 Visible and infrared images of the evacuation tunnel after blasting. (**a**) Visible image. (**b**) Infrared thermographic image

the naked eye image is quite blurred. This is a very important feature of infrared thermographic imaging technique, which is not affected by the dust cloud soon after the excavation. The cooler areas are likely to be corresponding to de-stressed zones at the tunnel face and the groundwater seepage locations. The remains of blasted holes are clearly observed in infrared images as hot spots as seen in Fig. 33.



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 33 Infrared images of remains of blasted holes



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 34 Infrared thermographic images of tunnel face indicating seepage locations and water drops at the tunnel wall. (**a**) Seepage locations in main tunnel. (**b**) Water drop and seepage

Figure 34a shows the seepage conditions at the main tunnel on September 27, 2014. As noted from the figure, the seepage locations are well recognized from the infrared thermographic images. Additional observations showed that it was also possible to recognize even water drops seeping into the excavation space as seen in Fig. 34b.

Future Directions for Possible Applications in Geoengineering and Geosciences

The infrared thermographic imaging technique utilized to show how mechanical energy is transformed into heat during deformation and fracturing of minerals and rocks in compression and tensile experiments and compression of quartz fault gouge in addition to measurable parameters such as deformation, load, displacement, acoustic emissions, acceleration, magnetic field, electrical resistivity, and electrical potential, which are called multiparameters [12, 16, 19]. The experiments clearly show that high-temperature bands are observed along the potential facture zones. Particularly, the failure of quartz crystal was like an explosion, which was not reported in geomechanics so far. Infrared thermographic imaging technique together with monitoring of multiparameter thermodynamic responses of crystals and rocks ranging from soft to hard implied that multiparameter responses may be of great value for predicting the failure phenomenon, and this fact may be very important in the fields of the rock engineering and geoscience.

Laboratory tests on rock specimens and large rock blocks having a circular hole under compression environment showed that high-temperature zones are directly associated with plastic straining and crack formation in relation to the deformation and rupture processes. Such zones appear before the macroscopic failure, and it should be of great value to assess the real-time stability of underground excavations. The infrared thermographic imaging technique is used to observe the condition of the face of Taru-Toge Tunnel before and after the blasting operations. As the tunnel behaved in a stable manner under the present overburden conditions, the high stress condition was not observed. Nevertheless, it was possible to identify de-stressed regions at the tunnel face. Compared to naked eye observations soon after the blasting, the infrared thermographic imaging technique is not affected by the dust cloud soon after the excavation so that it is possible to observe the condition of the tunnel face clearly soon after blasting. The groundwater seepage conditions, which may lead to some local failures, can also be easily identified. Furthermore, the quality of shotcrete and its hardening process may be easily evaluated using the infrared thermographic imaging technique.

Rockbursts in deep mining and deep tunnels are very dangerous for workers at the stopes in mines and tunnel faces. The observations during experiments clearly showed that it is possible to locate high heat spots in the samples, which may directly imply that one should be possible to locate the possibility of rockburst and their locations. It is well known that the rockburst occurs soon after the blasting operations within a short distance from the excavation face. The utilization of infrared thermographic imaging cameras soon after the blasting near the excavation faces in rock with a potential rockburst should be quite useful for the safety of workers at the vicinity of the working site.

Earthquake prediction is also one of the most common topics for geoscientists. Despite many attempts, there is no successful method prediction yet. The observations during experiments clearly showed that high heat bands before rupturing in the samples may directly imply that one should be possible to locate the possibility of locations of earthquakes and probably the time of earthquake. This might be quite important particularly for very short-term earthquake prediction.

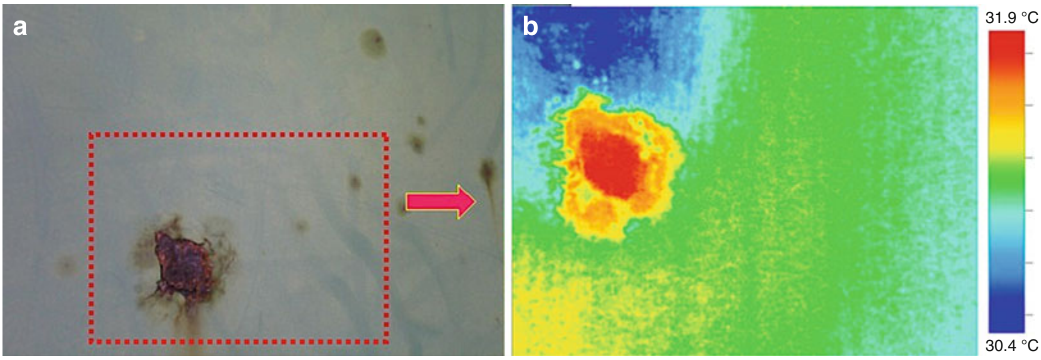
Dynamic shearing process induces temperature rises along discontinuities and adjacent rock mass. Temperature rise depends on the dynamic shearing rate, normal stress, and frictional properties of discontinuities as well as thermal properties of adjacent rocks. The increase of normal stress, dynamic shearing rate, and frictional properties proportionally increases the temperature rise. Particularly normal stress and dynamic shearing rate have a great influence on the overall rise of temperature. The responses observed throughout dynamic shearing experiments can be explained through the consideration of the energy conservation law of the continuum mechanics (i.e., Aydan [3, 11, 13, 21]). The energy conservation law for shearing experiments may be written as

$$\rho c \frac{\partial T}{\partial t} = -\nabla q + \tau \dot{\gamma} \tag{2}$$

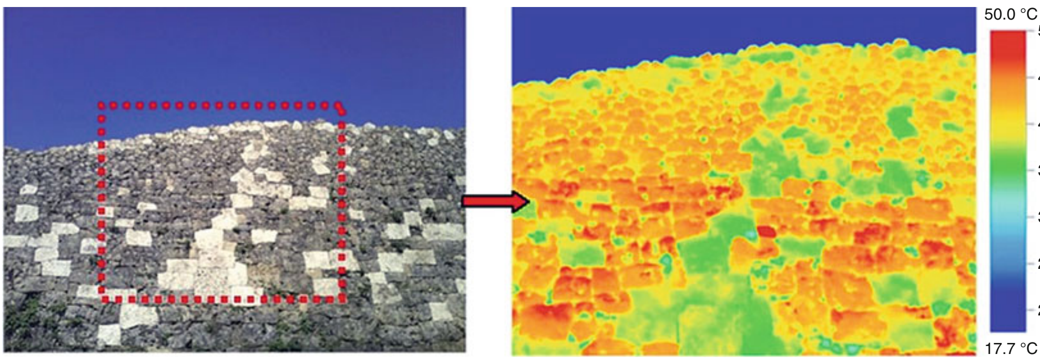
where $\rho, c, T, q, \tau, \dot{\gamma}$ are density, specific heat, temperature, heat flux, shear stress, and shear strain rate, respectively. The heat flux is related to temperature through Fourier law, and it is given for one-dimensional case as

$$q = -k \frac{\partial T}{\partial x} \tag{3}$$

where k and x are thermal conductivity and physical space, respectively. Temperature rises observed in the experiments reported in this study can be easily evaluated using the imposed cyclic shearing condition, frictional characteristics of discontinuities, and thermal properties of adjacent rock blocks. Aydan [3] solved Eq. 2 for a creeping fault and sudden energy release along a fault and computed temperature responses. In view of computational results and experiments reported in this entry, it should be possible to observe heat release locations of the active faults with a high potential to



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 35 Visible and infrared thermographic images of a corroded part in a steel bridge. (**a**) Visible image. (**b**) Infrared thermographic image

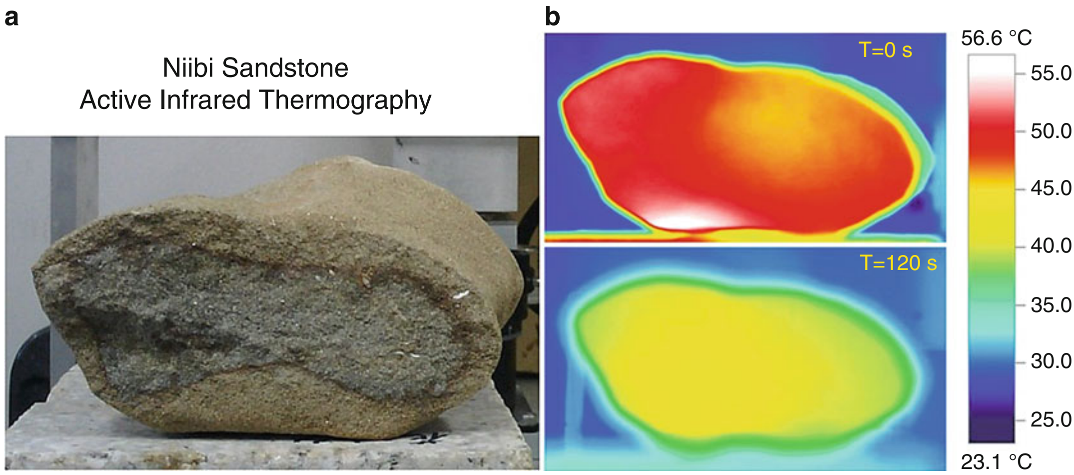


Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 36 Visible and infrared thermographic images of a reconstructed wall of the Nakagusuku Castle. (**a**) Visible image. (**b**) Infrared thermographic image

cause great earthquakes utilizing satellites equipped with infrared thermographic imaging cameras. The experimental results have also some implications in the science of earthquakes such as the illumination phenomenon. It is often reported that new hot springs appear soon after the earthquake, and some illumination of the sky, particularly at nights, sometimes occurs during and after earthquakes. The author has personally observed the same phenomena in the 1999 Kocaeli earthquake [20] occurs during and after earthquakes.

As pointed out in the previous section, active-type infrared thermographic imaging technique should be also useful to evaluate the soundness of various structures for maintenance purposes. In this respect, the utilization of artificial heat sources such as heaters, coolers, sunlight radiation, dynamic excitation, and wind may be useful. Figure 35 shows the visible and infrared thermographic images of a corroded spot in a girder of steel bridge. It is clearly observed that the heat absorption of the corroded spot is higher than surrounding noncorroded parts.

Figure 36 shows the visible and infrared thermographic images of the Nakagusuku Castle remain where castle walls are reconstructed utilizing original blocks and newly replaced blocks, which are made of Ryukyu limestone. As original blocks are partly weathered compared with the newly replaced limestone blocks, their heat absorption characteristics under the sunlight heating are different from each other. This observational fact clearly indicates that the active infrared thermographic imaging



Infrared Thermographic Imaging in Geoengineering and Geoscience, Fig. 37 Visible and infrared thermographic images of weathered Niibi sandstone. (**a**) Visible image. (**b**) Infrared thermographic images

technique could be quite useful particularly in the evaluation of the soundness of structures in geoengineering for maintenance purposes.

An active infrared thermographic imaging attempt was done on a weathered sandstone, locally known as Niibi stone. The main purpose was to see if the weathered zone could be distinguished from the non-weathered part. Figure 37 shows both visible and infrared thermographic images of the partially weathered Niibi sandstone. Infrared thermographic images correspond to the images just after and 120 s after the application of the heat shock. This experiment clearly indicates that weathered and unweathered parts can be distinguished and the rate of cooling, which could be determined from continuous monitoring, can yield the variation of their thermophysical characteristics. Therefore, the active infrared thermographic imaging technique may also be another efficient tool in characterizing the weathering of rocks and its thermophysical state.

Acknowledgments The author sincerely acknowledges the great help and support of Mr. T. Yaoita of Ken Automation (Yokohama, Japan) and Mr. A. Mochizuki of Mochizuki System (Shizuoka, Japan) on infrared thermography measurements during laboratory experiments and observations at Taru-Toge tunnel. In additon, Mr. T. Fuse of Testo Corporation (Yokohama, Japan) is also sincerely acknowledged for providing the infrared camera Testo 885 for some observations used in this chapter.

Bibliography

1. Aydan Ö, Manav H, Yaoita T, Yagi M (2014) Multiparameter thermo-dynamic response of minerals and rocks during deformation and fracturing. In: Proceedings of the 8th Asian rock mechanics symposium, Sapporo, pp 817–826
2. Aydan Ö, Fuse T, Ito T (2015) An experimental study on thermal response of rock discontinuities during cyclic shearing by Infrared (IR) thermography. In: Proceedings of the 43rd symposium on rock mechanics, JSCE, pp 123–128
3. Aydan Ö, Malistani N, Tokashiki N (2017) The possibility of infrared camera thermography for assessing the real-time stability of tunnels against rockburst. In: Proceedings of the 51st US rock mechanics/geomechanics symposium, San Francisco, ARMA 17-0479, p 6
4. Luong MP (1990) Infrared thermovision of damage processes in concrete and rock. Eng Fract Mech 35(1–3):291–301
5. Luong MP (1993) Infrared thermography of a rock salt specimen before failure. In: Proceedings of the international symposium on engineering in complex rock formations, Beijing, pp 232–238
6. Srajbr C, Dilger K (2012) Induction-excited thermography – a method to visualize defects in semi-structural adhesive bonds of car body structures. Weld World 56(3):126–132
7. Wu L, Liu S, Wu Y, Wu H (2002) Technical note, changes in infrared radiation with rock deformation. Int J Rock Mech Min Sci 39:825–830
8. Zweschper T, Dillenz A, Riegert G, Scherling D, Busse G (2003) Ultrasound excited thermography using frequency modulated elastic waves. Insight 45(3): 178–182

9. Weber W (1830) Über die spezifische Warme fester Körper insbesondere der Metalle. Ann Phys Chem 96:177–213
10. Thomson W (Lord Kelvin) (1853) On the dynamical theory of heat. Trans R Soc Edinburgh 20:261–83
11. Eringen AC (1980) Mechanics of continua, 2nd edn. Krieger Pub. Co., New York
12. Aydan Ö, Tokashiki N, Ito T, Akagi T, Ulusay R, Bilgin HA (2003) An experimental study on the electrical potential of non-piezoelectric geomaterials during fracturing and sliding, 9th ISRM congress, South Africa, pp 73–78
13. Aydan Ö (2010) Numerical modelling of heat flow in porous rock masses and its application to Armutlu geothermal area. Bull Eng Geol 30:1–16
14. Ohta Y, Aydan Ö (2010) The dynamic responses of geomaterials during fracturing and slippage. Rock Mech Rock Eng 43(6):727–740
15. Aydan Ö, Kumsar H (2005) Investigation of the Kuşini antique underground marble quarry in view of engineering geology and rock engineering. Bull Eng Geol Turkey 20:41–60
16. Aydan Ö, Daido M, Tokashiki N, Bilgin A, Kawamoto T (2007) Acceleration response of rocks during fracturing and its implications in earthquake engineering. 11th ISRM congress, Lisbon, vol 2, pp 1095–1100
17. Aydan Ö (2003) An experimental study on the dynamic responses of geomaterials during fracturing. J Sch Mar Sci Technol Tokai Univ 1(2):1–7
18. Imazu M, Ideura H, Aydan Ö (2014) A monitoring system for blasting-induced vibrations in tunneling and its possible uses for the assessment of rock mass properties and in-situ stress inferences. In: Proceedings of the 8th Asian rock mechanics symposium, Sapporo, pp 881–890
19. Aydan Ö, Ohta Y, Daido M, Kumsar H, Genis M, Tokashiki N, Ito T, Amini M (2011) Chapter 15: Earthquakes as a rock dynamic problem and their effects on rock engineering structures. In: Zhou Y, Zhao J (eds) Advances in rock dynamics and applications. CRC Press, Taylor and Francis Group, Boca Raton, pp 341–422
20. Aydan Ö, Ulusay R, Hasgür Z, Taşkın B (1999) A site investigation of Kocaeli Earthquake of August 17, 1999. Turkish Earthquake Foundation, TDV/DR 08-49, p 180
21. Aydan Ö (2001) A finite element method for fully coupled hydro-thermo-diffusion problems and its applications to geo-science and geo-engineering. 10th IACMAG conference, Austin, pp 781–786

Volcanoes of Mexico

Nick Varley
Colima Exchange and Research in Volcanology,
Faculty of Science, Universidad de Colima,
Colima, Mexico

Article Outline

Glossary

Andesite Magma of intermediate composition of silica, a common product of eruptions from stratovolcanoes.

Basalt Magma of a low silica composition has a low viscosity and usually is emplaced in lava flows or as scoria.

Dacite Magma of a composition higher in silica, between andesite and rhyolite.

Debris avalanche Large collapse event which produces an extensive deposit, often characterized by hummock-shaped hillocks. Occur from most likely all stratovolcanoes, due to their inherent instability.

Holocene The time period since the end of the last glaciation (11,700 years ago to present).

Lahar Mudflow with the remobilization of active volcanic pyroclastic deposits. Can be hot if occurring during or soon after the eruption.

Maar Explosive crater surrounded by a ring-shaped deposit or tuff ring of pyroclasts. Form due to phreatomagmatic interactions, often in basins with extensive aquifers.

Monogenetic field or Distributed volcanism Region featuring the evolution of cinder cones or maars, with each eruptive centre typically exhibiting only one eruption.

Phreatomagmatic Explosive eruption resulting from the interaction between magma and water. Results in the expulsion of juvenile material. Phreatic implies no juvenile component.

Pleistocene The epoch between 2.58 million and 11,700 years ago, which included the major period of glaciation.

Plinian Large explosive eruption producing a high column (>20 km) and resulting in multiple hazards: extensive tephra fall and the formation of pyroclastic density currents.

Pyroclastic flow or pyroclastic density current Gravity-driven mobile flow of hot gases mixed with ash and rocks, which descend volcano flanks at fast speeds posing a great threat to anything in their path.

Rhyolite Evolved magma with a high silica contents, often associated with large explosive eruptions.

Strombolian Explosive eruption of low magnitude, occur usually with low-viscosity basaltic magma. Activity that produces scoria cones: accumulation of vesiculated magma around the vent.

Surtseyan Eruption occurring through shallow water with the formation of an island from the deposition of pyroclastic material resulting from the highly explosive interaction between magma and seawater.

Tephra Pyroclastic material expelled into the air during explosive eruptions, largely comprising of ash <2 mm in diameter, but with larger

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0_477

Originally published in
R. A. Meyers (ed.), *Encyclopedia of Sustainability Science and Technology*, © Springer Science+Business Media LLC 2018
https://doi.org/10.1007/978-1-4939-2493-6_477-3

lapilli (between 2 and 64 mm) or bombs falling closer to the volcano.

Tuff ring or cone Low profile ring around maar or steep-sloped cone both formed from pyroclastic material expelled during phreatomagmatic eruptions.

Vulcanian Explosive eruption of medium magnitude: result from the failure of an impermeable layer above the magma column, which restricts the process of magma degassing. Produce a lot of ballistics.

Xenolith Material of external origin and different composition incorporated within magma. They are often pieces of upper mantle or lower crustal rocks.

Definition of the Subject

Volcanoes represent one of the nature's most formidable yet beautiful spectacles. They represent an omnipresent threat in many parts of the world, but also attract an increasing number of visitors, who have the urge to scale their flanks and peer into the depths of their craters. This entry includes a brief introduction to volcanoes: the reason they form, where they are located, and the hazard presented by the products of different types of eruption. Mexico is one of the world's most volcanic regions, and a summary of the volcanoes of this country is presented. Included is a list of the active volcanoes of Mexico, defined for the first time using systematic criteria.

Introduction

In a small village somewhere in Mexico, a farmer awoke one morning and headed off to tend his crops, not expecting that anything out of the ordinary was going to happen that day. His farm included a field that was not any ordinary cornfield, but a cornfield which all Mexican children now learn about in primary school. It was a cornfield which on that particular day was to become host to the birth a volcano. That farmer was to become one of the few people to witness the first moments of one of nature's most spectacular creations.

The eruption of Paricutín took place between 1943 and 1952. It is the youngest of at least 1040 volcanoes that are located in the Michoacán-Guanajuato volcanic field. Perhaps there is no better country than Mexico to study this type of volcanism. The country hosts at least 13 major monogenetic fields, each with countless scoria cones and other features. Many new fields are only now being defined.

The Basics of Volcanism

Volcanoes are located within certain regions on the Earth's surface, related either to plate tectonics or to so-called hot spots, which represent locations below which convection-driven mantle plumes carry hotter material toward the surface. Plate boundaries can be divergent, such as midocean ridges, which host the largest numbers of volcanoes on the planet, or rift zones, such as in East Africa. Other boundaries are convergent, which results in subduction, if one or both plates are thinner oceanic plates. In this process, one plate descends below the other carrying ample quantities of water within the rocks' structure. The water is a key ingredient to the formation of volcanoes since it lowers the melting point of the mantle, which becomes buoyant and rises toward the surface. A variable amount of the subducted plate is carried with it. Rather than erupt at the surface, the majority of magma accumulates within the crust after it loses its buoyancy, then cools with crystals slowly forming to create intrusive igneous rock. Some magma, however, will make it to the surface producing an eruption.

Volcanic eruptions can be divided into two broad categories: explosive or effusive. Various factors combine to determine how the magma emerges, the key ingredient again is water. Magma contains typically a few percent of water by weight, which is adequate to cause violent explosive eruptions if the magma rises to the surface sufficiently quickly. During ascent, decompression means that the dissolved water (and other gases) starts to form bubbles, which grow, and in the case of explosive eruptions, the volume of gas becomes much larger than the residual liquid magma. This gas can expand at an alarming rate, which can produce the largest eruptions known as Plinian. Plinian eruptions produce columns of gas and ash, which can rise to 20 km or more above the surface. The larger eruptions produce enormous clouds of ash, which enter the stratosphere, are dispersed widely by winds, and sometimes promote significant reductions in atmospheric temperatures. If the volume of magma emerging is large and its ascent rapid, the evacuation of the magma chamber below can promote collapse and the formation of a caldera. Some historical caldera-forming eruptions have had a considerable impact on the world's climate (e.g., the volcano Tambora, Indonesia, in 1815).

Less explosive types of explosive eruption, such as Strombolian, result from the ascent of less viscous magma (such as basalt). Bubbles of volcanic gas are able to flow through the magma, which prevents the build-up of extremely high pressures. This type of activity can result in the construction of a scoria cone, a common feature of the Mexican landscape. Larger stratovolcanoes are constructed over thousands of years with the erupted magma accumulating to form the volcanic edifice. The interaction of the rising magma with the ocean, a lake, or groundwater can escalate the explosivity, through the formation of rapidly expanding steam, and produce phreatic or phreatomagmatic eruptions. On land, the result can be a maar.

Effusive eruptions, on the other hand, produce lava domes and flows. In this case, the most important characteristics of the ascending magma: its temperature, viscosity, volatile contents, and amount of crystals, combine to produce this less hazardous type of eruption. Lava flows can displace human settlements and destroy the cultivated land, such as what happened at Paricutín, but they seldom endanger lives.

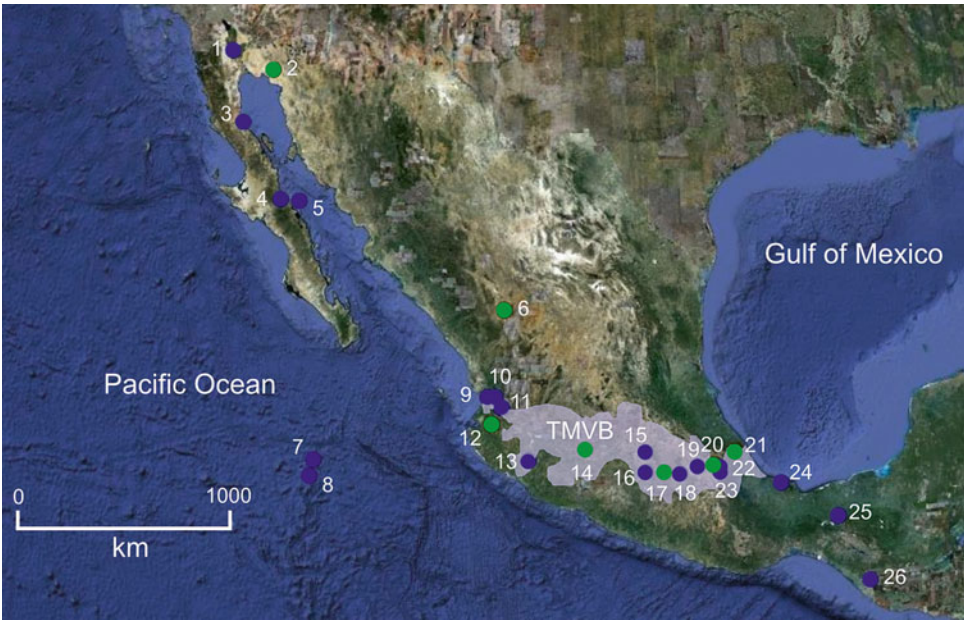
The products of explosive eruptions consist of ballistic rocks, ash, and pyroclastic density currents. The secondary remobilization of the pyroclastic deposits by water can produce immense mudflows or lahars. With the exception of ballistics, which do not reach very far from the volcano, each of these hazards can present a major risk to the population living close-by. Volcanologists need to investigate the deposits of previous eruptions to understand what might happen in the future. Hazard maps can be created to indicate where the eruptive products might accumulate. Several of Mexico's volcanoes have hazard maps, and for several others, maps are in progress.

Importance of Mexico's Volcanoes

Figure 1 shows the location of the major volcanoes of Mexico. The majority are located in the Trans-Mexican Volcanic Belt (TMVB), which extends across the country and has been described as the world's largest intracontinental volcanic arc. Over 40% of the population lives in this zone, which includes the large cities of Mexico City, Guadalajara, and Puebla. This makes volcanic risk an important issue. Enormous stratovolcanoes lie here within, like Nevado de Toluca, Popocatépetl, and the tallest volcano in North America: Citlaltépetl (or Pico de Orizaba, the name given by the Spanish colonialists). Construction of these edifices took place from the late Pleistocene onward. Figure 2 illustrates the magnitude of the most important eruptions of Mexico's active volcanoes.

Apart from the eruption of magma, stratovolcanoes can drastically influence their surroundings in another form: debris avalanches. These are huge collapses of a large proportion of the volcanic edifice and are common within the lifetime of a large volcano, given its unstable nature. Mexico has many examples with extensive deposits [1], which represent an important contribution to the shaping of the landscape.

The TMVB also includes nine identified large calderas with their associated ignimbrite deposits, such as Los Humeros or La Primavera, with 182 other circular features identified on satellite images [2]. Most likely many of these are collapse caldera structures; some play an important role in the energy supply in Mexico. Currently, some 1069 MW of geothermal power are produced, which puts the country in fourth place in the world's rankings [3]. Production is dominated by Cerro Prieto in Baja California, with Los Azufres and Los Humeros contributing in the TMVB. There is a large potential for further development within Mexico.

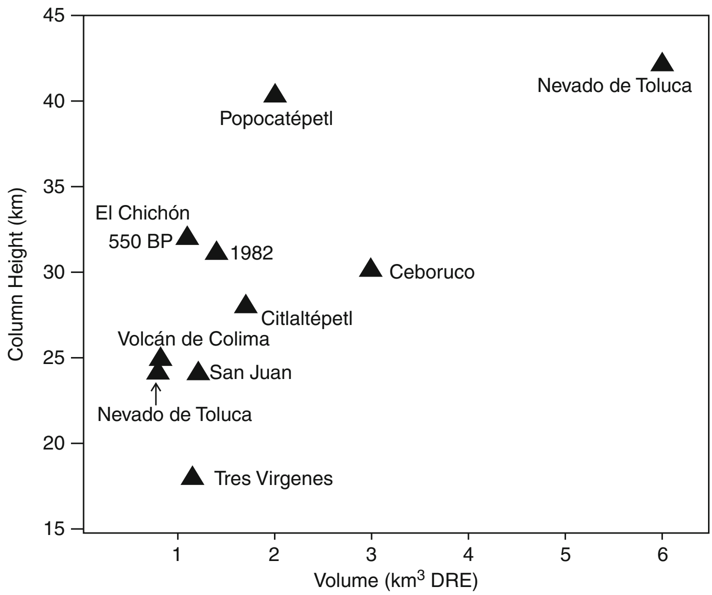


Volcanoes of Mexico, Fig. 1 Map of Mexico showing the locations of the active volcanoes or volcanic fields. *TMVB* Trans-Mexican Volcanic Belt. 1 – Cerro Prieto; 2 – Pinacate; 3 – Isla Tortuga; 4 – Tres Virgenes; 5 – Isla San Luis; 6 – Durango Volcánc Field; 7 – Bárcena; 8 – Isla Socorro; 9 – San Juan; 10 – Sangangüey; 11 – Ceboruco; 12 – Mascota Volcanic Field; 13 – Volcán de Colima; 14 – Michoacán-Guanajuato Volcanic Field; 15 – Jocotitlán; 16 – Nevado de Toluca; 17 – Chichinautzin Volcanic Field; 18 – Popocatépetl; 19 – La Malinche; 20 – Serdán-Oriental volcanic field; 21 – Naolinco volcanic field; 22 – Las Cumbres; 23 – Citlaltépetl; 24 – San Martín; 25 – El Chichón; 26 – Tacaná

Volcanoes make their appearance throughout the different chapters of Mexican history. Eruptions of Popocatépetl drove the population out of prehispanic cities such as Cuicuilco, on the outskirts of Mexico City. The Aztecs reported various eruptions of Popocatépetl, which means "the smoking mountain" in the Náhuatl language. Other sacred centres, such as La Campana in Colima, show evidence that the construction of certain pyramids tried to mimic distant volcanic peaks. Vast valleys are filled with soils that owe their fertility to the outpourings of countless eruptions, and other value has been reaped, such as the hugely important volcanic product that is obsidian. This volcanic glass, usually shiny black in colour, can be fashioned without much skill into knives and weapons. It was clearly important in the development of the early Mexican people. Cortes during his conquest of Mexico had run out of gunpowder. Unfortunately for countless indigenous people, he was able to get the sulphur to replenish his stocks by sending his men into the crater at Popocatépetl. The Paso de Cortes, situated high up between Popocatépetl and its older neighbor Iztaccíhuatl, is host to the only statue to Cortes in the whole of Mexico.

During the 19th century, various pioneers of exploration and scientific observation made important discoveries, and their documents along with the paintings of various travelers are important documentation of the early eruptive record of many volcanoes. Pioneers, such as Alexander von Humboldt, explored the vast territory and was the first to study the relative newcomer to the volcano world, Jorullo in Michoacán.

Apart from having some of the world's most dominating volcanic bodies with ten reaching over 4000 m in height, Mexico also hosts what is incorrectly declared as the "smallest volcano in the world." Cuexcomate "volcano" takes centre stage within the plaza of a suburb of the city of Puebla, and at 13 m tall and with spiral descending



Volcanoes of Mexico, Fig. 2 Eruption column height v. erupted volume of magma for Mexico's most recent large (Plinian) eruptions. *DRE* Dense rock equivalent, thereby taking into account any vesicles or pores within the deposits. Eruption ages: El Chichón 1982 and 550 BP; Volcán de Colima 1913; Ceboruco 1060 ± 55 BP; Popocatépetl 4965 ± 65 BP; Tres Vírgenes 6500 BP; Citlaltépetl 8500–9000 BP; Nevado de Toluca 10,500 BP & 21,700 BP; San Juan 14,770 ± 480 BP (figure modified from Arana-Salinas et al. [48]; BP = years before present)

into its "crater," it is quite the attraction. However, rather than being a small cinder cone, it is actually the remnants of a geyser, which was last active in 1664 [4].

Volcanology in Mexico is a young science, having only taken off following the 1982 eruption of El Chichón. An enormous amount of work remains to be done: identification of deposits; geochemical analyses to determine details of magma ascent and emplacement; geophysical work to establish the location of magma chambers, hydrothermal systems, etc.; risk assessments and the creation of hazard maps. There are countless young lava flows, some possibly matched by accounts of activity from prehispanic stories or texts from European missionaries or explorers. But many question marks remain regarding the date of the most recent activity for many centres.

Definition of Mexico's Active Volcanoes

The number of volcanoes described as being "active" varies depending upon the text being consulted. Most frequently lower total numbers are given and the difference with monogenetic fields or distributed volcanism is not correctly considered. The list in Table 1 now includes 27 volcanoes or volcanic fields ascribed as being active. The list published in the first edition of this book is being augmented, with the addition of the Los Humeros caldera and Valle de Bravo Volcanic Field. The list continues to be unique having employed a systematic approach to create a definitive list.

An active volcano is one that could erupt again, which implies the presence of a magma body at some depth, with the possibility of triggering renewed ascent, in the case of standard polygenetic volcanoes. In the case of volcanic fields or distributed volcanism, the situation is more complex, with the regional stress distribution largely controlling magma ascent [5]. The presence of magma should give certain tell-tale signs, such as seismicity, heat dissipated through the interaction with groundwater, resulting in hot springs, or direct degassing in the form of fumaroles. A seismic network may not be adequate to detect any movement, and thermal manifestations will not always be present. Medina et al. [6] in their

Volcanoes of Mexico, Table 1 Active volcanoes of Mexico, in order starting with the most recently active. 27 are included; *currently with active eruptive episodes during 2017

Volcano/field	Date of last eruption	Details	References
Volcán de Colima	Currently erupting*	Cyclic effusive and Vulcanian explosions	[41]
Popocatépetl	Currently erupting*	Cyclic effusive and Vulcanian explosions	[47]
Isla Socorro	1993	Submarine flank eruption	[22]
Tacaná	1986	Phreatic explosion	[81]
El Chichón	1982	Plinian eruption	[74]
Bárcena	1953	Explosive eruption	[23]
Michoacán-Guanajuato volcanic field	1943	Formation of cinder cone – Paricutín	[59]
Ceboruco	1875	Dacitic lava flow	[30]
Citlaltépetl	1846	Explosive	[54]
St. Martin Tuxtla	1793	Summit cinder cones	[72]
Tres Virgenes	1746	Explosive though unconfirmed	[18]
Sangangüey	1742	Unconfirmed eruption of flank cone	[31]
Jocotitlán	1270	Explosive	[46]
Chichinautzin volcanic field	340 AD	Xitle cinder cone and flow	[67]
Las Cumbres volcanic complex	1965 BP	Cinder cone	[57]
La Malinche	1170 BC	Ashfall and pyroclastic flow	[53]
Naolinco volcanic field	1200 BC	Cinder cone and lava flow	[71]
Isla San Luis	Recent but undefined age	Rhyolitic obsidian domes	[19]
Nevado de Toluca	3300 BP	Pyroclastic flow and surge	[45]
Valle de Bravo volcanic field	5000 BP	Cinder cone	[65]
Mascota volcanic field	Maybe 5600 BP	Lava flow and scoria	[64]
Los Humeros	6400 BP	Plinian eruption from caldera complex	[58]
Durango volcanic field	Few thousand years	Maar formation, scoria cones, and lava flows	[63]
Pinacate volcanic field	Holocene	Unknown	[13]
Isla Tortuga	Holocene	Unknown	[20]
Serdán-oriental volcanic field	Holocene or late Pleistocene	Las Derrumbadas – fumaroles	[70]
San Juan	Holocene or late Pleistocene	Plinian eruption 14,770 BP	[32]

early work considered the existence of fumarolic activity as a criterion for considering a volcano active. However, this easily observed characteristic cannot always be taken to indicate the possibility of a future eruption.

The other consideration is the repose interval since the last eruption. Mexico has many volcanoes with long repose intervals, often of several thousand years, hence a somewhat arbitrary limit can be taken as 10,000 years (following the definition of the Smithsonian Institute, Washington, D.C.). Since volcanic fields feature monogenetic volcanoes (those that erupt just once), these are taken as a single active entity. Several other volcanic fields, or scoria cone complexes not considered here, could also be active, in that they feature lava emissions that may have occurred within the Holocene. However, they have not been included in the list since no strong evidence exists to date. Examples are La Gloria field in Veracruz and Isla Isabel in Nayarit, where, as is the case at many locations, further studies are needed to clarify the situation. Three Pleistocene calderas in the TMVB still have fumarolic activity with extensive geothermal fields: La Primavera, Los Azufres, and Los Humeros. The first two have not been included as active, since the probability of a future eruption is regarded as negligible, given the long periods that have passed since their last eruptions. Prehistoric calderas sometimes feature persistent hydrothermal activity, which unlike the case of stratovolcanoes or domes, is not necessarily regarded as evidence for the likelihood of a future eruption. It is important to note that these are not related to a hot spot, as is the case with the Yellowstone caldera in the USA, which has huge eruptions separated by 100,000 s of years.

Monitoring has been expanded greatly in the last 20 years with sophisticated networks feeling the pulse of Popocatépetl and Volcán de Colima. Seismic and geochemical monitoring (water samples) is being carried out at El Chichón and Tacaná with some seismic monitoring of Citlaltépetl and Ceboruco. Much effort is being directed at increasing the quality and diversity of the information being generated, as well as the definition of models of the eruption mechanisms, thus advancing what can be interpreted from the data.

Tectonics

Volcanism in the north of Mexico can be clearly divided into various regions, each with its characteristic tectonic situation. Volcanism can be associated with extension; this gives rise to various volcanic centres in Baja California and Sonora. Further south the oceanic Rivera and Cocos plates collide with the continental North America plate; the resulting subduction gives rise to the majority of volcanoes which are located in the TMVB. Further to the south the Central American Volcanic Belt (CAVB) starts with the potentially dangerous volcano Tacaná, shared between Mexico and Guatemala. Whereas this belt, which extends through Central America, is parallel to the subduction trench, the TMVB clearly extends obliquely across the country, making an angle of about 15° with respect to the Middle American Trench. This intriguing characteristic has been explained using geophysical and geochemical observations that suggest a decreasing angle of subduction as one moves down the coast from the NW to SE [7, 8]. This results in an increasing distance between the trench, the point where the descent of the oceanic plate commences, and the zone of melting, which typically occurs at pressures corresponding to a 100 km depth.

One fascinating feature within the belt is the existence of several small chains of volcanoes, which demonstrate a decreasing age from north to south: Cántaro – Nevado de Colima – Volcán de Colima; Tlaloc – Itzaccihuatl – Popocatépetl; and Cofre de Perote – Las Cumbres – Citlaltépetl –Sierra Negra. Evidence of an overall southern migration of the magma chamber has been identified [9].

At each end of the TMVB, things get tectonically more complicated. Firstly to the east, between the TMVB and CAVB, the very active volcano El Chichón is found along with other extinct peaks forming the Pliocene to Recent age Chiapanecan Volcanic Arc (CVA). This rather esoteric arc stretches some 150 km in a NW to SE direction. The subducting Cocos plate changes from being a virtually flat slab in Central Mexico to a 45° dip angle beneath the CVA, with the slab being located an unusually deep 200 km below the arc [10]. At the northwestern extreme of the TMVB, evidence of rifting, or the separation of two plates, is present. Here magmas with a higher alkaline content have been erupted, which combined with geomorphological evidence, such as the alignment of scoria cones or fault scarps, suggests that the so-called Jalisco Block might one day separate from mainland Mexico [11, 12].

The following summary of Mexico's different volcanic centres starts in the NW of the country, continues with the TMVB, going from W to E, follows with monogenetic fields and finishes with the volcanoes of Chiapas. The account concentrates on active volcanic centres, though some important extinct volcanoes are included.

Baja California and Sonora

Sierra Pinacate Volcanic Field

This region includes more than 400 cinder cones and lava flows, and eight large maar craters of late-Pleistocene to Holocene age [13]. As with many cases of volcanic fields in Mexico, the cones follow alignments determined by the regional tectonics. Maars are the result of phreatomagmatic activity and this field has excellent examples. Crater Elegante is the largest with a diameter of 1.6 km and depth of 240 m. Eruptions here generally commenced with Stombolian eruptions and the construction of cinder cones. Unusually, the phreatomagmatism occurred later, probably a reflection of the aridity of the region. Despite a lack of any firm dates of Pinacate rocks placing the most recent activity from the field in the Holocene, the low level of erosion suggests activity within this period [6] and hence its inclusion in the list of active Mexican volcanic fields.

Cerro Prieto

This is the location of the most productive geothermal field in Mexico. The dacitic lava dome complex is located within an active continental rift, which marks the transition from the famous San Andreas transform fault system to the north to the spreading ridge of the East Pacific Rise in the Gulf of California to the south. Dating of the deposits has determined that activity at Cerro Prieto occupied a narrow period from 78–81,000 years ago [14]. As a result of this recent dating, Cerro Prieto has been removed from the first list of active Mexican volcanoes. Legends of the local Cucupas people describe hot rocks being thrown by a monster and fires coming from the soil [6] but this could be related to minor phreatic activity within the geothermal field.

San Quintín, Jaraguay, and San Borja Volcanic Fields

These three fields are not being included as active, although some early authors thought they are of Holocene age [6, 15], there is no clear evidence. Recent dating suggests the San Quintín field is no younger than 20,000 years [16]; however, the Jaraguay and San Borja fields have less vegetation on their flows, suggesting younger ages [17]. The three fields consist of cinder cones and lava flows, San Quintín with ten distinct complexes [15].

Tres Vírgenes

The only large stratovolcano in the Baja California region is the Tres Vírgenes complex. It consists of three cones, with a progression to younger ages to the SW, La Virgen being the youngest. There may have been an eruption in 1746, according to a record of observations, but there is no hard evidence [18]. The last major Plinian eruption occurred about 6500 years ago. Volcanism here is associated with dominantly extensional faulting.

Isla San Luis and Isla Tortuga

These two islands represent the youngest volcanism in the Sea of Cortés. The oldest deposits of Isla San Luis show that it was born like the island of Surtsey, in Iceland, with corresponding highly explosive eruptions producing pyroclastic surges [19]. This activity was followed by dacitic flows and the formation of tuff rings, then finally two rhyolitic domes were emplaced. The activity can be thus characterized by a successive eruption of progressively more differentiated lavas. Suggestions have been made that the two rhyolitic domes are less than 100 years old [6]. While this might be wishful thinking, they certainly are relatively young, warranting the inclusion of this island in the active volcanoes list.

Isla Tortuga is a shield volcano with young lava flows located further south. The latest stage of activity culminated in caldera collapse, extrusion of the surficial flows, and the formation of a lava lake. Medina et al. [6] defined it as being Holocene, though no evidence was given. The spatter cones within the caldera appear to be recent as do many of the flows that cover its flanks [20].

Isla Isabel

This small island off the state of Nayarit represents an emergent Surtseyan-type volcanic complex. Various craters on this island show evidence that there was a general migration of volcanic activity from northwest to southeast [21]. It is one locality where it is possible to find abundant mantle xenoliths. It has been suggested that the most recent eruptions were less than 10,000 years ago; though due to the lack of hard evidence, the island is not included in the list of active Mexican volcanoes.

Pacific Islands

The Revillagigedo Archipelago is volcanic with two active centres. It is located on the Mathematician Ridge, which was the location of an active spreading centre prior to its migration 3.15 Ma BP [22]. The volcano Bárcena on the small 4.5 km long unpopulated island of San Benedicto formed in an eruption in 1952–1953 [23]. It has been classified as a tuff cone, formed through the explosive interaction between the magma and seawater. The final eruptive phase was the emplacement of a short lava flow from its base. Since, it has remained quiet and unstudied.

Famous for its marine life and endemic species of fauna and flora, Isla Socorro represents the other active volcano of the archipelago. The only human occupancy is a naval base in the SE corner of the island. In Fig. 3 you can see a view from off the southern coast, showing domes and old lava flows, one with beautiful levees (ridges at the edge of the flow parallel to its flow direction). The last eruption occurred in 1993: a relatively small submarine event, which sent blocks of gas-filled pumice hurtling to the surface [24]. The summit area has not seen an eruption since at least 15,000 years ago [22], but has an extensive hydrothermal field, emitting gases which contain unusually high concentrations of methane and hydrogen [25]. The island is covered with extensive lava flows and domes, unusually of peralkaline composition, and more recent cinder cones in the SE part of the island. Here the ^{14}C method was used to date some lacustrine deposits, which are clearly older than nearby cones [26]. An age of about 5000 years was obtained. Further work is required to fully establish the eruptive history of Isla Socorro, an important requirement for the evaluation of the current risk. Although the recent activity has been of low intensity, deposits on the island, combined with the results of geophysical surveying, show the possibility of more than one larger eruption in the not too distant geological past [27].

Trans-Mexican Volcanic Belt

Tepic-Zacoalco Rift

This zone stretches from beyond Tepic southeast to Guadalajara. It includes a mesmerizing density

Volcanoes of Mexico, Fig. 3 View from the south side of Isla Socorro. Various domes and lava flows can be distinguished on the slopes of the volcano. The most recent activity from this volcano was in 1993 when a small submarine eruption occurred of volcanic features, including five andesitic stratovolcanoes, plus maars, domes, and countless scoria cones. Many groups of cinder cones are aligned in a NW-SE direction, the same as the rift. To the NE of the zone, the mountains represent the huge deposits of the Sierra Madre Occidental volcanic province, while to the SW the microplate known as the Jalisco Block is located [28]. The zone is dominated by several large stratovolcanoes, but also includes many fascinating features such as the maar Santa Maria del Oro, which contains a picturesque lake.

Ceboruco

The most recent activity in this region was a complex eruption of Ceboruco which lasted from 1870 to 1875. It resulted in a major dacite lava flow, which emanated from a vent high on the western flank. The eruption also produced explosions, forming various craters, and the emplacement of two fascinating domes. Heading around the edifice, a series of recent lava flows of either dacitic or andesitic composition can be observed, some originating from near the summit, others from lower on the flanks. Stunning views of these lava flows can be obtained by driving along the two roads that extend along the rift zone: one to the north, the other to the south of Ceboruco.

The last major explosive eruption occurred only 1000 years ago [29] and produced an extensive pyroclastic flow and pumice fall deposit. A caldera was formed with a 4 km diameter, which later hosted a complex sequence of activity with the emplacement of andesite and dacite domes and flows. Following, there was a period of about 500 years with the formation of the lava flows on the flanks [30]. Since then the volcano has been quieter, with just the one recent 7 km long lava flow.

Sangangüey and San Juan

Further to the NW, Sangangüey is an impressive mass with a spectacular lava spine at its summit and large collapse scars on its flanks. The spine can be seen in Fig. 4, which was taken from the summit of Ceboruco. Many cinder cones exist along the rift, all aligned in the NW-SE direction in five different linear groups. There are no records of historic activity from the peak, although historical accounts coupled with low levels of erosion suggest that some of the flank cinder cones have erupted less than 1000 years ago [31]. For this reason, the volcano is included in the list of active volcanoes of Mexico.

At the Tepic end of the rift, in fact adjacent to the city of Tepic, there lies San Juan, a volcano which is not termed active using the 10,000 year rule. However, it has been pointed out that its explosive past, with the last Plinian event occurring about 14,770 years ago, suggests that a reawakening should not be totally ruled out [32].

Volcán Tequila

Tequila, apart from being a notorious Mexican export, gives its name to a large stratovolcano, located further to the SE along the rift. The youngest associated volcanic feature (~60 ka) is a small andesitic vent called Cerro Tomasillo [33]. The region features various cones, domes, and flows, with a large variety of compositions from basalt to rhyolite. Volcán Tequila itself also has a summit spine; in this case, it is a 300 m high pinnacle, which dominates any view of the volcano from an east or north direction.

La Primavera

Very close to Mexico's second largest city, Guadalajara, there is a picturesque zone of obsidian and hot springs, otherwise known as Bosque La Primavera. A large 11 km diameter caldera was formed about 95,000 years ago, with the Tala Tuff being emplaced which has a volume of about 20 km^3 [34]. The caldera was then filled by a lake. Subsequent activity included the formation of many domes, which along with uplift ended the life of the lake. Its legacy can be seen in the form of a notorious deposit of the so-called giant pumices, some of which are more than 1 m across [35]. Remnant heat is evident given the extensive geothermal field with hot springs and abundant fumaroles. Its exploitation for geothermal energy was commenced but never made it to production, due to objections based on the perceived environmental impact.

Volcanoes of Mexico, Fig. 4 Sangangüey shown in the distance, looking from Ceboruco. Many scoria cones can be seen following the NW-SE trending faults. One is thought to have erupted during the 18th century

Western TMVB

Volcán de Colima

Being the most active volcano in North America, Volcán de Colima deserves special attention. It currently is undergoing its most active period since the last major eruption, which occurred in 1913. The Colima Volcanic Complex shows a southern migration of activity, the oldest edifice being Cántaro which was active between 1.6 and 1 million years before present. Volcanism then moved 15 km south to Nevado de Colima, and more recently to the current location, a further 6 km to the south [36]. The existence of a group of undated domes, called Los Hijos about 3 km further south, possibly corresponds to the first signs of further migration. Some major debris avalanches have occurred from the edifice, including a large one 18,500 years ago, which formed a dam across the Naranjo river [37]. The breaking of this dam produced an enormous debris flow which reached the Pacific Ocean, 120 km away. More recently, 3600 years ago, a smaller collapse went toward the SW, this time damming the Armería River [38]. The relatively large frequency of such events is demonstrated by evidence of a more recent event still, which has been dated at 2500 years ago. Emergency response plans cannot consider events of such low probability even though their magnitude is enormous. Sooner or later a debris avalanche will occur, if not at Colima, at one of Mexico's other sizable stratovolcanoes.

Since the most recent collapse, pyroclastic deposits and lava flows have been accumulating and building the current edifice. It has grown to a height of about 3860 m, though the current altitude is not precisely known, due to the frequent construction and destruction of summit domes. The 1913 sub-Plinian eruption produced a 23 km altitude column, which collapsed resulting in pyroclastic density currents that reached 15 km from the volcano [39]. The historical records show that Volcán de Colima has a large Plinian (or at least sub-Plinian) eruption every 100 years, more or less [40]. Previous to 1913, there were eruptions in 1818, 1690, and 1606. This is very frequent for a volcano and has recently increased levels of anxiety.

Volcán de Colima has produced a large number of andesitic domes and flows during recent years. There have been two episodes of activity, the first of which commenced in 1998 [41] and produced four different phases of dome growth (1998–1999, 2001–2003, 2004, 2007–2011). The effusive activity was interspersed with Vulcanian explosive eruptions, which peaked in 2005, when at least 30 Vulcanian eruptions occurred, each producing a pyroclastic density current resulting from column collapse [42]. The longest flow reached 5.4 km along a ravine to the SE of the volcano. An even longer flow occurred in October 2004 when a relatively large dome collapse produced a pyroclastic density current which headed down a ravine on the SW flank to a point 6.1 km from the volcano. From 2003 until the end of the episode, the volcano produced several Vulcanian explosions each day.

From June 2011 until the end of 2012, the volcano went surprisingly quiet. This was broken, however, at the start of 2013 when a new episode commenced, which continues until the present (August 2017). This period has featured complex variations with different phases of effusive activity producing domes and lava flows descending the volcano in different directions. Two periods were particularly important, firstly in July 2015 a batch of magma ascended very rapidly resulting in a multiple collapse event over 2 days. This produced the largest pyroclastic density current to have been emplaced since the 1913 sub-Plinian eruption. It moved with extraordinary mobility and reached a distance of 9.4 km. Fortunately there were no human victims, though some animals lost their lives. Strong winds carried the large ash cloud from the flow to the west and an evacuation was ordered. Monitoring equipment was destroyed, an example being shown in Fig. 5. The second critical period was late September 2016, when a further rapid ascent occurred of a magma batch. This time there was no explosion or dome collapse; however, a large mass of gas was released, which generated acid rain, leading to extensive loss of crops.

The population occupying the flanks of this volcano is much larger today than it was at the time of the last large eruption. As part of risk management, a monitoring network is maintained, which includes the application of a variety of techniques in the quest to identify precursors of any large-scale acceleration of activity. The seismic network had its first seismometers installed in the 1980s. Useful precursors have been identified like swarms of long-period events prior to Vulcanian explosions or effusion (2004 and 2005; [42]) and volcanotectonic events signaling the ascent of magma in 1997–1998. The recent introduction of monitoring methods such as the thermal monitoring of fumaroles [43], domes [44] and explosions have expanded the possibilities of generating models to explain the transition between different regimes of activity. Various cycles of activity can be observed: from the daily explosive cycle (Fig. 6), the several year cycle of effusive episodes, to the 100 year cycle of large explosions. The integration of monitoring data will be vital to try and determine whether a cataclysmic eruption is imminent.

Central Stratovolcanoes

The Eastern-central zone of the TMVB features five large stratovolcanoes, four of which are considered active, along with a number of smaller edifices like Jocotitlán, also considered active. Here the descriptions start to the west and head east.

Nevado de Toluca

Nevado de Toluca reaches 4680 m above sea level and features a pair of lakes within its caldera. Figure 7 shows the smaller Laguna de la Luna (moon) within the summit caldera. The lakes have proved to be interesting to archeologists, whose finds suggest they were used for prehispanic rituals. Its final episode of effusive activity 9100 years ago produced a small dacitic lava dome. This followed the most recent major Plinian eruption, which produced the Upper Toluca Pumice deposit 10,500 years ago [45]. This major eruption produced a column that is thought to have

Volcanoes of Mexico, Fig. 5 The remains of a monitoring station which used Doppler radar to measure the velocity of particles ascending from the crater. The station was destroyed in the July 2015 eruptions when large pyroclastic density currents descended the southern flank. It was thought that the station would survive anything less than a sub-Plinian eruption, which proved not to be the case

Volcanoes of Mexico, Fig. 6 Medium-sized Vulcanian explosion at Volcán de Colima, April 2016. Many thousands of these explosions have been witnessed during the last eruptive periods, however the majority have been smaller with less ash released. In the foreground, tree damaged by the July 2015 pyroclastic density currents can be observed. Showing remarkable resilience green leaves are again growing during the following spring

Volcanoes of Mexico, Fig. 7 Laguna de la Luna, within the crater of Nevado de Toluca at about 4200 m, making it, along with its partner the Laguna del Sol, the highest pair of lakes in Mexico. This volcano had a large eruption 10,500 years ago

reached 42 km in altitude and deposited about 14 km^3 of material. This was the largest eruption in Mexico in the last 15,000 years. An extensive record of large explosive activity and major collapses of the edifice can be observed in the surrounding region. The most recent deposits date about 3300 years before present and consist of a pyroclastic flow and surge emplaced on the NE flank.

Jocotitlán

Nearby the smaller Jocotitlán is a fine example of an edifice that has collapsed to produce a debris avalanche deposit. In contrast to the larger stratovolcanoes, it rises only 1300 m above the surrounding plains. Large conical hummocks have been identified with the collapse event that occurred 9690 years ago [46]; the 2.8 km^3 deposit reached a distance of 12 km. The most recent eruption of Jocotitlán was only 680 years ago; the material deposited from pyroclastic density currents on the upper flanks of the edifice makes it obligatory to include this peak in the list of active volcanoes of Mexico.

Popocatépetl

The potentially biggest volcanic threat in Mexico is Popocatépetl, which at 5472 m towers above the surroundings, and is only 65 km from Mexico City and 45 km from Puebla (Fig. 8). The population within a radius of 40 km is in excess of 1 million. Popocatépetl reawoke in December of 1994. Since then activity has been characterized by the growth of lava domes within the crater, periodic Vulcanian explosions and the release of large volumes of gas. This eruption has been the first since El Chichón in 1982 to produce fatalities. Unfortunately, five intrepid individuals decided to ignore official warnings and they climbed to the crater rim in April 1996. As well as ending their lives, an explosion sent cm-sized fragments of the destroyed lava dome to the nearest villages.

Volcanoes of Mexico, Fig. 8 Popocatépetl, which gets a frequent covering of snow during the rainy season. Its permanent glacier has greatly diminished in size over recent years. This is Mexico's most dangerous volcano since it has more than 1 million people living on its flanks

During the last 10,000 years, there have been at least three large Plinian eruptions [47]. 4965 years ago an eruption sent clasts with a diameter of up to 2.5 cm to a distance of 19 km from the volcano [48]. This implies that the height of the eruption column was 37–41 km. The eruption deposited 4.9 km^3 of material and was the second largest in Mexico within the Late Pleistocene–Holocene period. A similar eruption today would have an enormous impact on some 15 million people living in the surrounding region. More recently, two further large eruptions occurred 2150 and 1100 years ago. These two would have affected the local population centres at this time.

Various historical documents were used to recreate the activity record during the last few hundred years [49]. Several eruptions were identified in the 16th century, with another in 1664. None of these eruptions were Plinian, though some produced extensive ashfall deposits and some pyroclastic density currents. The current activity commenced in 1994, and although it has been relatively mild, it has created a variety of problems for the authorities given the large vulnerable population [50]. Cycles of dome growth and destruction from Vulcanian eruptions continue to date (August 2017) with 38 episodes occurring up to early 2016 [51].

For a large stratovolcano periodically depositing fresh material on its flanks, like the current activity of both Volcán de Colima and Popocatépetl, particularly when located in a climatic region characterized by intense seasonal rainfall, lahars represent an ongoing hazard. During the current activity of Popocatépetl, there have been two reasonably sized lahar events in 1997 and 2001 [52]. The first had more water with a partial melting of the glacier making a contribution. Both events reached the most vulnerable village on the flanks: Santiago Xalitzintla, some 15 km from the crater, though nobody was hurt.

La Malinche

The indigenous wife of Cortés certainly would not have guessed that her name, or the Mexican nickname for her, would one day be given to a large stratovolcano. Like many volcanoes in this country, activity is punctuated by large periods of rest [53]. At 4461 m, La Malinche is one of the large Mexican stratovolcanoes that has not appeared on lists of the country's active volcanoes, explained by its lack of a clear crater or fumaroles. However, the local land is covered by pyroclastic deposits with poor soil development on top, and one layer has been dated at 3100 years old. Given that over 2 million people live on its lower slopes, it needs to be considered carefully, with eruption scenarios determined [53].

Citlaltépetl

Citlaltépetl has the honor of being the highest active volcano in North America. Its prehispanic name is perhaps less used than the name which resulted from the Spanish invasion: Pico de Orizaba. At 5675 m, it is no mean feat to cross the glaciers and reach the rim of the crater. In recent years, it has not shown much activity; however, over 750,000 people live within a radius of 40 km. The last major eruption, which occurred 4100 years ago, produced a series of block and ash flows and lahars whose deposits have been found up to 28 km from the crater [54]. The most recent event was a smaller eruption that left deposits of tephra, having been dated at 690 years ago [55].

Like many of Mexico's volcanoes, widespread devastation is one possible future scenario in the event of a debris avalanche. There is evidence that Citlaltépetl has suffered many such events possibly without any eruptive activity [56], with some reaching the Gulf of Mexico, 120 km away. The edifice of this volcano has often been weakened by intense hydrothermal alteration, which combined with a large elevation difference of 4400 m dropping down to the Gulf Coastal Plain, have produced favorable conditions for these large-scale edifice collapses. This lack of precursory eruptive activity means that an event could occur one day without any warning.

Las Cumbres Volcanic Complex

Las Cumbres, located only about 10 km north of Citlaltépetl volcano, is an eroded stratovolcano which was once possibly as large as its neighbour [57]. It is the middle member of the enormous N-S chain of volcanoes, which starts with Cofre de Perote in the north and ends with Citlaltépetl to the south. Its last big eruption was 20,000 years ago and produced the widespread Quetzalapa pumice deposit. Completing the Las Cumbres Volcanic Complex are more recent dacitic domes and scoria cones. The youngest extrusion is the Yolotepec dome, which has been dated at less than 6000 years [57]. Four of the cones were dated within the Holocene, the youngest erupting only 1965 years ago. Large debris avalanches have also resulted from major collapses within this complex [56].

East TMVB

Los Humeros

One of the major calderas within the TMVB and one that currently represents a major source of geothermal energy. Although the caldera-forming event took place around 460,000 years ago, with a second smaller caldera forming 60–140,000 years ago, recent work has revealed that the youngest deposits associated with the caldera have an age of between 6000 and 7000 years [58]. They result from Plinian and Strombolian activity, which interestingly originated from multiple events with a large spatial separation. This relatively recent age means that Los Humeros should be considered as active.

Volcanic Fields

Michoacán-Guanajuato Volcanic Field

Driving through this large expanse of territory, one cannot escape the awe-inspiring impact of the extent of this volcanic field. In places, the cinder cones are so close they are touching one another. The recent addition of Paricutín is the most famous, but previously the birth of Jorullo was also witnessed in 1759. The field covers a vast area measuring 250 $\times$ 200 km in the two states which give the field its name, and it contains at least 1040 volcanic vents. The majority are cinder cones, but there are also small shield volcanoes, lava domes, and maars. In general, the cones are randomly distributed, although there are local areas where alignments can be identified [59]. The region with the largest density of cones is that of Paricutín, where the median spacing between each one is 1.15 km. The cones are not so large, the median height being 90 m with a basal diameter of 800 m. A visit to Paricutín today awards the visitor with views of the famous church, which was partly spared by the advancing lava flows. Figure 9 shows the church partially submerged.

The Tacámbaro-Puruarán area, which is located just to the NE of Jorullo, toward the southern extreme of the Michoacán-Guanajuato Volcanic Field, exhibits on the areas with the largest density of young scoria cones. Dating using radiocarbon and stratigraphy suggest that there are at least 13 volcanoes in this area that erupted in the Holocene [60].

Recent fieldwork at the shield volcano El Metate has produced some unexpected results [61]. This volcano, to the north of Uruápan, is the youngest of the shield volcanoes in the Michoacán-Guanajuato Volcanic Field with an age of only about 770 years. Estimations of the volume of lava flows emitted during its formation give a total of about 9.2 km^3 of magma, making it the most voluminous Holocene eruption in the whole of Mexico. Many other areas of the field remain little studied.

One northern region of the field is worthy of a mention: Valle de Santiago, which features the stunning beauty of its seven major maars. There are a total of 20 within a zone with dimensions of 7 × 50 km [62]. The youngest, La Alberca, has been dated at 73,000 years old. Unfortunately, overexploitation of the groundwater has meant that the maars have almost all lost their lakes over recent years.

Volcanoes of Mexico, Fig. 9 The church of San Juan Parangaricutiro, which was buried by the lava flows from Paricutín during the eruption that commenced in 1943. The cone can be seen in the background

Durango Volcanic Field

This large field covers some 2100 km^2 and contains about 100 cinder cones [63]. The La Berña-El Jagüey Maar Complex is one of the youngest centres in the field and the only section studied in detail. Being maars, they were formed by the interaction between magma and groundwater (phreatomagmatic eruptions), which produced large explosions. The final phase of the activity was the formation of several scoria cones within the crater and related lava flows. The age of this complex has been estimated at a few thousand years [63].

Mascota Volcanic Field

Several volcanic fields are located within the Jalisco Block, the youngest being close to the city of Mascota. This field is notorious for the geochemistry of its lavas, which are dominated by minettes, an unusual type of lava which, instead of the more common feldspar crystals dominating, contains large mica crystals [64]. The youngest flow of this field stands out through the lack of vegetation or soil on its surface. This led the authors to believe its age to be a few thousand years old. Further evidence of its age could be a correlation with scoria found in nearby lake deposits that were dated at less than 5600 years old.

Valle de Bravo Volcanic Field

The Valle de Bravo Volcanic Field contains some 120 cinder cones, 21 mainly dacitic lava domes and a shield volcano. The youngest date to be obtained is 5000 years before present [65] and it has been estimated that 29% of the cones are Holocene, based on their morphology [66]. The cone density in the Zitácuaro–Valle de Bravo region is 2.1/100 km^2, which is only slightly lower than the 2.6 cones/100 km^2 found in the Michoacán-Guanajuato Volcanic Field [65]. Interestingly, cinder cones in the field follow mainly a NE trend, whereas domes are aligned towards the NW [66].

Chichinauzin Volcanic Field

The extensive Chichinauzin field lies to the south of Mexico City and covers some 2500 km^2. It contains more than 200 monogenetic scoria cones and associated lava flows [67]. The last eruption was 1670 years ago [68] and given that its historic $<$1250 years reoccurrence time has been greatly exceeded, it represents one of the more likely regions for the birth of the next Mexican volcano. During the past 10,000 years, at least seven monogenetic eruptions have occurred [68, 69] (Jumento, Pelado, Cuauhtzin, Tláloc, Guespalapa, Chichinautzin, and Xitle). The cone of Xitle is a clear landmark within the southern limits of the city boundary, reminding us of its eruption 1670 years ago. Obviously, when the day arrives for the next eruption, the effect will be catastrophic and resulting lava flows and ashfall will paralyze the capital city.

Serdán-Oriental Volcanic Field

No evidence has been found for activity within the last 10,000 years; however, it has been included in Table 1 for two main reasons. Firstly, Las Derrumbadas are a pair of volcanic domes within this field. They still have active fumaroles outputting measurable quantities of sulphur dioxide. This volcanic field also hosts several maars, some with lakes, some without, other rhyolitic domes and scoria cones. An interesting dome named Cerro Pizarro shows evidence that it has produced multiple eruptions, unusual for this type of volcano, which normally is monogenetic [70]. Furthermore, the repose periods could be greater than 65,000 years, which is another reason for not declaring this field as no longer capable of producing an eruption.

Cofre de Perote Vent Cluster and Naolinco Volcanic Field

These recently identified mafic fields are located on the flank of the largely Pleistocene Cofre de Perote shield volcano, in the case of the Cofre de Perote Vent Cluster (CPVC) and to the north of Jalapa, Veracruz, the Naolinco Volcanic Field (NVF) [71]. The CPVC consists of an extensive lava field that covers $>$100 km^2. The most recent eruption was from the El Volcancillo scoria cone, and it produced an impressively large flow which travelled 50 km only about 870 years ago. The Rincón de Chapultepec scoria cone in the NVF produced a lava flow 2980 BP. Interestingly these flows are some of the largest in the whole TMVB, but remained almost totally ignored until within the last 10 years.

San Martín Tuxtla

The basaltic volcano San Martin Tuxtla is located near the coast of the Gulf of Mexico, in southern Veracruz. It actually represents the active centre within another volcanic field containing monogenetic volcanic cones, maars, and three other large volcanoes which have not shown evidence of activity in the Holocene [72]. It is so far unclear whether this volcano is the eastward end-member of the TMVB or related to extensional tectonics. Its most recent eruption in 1793 was a succession from phreatomagmatic explosions, to Strombolian explosions and then an effusive episode, which produced a 3 km lava flow. At least nine other eruptions took place within the last 6000 years of scoria cones and maars in this field.

Volcanoes of Chiapas

El Chichón

The largest historic eruption to occur in Mexico was the 1982 eruption of El Chichón, which resulted in the death of around 2000 people and impacted the world's climate through a measurable reduction in temperature. Prior to the eruption, it was almost totally unknown, with nobody aware of its potential for devastation. El Chichón is not a towering stratovolcano like those of the TMVB, but an unassuming hill, which is difficult to see until one has almost arrived at its lower flanks. Ironically, several months before the eruption, El Chichón was the subject of a geothermal prospecting trip. In a report, it was stated that the volcano needed to be studied and might reactivate in the future.

Three large explosions occurred within a week from 28 March to 4 April 1982. The authorities were taken by surprise [73]. The volcano had been asleep for 550 years previous to this Plinian eruption. Lack of experience largely affected the handling of the emergency, with a proportion of the deaths occurring due to the early return of evacuees. The eruption attracted a lot of international interest, one reason being that the magma had an unusually high sulphur content. This increased the impact on the world's climate from the eruption cloud, with aerosols in the stratosphere producing a decrease of 0.2–0.5 °C in temperature in the Northern Hemisphere [74].

The eruption at El Chichón formed a 1 km diameter and 180 m deep crater (Fig. 10). Geochemical and geophysical studies have attempted to define the hydrothermal system which includes the shallow crater lake, bubbling springs, and geysers [75, 76]. The lake can be observed to vary considerably in size, seemingly following a cycle that is independent of the local rainfall. Another interesting observation is that some hot springs found outside of the crater appear to be part of a system that was undisturbed by the Plinian eruption.

Research has now shown that the volcano has had a large eruption approximately every 800 years [77]. Studies looking at the stratigraphy of local deposits have deduced that during the past 8000 years it has produced at least 11 eruptions [78], with many being similar to that of 1982. Larger magnitude events occurred in the year 750 and 1450 AD. The former of the two coincided with the collapse of the Mayan civilization in that region; perhaps it played a role. A future large eruption would impact more than 70,000 people living in a radius of 35 km from the volcano.

Tacaná

This volcano sits right on the Mexico–Guatemala border and presents a largely unacknowledged hazard. The most recent eruption was in 1986; the phreatic explosion resulted in the formation of a new fumarole field below the summit. Although it was a small event, it indicates the proximity of a magma body to the surface. The study of the geochemical characteristics of a number of springs at different elevations has provided an interesting insight into the often little-understood interaction between magmatic gases and the aquifers within the volcanic edifice [79].

The new fumaroles are located in a scar that was left by a collapse that occurred about 10,610 years ago after the growth of a summit

Volcanoes of Mexico, Fig. 10 The crater of El Chichón taken on 5 May 2004. This was the site of Mexico's most recent large eruption, when some 2000 people lost their lives

dome [80]. More recently during the Holocene, Tacaná had a series of eruptions, both explosive and effusive, the most recent being about 1950 years ago [81]. Following the eruption, a series of lahars devastated the surrounding countryside. There is evidence that the construction of a nearby prehispanic settlement called Izapa was interrupted by these events. New activity of a similar magnitude could impact the approximately 300,000 people who live within 35 km of the volcano, an order of magnitude higher than its more famous neighbour in Chiapas.

Future Directions

Hopefully, this entry has provided a useful introduction to the many volcanoes of Mexico. The cultural and geographical richness of the country is exemplified in its volcanoes. For their study, Mexico offers a huge variety of landforms and endless geophysical and geochemical case studies. There is a considerable potential for future eruptions with a significant impact. The most dangerous volcanoes in Mexico for their potential for large eruptions are probably Popocatépetl, Volcán de Colima, Tacaná, Citlaltépetl, and Ceboruco. But monogenetic fields such as Chichinautzin could also wreak havoc with the birth of a new cinder cone. The reactivation of one of its other sleeping giants always remains a distinct possibility.

Studies are only just being initiated of many of the active or potentially active volcanic centres of Mexico. There are many areas of research to be explored, with many more waiting to be defined. In certain cases the origin of extensive deposits emplaced during historic or prehistoric eruptions is unknown. With others the eruption mechanism has yet to be fully understood, which creates a dilemma when attempting to define a monitoring strategy, while faced with the possibility of an impending cataclysmic event. This lack of understanding makes it difficult to set threshold levels for monitored parameters, such as gas flux or seismicity, which should trigger a change in the alert system or an action within the emergency risk mitigation plan.

Hazard maps have been published of Popocatépetl, Volcán de Colima, Citlaltépetl, El Chichón, and Nevado de Toluca. These represent vital tools for the mitigation of volcanic risk by improving land-use planning and for defining procedures during emergencies. These need to be regarded as dynamic, with constant updating as new information becomes available. Work is underway to construct risk maps for certain hazards, which not only consider the geographical extent, but also the probability of occurrence and the vulnerability of the local population. Similar maps are required for some of the other active volcanoes of Mexico.

After the large loss of life at the hands of El Chichón in 1982, much progress has been made to reduce the risk at many of Mexico's volcanoes. More needs to be done in areas such as education to prepare for future potentially devastating eruptions.

Bibliography

Primary Literature

1. Capra L, Macías JL, Scott KM, Abrams M, Garduño-Monroy VH (2002) Debris avalanches and debris flows transformed from collapses in the Trans-Mexican Volcanic Belt, Mexico – behavior, and implications for hazard assessment. J Volcanol Geotherm Res 113:81–110
2. Anguita F et al (2001) Circular features in the Trans-Mexican Volcanic Belt. J Volcanol Geotherm Res 107(4):265–274
3. Matek B (2016) 2016 Annual U.S. & global geothermal power production report. Geothermal Energy Association
4. Beraldi-Campesi H (2012) Cuexcomate: from the smallest volcano to the biggest Geyser on Earth. GSA section meeting – Cordilleran section – 108th annual meeting – session no. 22 – Geochemistry and petrology of igneous rocks (posters). https://gsa.confex.com/gsa/2012CD/finalprogram/abstract_201524.htm
5. Martí J, López C, Bartolini S, Becerril L, Geyer A (2016) Stress controls of monogenetic volcanism: a review. Fronti Earth Sci 4(106), https://doi.org/10.3389/feart.2016.00106. Retrieved 3 Sept 2017
6. Medina F, Suarez F, Espindola JM (1989) Historic and Holocene volcanic centres in NW Mexico. Bull Volcanol 51:91–93
7. Pardo M, Suarez G (1995) Shape of the subducted Rivera and Cocos plates in southern Mexico: seismic and tectonic implications. J Geophys Res 100(B7): 12357–12373
8. Ferrari L (2004) Slab detachment control on mafic volcanic pulse and mantle heterogeneity in central Mexico. Geology 32(1):77–80
9. Marquez A, Oyarzun R, de Ignacio C, Doblas M (2001) Southward migration of volcanic activity in the central Mexican Volcanic Belt: asymmetric extension within a two-layer crustal stretching model. J Volcanol Geotherm Res 112(1–4):175–187
10. Manea M, Manea VC (2008) On the origin of El Chichón volcano and subduction of Tehuantepec Ridge: a geodynamical perspective. J Volcanol Geotherm Res 175(4):459–471
11. Luhr JF, Nelson SA, Allan JF, Carmichael ISE (1985) Active rifting in southwestern Mexico: manifestations of an incipient eastward spreading-ridge jump. Geology 13:54–57
12. Frey HM, Lange RA, Hall CM, Delgado-Granados H, Carmichael ISE (2007) A Pliocene ignimbrite flare-up along the Tepic-Zacoalco rift: evidence for the initial stages of rifting between the Jalisco Block (Mexico) and North America. GSA Bull 119(1/2):49–64
13. Gutmann JT (2002) Strombolian and effusive activity as precursors to phreatomagmatism: eruptive sequence at maars of the Pinacate volcanic field, Sonora, Mexico. J Volcanol Geotherm Res 113(1–2):345–356
14. García-Sánchez L, Macías JL, Sosa-Ceballos G, Arce JL, Garduño-Monroy VH, Saucedo R, Avellán DR, Rangel E, Layer PW, López-Loera H, Rocha VS, Cisneros G, Reyes-Agustín G, Jiménez A, Benowitz JA (2017) Genesis and evolution of the Cerro Prieto Volcanic Complex, Baja California, Mexico. Bull Volcanol 79(6):44
15. Luhr JF (1995) San Quintin volcanic field, Baja California Norte, Mexico: geology, petrology, and geochemistry. J Geophys Res 100(B7):10353–10380
16. Ortega-Rivera A, Bohnel H, Lee J (2004) The San Quintin volcanic field – 40Ar/39Ar geochronology and paleomagnetism. Geological Society of America Penrose Conference. Metepec, Puebla, p 59
17. Rogers G, Saunders AD, Terrell DJ, Verma SP, Marriner GF (1985) Geochemistry of Holocene volcanic rocks associated with ridge subduction in Baja California, Mexico. Nature 315:389–392
18. Capra L, Macías JL, Espíndola JM, Siebe C (1998) Holocene plinian eruption of La Virgen volcano, Baja California, Mexico. J Volcanol Geotherm Res 80:239–266
19. Paz Moreno FA, Demant A (1999) The recent Isla San Luis volcanic centre: petrology of a rift-related volcanic suite in the northern Gulf of California, Mexico. J Volcanol Geotherm Res 93(1–2):31–52
20. Batiza R (1978) Geology, petrology, and geochemistry of Isla Tortuga, a recently formed tholeiitic island in the Gulf of California. Geol Soc Am Bull 89: 1309–1324

21. Housh TB, Aranda-Gómez JJ, Luhr JF (2010) Isla Isabel (Nayarit, Mexico): quaternary alkalic basalts with mantle xenoliths erupted in the mouth of the Gulf of California. J Volcanol Geotherm Res 197(1–4):85–107
22. Bohrson WA et al (1996) Prolonged history of silicic peralkaline volcanism in the eastern Pacific Ocean. J Geophys Res 101(B5):11457–11474
23. Richards AF (1959) Geology of the Islas Revillagigedo, Mexico 1. Birth and development of Volcan Barcena, Isla San Benedicto. Bull Volcanol 22:73–124
24. Siebe C et al (1995) Submarine eruption near Socorro Island, Mexico: geochemistry and scanning electron microscopy studies of floating scoria and reticulite. J Volcanol Geotherm Res 68:239–271
25. Taran YA, Varley NR, Inguaggiato S, Cienfuegos E (2010) Geochemistry of H2- and CH4-enriched hydrothermal fluids of Socorro Island, Revillagigedo Archipelago, Mexico. Evidence for serpentinization and abiogenic methane. Geofluids 10:542–555
26. Farmer JD, Farmer MC, Berger R (1993) Radiocarbon ages of lacustrine deposits in volcanic sequences of the Lomas Coloradas area, Socorro Island, Mexico. Radiocarbon 35(2):253–262
27. Paoletti V, Gruber S, Varley N, D'Antonio M, Supper R, Motschka K (2016) Insights into the structure and surface geology of Isla Socorro, Mexico, from Airborne Magnetic and Gamma-Ray Surveys. Surv Geophys 37(3):601–623
28. Ferrari L, Pasquarè G, Venegas-Salgado S, Romero-Rios F (1999) Geology of the western Mexican Volcanic Belt and adjacent Sierra Madre Occidental and Jalisco block. Cenozoic tectonics and volcanism of Mexico. Geol Soc Am Spec Pap 334:65–83
29. Gardner JE, Tait S (2000) The caldera-forming eruption of Volcán Ceboruco, Mexico. Bull Volcanol 62:20–33
30. Sieron K, Siebe C (2008) Revised stratigraphy and eruption rates of Ceboruco stratovolcano and surrounding monogenetic vents (Nayarit, Mexico) from historical documents and new radiocarbon dates. J Volcanol Geotherm Res 176(2):241–264
31. Nelson SA, Carmichael ISE (1984) Pleistocene to recent alkalic volcanism in the region of Sangangüey volcano, Nayarit, Mexico. Contrib Mineral Petrol 85(4):321–335
32. Luhr JF (2000) The geology and petrology of Volcán San Juan (Nayarit, Mexico) and the compositionally zoned Tepic Pumice. J Volcanol Geoth Res 95(1–4):109–156
33. Lewis-Kenedi CB, Lange RA, Hall CM, Delgado-Granados H (2005) The eruptive history of the Tequila volcanic field, western Mexico: ages, volumes, and relative proportions of lava types. Bull Volcanol 67(5):391–414
34. Mahood GA (1981) A summary of the geology and petrology of the Sierra La Primavera, Jalisco, Mexico. J Geophys Res 86(B11):10137–10152
35. Walker GPL, Wright JV, Clough BJ, Booth B (1981) Pyroclastic geology of the rhyolitic volcano of La Primavera, Mexico. Geol Rundsch 70:1100–1118
36. Luhr JF, Carmichael ISE (1980) The Colima volcanic complex, Mexico. 1. Post-caldera andesites from Volcán Colima. Contrib Minerol Petrol 71:343–372
37. Capra L, Maclas JL (2002) The cohesive Naranjo debris-flow deposit (10 km3): a dam breakout flow derived from the Pleistocene debris-avalanche deposit of Nevado de Colima Volcano (Mexico). J Volcanol Geotherm Res 117(1–2):213–235
38. Cortes A, Macías JL, Capra L, Garduño-Monroy VH (2010) Sector collapse of the SW flank of Volcán de Colima, Mexico: the 3600 yr BP La Lumbre-Los Ganchos debris avalanche and associated debris flows. J Volcanol Geotherm Res 197(1–4):52–66
39. Saucedo R et al (2010) Eyewitness, stratigraphy, chemistry, and eruptive dynamics of the 1913 Plinian eruption of Volcán de Colima, Mexico. J Volcanol Geotherm Res 191(3–4):149–166
40. Luhr JF (2002) Petrology and geochemistry of the 1991 and 1998–1999 lava flows from Volcán de Colima, Mexico: implications for the end of the current eruptive cycle. J Volcanol Geotherm Res 117(1–2):169–194
41. Zobin VM et al (2002) Overview of the 1997–2000 activity of Volcán de Colima, Mexico. J Volcanol Geotherm Res 117(1–2):1–19
42. Varley N, Arámbula-Mendoza R, Reyes-Dávila G, Stevenson J, Harwood R (2010) Long-period seismicity during magma movement at Volcán de Colima. Bull Volcanol 72(9):1093–1107
43. Stevenson JA, Varley N (2008) Fumarole monitoring with a handheld infrared camera: Volcán de Colima, Mexico, 2006–2007. J Volcanol Geotherm Res 177(4):911–924
44. Thiele ST, Varley N, James MR (2017) Thermal photogrammetric imaging: a new technique for monitoring dome eruptions. J Volcanol Geotherm Res 337:140–145
45. Arce JL, Macías JL, Vázquez-Selem L (2003) The 10.5 ka Plinian eruption of Nevado de Toluca volcano, Mexico: stratigraphy and hazard implications. GSA Bull 115(2):230–248
46. Siebe C, Komorowski J-C, Sheridan MF (1992) Morphology and emplacement of an unusual debris avalanche deposit at Jocotitlán volcano, central Mexico. Bull Volcanol 54:573–589
47. Siebe C, Abrams M, Macías JL, Obenholzner J (1996) Repeated volcanic disasters in Prehispanic time at Popocatépetl, central Mexico: past key to the future? Geology 24:399–402
48. Arana-Salinas L, Siebe C, Maclas JL (2010) Dynamics of the ca. 4965 yr 14C BP "Ochre Pumice" Plinian eruption of Popocatépetl volcano, Mexico. J Volcanol Geotherm Res 192(3–4):212–231
49. Martin-Del Pozzo AL, Rodríguez A, Portocarrero J (2016) Reconstructing 800 years of historical eruptive activity at Popocatépetl Volcano, Mexico. Bull Volcanol 78(3):1–13
50. De la Cruz-Reyna S, Tilling RI (2008) Scientific and public responses to the ongoing volcanic crisis at Popocatepetl volcano, Mexico: importance of an

effective hazards-warning system. J Volcanol Geotherm Res 170(1–2):121–134
51. Gómez-Vazquez A, De la Cruz-Reyna S, Mendoza-Rosas AT (2016) The ongoing dome emplacement and destruction cyclic process at Popocatépetl volcano, Central Mexico. Bull Volcanol 78(9):1–15
52. Capra L, Poblete MA, Alvarado R (2004) The 1997 and 2001 lahars of Popocatépetl volcano (Central Mexico): textural and sedimentological constraints on their origin and hazards. J Volcanol Geotherm Res 131(3–4):351–369
53. Castro-Govea R, Siebe C (2007) Late Pleistocene-Holocene stratigraphy and radiocarbon dating of La Malinche volcano, Central Mexico. J Volcanol Geotherm Res 162(1–2):20–42
54. Carrasco-Nüñez G (1999) Holocene block-and-ash flows from summit dome activity of Citlaltepetl volcano, Eastern Mexico. J Volcanol Geotherm Res 88(1–2):47–66
55. De la Cruz-Reyna S, Carrasco-Nüñez G (2002) Probabilistic hazard analysis of Citlaltépetl (Pico de Orizaba) Volcano, eastern Mexican Volcanic Belt. J Volcanol Geotherm Res 113(1–2):307–318
56. Carrasco-Nüñez G et al (2006) Multiple edifice-collapse events in the Eastern Mexican Volcanic Belt: the role of sloping substrate and implications for hazard assessment. J Volcanol Geotherm Res 158:151–176
57. Rodríguez SR (2005) Geology of Las Cumbres volcanic complex, Puebla and Veracruz states, Mexico. Rev Mex Cienc Geol 22(2):181–199
58. Dávila-Harris P, Carrasco-Núñez G (2014) An unusual syn-eruptive bimodal eruption: the Holocene Cuicuiltic member at Los Humeros caldera, Mexico. J Volcanol Geotherm Res 271(0):24–42
59. Hasenaka T, Carmichael ISE (1985) The cinder cones of Michoacán-Guanajuato, Central Mexico: their age, volume and distribution, and magma discharge rate. J Volcanol Geotherm Res 25:105–124
60. Guilbaud M-N, Siebe C, Layer P, Salinas S (2012) Reconstruction of the volcanic history of the Tacámbaro-Puruarán area (Michoacán, México) reveals high frequency of Holocene monogenetic eruptions. Bull Volcanol 74(5):1187–1211
61. Chevrel MO, Guilbaud M-N, Siebe C (2016) The ~AD 1250 effusive eruption of El Metate shield volcano (Michoacán, Mexico): magma source, crustal storage, eruptive dynamics, and lava rheology. Bull Volcanol 78(4):1–28
62. Uribe-Cifuentes RM, Urrutia-Fucugauchi J (1999) Paleomagnetic study of the Valle de Santiago volcanics, Michoacán-Guanajuato volcanic field, Mexico. Geofis Int 38(4):217–230
63. Aranda-Gómez JJ, Luhr JF, Pier G (1992) The La Breña – El Jaguey Maar Complex, Durango, México: I. Geological evolution. Bull Volcanol 54(5):393–404
64. Carmichael ISE, Lange RA, Luhr JF (1996) Quaternary minettes and associated volcanic rocks of Mascota, western Mexico: a consequence of plate extension above a subduction modified mantle wedge. Contrib Mineral Petrol 124(3):302–333
65. Blatter DL, Carmichael ISE, Deino AL, Renne PR (2001) Neogene volcanism at the front of the central Mexican volcanic belt: basaltic andesites to dacites, with contemporaneous shoshonites and high-TiO2 lava. Geol Soc Am Bull 113(10):1324–1342
66. Aguirre-Díaz GJ, Jaimes-Viera MC, Nieto-Obregón J (2006) The Valle de Bravo Volcanic Field: geology and geomorphometric parameters of a quaternary monogenetic field at the front of the Mexican Volcanic Belt. In: Siebe C, Macías GJL, Aguirre-Díaz J (eds) Neogene-Quaternary continental margin volcanism: A perspective from Mexico. Geological Society of America, Boulder, pp 139–154
67. Siebe C, Rodríguez-Lara V, Schaaf P, Abrams M (2004) Radiocarbon ages of Holocene Pelado, Guespalapa, and Chichinautzin scoria cones, south of Mexico city: implications for archaeology and future hazards. Bull Volcanol 66:203–225
68. Siebe C, Arana-Salinas L, Abrams M (2005) Geology and radiocarbon ages of Tlaloc, Tlacotenco, Cuauhtzin, Hijo del Cuauhtzin, Teuhtli, and Ocusacayo monogenetic volcanoes in the central part of the Sierra Chichinautzin, Mexico. J Volcanol Geotherm Res 141(3–4):225–243
69. Arce JL, Muñoz-Salinas E, Castillo M, Salinas I (2015) The ~ 2000 yr BP Jumento volcano, one of the youngest edifices of the Chichinautzin Volcanic Field, Central Mexico. J Volcanol Geotherm Res 308:30–38
70. Carrasco-Nüñez G, Riggs NR (2008) Polygenetic nature of a rhyolitic dome and implications for hazard assessment: Cerro Pizarro volcano, Mexico. J Volcanol Geotherm Res 171(3–4):307–315
71. Siebert L, Carrasco-Nüñez G (2002) Late-Pleistocene to precolumbian behind-the-arc mafic volcanism in the eastern Mexican volcanic belt; implications for future hazards. J Volcanol Geotherm Res 115(1–2):179–205
72. Espíndola JM, Zamora-Camacho A, Godinez ML, Schaaf P, Rodríguez SR (2010) The 1793 eruption of San Martín Tuxtla volcano, Veracruz, Mexico. J Volcanol Geotherm Res 197(1–4):188–208
73. Tilling RI (2009) El Chichón's "surprise" eruption in 1982: lessons for reducing volcano risk. Geofis Int 48(1):3–19
74. Luhr JF, Carmichael ISE, Varekamp JC (1984) The 1982 eruptions of El Chichón Volcano, Chiapas, Mexico: mineralogy and petrology of the anhydrite bearing pumices. J Volcanol Geotherm Res 23(1–2):69–108
75. Rouwet D, Taran Y, Inguaggiato S, Varley N, Santiago Santiago JA (2008) Hydrochemical dynamics of the "lake-spring" system in the crater of El Chichón volcano (Chiapas, Mexico). J Volcanol Geotherm Res 178(2):237–248
76. Jutzeler M, Varley N, Roach M (2011) Geophysical characterization of hydrothermal systems and intrusive bodies, El Chichón volcano (Mexico). J Geophys Res 116(B4):B04104

77. Macías JL et al (2008) Hazard map of El Chichón volcano, Chiapas, Mexico: constraints posed by eruptive history and computer simulations. J Volcanol Geotherm Res 175(4):444–458
78. Espíndola JM, Maclas JL, Tilling RI, Sheridan MF (2000) Volcanic history of El Chichón volcano (Chiapas, Mexico) during the Holocene, and its impact on human activity. Bull Volcanol 62(2):90–104
79. Rouwet D, Inguaggiato S, Taran Y, Varley N, Santiago SJ (2009) Chemical and isotopic compositions of thermal springs, fumaroles and bubbling gases at Tacaná volcano (Mexico-Guatemala): implications for volcanic surveillance. Bull Volcanol 71(3):319–335
80. Macías J et al (2010) Late-Pleistocene flank collapse triggered by dome growth at Tacaná volcano, Mexico-Guatemala, and its relationship to the regional stress regime. Bull Volcanol 72(1):33–53
81. Macías JL et al (2000) Late Holocene Peléan style eruption at Tacaná volcano, Mexico-Guatemala: past, present, and future hazards. Geol Soc Am Bull 112:1234–1249

Index

J. W. LaMoreaux (ed.), *Environmental Geology*,
https://doi.org/10.1007/978-1-4939-8787-0