Coastal Research Library 18

Xin Zhang
Lei Wang
Xiaoyi Jiang
Changming Zhu *Editors*

Modeling with Digital Ocean and Digital Coast

Springer

Coastal Research Library

Volume 18

Series Editor
Charles W. Finkl
Department of Geosciences
Florida Atlantic University
Boca Raton, FL 33431, USA

The aim of this book series is to disseminate information to the coastal research community. The Series covers all aspects of coastal research including but not limited to relevant aspects of geological sciences, biology (incl. ecology and coastal marine ecosystems), geomorphology (physical geography), climate, littoral oceanography, coastal hydraulics, environmental (resource) management, engineering, and remote sensing. Policy, coastal law, and relevant issues such as conflict resolution and risk management would also be covered by the Series. The scope of the Series is broad and with a unique cross-disciplinary nature. The Series would tend to focus on topics that are of current interest and which carry some import as opposed to traditional titles that are esoteric and non-controversial. Monographs as well as contributed volumes are welcomed.

More information about this series at http://www.springer.com/series/8795

Xin Zhang • Lei Wang • Xiaoyi Jiang
Changming Zhu
Editors

Modeling with Digital Ocean and Digital Coast

Springer

Editors
Xin Zhang
State Key Laboratory of Remote Sensing Science
Institute of Remote Sensing and Digital Earth
Chinese Academy of Sciences
Beijing, China

Xiaoyi Jiang
National Marine Data and Information Service
State Oceanic Administration of China
Tianjin, China

Lei Wang
Institute of Remote Sensing and Digital Earth
Chinese Academy of Sciences
Beijing, China

Changming Zhu
Department of Geography and Environment
Jiangsu Normal University
Xuzhou, Jiangsu, China

ISSN 2211-0577 ISSN 2211-0585 (electronic)
Coastal Research Library
ISBN 978-3-319-42708-9 ISBN 978-3-319-42710-2 (eBook)
DOI 10.1007/978-3-319-42710-2

Library of Congress Control Number: 2016951456

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

The Earth is the only known blue planet in the universe. Approximately 71 % of the Earth's surface is covered by oceans, which operate as an elementary component of the global life support system and act as a balance for the human resource treasury and environment. Coastal regions all over the world are densely populated. Oceans and coastal regions are changing at faster rates, over broader scales, than ever before and in fundamentally new ways. Digital analysis based on multisource data can greatly improve the knowledge about the oceans and coasts. The data are obtained by diverse observation systems such as satellites, airplanes, ships, high-frequency ground-wave radars, buoys (moored and drifting), and land-based stations. Though there are many global information systems, such as Google Ocean, which can provide information, such as videos of ocean life, and allow the public to watch unseen footage of historic ocean expeditions, more studies should be carried out to meet the requirements of scientific research and governmental applications, especially from the Digital Ocean and Digital Coast perspective.

Based on various types of data analysis of the ocean, including survey and evaluation data, historic data, basic geographic data, remote sensing data, socio-economic data, and model estimate data, the specific content of the Digital Ocean and Digital Coast system construction modeling theories and technologies is included in this book. The modeling theories and technologies in this book are described from data, computation, analysis, application, and decision-making perspectives.

As modeling with Digital Ocean and Digital Coast relates to a number of research areas on a technical level, such as remote sensing, geographical information systems (GIS), virtual reality, scientific data visualization, computer network, geodesy, and data warehouse, the study content of modeling with the Digital Ocean and Digital Coast is introduced in Chap. 1 by Xin Zhang, Lei Wang, Xiaoyi Jiang, and Changming Zhu. Then ocean big data characteristics, acquisition, integration, and web service technologies are introduced in Chaps. 2 and 3 by Xin Zhang. In Chaps. 4, 5 and 6, the coastal flood forecasting modeling and analysis, coastal flood frequency modeling, spatial decision-making, and analysis theories and

technologies are introduced by Lei Wang and Xin Zhang. In Chap. 7, the ocean and coast disaster data modeling technologies are introduced by Xin Zhang. Chapter 8 by Changming Zhu and Xin Zhang investigates several new methods, including coastline automatic extraction, intertidal zone identification, coastal wetland classification, and coastal invasive plant detection using remote sensing. In the end, some applications and practical achievements of Digital Ocean and Digital Coast study in the China Offshore Digital Ocean Information Infrastructure Program are introduced by Xiaoyi Jiang.

In terms of modeling with Digital Ocean and Digital Coast, these preliminary results and applications push Digital Ocean a step forward from an unrealized concept to realistic systems. For more powerful Digital Ocean and Digital Coast studies and applications, more in-depth future research is needed.

The study is funded by the Great Research Project, the Free Exploring Program of the State Key Laboratory of Remote Sensing Science of China (No. 14ZY-03), the National Science & Technology Pillar Program (2015BAJ02B01, 2015BAJ02B02) the Special Research Project for the Commonweal of the Ministry of Water Resources of the People's Republic of China (grant no. 201201092), the National Natural Science Foundation of China (grant no. 61473286 and 61375002).

Chaoyang District, Beijing, China — Xin Zhang
Chaoyang District, Beijing, China — Lei Wang
Hedong District, Tianjin, China — Xiaoyi Jiang
Xuzhou, Jiangsu Province, China — Changming Zhu

Contents

Chapter 1
Introduction

Xin Zhang, Lei Wang, Xiaoyi Jiang, and Changming Zhu

Abstract The oceans are an important part of the earth that is a treasure house of resources and an important regulator of the global environment. There is a great part of the human being living in the coastal region all over the world. Oceans and coastal regions are changing at faster rates, over broader scales, than ever before and in fundamentally new ways. Digital analysis based on multisource data can greatly improve the cognition about the oceans and coasts, which are from diverse observing approaches such as satellites, airplane, ship, high frequency ground wave radar, buoys (moored and drifting) and land-based stations. This chapter briefly describes the concepts of the digital ocean and digital coast (DO&DC), discusses the modeling and visualization technologies for the realization of the DO&DC system, and notes the important roles of the DO&DC in digital earth development.

Keywords Digital ocean • Digital coast • Digital earth • Modeling

X. Zhang (✉)
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zhangxin@radi.ac.cn

L. Wang
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

Hainan Province Key Laboratory of Earth Observation, Sanya Institute of Remote Sensing, Sanya 572029, Hainan Province, China
e-mail: wanglei98@radi.ac.cn

X. Jiang
National Marine Data and Information Service, State Oceanic Administration of China, Tianjin 300171, China
e-mail: andyjiangxy@126.com

C. Zhu
Department of Geography and Environment, Jiangsu Normal University, Xuzhou 221116, Jiangsu, China
e-mail: ablezhu@163.com

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_1

1.1 Digital Earth

Advances in the past few years have demonstrated that many aspects of the Digital Earth (DE) that Al Gore envisioned in 1992 and later described in his 1998 speech are now technically feasible (Goodchild 2008). Chen and Genderen (2008) state that DE is an integrated approach to build the next level of scientific infrastructure to support global change research. They present a number of examples, including cover forest and grassland fires, desertification and sandstorms, deforestation, forest carbon sequestration, wetlands conservation, the observation of migratory birds for the spread of avian influenza (bird flu), the Tibet Plateau uplift, the rise of sea levels, and underground coal fires. In terms of DE's potential fields of applications, Guo et al. (2009, 2010) studied the DE prototype system DEPS/CAS and, through this research, defined DE systems as either scientific systems (such as the World Wind of the USA, the DE Prototype system/Chinese Academy of Sciences (DEPS/CAS) of China, the Blue Link and the Glass Earth of Australia, and the Earth Simulator (ES) of Japan, among others) or commercial systems (such as Skyline and Google Earth) and proposed that DE is a comprehensive platform for the integration of future information resources. Wright et al. (2010) presented a methodology for visualizing reconstructed plumes using virtual globes, such as that of Google Earth, which allows the animation of the evolution of a gas plume to be easily displayed and shared on a common platform. Yasuko et al. (2010) developed a visualization system (KML generators) for multidisciplinary geoscience data that visualizes seismic tomographic models, geochemical datasets of rocks, and geomagnetic field models by exploiting Google Earth technologies. Chen et al. (2009) put forward a solution to render the vertical profiles of atmospheric data from the A-Train satellite formation in Google Earth, using data from the NASA Cloud satellite as a proof-of-concept. However, these authors have not yet solved the problems facing the exploration and visualization of three-dimensional ocean data on visual globes.

DE is a powerful digital information world that can be accessed of space-time data of the earth in a virtual form (Vahidnia and Alesheikh 2013). With the development of technology, science has entered into an era of big data in which early research on the earth does not meet development of the earth's needs (Chaowei Yanga et al. 2013; Guo 2013). In the digital information era, huge volumes of geo-referenced data can be transformed into useful information that can be analyzed, visualized and shared and is of great interest to scientists (Manfred Ehlers and Peter Woodgate 2014). Guo (2013, 2014) noted that the use of DE technology can address real world human challenges (e.g., climate change, natural disasters, and urban expansion, and in the cloud computing platform), support dust storm forecasting, soil erosion monitoring, and forest disaster prevention (Peng Yue and HongxiuZhou 2013; Ick-Hoi Kim and Ming-Hsiang Tsou 2013). Lars Bernard and Stephan Mäs (2014) noted that scientific GDI (geodata infrastructures) could become one of the core components in the future DE. Yingjie Hu and Zhenhua (2015) proposed making 3D models, virtual earth map layers, remote sensing

images, digital elevation models (DEMs), and other data formats to build the DE system in a unified virtual environment.

DE is becoming a global challenge and a widespread leading field in science and technology. It is a comprehensive embodiment of a country's science and technology, economic strength, and national security safeguarding ability is and also one of the important signs of a country's comprehensive national strength. DE has a very vigorous life-force and is a global strategic goal of science and technology development. DE is a comprehensive platform and integration of future information resources. In many ways, DE, as a collection of technologies, has integrated the huge and valuable body of geo-data resources. It will prove to be an important area of research in the coming years (de By and Yola Georgiadou 2014; Yingjie Hu and Zhenhua 2015), but we are also facing a series of challenges such as big data storage, communication and processing (Guo 2014, 2015; Chris Pettit and Arzu Coltekin 2015).

1.2 Digital Ocean

The Earth is the only known blue planet in the universe. Approximately 71 % of the earth's surface is covered by ocean, which operates as an elementary component of the global life support system and acts as a balance for the human resource treasury and environment. Oceans and, indeed, the entire planet, are changing at faster rates, over broader scales, than ever before and in fundamentally new ways. In a very short period of time, the bounty of oceans has been depleted, and the ocean ecosystems have become seriously disrupted (Wright et al. 2007). The bounty and circulation of the oceans makes the study of the problems of rising sea levels, warming seawater, increasing storm intensities and other issues important from a global perspective. Since 2005, the Global Earth Observation System of Systems (GEOSS) has been implemented to achieve comprehensive, coordinated and sustained observations of the Earth to increase our understanding of the Earth's processes and to enhance the prediction of the behavior of the Earth system (GEO 2005). As the oceanographic component of GEOSS, one objective of the Global Ocean Observing System (GOOS) is to foster the development of data management systems that allow users to exploit multiple data sets from many different sources through "one-stop shopping" (Thomas 2003). The purpose of the Integrated Ocean Observing System (IOOS) is to make a more effective use of existing resources, new knowledge and advances in technology to provide the data and information required for global or regional scientific studies (Ocean.US 2002). As the National Science Foundation's contribution to the U.S. IOOS, the Ocean Observatories Initiative (OOI) will construct a networked infrastructure of science-driven sensor systems to measure the physical, chemical, geological and biological variables of the ocean and seafloor. Greater knowledge of these variables is vital for the improved detection and forecasting of environmental changes and their effects on biodiversity, coastal ecosystems and climate (COL 2009). In the future, ocean

information will become an increasingly valuable commodity worldwide because of the role of maritime commerce and new ocean-related investments, vulnerability to ocean-related natural disasters, the need to provide security for coastal populations, and the challenges of providing food and water to more people (Interagency Ocean Observation Committee 2013).

The term of DO emerged after the DE program was proposed by Al Gore (1998). The DO is the embodiment and re-innovation of "DE" theory and technology for the oceans of the world (Hou 1999). Patrikalakis et al. (2000) studied a knowledge network of distributed heterogeneous data and software resources for multidisciplinary ocean research that brought together advanced modeling, observation tools and field estimation methods. However, these authors did not discuss the integration of the technologies based on the global visualization of the Earth. Su et al. (2006a, b, c) studied the technologies of ocean GIS and proposed the benchmarks and the key technologies of a China DO Prototype System. As a public application platform, Google Ocean (www.googleearth.com) can provide much educational information, such as videos of ocean life, and allow the public to watch unseen footage of historic ocean expeditions, but the site has many shortcomings for scientific research and governmental applications. There have been many visualization systems for regional ocean data developed using OpenSceneGraph, including the Regional Ocean Modeling System (ROMS) (Shen et al. 2007), but there are many other technologies that can be researched from the DO perspective.

Based on various types of data about the ocean, including survey and evaluation data, historic data, basic geographic data, remote sensing data and business data, the specific content of the DO includes the researching and developing the DO sphere system for social and business management services based on the earth sphere model to achieve the interactive 3D visualization, reproduction and prediction of various ocean subsystems (such as the seabed, water, and sea surface) and phenomena. The DO system can be divided into two versions, "management" and "public service", based on the analysis of the demand characteristics and application modes of different user groups.

The DO system can complete the integration of small-scale geographic data and thematic data of the ocean based on the earth sphere model. It can also realize the dynamic update of the ocean monitoring and forecasting information. The DO system can be used to achieve a variety of expressions of regional data and query and retrieval of ocean information and data. It can also display and distribute business information generated in the process of ocean management.

The DO public service system is designed to satisfy public information inquiries and share science knowledge and information products on Internet. They depend on the data integration of large-scale ocean basic geographic data, thematic data and real-time updated observing and forecasting information. The DO public service system applications span a variety of aspects, including sea and island management, ocean environmental protection, disaster prevention and mitigation, economy planning, law enforcement monitoring, rights and interests maintenance, science and technology management, and fisheries.

1.3 Digital Coast

The coastal zone is a special area that is under the effect and interaction of land and sea. Under the background of the global sea level rising, the coastal disasters seriously threaten the survival and living environment of the coastal zone living things. The dynamic monitoring of the coastal zone has gradually become a hot research topic in related research fields.

The coastal zone has abundant energy resources such as tidal energy, wave energy, and wind energy et al. It also has many biological resources such as reeds, seaweed, ocean microorganisms, mangroves and fish and many natural resources such as groundwater, seawater, minerals, beaches, and shoreline et al. However, the coastal zone can be subject to many natural disasters such as storm surge, typhoons and slow-onset disasters including coastal erosion, sea water intrusion, land subsidence, sea surface elevation, and estuary and harbor sediment deposition. It is difficult to use traditional methods to monitor the environment, especially large-scale coastal monitoring because of the characteristics of uncertainty in the surrounding coastal environment. With the development of remote sensing technology, more and more fields have been introduced. Remote sensing technology is large-range, short period and multi time phase, making coastal zone dynamic monitored in convenient and quick style.

Digital coast is an extension and application of the DE theory and technologies in the coastal zone, and it is a product of information technology development and application.

Under a broadband, high-speed computer information network environment as an information infrastructure, Digital Coast system takes a mass storage and distributed computing system to manage and process the ocean data.

As for the theory study, Digital Coast should study the method of water quality monitoring and analysis technologies systematically. It should research the change and development trends of coral reefs, mangroves and coastline, as well as the ocean oil spill pollution monitoring.

Digital Coast system can realize 3D visual expression and localization of high-precision satellite images, vectors data and terrain data of the offshore area and islands. It can take advantage of the ocean scene simulation function to show the real-world islands by integrating ocean water, ocean scene, coastal landscape, seabed topography and so on.

From the query and display of ocean usage information about the current situation, historical situation, statistics, and functional region division information in the different theme modules, we can achieve the functions of query, retrieval, location and statistical analysis of all types of information.

Based on the survey data from the islands and coastal zones in the ocean island theme, we can fulfill a variety of island information queries (such as location, area, length of coastline, and population), as well as the visualization of island images and terrain. Aimed at the ocean rights and interests, we can demonstrate the political, economic and military interests of a country and its neighboring countries

from the aspects of the political situation, economic situation, military situation. We can also browse and view the interests of the island terrain, high-precision images, profile information, multimedia information and its relevant attribute information. For the polar ocean theme, we can achieve the display, query and 3D landscape simulation of the polar expeditions. The previous tracking information, thematic polar information, and research stations visualization can also be viewed.

1.4 Modeling with DO&DC

Modeling with DO&DC is the important content to build a digital information system. The process of using a model to describe the causal and interrelation is called modeling. The means and methods to achieve this process are varied, as there are different ways of describing the relationship.

Modeling with DO&DC mainly focuses on system applications and data modeling. The main contents of the modeling include determining the data and related processes, defining the data structure, numerical simulation and forecasting modeling, data organization and storage modeling, system application modeling, and so on.

Modeling with DO&DC relate to a number of related research areas on the technical level, such as remote sensing, GIS, virtual reality, scientific data visualization, computer network, geodesy, and data warehouse.

Modeling with the DO&DC includes:

1. DO&DC is a new concept and application model from the application level of the earth sphere model. It requires DO&DC transition from theory to practice and a break from concept to engineering application entity;
2. From the perspective of data integration in the global sphere, DO&DC cover the major data types in the ocean areas, such as ocean survey data, model forecast data, pattern prediction data, space remote sensing data, aerial remote sensing data, basic geographic data, business approval data, and statistical data;
3. In the space region, DO&DC covers oceans, the polar regions, key ocean areas, the islands and coastal zones, and the vertical dimension of the space related to the expression of the main elements (seabed, water, sea surface, and sea islands);
4. In the application services, DO&DC covers the major areas, including ocean comprehensive management and macro decision-making, the ocean economy, ocean disaster prevention and mitigation, ocean fisheries, ocean dynamic environment, and ocean homeland security and maintenance;
5. Aimed at the two demand groups of ocean thematic business and information services groups from the perspective of application service groups, it needs to put forward unique system programs of research and development, which have a strong relevance and practicality;

6. The system development should be based on the requirements of business applications. It has realized the development of various elements' multi-dimensional expression to study the visual expression of different types of ocean and ocean phenomena;
7. Research on the true 3D Earth sphere model space parameters integrated the feature in electronic chat and its construction technology;
8. Research on ocean real-time monitoring data management is based on the earth sphere model and application integration technology to achieve a unified space-time management and application integration of multi-source monitoring data;
9. Study of the data model of the process of the typical EI Nino phenomenon and the sea level rise based on the earth sphere model;
10. Research data loading and sphere integration techniques based on the cloud service architecture under a distributed network environment;
11. Research on the integration, process analysis, feature mining, and application service technology of large ocean space-time data.

1.5 Virtual Visualization Application in DO&DC

Virtual visualization is the interactive processing theory, algorithm and technology research branch of visualization technology, which is based on computer image processing and graphics technology converting the scientific computing data, engineering calculation data and measurement data into graphics or images drawn on the screen. Using virtual visualization technology, we can transform the spatial data into a graphic image, to help people to address and identify its characteristics and rules and to understand the nature of laws. The virtual visualization can be applied in many areas of ocean research.

1.5.1 Research of Global Climate Change

The visual simulation of ocean circulation phenomena is crucial for the analysis and prediction of global climate change and the effect on the oceans. The oceans regulate the climate of the earth's land and the height of the oceans, the air dry or wet effect, and control of the wind speed and direction. Thus, the modeling and visualization of ocean temperature changes can help scientists understand the impact of climate change and predict future impacts on human activities. The research results of the modeling, simulation and visualization of the spatial analysis of global warming, El Niño, sea level change, carbon cycle, ocean circulation variability and other global climate changes in the research field could provide support services for international negotiations on climate change.

1.5.2 Management of Fisheries and Ocean Biology

The DO is capable of showing the ocean nature to reveal the living environment and ocean organism behavior. It can also be used to study the ocean habitats and explore the ocean resources.

1.5.3 Ocean Emergency Decision-Making

Under the information of wind field, flow field and the temperature field emergency incident background information, real-time monitoring information, all kinds of statistical information (surrounding economic losses, casualties and ecology destruction), we can build all types of ocean emergency (oil spill, hazardous chemicals leakage, leakage of nuclear, maritime) contingency plans and impact assessment models to realize the dynamic visualization expression to evaluate the emergency influence and assess the thematic information products made from the impact of unexpected events. By assessing the impact of emergency situation, it is possible to make an assessment of thematic information products.

1.5.4 Ocean Scientific Research Field

Using DO&DC data resources and computing resources integrated ocean numerical model, GIS spatial analysis model, ocean phenomena visualization model and other ocean research and development model to provide a virtual platform for scientific research users to load and test the mechanism of the ocean phenomenon, simulation and comprehensive analysis of scientific research results.

1.6 Conclusions

DE could provide support to social and economic development. On the one hand, it can support the overall sustainable development of a country, and on the other hand, it is closely linked to the integration of research on global change, resources, and environment as well as on global economic unifying processes.

DO&DC construction is a long-term engineering with strategic and prospective, and the modeling theories and technologies research is the core of the whole engineering. In future work, we should actively carry out strategic planning on DO development and establish the management and service system of ocean and coast information resources, as well as build up the information exchange and share channels.

Acknowledgments The study is funded by the Free Exploring Program of the State Key Laboratory of Remote Sensing Science of China (No. 14ZY-03), the Special Research Project for the Commonweal of the Ministry of Water Resources of the People's Republic of China (grant no. 201201092), the 908 Project of the State Oceanic Administration, China (No. 908-03-03-02), the National Natural Science Foundation of China (grant no. 61473286 and 61375002), and the National Science & Technology Pillar Program (2015BAJ02B01, 2015BAJ02B02). All memberers of the Research Team are gratefully acknowledged for their contributions to the work carried out in this chapter in recent years. In particular, special thanks are given to Miss Qiong Zheng, Mr. Yongxin Chen and Mr. Yuqi Liu for their contributions in the document revision.

References

Al Gore (1998) The DE: understanding our planet in the 21st century [online]. Available from: http://portal.opengeospatial.org/files/?artifact_id=6210. Accessed 14 July 2010

Chaowei Yanga, Yan Xu, Douglas Nebert (2013) Redefining the possibility of digital Earth and geosciences with spatial cloud computing. Int J Digital Earth 6(4):297–312

Chen SP, Genderen J (2008) Digital Earth in support of global change research. Int J Digital Earth 1(1):43–65

Chen AJ, Leptoukh G, Kempler S, Lynnes C, Savtchenko A, Nadeau S, Farley J (2009) Visualization of A-Train vertical profiles using Google Earth. Comput Geosci 35(2):419–427

Chris Pettit, Arzu Coltekin (2015) Geovisual analytics: design and implementation. Int J Digital Earth 8(7):517–521

Consortium for Ocean Leadership (COL) (2009) Project execution plan [online]. Available from: http://www.oceanleadership.org/wp-content/uploads/2009/04/. Accessed 14 July 2010

de By* RA, Georgiadou Y (2014) Digital earth applications in the twenty-first century. Int J Digit Earth 7(7):511–515

GEO (2005) The Global Earth Observation System of Systems (GEOSS) 10-year implementation plan [online]. Available from: http://www.earthobservations.org/documents/. Accessed 14 July 2010

Goodchild MF (2008) The use cases of digital Earth. Int J Digital Earth 1(1):31–42

Guo HD (2013) Making digital Earth on Earth. Int J Digital Earth 6(1):1–2

Guo HD (2014) Digital Earth: big Earth data. Int J Digital Earth 7(1):1–2

Guo HD (2015) Big data for scientific research and discovery. Int J Digital Earth 8(1):1–2

Guo HD, Fan XT, Wang CL (2009) A digital Earth prototype system: DEPS/CAS. Int J Digital Earth 2(1):3–15

Guo HD, Liu Z, Zhu LW (2010) Digital Earth: decadal experiences and some thoughts. Int J Digital Earth 3(1):31–46

Hou WF (1999) Tentative idea for development of "Digital Ocean". Aviso of Ocean 18(6):1–10

Ick-Hoi Kim, Ming-Hsiang Tsou (2013) Enabling digital Earth simulation models using cloud computing or grid computing-two approaches supporting high-performance GIS simulation frameworks. Int J Digital Earth 6(4):383–403

Interagency Ocean Observation Committee (2013) U.S. IOOS Summit Report: A new decade for the Integrated Ocean Observing System, [online]. Available from: http://www.iooc.us/wp-content/uploads/2013/01/U.S.-IOOS-Summit-Report.pdf. Accessed 14 Nov 2013

Lars Bernard, Stephan Mäs (2014) Scientific geodata infrastructures: challenges, approaches and directions. Int J Digital Earth 7(7):613–633

Manfred Ehlers, Peter Woodgate (2014) Advancing digital Earth: beyond the next generation. Int J Digital Earth 7(1):3–16

Ocean. US (2002) Building consensus: toward an integrated and sustained ocean observing system [online]. Available from: http://www.ocean.us/documents/docs/Core_lores.pdf. Accessed 14 July 2010

Patrikalakis NM, Bellingham JG, Mihanetzis KP (2000) The Digital Ocean. Computer Graphics International 2000 (CGI'00) (1):45–48

Peng Yue, HongxiuZhou (2013) Geoprocessing in cloud computing platforms-a comparative analysis. Int J Digital Earth 404(4)

Shen YZ, Austin JA, Crouch JR, Dinniman MS (2007) Interactive visualization of regional ocean modeling system. In: Proceedings of the IASTED International Conference on Graphics and Visualization in Engineering., pp 74–82

Su FZ, Du YY, Pei XB (2006a) Constructing digital sea of China with the Datum of coastal line. Geo-Inf Sci 8(1):12–15

Su FZ, Yang XM, Xu J (2006b) Basic theory and key technologies for ocean geographic information system. Acta Oceanol Sin 25(2):80–86

Su FZ, Zhou CH, Zhang TY (2006c) Constructing a raster-based spatial-temporal hierarchical data model for ocean fisheries application. Acta Oceanol Sin 25(1):57–63

Thomas CM (2003) The coastal module of the Global Ocean Observing System (GOOS): an assessment of current capabilities to detect change. Ocean Policy 27(3):295–302

Vahidnia MH, Alesheikh AA (2013) Modeling the spread of spatio-temporal phenomena through the incorporation of ANFIS and genetically controlled cellular automata: a case study on forest fire. Int J Digital Earth 6(1):51–75

Wright DJ, Blongewicz MJ, Halpin PN, Breman J (2007) Arc Ocean: GIS for a Blue Planet. ESRI Press, Redlands

Wright TE, Burton M, Pyle DM, Caltabiano T (2010) Visualising volcanic gas plumes with virtual globes. Comput Geosci 35(3):1837–1842

Yasuko Y, Yanaka H, Suzuki K, Tsuboi S, Isse T, Obayashi M, Tamura H, Nagao H (2010) Visualization of geoscience data on Google Earth: development of a data converter system for seismictomographic models. Comput Geosci 36(2):373–382

Yingjie Hu, Zhenhua LV (2015) A multistage collaborative 3D GIS to support public participation. Int J Digital Earth 8(3):212–234

Chapter 2
Ocean Big Data Acquiring and Integration Technologies

Xin Zhang, Lei Wang, Xiaoyi Jiang, and Changming Zhu

Abstract Digital Earth is an integrated approach building scientific infrastructure. The ocean data is a typical big data, which can be seen from the data volume, velocity, variety, and value perspectives. The Digital Earth systems provide a three-dimensional visualization and integration platform for ocean big data, which include ocean management data, in situ observation data, remote sensing observation data and model output data. Based on the analysis on the characteristic of ocean big data, this chapter studies the ocean big data acquire and integration technology that is based on the Digital Earth system. Firstly, the construction of the Digital Earth based three-dimensional ocean big data integration environment is discussed. Then, the ocean management data integration technology is presented which is realized by general database access, web service and ActiveX control. Third, the in situ data stored in database tables as records integration is realized with a three-dimensional model of the corresponding observation apparatus display in the Digital Earth system using a same ID code. In the next two parts, the remote

X. Zhang (✉)
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zhangxin@radi.ac.cn

L. Wang
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

Hainan Province Key Laboratory of Earth Observation, Sanya Institute of Remote Sensing, Sanya 572029, Hainan Province, China
e-mail: wanglei98@radi.ac.cn

X. Jiang
National Marine Data and Information Service, State Oceanic Administration of China, Tianjin 300171, China
e-mail: andyjiangxy@126.com

C. Zhu
Department of Geography and Environment, Jiangsu Normal University, Xuzhou 221116, Jiangsu, China
e-mail: ablezhu@163.com

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_2

sensing data and the model output data integration technologies are discussed in detail. The application in the Digital Ocean Prototype System of China shows that the method can effectively improve the efficiency and visualization effect of the data.

Keywords Digital earth • Digital ocean • Ocean big data integration

2.1 Introduction

Since the concept of Digital Earth was put forward by Al Gore in 1998 (2010), many studies have been carried out from the perspective of the Earth. In 2002, Skyline Inc. released the Skyline TerraSuite software. In 2004, NASA launched WorldWind version 1.1, marking the first Digital Earth platform software with a complete scientific research function, which provided scientists with a simulation and display platform for conducting Digital Earth research. In 2005, Google Inc. introduced Google Earth to the world, which raised the Digital Earth application to a new level. Google Earth supports the visualization of DEM data and realizes the virtual representation of the Earth. Similar software includes Leica Visual Explorer (LVE) form Leica Inc., the Visual Earth system from Microsoft Inc., ArcGlobe from ESRI Inc., the digital earth prototype system from the Chinese Academy of Sciences and so on (Guo Huadong 2009). In terms of Digital Earth's potential fields of application, Guo et al. (2009, 2010) studied the digital earth prototype system DEPS/CAS and, as a result of this research, defined digital earth systems as either scientific (such as those of World Wind of the USA, the Digital Earth Prototype system/Chinese Academy of Sciences (DEPS/CAS) of China, Blue Link and Glass Earth of Australia, and the Earth Simulator (ES) of Japan) or commercial (such as Skyline and Google Earth) and proposed that Digital Earth was a comprehensive platform for the integration of future information resources (Guo et al. 2009, 2010).

– With the development of ocean observation technologies, a substantial amount of oceanic data and model products are produced from the three-dimensional ocean observation system, which is composed of diverse monitoring sources such as satellite, airplane, ship, high frequency ground wave radar, buoys (moored and drifting) and land-based stations (Ocean 2010). How to efficiently and effectively integrate the data has become an urgent problem because these heterogeneous and widely distributed data and model products are usually collected in a project-based fashion. So far, the primary ocean observation data integration technologies are GIS (Geographic Information System) and web applications (Yingqi Tang and Wong 2006). In America, portals such as the Oregon coastal atlas, SCCOOS portal (Chongjie Zhang et al. 2007), provide clearinghouses for common decision-support tools, as well as data, maps and ancillary information. In Australia, there are several portals: the Oceans Portal proposed by the Australian National Oceans Office (NOO), Australian Ocean Boundaries Information Systems (AMBIS), and CSIRO's ocean data directory – Marlin (Strain et al. 2006). Global-level ocean portals are being developed as

well, such as the Oceans Biogeographic Information System (OBIS) – a virtual repository of oceanographic and biogeographic information (Malone 2003). Since 2005, the Global Earth Observation System of Systems (GEOSS) has been implemented to achieve comprehensive, coordinated and sustained observations to improve the monitoring of the state of the Earth, to increase understanding of the Earth's processes and to enhance the ability to predict the behavior of the Earth ocean system (GEO 2005). As the oceanographic component of GEOSS, one objective of the Global Ocean Observing System (GOOS) is to foster the development of data management systems that allow users to exploit multiple data sets from many different sources through "one stop shopping" (Thomas 2003). The purpose of the Integrated Ocean Observing System (IOOS) is to make more effective use of existing resources, new knowledge and advances in technology to provide data and information for global and regional scientific study (Ocean.US 2002). The National Science Foundation's contribution to the U.S., IOOS, the Ocean Observatories Initiative (OOI), will construct a networked infrastructure of science-driven sensor systems to measure the physical, chemical, geological and biological variables of the ocean and seafloor. Greater knowledge of these variables is vital for the improved detection and forecasting of environmental changes and their effects on biodiversity, coastal ecosystems and climate (Consortium for Ocean Leadership (COL) 2009). These achievements have provided data and product resources as well as the communications environment for information sharing and scientific research.

– The ocean management data, in situ observing data, remotely sensed data, and model output are produced in a distributed geographical environment. These data have significant scientific value for scientists and decision-makers in government. There are some studies on distributed geographic information processing that focus on geographic information processing from the Geographical Information System (GIS) view (Yang and Raskin 2009; Yue et al. 2009; Friis-Christensen et al. 2009; Zhang Tong and Tsou Ming-Hsiang 2009; Wang Shaowen and Liu Yan 2009). However, given the relatively high temporal frequency and the intrinsic spatial nature of the data, ocean data integration technology based on the Digital Earth system has not been widely implemented. This study is based on ongoing research in China that seeks to construct the China Digital Ocean Prototype System (CDOPS) as part of the China Digital Ocean Information Basic Framework. The experience with the development of the Digital Ocean prototype system in China is relatively rich and thus helpful for addressing this question.

2.2 Acquiring Ocean Big Data

There are many kinds of ocean data produced every day, which operates as an elementary component for the study and application on the oceans. In this chapter, the big data cognition for ocean is studied from four perspectives. Form the data

volume perspective, the ocean data acquire and analysis can produce big volume data. Form the data velocity perspective, the ocean data is collected from the eyes in the sky and objects on-the- ground networks, together with demographic, geologic, and socio-economic data and model estimates. Form the data variety analysis perspective, the database stored data and the unstructured data and model estimates are included. Form the data value mining perspective, there are many kinds of knowledge can be mined. Among the ocean big data, the three-dimensional data play an important role.

The three-dimensional data includes Digital Elevation Model (DEM) data for the seafloor and coast, in situ observational data, remote sensing data and model output data.

The DEM data are acquired from single- or multi-beam ship-borne echo sounders, which are the traditional systems used to map the seafloor topography with high precision results; in addition, data are produced from airborne AIRSAR/POLSAR synthetic aperture data or other methods (Maged et al. 2009). The DEM data from below the sea surface are used to construct the three-dimensional seafloor model as well as the model of the sea surface. The DEM data from the coast are used in land surface three-dimensional modeling, remote sensing data with different spatial resolutions represent as the surface texture.

Station observation data better reflect the environmental condition of the waters within their zones, and the changes in data offer a certain representativeness. Station observation also offers the characteristics of continuity, accuracy, and timeliness and mainly reflects the two aspects of continuous time and continuous space; a reasonable measure of space is continuous in the horizontal direction of the site layout and the vertical direction of the air, surpolygon and subsurpolygon. With regard to ocean profile measurement; continuous time refers to the various processes that can be captured by long-term continuous observation data; its accuracy is derived from the use of the infinite sequence of sample recovery. We care about the continuous process of change. With regard to the station system, on the one hand, the field observation data are a fast, accurate and reliable means of communicating in real time with ocean forecast and other departments for the purpose of controlling the ocean environmental features and evolution process. On the other hand, the field data are a historical resource that can be stored permanently.

The continuous space means in the horizontal direction and vertical direction of the air, surface and subsurface, and ocean profile measurement. The continuous time refers to the various processes in long-term continuous observation.

A self-propelled type of ocean platform is mainly used for underwater, unmanned, wide range, extended underwater environmental monitoring, including the physical parameters, the ocean geology and geophysics, and the ocean chemistry and biology parameters as well as aspects of ocean engineering – all of which can be performed close to the observation area. Its features are the following: low cost; environmentally adaptable; able to surpass the artificial diving limit and enter the field observation area; small size; easy to use; easy to wash; operates according to the acoustic signal remote control or preset program control; built according to

the requirements for related observation projects; independent power and relatively long underwater running time; noise is low; and can be hidden from observation.

The data buoy is anchored or floating on the ocean observation platform; plays an important role in the ocean observation system. Although remote sensing by air and satellite can be performed at great speed and over a wide area, only the surpolygon data are accessible; the unattended buoy submerged in an ocean environment works continuously for a long period of time in combination with other buoys for a comprehensive, profound, all-weather assessment of the ocean environment and the changes that it undergoes.

The main technical instruments for field observation include the following: a specially designed oceanographic vessel and salinity (conductivity)–temperature depth gauge (CTD), acoustic Doppler velocity profiler (ADCP), profiler, side sonar, an underwater vehicle and underwater laboratory, underwater robots, and equipment for seabed deep drilling. Direct observation of the data and mathematical models used to provide reliable reference can also be verified by the results of the experiment and mathematical methods. In fact, the use of advanced research vessels, test equipment and technical facilities for direct observation have indeed promoted the development of ocean science; especially since the 1960s, almost all of the major progress in this field has been closely related to the use of these resources.

Direct observation data can either be used as reliable information for the experimental study and mathematical model or to verify the results of the experiment and mathematical methods. The basic features are direct observation data authenticity and discreteness. Direct observation data are the real basis for understanding complex ocean phenomenon and the calibration of the model test; in addition, the remote sensing data application plays an irreplaceable role. In other words, these are the most basic data, and they serve as a reference for other forms of data and theoretical results. The development of monitoring technology for the ocean environment improved monitoring, prediction and forecast ability and therefore promoted the development of the ocean and coastal economy.

The in situ observational data are being collected from different types of sensors on a range of time scales, such as the data produced by the Argo buoy. The data-obtaining devices include oceanic optical buoys, seafloor moored instrumentation (for dynamic factors), self-locating sub-water tide monitors, sea sound detection buoys, sub-water tide comprehensive admeasuring apparatuses, cruises and others. Real-time and delayed-mode monitoring data collected from the above equipment are transmitted to a local data center. The characteristics of the in situ observational data are that it is related to a fixed x, y, z location in space and has attributes that are attached or related to the location. Most of the data have the required temporal attributes.

– Satellite remote sensing covers most of the ocean environment parameters and information, including sea surpolygon temperatures, ocean currents, sea surpolygon wind fields, concentrations of chlorophyll, suspended mass concentrations, sea levels, gravity anomalies, ocean optical parameters, atmospheric

aerosol ocean rainfall, direction of the wave spectrum, and sea pollution in the surpolygon. Satellite ocean remote sensing involves the electromagnetic wave scope of visible light, infrared waves and microwaves. Visible light remote sensing using the sun as a light source, thermal infrared remote sensing using the sea, and passive microwave remote sensing can be divided into the surpolygon of the microwave radiation source and spaceborne microwave remote sensing active source. Satellite remote sensing instruments presently use radar scatterometers, radar altimeters, synthetic aperture radars, microwave radiometers, visible light/infrared radiometer ocean color scanners, etc.

- The radar scatterometer is an active microwave device; strabismus observations can show the surpolygon wind speed, wind direction and wind stress and sea waves. Using the scattering of the wave field is a reliable basis for sea condition forecasting.
- Spaceborne radar altimeters are also a type of active microwave sensor; they can measure geoids, sea ice, tides, water depths, sea surpolygon wind intensity and effective wave heights, and El Nino phenomena.
- Synthetic Aperture Radar (SAR) is a high azimuth resolution type of coherent imaging radar; it uses phase and amplitude information and is a type of holographic system that can be divided into side, strabismus, Doppler sharpening and bunching of surveying and mapping as well as other applications. According to the difference in SAR image brightness, the sea ice ridge, thickness and distribution, water and ice boundary, tip height and other important information can be extracted. By using SAR images, not only large-area oil pollution can be found in a timely fashion, but sudden pollution incidents can also be detected.
- The microwave radiometer is a passive microwave sensor that conducts remote sensing by measuring the thermal radiation temperature from the sea surpolygon temperature. Sea surpolygon temperature tests one of the most basic parameters, water temperature, which is one of the main measures for determining water mass and is also used to analyze the ocean front and flow. NOAA – 10, one of the last three satellites in the United States, is an advanced, very high resolution radiometer (AVHRR); an image represented by the sensor can be accurately mapped to the surpolygon with a resolution of 1 KM, and 1 °C temperature accuracy is possible. Satellite remote sensing of the sea surpolygon temperature of the global ocean isotherm distribution reveals complex phenomena that conventional methods were unable to discover and has even corrected previous findings.

The visible/near infrared bands of the multispectral scanner and coastal zone color scanner act as passive sensors; these devices measure ocean color, suspended sediment, water quality, etc. Using multi-spectral information and reflectance, suspended load concentration and migration can be extracted. Satellite remote sensing images can display frontal systems, eddy currents, and bodies of water, e.g., mesoscale ocean phenomena, and, combined with other satellite data research, can reveal the phenomena of many dynamic ocean mechanisms and processes.

– Compared with the traditional ship-based monitoring and buoy data, satellite synchronous measurement has major advantages in large areas, higher spatial resolution, long phase measurement, and the combination of multiple platform observations can capture various types of regional phenomena, thus affecting change research and the demand for global change research. In terms of relative sea direct observation, the cost of data is also very low. Satellite remote sensing has obvious disadvantages: observation is only possible on the surpolygon of the sea surpolygon, while visible light remote sensing reaches tens of meters into the sea surpolygon, but this is not sufficient; Various sensors have inherent defects, e.g., infrared remote sensing and spatial synchronization coverage problems in ocean color remote sensing.

Ocean numerical model data is also one of ocean big data sources. With the rapid development of modern numerical computing technology, more and more complex, accurate and efficient ocean numerical models have been developed.

The ocean numerical model is also a data source. Describing the ocean is very complex, and there are numerous forms of mathematical equations, but only in rare cases is it possible to obtain an accurate solution; however, numerical equations and methods can solve these problems to a great extent. With the rapid development of modern numerical technology and computer technology, increasingly complex, accurate and efficient ocean numerical models were created. The Princeton University POM model is one of the most widely used numerical models. MM5 adoption provides much useful data for the research on the atmospheric environment of the ocean. These powerful numerical models can obtain accurate ocean movement scenarios through measured data validation. The ocean model also produces a wide range of data, which can be used for further research.

Ocean large field data commonly uses the form of grid data model. Advantages in the form of a grid data model are as follows.

The main numerical model output products and all kinds of remote sensing observation data are in the form of grid.

The original format of the data, in line with the data acquisition and storage, are studied. The current large data are obtained in the form of grid, the main source of numerical model output data products and all types of remote sensing observation data.

The form of the grid is simple and intuitive and conducive to application. In view of the current data used to evaluate this grid form, it is simple and practical but still maintains a variety of scalable features; it is sufficiently flexible to meet current needs. There are simple topological relations; thus, creating, recreating and all the basic operations are easy.

– The grid is flexible and efficient. Compared with the vector format of geographical information systems, the grid data operation is very easy but not inferior to the vector format with respect to many functions; the grid form can undertake all types of space operations and complete many types of spatial analysis functions, and the straightforward features make it consumer friendly. The model data are usually stored as a multilayer stack of column and row data. The structure can be

defined as regularly or irregularly spaced grid units, with discrete node locations defined in the x, y and z dimensions. The grid nodes or cell values vary over time, and each layer represents a unique slice in space with a given depth.

Heterogeneous data formats increase the difficulty of data sharing and application, prompting the formation of data standards. Currently, ocean field data can be stored in more common data formats, such as the HDF (Hierarchy Data Format) and NetCDF (Network Common Data Form) formats,

Binaries with a small footprint, which offers the advantage of a fast read speed, are more common in the ocean business. As business needs or other application requirements differ, the internal data organization are not the same, and the actual application requirements for the organization of data for analysis is a prerequisite for the processing of the data.

Routine hydrological elements are seawater temperature, salinity, density and sound velocity, all of which contribute to the flow field numerical model for 3-D mesh data. According to the data grid resolution, they can be divided into uniform grid data or variable grid data. Grid data adopt an isometric grid, and the grid resolution will be according to the calculation model; different bands have different resolutions.

Two-dimensional grid data include tidal field water-level data, storm surge water-level data, and surface flow data; this is the basis for the study of ocean phenomena regarding spatial-temporal processes. In terms of the data for the ocean surface, there is no direct concept for the "deep layer", and we enter it as two-dimensional grid field data. However, in the process of time and space analysis of the data, the data and time dimension can also be structured for 3-D data.

Other grid field data file formats are text files, tables, etc.,; the text file is not widely used only because it occupies more space, often as an intermediate data format for quality checking.

Using a numerical model to provide data for ocean research generally involves the integration of fairly measured data; in ocean science, this is often called "data assimilation"; using measured data (including mature remote sensing data), improves upon the drawbacks of the numerical method; the process of calculation can amount to continuously effective intervention. Generally, export data include two categories: data from a variety of data sets that produce export data, which are usually required to pass through the interpolation algorithm that is plugged into the grid-type data.

The ocean big data and information described above are archived or stored in more than one department because they are usually collected or produced based on a project-based approach at different times. At the same time, interdisciplinary research requires different types of data at different times. To effectively and efficiently integrate a substantial amount of ocean big data, one global visualization system should facilitate the sharing, accessing, exploration and visualization of these datasets for scientists, users and decision-makers.

2.3 Characteristics of Ocean Big Data

The ocean is dynamic, and continuous, with fuzzy boundaries of time and space; compared to other industries, ocean data have obvious features. In general, the data characteristics of the ocean are the following.

2.3.1 Acquired from Multiple Sources

Here, data sources are addressed, including data acquisition means and methods. The common marine data types include the in situ observation data, remote sensing observation data and model output data and so on. For example, the ocean environmental elements data can be obtained from different survey platforms, such as CTDs, drifting buoys, sea buoys, ARGOs, and satellite remote sensing, etc.

2.3.2 The Variety of the Data Content

Data reflect the diversity of content including phyletic and various – such as basic hydrological – elements such as temperature, salinity, density, sound velocity, and current; ocean phenomenon data include tide, tidal current, and storm surge; meteorological data involve sea surpolygon wind fields, air pressure, precipitation, relative humidity, etc. as well as biochemical data, geophysical data, etc. Different categories of data on the properties, management and application are not entirely the same. These different uses of the data content as well as the data management, maintenance and solutions to problems, especially in terms of analyzing and accurately obtaining relevant information, are very difficult and require a great deal of manpower and time.

2.3.3 The Heterogeneity Storage Format

Different measuring instruments, calculation methods and tools, data standards have resulted in heterogeneous data storage formats. Different production sectors use custom formats; in addition, semantic heterogeneity, heterogeneous coding, accuracy and data processing heterogeneity and a series of heterogeneous features.

Data from different sources are generally not the same in storage format, even if they are the same type of data, because of the variety of measuring instruments; measuring instruments also have different storage formats, for example, buoys (including mooring buoys and drifters), Nansen stations, different CODAS,

CTDs, ADCPs, and observation ships; remote sensing, satellite and other means of observation have different accuracies, and the data are in different formats.

The people dealing with these methods and tools for ocean data platforms formed a different storage format. In the past, file formats were used to store data; even now, the file format is an important means of data storage. However, the database, with its clear structure, convenient operation, ease of sharing, support for large amounts of data, and other incomparable advantages, has gradually become the most common type of data storage. File data, e.g., the most general text files, binary files, and ranks in the spreadsheet file (such as EXCEL spreadsheets), are suitable for a network with XML files, etc. The database data includes both small- and medium-size databases, such as Access, My SQL, and SQL Server, as well as large databases, such as Oracle, DB2, Informix, and Sybase. For basic geographic data, due to the use of different GIS platforms, the data formats are Shp, Coverage, Map, Tab, Mif, etc.

Heterogeneous data formats result in the difficulty in data sharing, exchange, and applications. There are many general Data Formats such as Hierarchy Data Format (HDF) and Network Common Data Form (NetCDF). The HDF file format is a hypertext document format developed by the University of Illinois. The HDF is suitable for a variety of computer platforms, easy to expand, mainly used to store by different computer platforms to produce various types of scientific data. The NetCDF was created by the University Corporation for Atmospheric Research (UCAR) under the Unidata program according to the characteristics of scientific Data; this format puts forward a kind of oriented array type, suitable for the network sharing. Additional Extensible Markup Language (XML) data are necessary because of its strong data description capabilities, it has in fact gradually become the network data exchange standard. Oceans in such locations as the United States, Australia, Russia and Europe are suitable for the Ocean XML application, and a great deal of development work has successful applied the XML data exchange, data processing and storage, daily management, etc. Because of the different data standards, the data storage format often differs, and different data are used according to the needs of the business; the production department will design their own data format specification. The ocean data format for heterogeneous data is a solution to data sharing.

In the actual data process, we should adopt a unified data structure, to achieve a variety of compatible purposes.

2.3.4 The Large Data Volume

The accumulation of observation involves progress and time; each department has accumulated an increasing amount of historical data; for example, satellite remote sensing for the study of ocean phenomena provides a new set of data, amounting to more than a century of data when combined with ship and buoy data.

After years of ocean observation and special surveys regarding the ocean environment, information has accumulated in areas such as physical oceanography (hydro), ocean surpolygon meteorology, ocean biology, ocean chemistry, ocean environmental quality (pollution), ocean geology, ocean Earth content physics, ocean aviation and satellite remote sensing – amounting to hundreds of gigabytes of data regarding the ocean environment in various disciplines.

2.3.5 The Velocity of Data

Ocean data has the characteristics of spatial and temporal dynamic. The ocean is dynamic and changing, and the dynamic nature of this environment is unlike land; in general, the dynamics of land do not involve the movement of an entire field, for example. Often change is partial and involves only a small area or boundary but lasts for a long period of time; the ocean, however, is changing every moment on a global basis.

Features of time and spatial characteristics and attributes are the three basic characteristics of temporal geographic data, and they are always changing. In time, the state of the space (including the actual space and shape of the space as well as spatial relationships) and its attributes will change. Basic hydrological factors (temperature, salinity, density and sound velocity, current, etc.) of time and space change dynamically, and this is one of the most properties of the ocean. In general, the ocean can be viewed as changing all the time, over a period of a few hours or even a few minutes. A number of attribute values change, which cannot be ignored. The basis of hydrological research is the study of elements of an area of the sea whose common attribute values change. The most obvious ocean data are related to ocean phenomena (storm tides, typhoons, tides) dynamics; the dynamic changes in time and space of ocean phenomena and attribute changes are given equal attention; for example, storm surge processes over time, storm surges in spatial locations (spatial information) and water level values (attribute information) are always changing.

Usually, a series of discrete time data are used to record the changes in hydrological elements or ocean phenomena through the dynamic process of interpolation to simulate changes in time and space. The temporal and spatial characteristics of ocean disasters receive the highest number of requests. The research in the field of geographic information systems has always been a problem that is difficult to overcome and requires constant exploration.

2.3.6 Multiple Temporal and Spatial Scales

From the perspective of oceanography, in terms of the time scale, the ocean environment and climate change characteristics for both information scales (i.e.,

for chronological or dated information) change with the seasonal variation in characteristics; there are synoptic scales that describe the variation in the time scale in days or hours. Thus, time scales range from hours to days, half-days, months, 10-day periods, seasons, years, and other scales. On the spatial scale, environmental information can be divided into large, medium and small categories; this includes phenomena that involve areas of 1000 km or more, such as the Kuroshio, areas measured in kilometers, such as the mesoscale vortex, and areas measured in meters, such as the mixed layer. In addition, spatial data production is the process of determining multiple spatial and temporal scales for ocean spatial data. Time scale performance data are produced in different time sequences. Some stations measure observation data once every hour, and some test once every day or even longer. Numerical model export data are measured according to different purposes. Time scale forecast data are typically measured in hours, and time scale statistical analysis of the data involves years, quarters, or months. Spatial multi-scale performance involves precision measurement of regional research and data, e.g., ocean remote sensing data from different satellites. Different spatial resolutions result in multiscale spatial data; examples include the Taiwan Strait and Chinese offshore waters; areas on a global scale, such as the Taiwan Strait, have a complicated flow architecture. In terms of areas prone to coastal storm surges, the spatial scale of the research is generally the Taiwan Strait.

2.3.7 Multi-Level in Depth

According to the ocean study objectives, the ocean data are required in different levels in depth. The ocean surface wind and wave data are mostly used for the mankind applications. While the water temperature, density, currents, chemical composition are different in different depth.

2.3.8 Multiple Levels of Data Users

From the point of view of data applications, the types of users differ because the task-level difference is substantial for the requirements of ocean disaster data. The end user data has three main categories: (1) the public; (2) government management departments; (3) and experts and scholars engaged in scientific research. Clearly, they care about the data product level, which can make a large difference in the corresponding data access levels; experts engaged in scientific research and scholars need primitive-level and feature-level information; government regulators need decision-level information; public policy makers require information at the decision-making level, i.e., conclusive information. In the data production and distribution process, different levels of data content, quality, and additional information make a large difference. Data for production department real-time

performance, quality, and confidentiality, as well as the different needs of users should be introduced into aspects of various levels of data in different production cycles.

2.3.9 The High Potential Value

The data obtained by numerical simulation or observation equipment cannot directly address the complicated relationship between each element; only through deeper analysis, calculation and visual expression can we better understand the data and phenomena and then improve the level of prediction and forecasting. Ocean disaster data are in the form of points, lines, polygons and bodies. In computing visualization, data can be divided into vector data and scalar data. The elements of visual expression are now concentrated in a single static expression, and ocean data are characterized by degeneration; again, there is a mutual influence involved in the interaction between the elements of the relationship; thus, there is an urgent need to study more elements as well as a need for the dynamic visualization of visual expression.

2.4 Primary Study on Ocean Big Data Integration Technology

The Digital Earth System based three-dimensional ocean environment is the basis for the integration of ocean big data. The three-dimensional ocean environment is composed of three parts: the three-dimensional coast, the ocean surface and the three-dimensional ocean floor. The three-dimensional coast is constructed from the DEM data and the land remote sensing data with multi-resolutions. The ocean surface is constructed by pictures with lumpy textures in commercial Digital Earth System platforms such as Skyline. The three-dimensional ocean floor is constructed by the DEM data of the seabed with gradual color rendering. In the three-dimensional ocean environment, a user can zoom all the way from space right down to street level on the coast (with images that in some places are sharp enough to show individual houses or even people) and through the ocean surface to the ocean floor. Some three-dimensional biological or man-made establishment models can also be added between the ocean surface and the three-dimensional ocean floor. Because actual commercial modeling platforms are mainly used for land environments, they have some shortcomings in digital ocean applications, especially in the construction of the three-dimensional ocean environment. Firstly, because the surface is constructed by pictures with a lumpy texture, the region representing the borders of different pictures looks discontinuous. Secondly, the visual resolution can change when users zoom from space right down to sea surface level.

Thirdly, the system running efficiency falls with the accumulation of pictures. Fourthly, it cannot simulate the actual scene to include features such as the ocean current direction. The reason for the shortcomings above is that these commercial systems don't provide programming interfaces for the ocean application. An independent scientific digital earth system can make up for these shortcomings through special arithmetic research and through software development. Because the Component Object Model (COM) has nothing to do with programming languages and can be issued by a binary specification, the scientific digital earth system with the global ocean current presentation function is designed to be developed with COM technology. Consequently, this system can be integrated with a commercial system like Skyline.

- The ocean management data are usually stored in tables of relational database management systems (RDBMS) such as Oracle or SQL Server and the like. And a database record is linked with a geographical feature in a map layer using the same ID number code as the attribute data. As the map layers can be added and visualized on Digital Earth systems, the ocean management data can accessed and viewed through selecting of the corresponding ocean management object such as a fishery region and the like. If the ocean management database is located in the same Intranet network environment with the Digital Earth system application server, the data are accessed with the standard interfaces which include ADO, OLEDB and others. If the ocean management database is located in the distributed network environment with the Digital Earth system application server, the data are accessed through web services. The main style of web services is Simple Object Access Protocol (SOAP). SOAP provides a standard message protocol for communication based on Extensible Markup Language (XML). SOAP web services have two main conventions: any non-binary attachment messages must be carried by SOAP and the service must be described using Web Service Description Language (WSDL). Because the Digital Earth systems have shortcomings in spatial data editing and analysis, it must be used through integration with the two-dimensional GIS system such as ArcGIS and the like to meet the requirement of ocean management. By controlling the display of two web pages which are embedded the ActiveX control of the Digital Earth system and the two-dimensional GIS system, the switch can be realized between three-dimensional Digital Earth display and two-dimensional display. The communication is realized by XML coding to keep the synchronization in spatial location of the two systems. The data synchronization is realized through sharing and access the same database.
- For the in situ observation data, spatial location (x, y, z) and time (t) are critical to the representation and exploration of the data and for further analysis. The location (x, y, z) is used by the Digital Earth systems to display the three-dimensional models of the observation apparatus such as buoys and the like, which are constructed by three-dimensional (3D) modeling software systems such as 3DMAX, Multigen Creator and so on. The measurement data at different time are stored in database tables as records and integrated with

three-dimensional model of the corresponding observation apparatus display in the Digital Earth system by a same ID code. The parameter introduction information can also be included in the web page linked with the apparatus by hot link As the number of the 3D models can influence the system running efficiency, the integration technology mentioned above is suitable for the observation apparatus which are moored at a fixed location. For the drifting observation apparatus, the display on the Digital Earth systems should be used by the two-dimensional symbols and the database tables design should be added the spatial location (x, y, z) information with each record.

– For remote sensing observation data, three integration technologies are used according to different data storage formats. If the remote sensing information products are in format of JPEG or GIF, they can be added to the Digital Earth systems directly as data files. If the remote sensing data are stored in a Spatial Data Engine (SDE) such as ArcSDE and the like, they can be added to the Digital Earth systems through SDE access interface. If the remote sensing data are stored in the Hierarchical Data Format (HDF), they can be added to the Digital Earth systems through Web Map Services (WMS) which is standard web service conformed to the standards set by the Open Geospatial Consortium (OGC).
– According to the spatial area of the model output data, the model output data are divided into two types, i.e. the global data and the regional data. For the global data, it is appropriate to visualize in the globe style. The two-dimensional map is overlaid on the three-dimensional visualizing globe of the Digital Earth systems. For the regional model output data, a special visualization ActiveX control is developed which can be integrated with the Digital Earth systems. Because the commercial Digital Earth systems are not suitable for display multi-layer at different depths, the web control integration technology can make up the shortcoming of it. When a kind of model output data need to be visualized, a region filled with given color is displayed firstly, then the ActiveX control interface is displayed and the model output data can be visualized.

2.5 Applications

The integration technologies discussed in above section are used in the Digital Ocean Prototype System of China. Figure 2.1 shows the system initialization interface and Fig. 2.2 shows the seabed terrain data visualization interface which is a part of the Digital Earth based three-dimensional integration environment.

The ocean management data integration content includes the island management data, economic statistical data, and ocean application region data. Figure 2.1 shows an ocean management data integration example, which is an ocean region application feature, and its attribute information display Fig. 2.3.

The in situ observing data are acquired from different devices which include large-scale environment observing buoys, high frequency ground wave radar, and

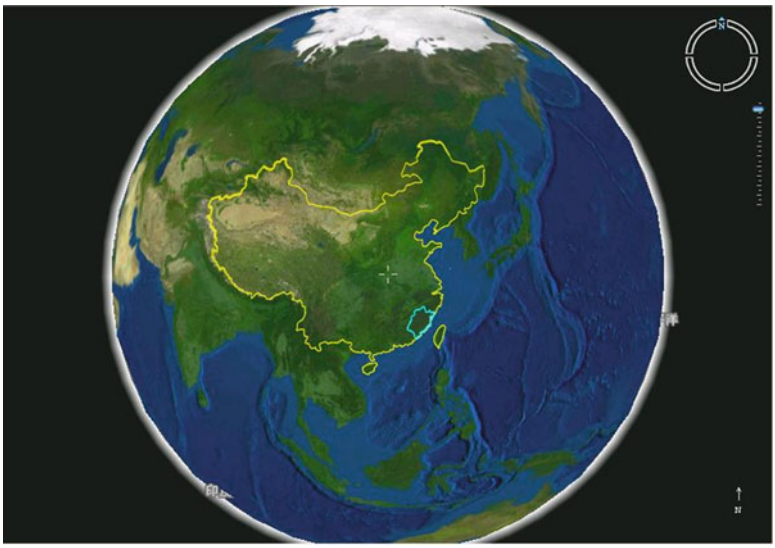


Fig. 2.1 System initialization interpolygon

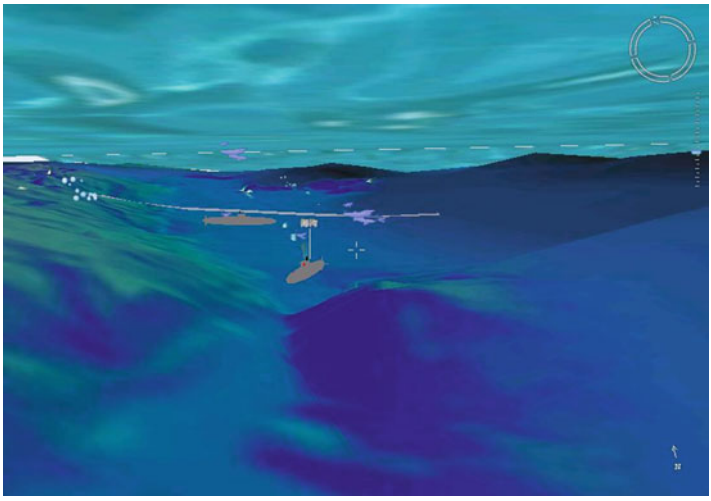
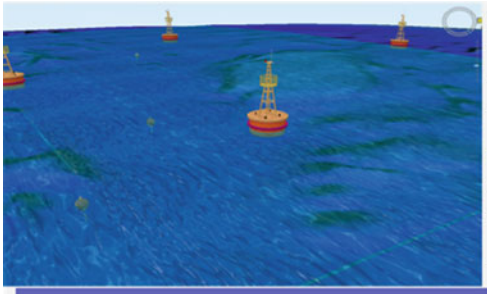


Fig. 2.2 Digital Earth based three-dimensional integration environment

shipping observing. The observing data are stored in a distributed environment and integrated by a Digital Earth system such as Skyline. The system can help users to access the collected data and devices information. As shown in Fig. 2.4, the key components that make up the data service interface are the map area, the introduction area and data display and the download area. The map area displays the geographical location of the observing device. The introduction area provides the general descriptions of large-scale environment supervision buoys. By selecting the



Introduction:

The variables include current direction, current speed, wind direction, wind speed, wave height, air temperature, water temperature at various depths and so on.

Data Browse & Download:

Start Date : 2004-12-30

End Date : 2007-12-30

Data Browse | Download

ID	Time	Air Pressure (hPa)	Air Temperature (℃)	Wind Speed (m/s)	Wind Direction (°)	Effective Wave Height(m)	Max Wave Height (m)	Water Temperature (℃)	Salinity (‰)	Current Speed (cm/s)	Current Direction (°)
1	2006-09-30 20:00	1013.3	26.9	12.2	215	1.7	2.6	25.4	14.0	06.8	243
2	2006-09-30 14:00	1012.5	27.1	11.5	208	1.9	2.6	25.5	13.4	05.3	30
3	2006-09-30 11:00	1014.0	26.7	12.3	210	2.1	3.6	25.4	13.0	15.2	68
4	2006-09-30 08:00	1014.7	26.5	11.3	211	2.6	3.7	25.4	13.0	03.3	207
5	2006-09-29 17:00	1012.1	27.5	13.1	218	2.5	3.3	25.5	13.4	14.8	226

>>Paging 1/138pages sum: 825 First Previous Next Last Goto Details

Fig. 2.3 The large-scale environment observing buoy data integration interpolygon

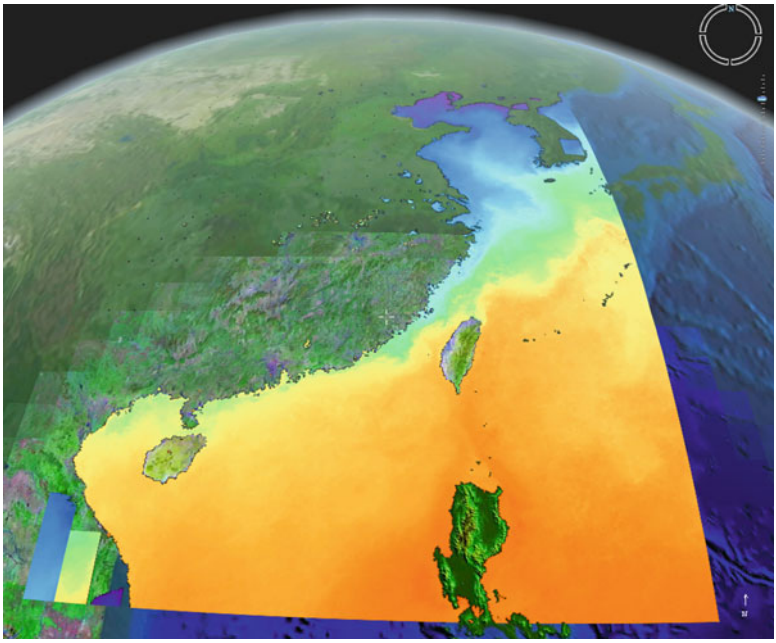


Fig. 2.4 The sea surface temperature (SST) product of MODIS integration

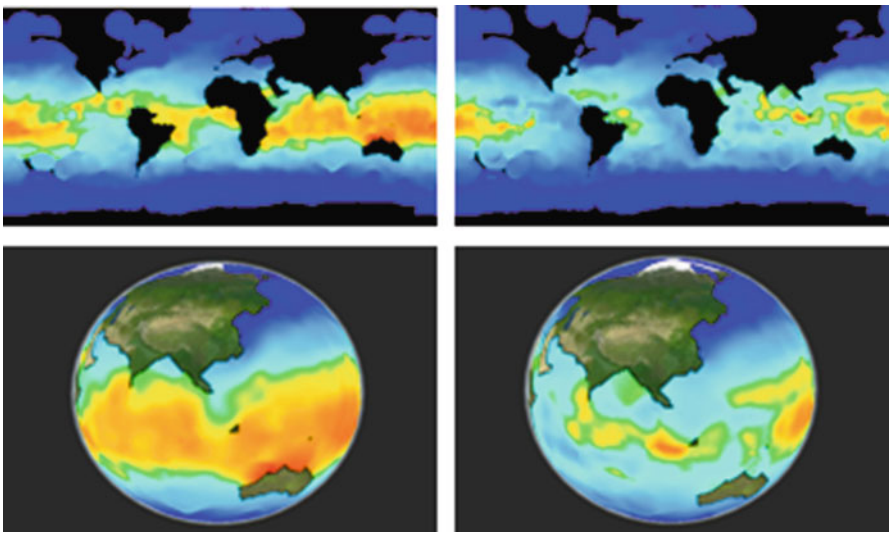


Fig. 2.5 The Integration of the ocean salinity statistical analysis data on different depth dimension

start date and the end date, historical data can be made available to the users in the data display area. The variables include air pressure, air temperature, wind speed, wind direction, effective wave height, maximal wave height, water temperature, salinity, ocean current speed, and ocean current direction. Real time data are updated every 3 h. The data displayed in the table can be downloaded in and XML file. XML was chosen because it is an open standard, and is platform independent and enables easy exchange of information with other organizations.

The satellite observing data integration provides a number of real-time and archived satellite derived products for the Chinese ocean region. The products collected from NASA's MODIS include 250 m RGB, sea surface temperature imagery, chlorophyll imagery and chlorophyll fluorescence. The QuikSCAT wind products and QuikSCAT Level 3 winds products are also available. If the products are available to all users, they can be obtained via direct download from the system. Figure 2.5 shows the sea surface temperature (SST) product of the MODIS integration interface.

Model computing products provide the main ocean data, which include statistical analysis data, ocean models such as the Princeton Ocean Model (POM) computing product data and so on. The data cover the whole or a region of Earth are divided into different levels of vertical dimensions. The original data format is in the grid style.

On the whole, model computing products have some common characteristics.

1. Data includes the longitude, latitude, depth and time, i.e., four dimensions. Longitude, latitude, and depth constitute the element distribution in three-dimensional space; the time dimension determines the time scope of the data set records.

2. The time dimension resolution of data varies because data sources vary, but at equal intervals, it chronicles the elements of the field distribution or elements of the statistical distribution of a certain period of time (mean, median, mode value etc.); and at different times of the nodes on the data, the space characteristics are consistent for those that have the same spatial range and spatial resolution.
3. The depth dimension resolution of the data is determined by access, the latitude and longitude data for different depth ranges and the resolution of two-dimensional spatial characteristics.
4. Data at the same depth layer and the same time point is two-dimensional grid data and applies more in the case of the uniform grid; it is the alternation of longitude and latitude on spacing as a fixed value, but the grid length-width ratio was not necessarily 1.
5. Within a certain area for each grid, it is necessary to record a number, and for the land in the area corresponding to the grid, it is necessary to use a fixed numerical identifier.

According to the above data Ccharacteristics analysis, based on the time dimension, the data can be divided into a series of identical space structure identical three-dimensional data,; using itused in accordance with the depth dimension, they can be further divided into a series of two-dimensional grid data structures, which is consistent with the rules.

From the above analysis, the ocean field element data comprises the three-dimensional space dimensions in addition to four-dimensional time data, but the time dimension can be viewed as the same spatial structure over time, so the time data analysis can be seen as the same operation in different three-dimensional space repeated over time. Its focus lies in the design of the three-dimensional data model.

Due to the need for field element data and general ocean application analysis, the existing three-dimensional data model cannot satisfy their needs, and therefore, it is necessary to study the characteristics and application analysis of the characteristics of ocean scalar fields of feature data based on the design of an appropriate data model.

The data for each depth layer is composed of a regular rectangular grid, and the same is true in the horizontal plane of projection; therefore, the region corresponding to the upper and lower interface can form a rectangular space. Due to the spacing between adjacent layers of uncertainty, rectangular hight can be seen as a variable.

The ocean field data three-dimensional space model can be viewed as the inverse process of data structure analysis. The first step is to build an interface rectangular mesh. The second step is to build a cuboid. The elevation value is only needed because of the single layer of data; to construct the cuboid, it is only necessary to search through a point with latitude and longitude coordinates and compare the elevation values that determine the up and down interface.

At present, the most commonly used approaches in the analysis of the ocean elements includes the following: single point information, profile analysis, plane distribution, and cross-section distribution analysis. The high rectangle model data

structure is one of the foundations of single point information analysis, according to the theory of cuboid model vertex information analysis. Profile analysis can be seen as a vertex information analysis. In addition, the essence of plane and profile analysis is three-dimensional space analysis of the cutting.

A higher cuboid modeling method is used for cutting analysis and will be applied to any area within the scope of the cuboid element analysis. The cutting method can be divided into three types: horizontal cutting, vertical cutting and cutting in an arbitrary direction; In essence, the first two are a special case of the third. Horizontal cutting corresponds to plane analysis, vertical cutting and cross section analysis. Because the analysis of the ocean elements is mostly based on two-dimensional data, any direction of cutting in the analysis of actual business is seldom used.

Because ocean data with a spatial structure can be divided into scalar field elements consisting of the characteristics of a planar grid, the first two special cases of cutting can be easily converted to the simple intersection of plane geometry problem solving and data interpolation.

As a result of current 3-D GIS platforms, terrain scenes and visualization of 3-D model objects and support technology is mature; the visualization technology and 3-D GIS platform can be set up with the aid of this advantage. On this basis, the terrain scene elements of ocean visualization can be more accurate; the image and intuitive data are contained in the message. The general 3-D GIS platform provides the space point, line, and polgon related interface but not special data model analysis and visualization of the interface. Therefore, in the existing 3-D GIS platform, to realize the ocean data visualization of scalar field elements, the key lies in how to make the results of the model analysis and platform point and polgon each other in a relationship that resembles an interface. Based on the above data model, there are two types of analysis.

The two main forms of analysis for the resulting points are numeric and chart. The discovery of some temperature value returns a value; and an analysis of the trend returns is a two-dimensional trend line chart (which may also be other types of diagrams).

For point objects, a three-dimensional GIS platform provides an interface for basic spatial location information but also provides the operational attribute information and type of attribute information for the acceptable data types.

Therefore, in view of the analysis, the analysis results can be used as point object attribute data correlation. Select any point on a sphere, and corresponding point pairs are dynamically created, and this provides the background model analysis results as the attribute data for the corresponding property field; in turn, by its attribute operation interface, visualization is achieved.

In view of the body analysis results for different types of planar space distribution, form a grid (using color for said numerical changes) and vector (in the form of a contour distribution).

For a surface, the 3-D GIS platform provides a very important attribute – the texture attribute.

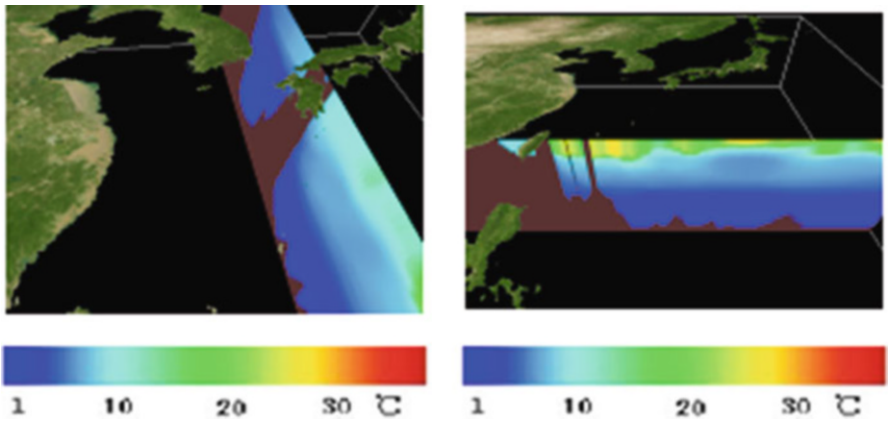


Fig. 2.6 The vertical data visualization integration of a section

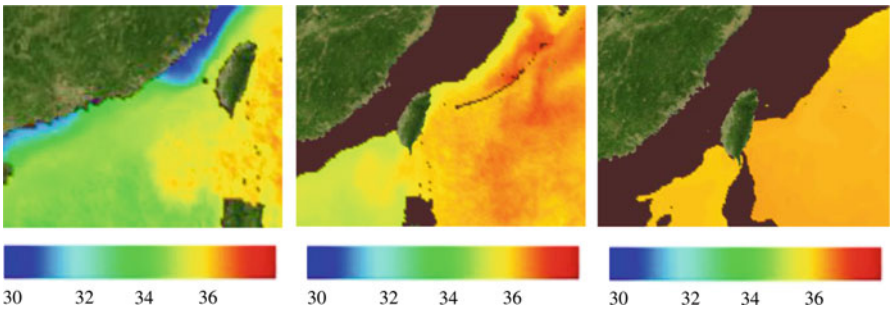


Fig. 2.7 Visualization of model output data of regional seawater salinity at different depth

Thus, for body analysis, in fact, it can be converted into space on a GIS platform that dynamically creates the face image and rendering textures.

Figure 2.6 shows the integration result of the ocean salinity statistical analysis data at the polgon and −75 m. Figure 2.7 shows the vertical data visualization integration result of a section.

2.6 Conclusions and Future Work

The ocean data is a typical big data, which can be seen from the data volume, velocity, variety, and value perspectives. This chapter studies Digital Earth system-based ocean data integration technology, which includes ocean management data, in situ observation data, remote sensing observation data and model output data. The application in the Digital Ocean Prototype System of China shows that the method can effectively improve the efficiency and visualization effect of the data.

In future work, we will study more international standard interpolygons or protocol-based integration technologies to achieve interoperability, which includes OGC-SWE and the IEEE 1451 standard.

Acknowledgments The study is funded by the Free Exploring Program of the State Key Laboratory of Remote Sensing Science of China (No. 14ZY-03), the Special Research Project for the Commonweal of the Ministry of Water Resources of the People's Republic of China (grant no. 201201092), the 908 Project of the State Oceanic Administration, China (No. 908-03-03-02), the National Natural Science Foundation of China (grant no. 61473286 and 61375002), and the International Science & Technology Cooperation Program of China (grant no. 2010DFA92720). All members of the Research Team are gratefully acknowledged for their contributions to the work carried out in this chapter in recent years. In particular, special thanks are given to Miss Qiong Zheng, Mr. Yongxin Chen and Mr. Yuqi Liu for their contributions in the document revision.

References

Al Gore (2010) The digital Earth: understanding our planet in the 21st century [online]. Available from: http://portal.opengeospatial.org/files/?artifact_id=6210. Accessed 14 July 2010

Chongjie Zhang, Chirag Dekate, Gabrielle Allen, Ian Kelley, Jon MacLaren (2007) An application portal for collaborative coastal modeling. Concurr Comput: Pract Exp 19(12):1571–1581

Consortium for Ocean Leadership (COL) (2009) Project execution plan [online]. Available from: http://www.oceanleadership.org/wp-content/uploads/2009/04/. Accessed 14 July 2010

Friis-Christensen A, Lucchi R, Lutz M, Ostländer N (2009) Service chaining architectures for applications implementing distributed geographic information processing. Int J Geogr Inf Sci 23(5):561–580

GEO (2005) The Global Earth Observation System of Systems (GEOSS) 10-year implementation plan [online]. Available from: http://www.earthobservations.org/documents/. Accessed 14 July 2010

Guo Huadong (2009) Digital Earth: ten Years' development and prospect, advances in Earth Science. Int J Digital Earth 24(9):955–962

Guo HD, Fan XT, Wang CLA (2009) Digital Earth prototype system: DEPS/CAS. Int J Digital Earth 2(1):3–15

Guo HD, Liu Z, Zhu LW (2010) Digital Earth: decadal experiences and some thoughts. Int J Digital Earth 3(1):31–46

Malone TC (2003) The coastal module of the Global Ocean Observing System (GOOS): an assessment of current capabilities to detect change. Ocean Policy 27(4):295–302

Ocean.US (2002) Building consensus: toward an Integrated and Sustained Ocean Observing System [online]. Available from: http://www.ocean.us/documents/docs/Core_lores.pdf. Accessed 14 July 2010

Ocean.US (2010) The National Office for Integrated and Sustained Ocean Observations, first annual Integrated Ocean Observing System (IOOS) development plan-a report of the National Ocean Research Leadership Council Prepared By Ocean. US [online]. Available from: http://www.ocean.us. Accessed 14 July 2010

Strain L, Rajabifard A, Williamson I (2006) Ocean administration and spatial data infrastructure. Ocean Policy 30(4):431–441

Thomas CM (2003) The coastal module of the Global Ocean Observing System (GOOS): an assessment of current capabilities to detect change. Ocean Policy 27(3):295–302

Wang Shaowen, Liu Yan (2009) TeraGrid GIScience Gateway: bridging cyberinfrastructure and GIScience. Int J Geogr Inf Sci 23(5):631–656

Yang C, Raskin R (2009) Introduction to distributed geographic information processing research. Int J Geogr Inf Sci 23(5):553–560

Yingqi Tang, Wong DW (2006) Exploring and visualizing sea ice chart data using Java-based GIS tools. Comput Geosci 32(6):846–858

Yue P, Di L, Yang W, Yu G, Zhao P, Gong J (2009) Semantic web services-based process planning for earth science applications. Int J Geogr Inf Sci 23(9):1139–1163

Zhang Tong, Tsou Ming-Hsiang (2009) Developing a grid-enabled spatial web portal for Internet GIServices and geospatial cyberinfrastructure. Int J Geogr Inf Sci 23(5):605–630

Chapter 3
Digital Ocean and Digital Coast Data Web Service Modeling

Xin Zhang

Abstract This chapter mainly focuses on the network service technology for ocean data application. This chapter starts with an introduction of two classes of traditional network service composition technology, all of which have limitations. To satisfy the increasing requirements of the user, the chapter introduces a new type of dynamic service combination technology. Then, this chapter introduces the modeling process of the technology, the network data service validation process and the implementation process of network data services. This new technology enables the rapid extension of the ocean information system and allows the system to rapidly design an individual response to user demand. Thus, the prospect of this application is notably good.

Keywords Web service • Dynamic combination

3.1 Introduction

In recent years, Web Services technology and the Service-Oriented Architecture (SOA) have made great developments and have had broad applications in the practice of information technology. On one hand, researchers continually develop new service-oriented computing theories, methods and models. The industry's major academic institutions and standardization organizations such as Organization for the Advancement of Structured Information Standards (OASIS) and World Wide Web Consortium (W3C) also establish service-oriented seminars for computing services and introduce SOA-related technical standards and regulations. On the other hand, the world's major IT giants, who announced their plan to fully support the SOA, have launched their own service-oriented application platform. Many enterprises have also adopted the SOA to build a more flexible, extendible enterprise information system. In the application of ocean technology, the level of informatization is relatively low. Various existing ocean information systems

X. Zhang (✉)
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zhangxin@radi.ac.cn

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_3

usually adopt the traditional closed system architecture. With the improvement in ocean information acquisition ability and the development of globalization, regionalization and integration research in the ocean field, the traditional system construction method gradually exposed the contradiction between the users' demand for higher comprehensive application and analysis and the system builders' growing professional knowledge structure. To satisfy the requirements of interdisciplinary comprehensive analysis based on the existing construction of ocean information, we must build a dynamic, open method of resource integration to provide the users with an integrated platform that flexibly adapt to all types of application requirements. Web service technology and the combination of these technologies can dynamically integrate resources and adapt to the changing requirements. However, in the current process of building an ocean information system, the static Web service combination approach based on the requirement limits the realization method and scope of the demand. When the user introduces a new application demand, the designer must implement the process based on the design under the system architecture, complete the corresponding service work, and provide the user with a new system functionality package. The problem of this approach is that the user's requirements cannot be quickly achieved even when the users want to add some slight changes or combine some existing functions. The system cannot automatically build a combination of functions or services. Thus, the dynamic combination technology of Web services gradually shows its advantage.

Web service combination satisfies the user's specific requirements by synthesizing fine-divided services as coarse-divided services. Different scholars proposed various methods of service combination from different angles. These methods can be divided into two categories: a service combination method based on the workflow or artificial intelligence (AI).

The service combination method based on workflow is proposed based on the process. The method requires the process structure, function and purpose of each activity in the process. Many different projects of service combinations are proposed according to different application purposes, such as the service combination based on the establishment and operation stage of the dynamic workflow, distributed dynamic service combination method based on roles, hierarchical model of the business process based on the hierarchical analysis method, and dynamic service combination model based on the workflow, etc. Currently, the combinations based on the workflow are mostly semi-automatically combined according to fixed business process models or templates. In the combining process, process modeling, service binding and parameter settings are performed by human intervention, which reduces the automatic degree of service combination.

The combination based on artificial intelligence (AI) service regards the Web service as an action in AI. The method describes the web services through the input and output parameters, i.e., the premise and the results. In the combination process, the method maps these descriptions into formalized descriptions of the action. In the Web service space, the service combination plan is based on the sequence of web services, which is dynamically formed by reasoning the formalized description of the action. Although it has a higher degree of automation, the service

combination method based on AI is limited by accidental combination. The objects in the combination are produced at run time, which may generate many branches in the process. After the service sets that are referred in the combination of services become notably large, the reasoning or query process becomes extremely difficult and cannot be completed in an acceptable time. In addition, the methods are usually too difficult to realize. Although two types of methods can dynamically combine services, both types have defects. The methods based on the workflow require clear business logic and is more dependent on the participation of humans. The methods based on AI have a high degree of automation but lack the necessary combination constraint, which may produce many combination results. These methods are difficult to satisfy the users' goal of building a dynamic service combination and efficient ocean applications.

3.2 Modeling of the Web Data Service

To prevent frequent updates of the ocean information system and system functional understanding discrepancies between the user and the system designer, a method to construct the ocean-monitoring data service chain is proposed. The method can rapidly respond to the user requirement, provide the users with a type of interactive, dynamic strategy to create and realize a new application.

The specific technical content of the method includes the following:

1. Abstract atomic services: Abstract atomic functions of the ocean information system services refer to the relevant Web service standards and protocols. Define the standardized service atom interface so that all ocean-monitoring data that apply atomic services can be automatically identified and managed.
2. Abstract model elements of the service chain: Analysis and decomposition the service chain. Define the visualization, description and constrain pattern of the elements. The user can understand the meaning of the elements of a formal model using these definitions. Meanwhile, the system can identify the description of the model elements.
3. Creation of the service chain: Analysis logic process of ocean stereo monitoring data. Select and combine the corresponding model elements based on the visual interface. Form a specific ocean service chain model of three-dimensional monitoring data.
4. Validation of the service chain: Validate the effectiveness and logic of the application service chain of ocean stereo monitoring data, which are created by the users according to the specific rules. Ensure that the service chain model that is used to perform is integrated.
5. Management of the service chain: Monitor and manage the composition, order of execution and service status of each atom in the chain. Ensure that the service chain information is available and can be controlled by the user.

Steps (1) and (2) are the basis of the implementation of the method. Define the service chain elements and function service of three-dimensional ocean-monitoring data in the building process. If the user wants to add some new functions, create a service chain to implement the new requirements and repeat steps (3)–(5).

Step (1) uses XML to describe the ocean-monitoring atomic data services. This step defines the service classification service and the referred input and output service interfaces to abstract and reference the WSDL definition of Web services model.

Step (2) abstracts the service chain model to 10 basic elements, i.e., nodes, according to the basic structure, combination of service chain and atomic services in the service chain model. This step also defines the name, attribute, and visualization pattern of each element.

Step (3) provides a visualization interactive interface of the service chain. The users select the model feature service chain and confirm the attribute value of each element. This step regulates the order of each element based on the logic of the ocean-monitoring data application and constructs the service chain model of ocean-monitoring data.

The validation process of the service chain in step (4) includes the parameter matching validation and model logical validation.

Step (5) is responsible for transforming the visualization of the service chain model into the service chain model description file, which the computer can identify, and storing and managing the service chain model, such as the management of model and execution. Model management includes saving, viewing, modifying, and applying the model information. Model execution management includes the management of execution status, the monitoring of the model process, etc.

Compared with the prior technique, this model has the following advantages. (1) The visualized service chain model of three-dimensional ocean-monitoring data provides the users with an intuitive and efficient method to monitor 3D ocean data. The user can use the existing ocean three-dimensional atomic service or a combination of dynamic monitoring data applications to build a specific service chain that satisfies the user's special requirements. (2) The application of atomic services for the three-dimensional ocean-monitoring data automatically identifies and manages ocean-monitoring data in the system. The system can recognize all services that the three-dimensional atomic ocean-monitoring data services have defined. (3) The abstraction of basic elements of the ocean-monitoring data service chain model make the ocean-monitoring data application service chain information easily understood and recognized by the users and computers, which is particularly advantageous for the management and application of the service chain model.

3.3 Validation of the Web Data Service

3.3.1 Matching Parameter Validation

The essence of model parameter matching is matching parameters of all service nodes in the process model and accepting input parameters with the data of the design services. When all service nodes in the process satisfy the following conditions, we believe that the matching parameters are successfully verified.

- The number of node service instance parameters PN_i = The number of input node service design parameters PN_d;
- The parameter types of node service instantiation TC_i = The input parameter types of node service design PC_d;
- The total number of nodes that correspond to the input DN_i = The number of node service design input data
- The type of input data that correspond to node DT_i = The type of node service design input data DT_d;

1. When to instantiate a node depends on the type of the node. If the node is a service node, execute steps (2)–(5); if not, directly jump to step (5);
2. Prompt the user to enter the service instance parameter;
3. According to the service identification of instantiated parameters, obtain the description of the corresponding service information;
4. Judge the above four matching conditions. Prompt the user when either condition does not satisfy the parameter mismatch and return to step (1) to instantiate this node again until all conditions satisfy the standards.
5. Instantiate the next node. The process model that includes the composite nodes can be verified by layer. Verify whether the complex internal nodes match the sub-process parameters and subsequently verify the upper layer until the top layer.

3.3.2 Model Logic Validation

Model logic verification refers to analyze the logical structure of the model. Judge whether the model is complete, bounded, retriggerable and executable.

A full service process model can be divided into a quintuple:

$$SN = (P, T, W, i, o)$$

where:

P: a collection of Web services states;
T: the services and collection of other necessary operations such as the user's confirmation;

W: a collection of the directed arc, acollection of causal association between P and T;
i: the initialization state;
o: the end of the state.

With this definition, we can describe the reachability, boundedness, retrigger and execution of the model as follows:

1. The model reachability indicates that the model starts from Pi. Any element Tn of the collection of changing model is accessible by triggering a series of changes. From the graphics, these elements indicate that any service or operation nodes can be reached through uninterrupted directed arcs from the initial state in the flow chart.
2. The boundedness model indicates that the entire model has unique start and end nodes. Any element Pn or Tn in P and T belongs to the reachable set from Po to Pi. The graphics indicate that a flowchart has only one start node and one end node, and all nodes are located in a path of directed arcs from Pi to Po.
3. If the model can be executed, from the starting node, every trigger will gradually lead to the end node. The reachable tree of the model does not have an infinite loop. From the flow chart, there is no endless loop in the model.
4. If the model can be triggered, any element Tn can be reached through a series of changes. The model can be completely executed.
5. Integrity of the model: for the elements in P and T, the model is accessible.

The service chain model that simultaneously satisfies these 5 points does not appear stagnant. All waiting news and the state are limited. There is no deadlock in the implementation process.

3.4 Implement Process of the Web Data Service

Specific implementation process (see Fig. 3.1) as follows.

In this instance, the service chain contains four basic atomic services: atomic service integration through a registered service in the service bus, construction of the service chain model through interactive service chain model modules, calling atomic services and status monitoring in the model management process and application process through the service management layer, and returning the results of the service chain execution to the user through the result display module.

The specific steps to create and implement a three-dimensional ocean-monitoring data application service chain model are as follows (see Fig. 3.2).

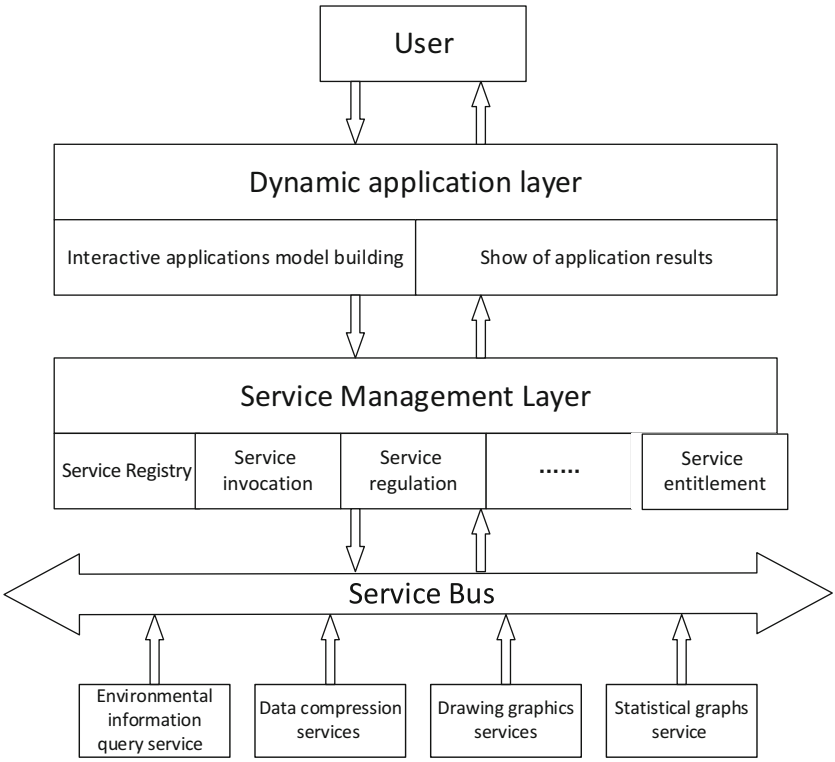


Fig. 3.1 Overall structure

3.4.1 The Service Model Creation

According to the logic order of application implementation, the user selects the abstraction elements of the service chain model and combines them. A specific ocean service chain model of three-dimensional monitoring data applications is formed. The basic elements of the service chain model belong to 9 classes: start nodes, end nodes, simple, consolidation, etc., as shown in Fig. 3.3.

To apply the information query of the ocean environment elements, the system requires the combination of four atomic services: query service of ocean environmental factor information (S_Dat_search), data compression service (S_Dat_reduce), graphics rendering service (S_Col_draw) and statistics generation services (S_Tab_stat).

This realization process is controlled by a combination of a sequential process and a selection process. The service chain model is shown in Fig. 3.4. Pi represents the initial state of the service, which includes the user's requirements for the display of ocean environmental factor information. T1 indicates the service S_Dat_search. P1 indicates that T1 correctly executes and maintains the data query results. T2 indicates the selection operation. According to the result of the user's selection, the

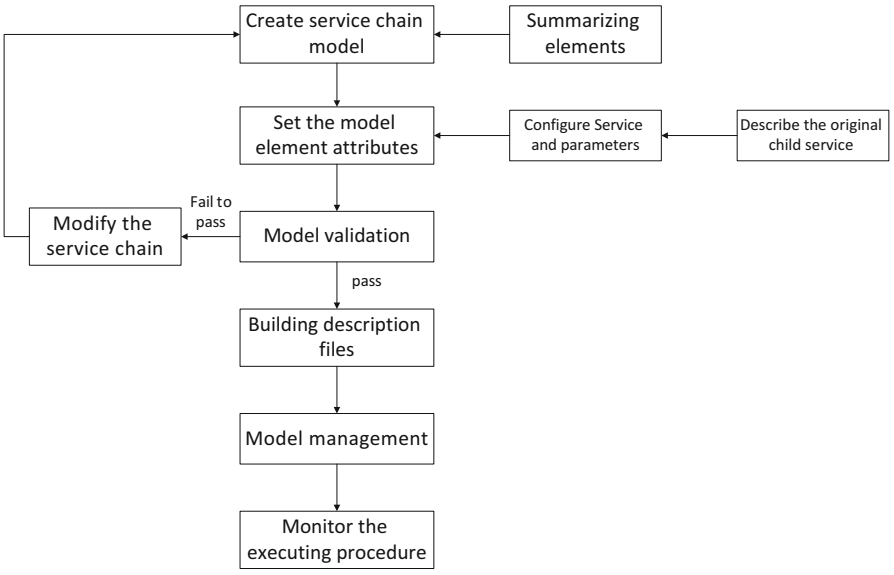
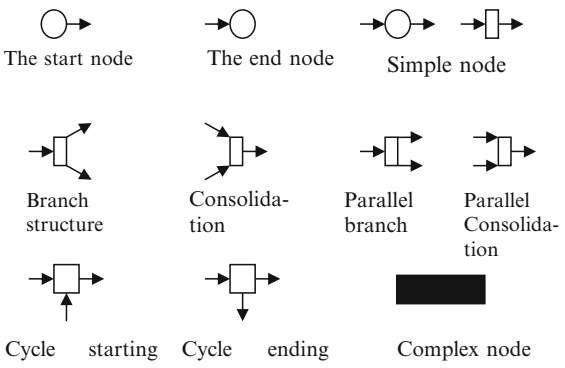


Fig. 3.2 Technical process

Fig. 3.3 Basic elements of service chain model



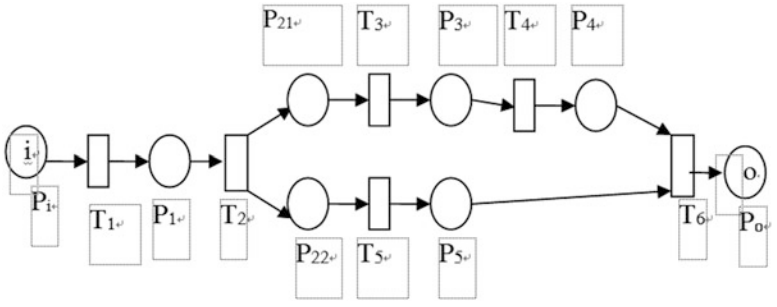


Fig. 3.4 The service chain model of ocean environment information query

system determines the follow-up service execution direction. P21 and P22 represent two different choice results of the users: graphical display and statistics show, respectively. T3 represents the service S_Dat_reduce. P3 represents that T3 correctly executes and maintains the data compression results. T4 represents the service S_Col_draw. P4 represents that T4 correctly executes and forms a graphical display of the results. T5 represents the S_Tab_stat services. P5 represents the correct completion of T5 and forms statistical tables to show the result. T6 represents the back-operation results. It is used to produce the display results that are returned to the user. Po represents that T6 properly completed and the entire process correctly terminates.

3.4.2 The Configuration of Model Element Attribute

After the construction of a process chart completes, the system must set the corresponding attribute for each node. The node attributes contain the operations or service name that correspond to the node. The system must output the necessary parameters to instantiate the service.

The attribute "property value" of the state nodes records the completion status information of a previous operation or service. There are two property values: 1 and 0, which represent complete and unfinished, respectively.

The attribute "property value" of service nodes records the necessary parameter information for the service instantiation. Different services have different number and content of parameters. Property information of the service nodes is mainly used to monitor or call the atomic services. Thus, the content of node information are consistent with the information of atomic services. According to Fig. 3.5, the atomic service model mainly consists of functional services, input, output, and

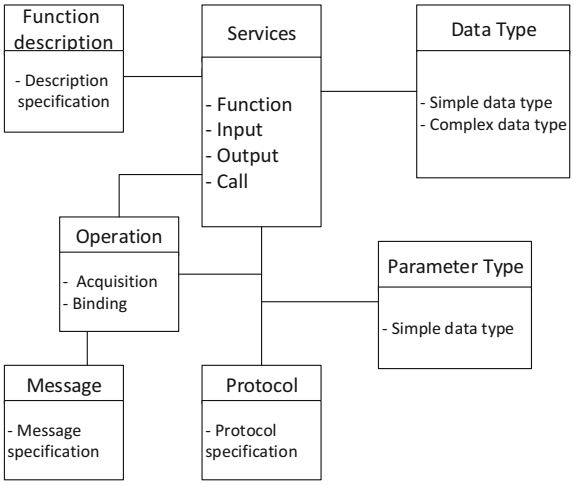


Fig. 3.5 Atomic service model

call information. The functional description corresponds to the described atomic services in the regulatory requirements. The input, output parameters and data are consistent with the type and data format definition. The call operations comply with the agreement of operator interface and information exchange.

The attribute "property value" of the operation node records the type of operation that must be performed by the user, such as the input section of information, mouse operation, etc.

3.4.3 The Validation of Service Chain

After setting the attribute information of the node in each service chain model, the system must validate the service chain model that is created by the user. The validation of the process model can be divided into two steps. First, match the parameters in the model building process. Second, verify the logic of the model after the modeling completes.

In the process of information query for the elements of the ocean environment function service chain model, parameter matching is the first step. The type and instantiation information of the node are accessible when the model node attribute information is set. The numbers of four atom service nodes in the service chain and related information are shown in Table 3.1 and Table 3.2. (The node numbers are shown in Fig. 3.6).

The identification of service instantiation can be used to query the description information of each atomic service. The system can obtain the input and output information of each service, as shown in the following table.

Table 3.1 Attribute information table of the atomic services node

Node	The identification of service instantiation	Number of input nodes	Parameter instantiation
2	S_Dat_search	1	SST,2008-04-04, (110.0,30.0,0,120.0,20.0,0)
6	S_Dat_reduce	1	100
8	S_Col_draw	1	red_blue
11	S_Tab_stat	1	2

Table 3.2 Input and output information table of the service design

Service identification	The input parameter number	Parameter type	Number of input data	Input data type	Output data type
S_Dat_search	3	string, time, float_arr[6]	1	1	1
S_Dat_reduce	1	int	1	1	1
S_Col_draw	1	string	1	1	1
S_Tab_stat	1	string	1	1	0

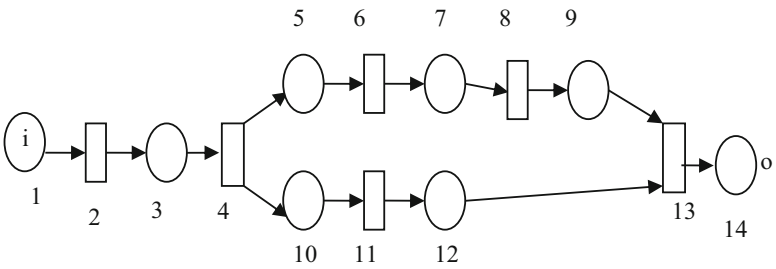


Fig. 3.6 The chain node number of ocean environment factor information query expression application service

Judge the above two tables according to four conditions of the parameter-matching information. If the four service nodes of the model simultaneously satisfy the four parameter-matching conditions, the system determines that the model parameter-matching test is passed.

1. Validation of the model logic. Model logic validation refers to the analysis of the logical structure of the model and the determination of whether the model is complete, bounded, retriggerable and executable. The reachability tree is used to validate the logical correctness of the model. Generate the ocean environment elements to express the function of the service chain model using the algorithm of a reachability tree. As shown in Fig. 3.7, the validation logic of the model includes the following steps:
2. Validation of model reachability. The reachable state set starts from M_0: R (M0) = {M0,M1,M2,M3,M4,M5,M6,M7,M8}. The process state set is MS ∈ {M0,M1,M2,M3,M4,M5,M6,M7,M8}. For any change Tn, its triggered state belongs to R (M_0), which is the process of any state that is reachable from M_0. The trigger state of any change Tn is Mn. If any Mn does not belong to R (M_0), this model cannot be reached.
3. Model boundedness validation. The reachability tree begins with a start trigger node and ends with an end trigger nodes. The tree begins with M_0 and ends with M5 and M8, and M5 = M8. All services and operations in the process are in the tree. The state of Mn belongs to R (M_0). If M5 < > M8, or Mn does not belong to R (M_0), this model is not bounded.
4. Validation of whether the model is retriggerable. As observed from the reachability tree, any element Tn from M_0 in T can be triggered from the beginning of the sequence of M_0. Thus, the model is retriggerable.
5. Validation of the executable model. In the reachability tree, if there is no meaningless cycle between any two states, the model can be executed.
6. Model int'egrity verification. In the reachability tree, if all states of a composite service from M_0 can reach and be triggered to terminate state M_S, the model is complete.

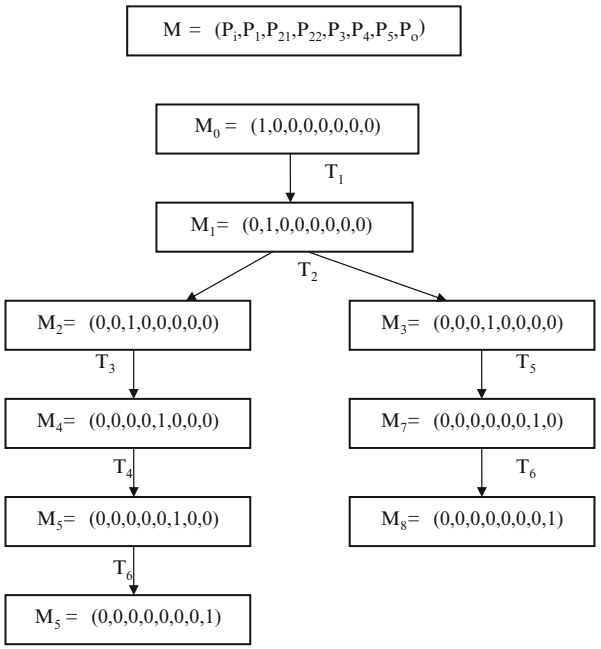


Fig. 3.7 The reachability tree of ocean environment factor information query expression application service chain

This service chain model does not contain complex nodes. The multiple-process model, which contains the complex node, should be validated by layer from top to bottom. First, validate the logic of the entire process of the top layer. Regard the complex node as a normal node. When the overall process is validated, judge the sub processes of the node. Repeat this process till the bottom layer of the process model. In the verification process, if there is any sub process is not logically correct, the entire process model is not logical. Prompt the user to correct.

3.4.4 The Construction of Model Description

By validating the service chain model, the system automatically builds the service chain model description file based on the attribute information. The system uses XML as the model description language.

3.4.5 The Model Management

When the service chain model description files are generated, these files contain all information of the service chain model. The system can reconstruct the service chain model based on these files. The essence of the service chain model management is the management of description files. The system will store, read, modify,

and delete the description files according to the user's operation. In the read operation, the system first analyzes the service chain model description files. Then, the model elements recur according to the recorded amount in the files and the corresponding relations among the node attribute information. The refactoring graphics show the service chain model. Repeat steps (1)–(4) when the model is updated. Replace the old description file by the new service chain model description files.

3.4.6 The Model Monitoring

Execution monitoring of the service chain model includes triggering the model, management of the model node lifecycle, management of the model node states, management of the model running state, etc. The user can launch and implement the function of the service chain by monitoring the model's execution and obtaining the state information of the service chain model.

Assuming that the attribute set of model node statuses is P = (Pi, P1, P2, P3, P4, P5,....., Pn,, Po). Pi indicates the model execution state of any time. In the execution process of the model, the execution state Pi has only one non-zero element. The positions of non-zero elements represent the progress of the current execution model. The value of non-zero elements represents the state of the model execution. For example, for the model instance, when R = (0,0,0,0,1,0,0,0), the user selects a graphical representation, and the dichroic drawing data (S_Col_draw) service process is implemented. If R = (0.0,0,0,2,0,0,0), the user selects a graphical representation, but dichroic drawing data (S_Col_draw) service process implement the error after the data compression (S_Dat_reduce) service. When R = (0,0,0,0,3,0,0,0), the error occurs in the data compression process ((S_Dat_reduce) service).

Steps (1)–(4) are the basic steps to dynamically build a service chain model. Each time a new ocean 3D monitoring data service chain model is built, steps (1)–(4) should be sequentially executed. Steps (5) and (6) are executed according to the user operation. When the user selects an operation, continue with the matching content based on the first four steps.

Compared with the existing technology, the scientific and advanced characteristics of the above method are as follows. (1) The construction to visualize the ocean service chain model provides the users with an intuitive and quick method to achieve the ocean application requirements. The user can use the existing ocean application function of atomic service or a combination to build a service chain that satisfies special requirements. (2) The abstract of ocean atomic services solves the automatic identification and management problems of the ocean service system. Services that conform to the definition of the ocean application service can be identified and applied by the system. (3) The abstract of basic elements of the ocean service chain model makes the ocean service chain information simultaneously understood and recognized by the users and computers, which is advantageous to the management and application of the service chain model.

Acknowledgments The study is funded by the Free Exploring Program of the State Key Laboratory of Remote Sensing Science of China (No. 14ZY-03), the Special Research Project for the Commonweal of the Ministry of Water Resources of the People's Republic of China (grant no. 201201092), the 908 Project of the State Oceanic Administration, China (No. 908-03-03-02), the National Natural Science Foundation of China (grant no. 61473286 and 61375002), and the International Science & Technology Cooperation Program of China (grant no. 2010DFA92720). All members of the Research Team are gratefully acknowledged for their contributions to the work carried out in this chapter in recent years. In particular, special thanks are given to Miss Wen Dong, Miss Qiong Zheng, Mr. Yongxin Chen and Mr. Yuqi Liu for their contributions in the document revision.

Chapter 4
Coastal Flood Forecasting Modeling and Analysis

Lei Wang and Xin Zhang

Abstract The mechanism of flood forecasting is a complex process, which involves precipitation, drainage-basin characteristics, land use/cover types, and runoff discharge. Because of the complexity of flood forecasting, hydrological models and statistical models need to be developed for flood frequency analysis, river runoff prediction, and flood forecasting. In this chapter, Soil Conservation Service (SCS) Curve Number (CN) model is applied for river runoff prediction in the Oak Ridges Moraine (ORM) area, southern Ontario. The historical data for the past several decades (river gauging, precipitation, ground water, census and land use) are used to model the relationship among the stream runoff, precipitation and hydrological-geographical features to apply SCS CN model for river runoff prediction.

Keywords Flood forecasting • Model

4.1 US-SCS Curve Number Method

The Soil Conservation Service (SCS) Curve Number (CN) (USDA-SCS 1972) method is widely used for estimating runoff (Viessman and Lewis 2003). The method was originally described in 1954 and been revised in 1956, 1964, 1965, 1971, 1972, 1985 and 1993 (Ponce and Hawkins 1996). The SCS method associates runoff (Q) and precipitation (P) in watershed context with a parameter CN that is influenced by land use type in the drainage basin. The SCS CN method is one of the

L. Wang (✉)
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

Hainan Province Key Laboratory of Earth Observation, Sanya Institute of Remote Sensing, Sanya 572029, Hainan Province, China
e-mail: wanglei98@radi.ac.cn

X. Zhang
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zhangxin@radi.ac.cn

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_4

most popular methods for computing runoff volume from a rainstorm, and it is probably one of the most widely used methods in the United States for estimating floods on rural and urban drainage basins (Maidment 1993)

Recent advances in computers and the availability of GIS tools have made a drastic change in hydrological modeling and development with the application of SCS-CN technique (Nageshwar et al. 1992; Arnold et al. 1993). Although the model has its own limitations (Ponce and Hawkins 1996; Choi et al. 2002; Mishra and Singh 2003), it has been widely used in numerous models (Geetha et al. 2008). And it has been approved applicable and effective for estimating direct runoff volume (Hawkins 1993; Steenhuis et al. 1995). Shi et al. (2007) used the SCS model for surface runoff and flood discharge simulation and found the relative error was 5–9 %. Liu and Li (2008) used the SCS model for different flood event runoff volume simulation and found that the relative error between estimated runoff and observed runoff ranged from 6.68 to 23.34 %, which is within the permissible limit.

Maidment (1993) described the derivation of the basic equation for estimating the volume or depth of runoff for accumulated volumes during a storm. This study repeats the derivation process here in order to clearly explain and understand the basic equation of the SCS CN model.

The general form of the relation is well established by both theory and observation. No runoff occurs until rainfall equals an initial abstraction Ia. After allowing for Ia, the depth of runoff Q is the residual after subtracting F, the infiltration or water retained in the drainage basin (excluding Ia) from the rainfall P. The potential retention S is the value that (F + Ia) would reach in a very long storm (Maidment 1993).

If Pe is the effective storm rainfall equal to (P−Ia), the basic assumption in the method is

$$\frac{F}{S} = \frac{Q}{P - I_a} \tag{4.1}$$

where Q — runoff depth (inch)
P — rainfall (inch)
S — water potential retention maximum of watershed (inch)
Ia — initial abstraction of rainfall by soil and vegetation (inch)
F — actual retention

After runoff starts, all excess rainfall becomes either runoff or actual retention. That is

$$F = P - I_a - Q \tag{4.2}$$

Replacing F in Eq. 4.1, then

$$Q = \frac{(P - I_a)^2}{(P - I_a + S)} \qquad (P > I_a) \qquad Q = 0 \qquad (P \le I_a) \qquad (4.3)$$

The empirical relation Ia = 0.2S was adopted as the best approximation from observed data, so

$$\begin{aligned} Q &= \left[\frac{(P - 0.2S)^2}{(p + 0.8S)}\right] \quad \text{for } P > 0.2S \\ Q &= 0 \qquad\qquad\qquad\quad \text{for } P \le 0.2S \end{aligned} \qquad (4.4)$$

For convenience and to standardize application of this equation, the potential retention S is expressed in the form of a dimensionless runoff Curve Number (CN), where

$$CN = \frac{1000}{S + 10} \qquad (4.5)$$

Eliminating S from Eqs. 4.4 and 4.5 gives the basis SCS relationship for estimating Q from P and CN, which has the advantage of having only one parameter.

From Eqs. 4.4 and 4.5, one can calculate CN from given P and Q as follows (Ko 2004):

$$CN = \frac{25400}{S + 254} = \frac{25400}{254 + \frac{(0.4P + 0.8Q) - \sqrt{(0.4P + 0.8Q)^2 - 4 \times 0.04\left(P^2 - PQ\right)}}{2 \times 0.04}} \qquad (4.6)$$

The value of CN depends on the soil, cover and hydrological condition of the land surface (Maidment 1993). CN also depends on the antecedent wetness of the drainage basin, and three classes of antecedent moisture condition (AMC) defined: dry, average, and wet (AMC I, AMC II, and AMC III). Selection of the runoff curve number is dependent on antecedent conditions and the types of land cover (Viessman and Lewis 2003).

4.2 Data Preprocessing

Modeling the rainfall-runoff process in the ORM area includes three steps: first, ArcHydro tools use DEM data to delineate the watershed boundary. Then the historical data for the past several decades (river gauging, precipitation, and land cover) are used to model the relationship among the stream runoff, precipitation and hydrological-geographical features to get the CN value for each watershed. Finally, CN value and precipitation are used for river runoff prediction.

4.2.1 Thiessen Polygon to Associate Runoff and Precipitation Stations

In order to model the rainfall-runoff process, stream gauging stations and weather stations need to be associated within the same watershed.

The selected weather stations and the stream gauging stations are mapped using ArcGIS (Fig. 4.1). There are 25 stream gauging stations within the study area, which record daily flow data and 131 weather stations within the study area, which record daily precipitation data.

Since there are more than one weather station located within the same watershed, in order to compare the time series runoff and precipitation, thiessen polygons are applied to associate the weather stations with the stream gauging stations within the same watershed boundary (Fig. 4.2). First, the thiessen polygon method divides the whole space into thiessen polygons by the perpendicular bisectors between those weather stations, and then these thiessen polygons are intersected with the watershed boundary. Finally, the percentage of area of each thiessen polygon within the watershed boundary is used as the weights to calculate the area-weighted average precipitation.

By combining these two layers, each stream gauging station has a pair of time series showing both the stream flow record and the precipitation record covering the same time period.

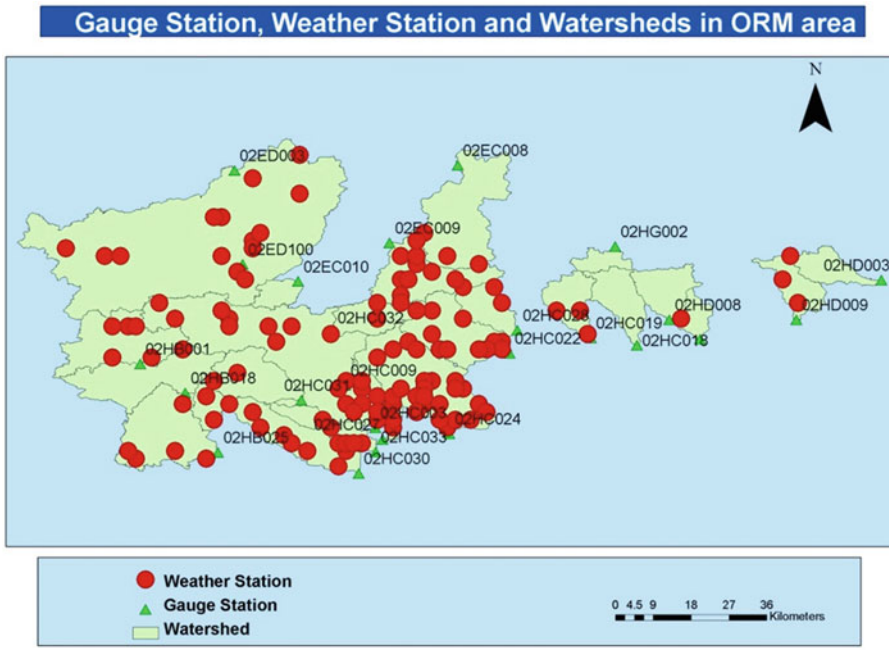


Fig. 4.1 Watersheds, stream gauge stations and precipitation stations in the ORM area

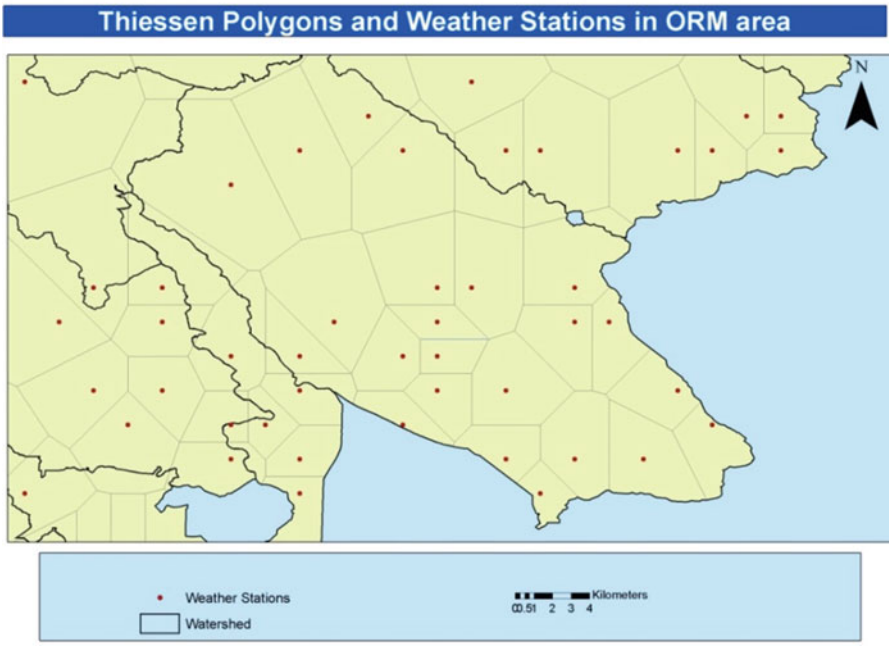


Fig. 4.2 Thiessen polygons and weather stations in the ORM area

The historical data for the past several decades (river gauging, precipitation, and land cover) are used to model the relationship among the stream runoff, precipitation and hydrological-geographical features to obtain the CN value for each watershed. The whole process includes three steps. First, base flow separation, since the SCS model is designed to model the relationship between direct flow and precipitation, so the first step is to study the method to separate base flow from runoff to get direct flow, and therefore each watershed has pairs of rainfall and direct flow depth, and these pairs of rainfall and direct flow depth can be used to calculate pairs of composite CN value. Second, the land cover classification map is interacted with the watershed boundary map to get the land cover percentage in each watershed. Then multiple regression is applied to those pairs of composite CN values and land cover percentage in each watershed to get the CN values for different land covers in the ORM area. Finally, these CN values are used to calculate the composite CN values for each watershed, and these new composite CN values are used for river runoff prediction.

Because the SCS CN method is designed for computing runoff volume from a rainstorm, in order to avoid the mixing of river flow due to rainfall and snow melt, only the flood events that occurred from May 1 to November 30 are used to model the relationship among the stream runoff, precipitation and hydrological-geographical features to apply the SCS CN model for river runoff prediction in the ORM area.

4.2.2 Base Flow Separation

Most of techniques used to separate a total runoff into direct surface runoff and base flow components are based on analysis of groundwater recession curves. A groundwater recession is characterized by a gradually decreasing rate of base flow. The recession curve shape has been found to approximate an exponential function (Viessman and Lewis 2003).

If there is no added inflow to the groundwater, and if all groundwater discharge from the upstream area is intercepted at the stream gauging station, then the groundwater discharge recession can be described by (Viessman and Lewis 2003):

$$Q_t = Q_0 e^{-\theta t} \tag{4.7}$$

where Q_0 — a specified initial discharge
Qt — the discharge at any time t after flow Q0
θ — the recession constant
e — base of the natural logarithm

Time units frequently used are days for large watersheds and hours for small basins.

The May 1974 flood event at the 02HB001 gauging station is used as an example for base flow recession and separation analysis. The May 1974 flood event started on May 14 and reached peak flow on May 17. After that, there was no rainfall until May 27. The river flow discharge and precipitation observed at 02HB001 gauging stations during this flood event are listed in Table 4.1.

We can apply log transform of river flow discharge for the May 1974 flood to study base flow recession (Fig. 4.3). Since the log value of the river flow discharge approximately fitted a straight line from May 20 to May 28, the direct flow was assumed to end on May 20, and the river flow was composed of base flow only from May 20 to May 28. The base flow recession coefficient was 0.088.

Using base flow recession Eq. (4.7), the base flow can be calculated as follows: Since the river flow was composed of base flow only on May 20 (3.74 m^3/s), the base flow for May 19 and May 18 can be calculated inversely by the base flow recession Eq. (4.7), with a recession coefficient of 0.088. Since the flood event started on May 14, the river flow on May 14 was also considered to be composed of base flow only, and the base flow from May 15 to May 17 can be calculated using a linear interpolation method. Then the direct flow can be calculated by separating base flow from runoff, and the direct flow depth can be calculated by dividing the total direct flow amount by the drainage basin area (205 km^2).

The results of base flow separation and direct flow depth obtained for the May 1974 flood event are listed in Table 4.2.

Similarly, pairs of direct flow depth and precipitation of each gauging station can be calculated for different flood events and these pairs of direct flow depth and precipitation are used for multiple regression analysis to get the CN value for different soil types (land covers) in the ORM area.

Table 4.1 River flow discharge and precipitation observed at 02HB001 gauging stations

Date	River flow discharge (m^3/t)	Precipitation (mm)
5-14-1974	2.940	11.2
5-15-1974	3.450	4.1
5-16-1974	4.530	36.3
5-17-1974	15.300	0
5-18-1974	9.120	0
5-19-1974	5.270	0
5-20-1974	3.740	0
5-21-1974	3.170	0
5-22-1974	2.830	1.5
5-23-1974	2.890	0
5-24-1974	2.550	0
5-25-1974	2.240	0
5-26-1974	2.000	0
5-27-1974	1.940	0.6
5-28-1974	1.830	3

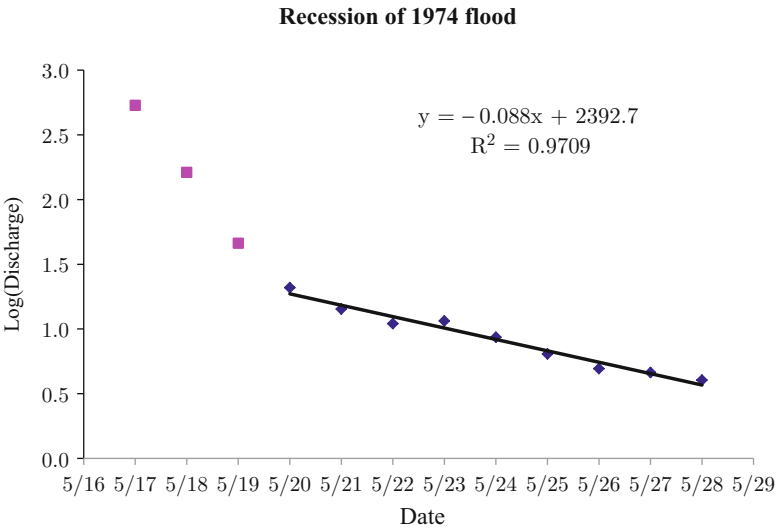


Fig. 4.3 Log transform of hydrograph of May 1974 flood

Table 4.2 Precipitation, direct flow depth for May 1974 flood event at gauging stations in the ORM area

Gauge station	Precipitation (mm)	Flow depth (mm)
02EC008	–	–
02EC009	31.9	8.1
02EC010	32.9	6.2
02ED003	30.5	6.8
02ED100	48.2	6.1
02HB001	53.1	7.6
02HB018	–	–
02HB025	–	–
02HC003	43.7	20
02HC009	33.6	9.3
02HC018	–	–
02HC019	25.4	5.5
02HC022	29.7	6.8
02HC024	41.6	20.3
02HC027	52.4	31.6
02HC028	31.8	13.2
02HC030	67.2	39.2
02HC031	–	–
02HC032	29	5.9
02HC033	60.3	31
02HD003	31.4	7.6
02HD008	36.5	8.8
02HD009	32.9	5.9
02HD013	–	–
02HG002	–	–

4.2.3 Antecedent Moisture Condition (AMC)

As described in Sect. 4.2.3, the Antecedent Moisture Condition (AMC) refers to the wetness of the soil surface or the amount of moisture in the soil (Mishra and Singh 2003). If the soil is fully saturated, the entire amount of rainfall will directly convent to runoff without infiltration losses and if the soil is fully dry, it is possible that there is no surface runoff because the whole rainfall amount is absorbed by the soil. The AMC has significant effect to the process of rainfall runoff (Mishra and Singh 2003).

AMC is based on the amount of antecedent rainfall, and it varies from previous 5 to 30 days. However, there is no explicit guideline available to define the soil moisture using the antecedent rainfall. The National Engineering Handbook (USDA-SCS 1968) uses the antecedent 5-day rainfall for AMC and it is generally used in practice (Mishra and Singh 2003). In this study, the antecedent 5-day rainfall for AMC is used for AMC analysis.

4.3 CN Value

The runoff estimation method used in this study is the SCS model developed by the US Soil Conservation Service. The SCS method associates Runoff (Q) and Precipitation (P) in the watershed context with a parameter CN that is influenced by land use types in the drainage basin. The SCS curve method integrates the contribution of precipitation and soil properties to estimate the storm runoff volume.

The CN values are the empirically derived values for different land types. CN values depend on the surface cover and soil moisture of a particular place. SCS recommend CN value for different antecedent conditions and the types of land cover. The suggested values could be found in National Engineering Handbook (USDA-SCS 1985). However, this recommended value may or may not fit for the ORM area.

4.4 Composite CN Value

Grove et al. (1998) suggested that the average CN for a watershed can be calculated in two ways: the "composite method" and the "distributed method." In the composite method, the CN values for each soil type are identified and the average CN for the watershed can be calculated as the area weighted average CN of each soil type occupying the watershed. Conversely, the distributed method identifies the different land uses and computers the runoff volume for each land use and the total runoff volume is the sum of all runoff volume for each land uses type. In this study, the composite method is used to calculate the average CN.

In this study, the National-Scale Ontario Land Cover Map (Ontario Ministry of Natural Resources 1999) was used to calculate the area weighted average CN of each watershed. This may have some impact to model simulation because the 1999 land cover may not represent the actual types of land cover when different flood events happened. However, because of data limitation, there are no other land cover maps available, so the National-Scale Ontario Land Cover Map is used as land cover map in this study. Figure 4.4 shows the land cover map, which shows a total of six land cover classes: water, forest, wetlands, agriculture, bare soil, and urban.

Geoprocessing (such as intersect) is then applied to get the area percentage of different land cover classes at each watershed. Table 4.3 shows the area percentage of different land cover classes at each watershed.

A composite curve number (CN) for a watershed having more than one land use or soil type can be obtained by weighting each curve number based on its area percentage. Although there are six land cover classes at the National-Scale Ontario Land Cover Map, the area percentage of water body, grass land and bare soil are

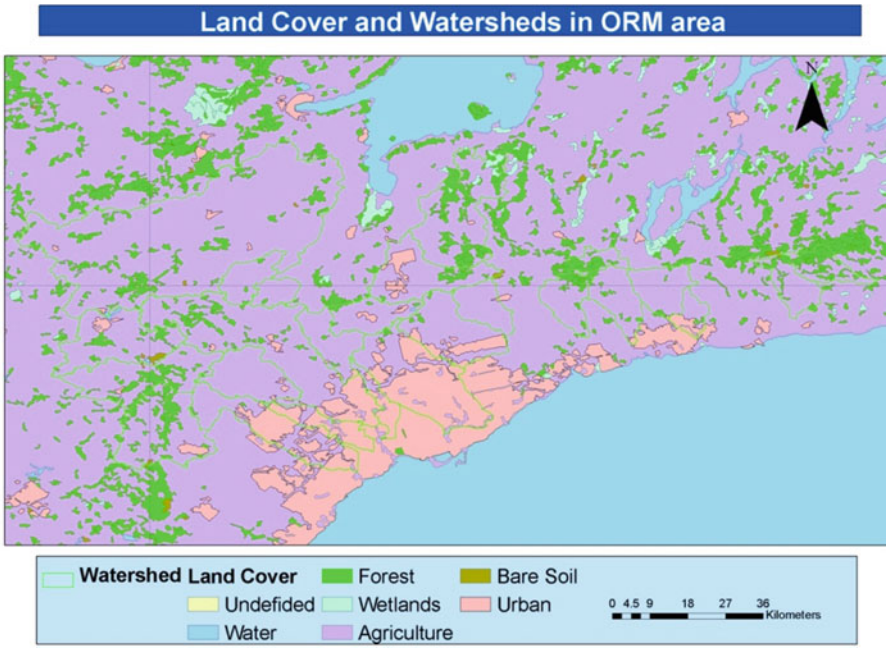


Fig. 4.4 Land cover and watersheds in the ORM area

very small for the study area watershed, so the composite CN for each watershed in the ORM area can be calculated as:

$$\begin{aligned}\text{Composite CN} = {} & \%\,\text{area of forest} * \text{CN forest} + \%\,\text{area of agriculture} * \\ & \text{CN agriculture} + \%\,\text{area of urban} * \text{CN urban}\end{aligned} \tag{4.8}$$

The SCS recommended CN value for different antecedent conditions, and the types of land cover from the National Engineering Handbook (USDA-SCS 1985) might need to be adjusted to fit the ORM area. The CN values involved in the model are determined by using multiple regression method.

The whole process includes three steps: first, pairs of direct flow and precipitation of each watershed can be calculated for different flood events and these pairs of direct flow and precipitation can be used to calculate the composite CN values for different flood events and watersheds using Eq. (4.6). Then these pairs of calculated composite CN values are classified into three classes: AMC I, AMC II, AMC III. Among them AMC II is used as case study to illustrate the CN values calculation process for different land cover type. Finally, multiple regression is applied to get the CN values for different land cover types including agriculture, forest and urban.

Table 4.4 lists the selected flood events in southern or central Ontario from May to November. However, not all flood events have enough rainfall-runoff data

Table 4.3 Area percentage of different land cover classes at each watershed in the ORM area

Watershed	Water body	Forest	Grass land	Agriculture	Bare soil	Urban
02EC008		25.93	2.55	71.20		0.32
02EC009		14.29		69.60	2.55	13.56
02EC010		1.86		98.14		0.00
02ED003		4.66		95.17		0.18
02ED100		3.15		96.85		0.00
02HB001	0.68	13.27		80.81	5.23	0.00
02HB018		20.81		78.05	0.33	0.80
02HB025		25.00		71.75		3.25
02HC003		0.00		53.33		46.67
02HC009		1.64		90.17		8.18
02HC018		4.69		85.88		9.43
02HC019		24.95		75.05		0.00
02HC022		1.64		75.07		23.29
02HC024		0.00		8.47		91.53
02HC027		0.58		21.23		78.19
02HC028		0.00		99.43		0.57
02HC030		0.00		29.16		70.84
02HC031		0.37		96.53		3.09
02HC032	0.51	4.68		87.47		7.35
02HC033		0.00		25.13		74.87
02HD003		45.72		54.28		0.00
02HD008		2.27		96.83		0.90
02HD009		14.61		85.39		0.00
02HD013		3.06		60.75		36.18
02HG002		37.03		62.97		0.00

Table 4.4 Selected flood events in southern or central Ontario from May to November

Rec	Year	Month	Day
1	1954	10	15
2	1956	08	30
3	1960	05	09
4	1974	05	17
5	1979	05	12
6	1992	07	31
7	1992	08	28
8	2002	05	13

Source: Canadian Disaster Database

available for modeling analysis. For example, for 1954, 1956 and 1960 flood events, only 5 gauging stations have rainfall-runoff data available; for May 1979 and July 1992 flood events, only 12 gauging stations have rainfall-runoff data available. So May 1974, Aug. 1992 and May 2002 flood events are selected to

Table 4.5 Composite CN and AMC types for May 2002 flood event

Watershed	Composite CN	Total 5-day antecedent rainfall (mm)	AMC type
02EC008	–	–	–
02EC009	69.5	5.6	AMCI
02EC010	66.33	17	AMCII
02ED003	83.47	29	AMCIII
02ED100	65.79	13.9	AMCII
02HB001	66.81	28.1	AMCIII
02HB018	68.35	22	AMCII
02HB025	67.72	19	AMCII
02HC003	77.04	22	AMCII
02HC009	71.09	26	AMCII
02HC018	70.57	15.6	AMCII
02HC019	67.62	14.9	AMCII
02HC022	84.94	15	AMCII
02HC024	81.45	19.2	AMCII
02HC027	81.67	17.7	AMCII
02HC028	83.49	23.8	AMCII
02HC030	84.33	17	AMCII
02HC031	81.12	22	AMCII
02HC032	–	–	–
02HC033	88.71	18	AMCII
02HD003	59.08	14	AMCII
02HD008	66.51	15.3	AMCII
02HD009	54.92	14.2	AMCII
02HD013	80.7	14	AMCII
02HG002	–	–	–

illustrate modeling process, these three flood events have 18, 21 and 22 gauging stations having rainfall-runoff data available for modeling analysis. May 2002 flood event is selected to apply multiple regression to get the CN values for different land cover types because the land cover map is developed in 1999, and May 1974 and Aug. 1992 flood events are used to verify the suitability of the CN values.

Table 4.5 is the composite CN values and AMC types for May 2002 flood event. Among them, those composite CN values that belong to AMC II type are selected for multiple regression analysis.

Since composite CN values can be consider as the area weighted average of CN values, so composite CN values can be expressed as a linear combination of land cover types and area percentage as Eq. (4.8). Using these paired composite CN values for different flood events and watersheds, the multiple regression can be applied to get the CN values for different land cover types including agriculture, forest and urban. Table 4.6 shows the results of multiple regression CN values.

Table 4.6 Results of multiple regression CN values

	Unstandardized coefficients		Standardized coefficients	
	B	Std. error	Beta	t
Forest	40.569	11.796	0.080	3.439
Agriculture	72.814	2.411	0.726	30.197
Urban	87.564	3.892	0.463	22.501

Table 4.7 Results of multiple regression CN values (second times)

	Unstandardized coefficients		Standardized coefficients	
	B	Std. error	Beta	t
Agriculture	69.628	2.339	0.736	29.773
Urban	88.818	4.412	0.498	20.130

For the multiple regression, R2 is 0.994 and the CN values of forest, agriculture and urban for AMC II are 41, 73 and 88 separately. The corresponding CN values for AMC I and AMC III can be obtained from "CN for Wet and Dry Antecedent Moisture Conditions Corresponding to an Average Antecedent Moisture Conditions" Table from (Viessman and Lewis 2003).

However, when apply these CN values for flood simulation, the relative error is still relatively large (the average relative error is near 26 % for May 1974 flood event and near 45 % for August 1992 flood event), and further investigation found that the CN value for forest is smaller than the recommended value (70) from SCS, this might because the area percentage of forest is small in most watersheds, which cause the contribution of forest for the regression to be relatively small. Thus the multiple regression needs to be re-applied again to modify the CN values for different land cover types. This time the recommended CN value for forest was not included in the multiple regression, only the CN values for agriculture and urban were included in the regression analysis. Table 4.7 shows the results of multiple regression CN values.

For the multiple regression, R2 is 0.990 and the CN values of forest, agriculture and urban for AMC II are 70, 70 and 89 separately.

Based on above two multiple regressions, the CN values of agriculture, forest and urban for AMC II were determined to be 70, 70 and 89 separately. The corresponding CN values for AMC I and AMC III can be obtained from "CN for Wet and Dry Antecedent Moisture Conditions Corresponding to an Average Antecedent Moisture Conditions Table" from (Viessman and Lewis 2003).

Table 4.8 shows the area percentage of different land cover clases and corresponding CN value for different Antecedent Moisture Condition (AMC).

After weighted area calculation, the composite CN value at each watershed is listed in Table 4.9.

Table 4.8 CN value for each watershed in the ORM area

Watersheds	Land use	Area percentage (%)	AMC I	AMC II	AMC III
02ec008	Forest	25.93	51	70	85
	Agriculture	71.20	51	70	85
	Urban	0.32	76	90	96
02ec009	Forest	14.29	51	70	85
	Agriculture	69.60	51	70	85
	Urban	13.56	76	90	96
02ec010	Forest	1.86	51	70	85
	Agriculture	98.14	51	70	85
	Urban		76	90	96
02ed003	Forest	4.66	51	70	85
	Agriculture	95.17	51	70	85
	Urban	0.18	76	90	96
02ed100	Forest	3.15	51	70	85
	Agriculture	96.85	51	70	85
	Urban		76	90	96
02hb001	Forest	13.27	51	70	85
	Agriculture	80.81	51	70	85
	Urban		76	90	96
02hb018	Forest	20.81	51	70	85
	Agriculture	78.05	51	70	85
	Urban	0.80	76	90	96
02hb025	Forest	25.00	51	70	85
	Agriculture	71.75	51	70	85
	Urban	3.25	76	90	96
02hc003	Forest		51	70	85
	Agriculture	53.33	51	70	85
	Urban	46.67	76	90	96
02hc009	Forest	1.64	51	70	85
	Agriculture	90.17	51	70	85
	Urban	8.18	76	90	96
02hc018	Forest	4.69	51	70	85
	Agriculture	85.88	51	70	85
	Urban	9.43	76	90	96
02hc019	Forest	24.95	51	70	85
	Agriculture	75.05	51	70	85
	Urban		76	90	96
02hc022	Forest	1.64	51	70	85
	Agriculture	75.07	51	70	85
	Urban	23.29	76	90	96

(continued)

Table 4.8 (continued)

Watersheds	Land use	Area percentage (%)	AMC I	AMC II	AMC III
02hc024	Forest		51	70	85
	Agriculture	8.47	51	70	85
	Urban	91.53	76	90	96
02hc027	Forest	0.58	51	70	85
	Agriculture	21.23	51	70	85
	Urban	78.19	76	90	96
02hc028	Forest		51	70	85
	Agriculture	99.43	51	70	85
	Urban	0.57	76	90	96
02hc030	Forest		51	70	85
	Agriculture	29.16	51	70	85
	Urban	70.84	76	90	96
02hc031	Forest	0.37	51	70	85
	Agriculture	96.53	51	70	85
	Urban	3.09	76	90	96
02hc032	Forest	4.68	51	70	85
	Agriculture	87.47	51	70	85
	Urban	7.35	76	90	96
02hc033	Forest		51	70	85
	Agriculture	25.13	51	70	85
	Urban	74.87	76	90	96
02hd003	Forest	45.72	51	70	85
	Agriculture	54.28	51	70	85
	Urban		76	90	96
02hd008	Forest	2.27	51	70	85
	Agriculture	96.83	51	70	85
	Urban	0.90	76	90	96
02hd009	Forest	14.61	51	70	85
	Agriculture	85.39	51	70	85
	Urban		76	90	96
02hd013	Forest	3.06	51	70	85
	Agriculture	60.75	51	70	85
	Urban	36.18	76	90	96
02hg002	Forest	37.03	51	70	85
	Agriculture	62.97	51	70	85
	Urban		76	90	96

Table 4.9 Composite CN value at each watershed in the ORM area

Watersheds	AMC I	AMC II	AMC III
02EC008	51	70	85
02EC009	55	73	86
02EC010	51	70	85
02ED003	51	70	85
02ED100	51	70	85
02HB001	52	70	85
02HB018	51	70	85
02HB025	52	71	85
02HC003	64	79	90
02HC009	53	72	86
02HC018	54	72	86
02HC019	51	70	85
02HC022	57	75	88
02HC024	76	88	95
02HC027	72	86	94
02HC028	51	70	85
02HC030	70	84	93
02HC031	52	71	85
02HC032	53	72	86
02HC033	71	85	93
02HD003	51	70	85
02HD008	51	70	85
02HD009	51	70	85
02HD013	61	77	89
02HG002	51	70	85

4.5 Runoff Simulation Using SCS CN Model in the ORM Area

The May 1954 and August 1992 flood events are simulated using the SCS CN model, and the results are shown in Tables 4.10 and 4.11 separately. Generally, it shows an average of 70 % accuracy for the flood simulation and 30 % of relative error (the average relative error is 17 % for May 1974 flood event and near 37 % for 1992 flood event). The reasons for the relative error may include: First, the accuracy of land cover map, it is a National Scale Land Cover Map of 1999 derived from a more detailed Provincial-Scale Ontario Land Cover data base by combining and redefining the original 28 classes to form 15 classes and by generalizing the original spatial resolution from 25 to 100 meters. Discrete features less than 50 hectares in size were eliminated, and the Ontario land cover classification reflects the nature of the land surface rather than the land use (Ontario Ministry of Natural Resources 1999). This land cover map may not represent the actual types of land cover. Second, some watersheds have a large area, however, the SCS CN model is

Table 4.10 Runoff simulation results for May 1974 flood in the ORM area

Watersheds	Event precipitation (mm)	AMC type	Calculated direct flow (mm)	Measured direct flow (Mm)	Relative error (%)
02EC008	–	–	–	–	–
02EC009	31.9	AMCIII	9	8.1	11.4
02EC010	32.9	AMCIII	8.3	6.2	35.5
02ED003	30.5	AMCIII	7	6.8	3.4
02ED100	48.2	AMCII	5.2	6.1	−15.4
02HB001	53.1	AMCII	7	7.6	−8.4
02HB018	–	–	–	–	–
02HB025	–	–	–	–	–
02HC003	43.7	AMCIII	22.1	20	10.3
02HC009	33.6	AMCIII	9.6	9.3	2.9
02HC018	–	–	–	–	–
02HC019	25.4	AMCIII	4.4	5.5	−19.8
02HC022	29.7	AMCIII	8.6	6.8	26.9
02HC024	41.6	AMCII	17.8	20.3	−12.6
02HC027	52.4	AMCIII	36.1	31.6	16.3
02HC028	31.8	AMCIII	7.8	13.2	−41
02HC030	67.2	AMCIII	48.2	39.2	23.1
02HC031	–	–	–	–	–
02HC032	29	AMCIII	6.8	5.9	16.8
02HC033	60.3	AMCII	30.5	31	−1.4
02HD003	31.4	AMCIII	7.5	7.6	−1.4
02HD008	36.5	AMCIII	10.6	8.8	20.1
02HD009	32.9	AMCIII	8.3	5.9	41.2
02HD013	–	–	–	–	–
02HG002	–	–	–	–	–

designed for small-area watersheds, which have a relatively simple land cover type. The large-area watersheds may not fit for the SCS model for rainfall-runoff modeling. These watersheds need to be separated into several small watersheds for SCS modeling analysis.

Similarly, the simulation results for August 1992 flood event are shown in Table 4.11.

4.6 Water Level Prediction

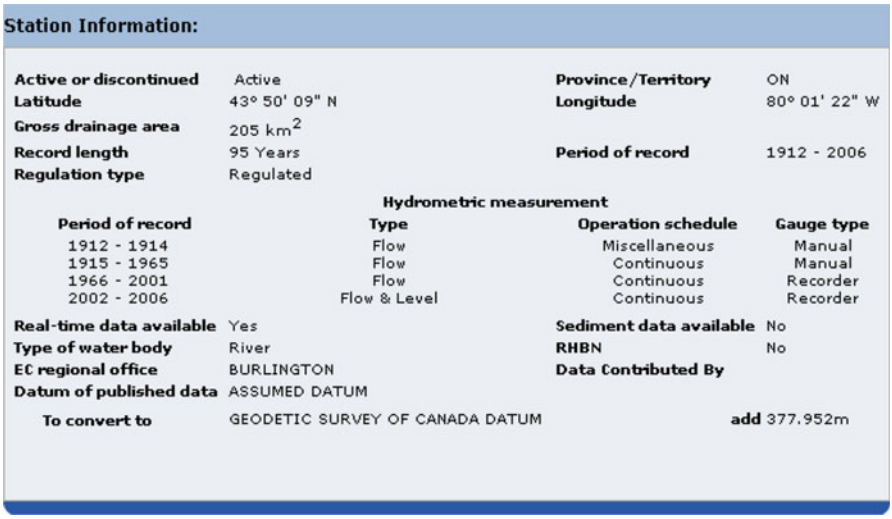
Water level is important for flood forecasting. Water Survey Canada (WSC) measures the water level (called stage) and uses a "Stage–Discharge Curves" (rating curve) to translate the stage into discharge for each gauging station. For

Table 4.11 Runoff simulation results for August 1992 flood in the ORM area

Watersheds	Event precipitation (mm)	AMC type	Calculated direct flow (mm)	Measured direct flow (mm)	Relative error (%)
02EC008	80.7	AMCI	3.8	2.6	−42.8
02EC009	64.1	AMCII	14.3	20.0	−28.4
02EC010	–	–	–	–	–
02ED003	47.3	AMCI	0.3	1.5	−78.4
02ED100	–	–	–		
02HB001	69.7	AMCI	0.8	3.8	−78.5
02HB018	69	AMCI	1.6	5.7	−71.5
02HB025	68.4	AMCII	1.8	4.0	−55.9
02HC003	61.6	AMCII	20.4	20	2.2
02HC009	65.4	AMCII	14.1	8.6	63.5
02HC018	67	AMCII	2.2	5.4	−60
02HC019	69.2	AMCII	14.4	8.6	67.4
02HC022	–	–	–	–	–
02HC024	63.5	AMCI	17.3	18.0	−3.8
02HC027	67.7	AMCII	34.4	29.3	17.4
02HC028	67.1	AMCII	13.4	16.9	−20.5
02HC030	65.7	AMCI	12.8	20.8	−38.7
02HC031	58.8	AMCII	9.9	13.1	−24.4
02HC032	51.9	AMCII	7.6	6.1	24.9
02HC033	66.2	AMCII	32.1	30.5	−5.1
02HD003	75.5	AMCI	3.3	5.3	−37.3
02HD008	65.5	AMCI	1.1	2.1	−46.8
02HD009	80	AMCI	4.2	3.3	28
02HD013	65.9	AMCI	5.6	6.8	−18.2
02HG002	–	–	–	–	–

flood forecasting purposes, the rating curve can also be used to translate the runoff prediction to water level prediction.

WSC supplies both water level and discharge data from 2002 to 2006 (Fig. 4.5). Before 2002, only flow discharge data was available, not water level data. The flow discharge and water level data from 2002 to 2006 can be used to model the relationship between water level and discharge to get a water level (stage) discharge curve, and this curve can be used to calculate the discharge from the water level or calculate water level from the discharge.



Station Information:

Active or discontinued	Active	**Province/Territory**	ON
Latitude	43° 50' 09" N	**Longitude**	80° 01' 22" W
Gross drainage area	205 km^2		
Record length	95 Years	**Period of record**	1912 - 2006
Regulation type	Regulated		

Hydrometric measurement

Period of record	Type	Operation schedule	Gauge type
1912 - 1914	Flow	Miscellaneous	Manual
1915 - 1965	Flow	Continuous	Manual
1966 - 2001	Flow	Continuous	Recorder
2002 - 2006	Flow & Level	Continuous	Recorder

Real-time data available	Yes	**Sediment data available**	No
Type of water body	River	**RHBN**	No
EC regional office	BURLINGTON	**Data Contributed By**	
Datum of published data	ASSUMED DATUM		
To convert to	GEODETIC SURVEY OF CANADA DATUM	**add**	377.952m

Fig. 4.5 Station information of O2HB001 gauging station (Environment Canada)

4.6.1 Stage Discharge Relation

"Stage-discharge relationship" is the relation between the water level (stage) and the discharge at a gauging station. Discharges in rivers are typically estimated by combining water level records with a functional relation, or suite of relations, describing variations in measured discharges with changing water levels. The functional relation between water level (or stage) and discharge is known as a stage-discharge curve, or rating curve (DeGagne et al. 1996). A series of national and international standards (ISO, 1981, 1982, 1983) have established procedures for measuring stage and discharge, as well as the development of stage-discharge relations.

A thorough understanding of the relationship between river stage and discharge is essential because one of the basic responsibilities of a hydrometric technician is to collect and compute daily discharge data for publication purposes. A stage-discharge relationship needs to be supported by real data. The more data points used to develop a graph, the better. Periodic checks of the discharge curve should be made after periods of flooding. The curve should be recalibrated if the periodic checks indicate the relationship has changed. Eventually, natural changes in the stream bottom will result in a change in the relationship between flow and gage height (Environment Canada 1999).

4.6.2 Stage–Discharge Curves

Stream gauging stations are used to measure the height (stage) and discharge in a stream. After a sufficient period of record, a rating curve can be derived, and subsequent discharge estimated given the river's stage.

Normally, a statistically based modeling method is used to model Stage–Discharge Curves. A major advantage of using statistically based modeling over conventional graphical methods of curve fitting is the ability to specify levels of accuracy to the curves, and the discharge estimate. A linear regression analysis is the modeling technique used to create the stage discharge relationship (DeGagne et al. 1996).

The stage-discharge relation may be expressed by an equation of the form (Herschy 1985):

$$Q = K(H + a)\, n \tag{4.9}$$

which is the equation of parabola where Q is the discharge in cubic meters per second

K is a constant
H is the gauge height
"a" is the gauge height at which discharge is zero

And n is an exponent.

Equation (4.9) can be transformed by logarithms to:

$$\text{Log } Q = \log K + n \log(H + a) \tag{4.10}$$

which is equivalent to the equation of a straight line $y = mx + c$, where $y = \log Q$, $c = \log K$ and $x = \log (H + a)$. Since "a" cannot be measured accurately, it must be determined by various numerical methods. With "a" determined, least squares regression is used to estimate K and n.

Plotting Q against H on log-log paper when "a" is not zero will produce a curve. If the value of "a" is known, Q can be plotted against $(H + a)$ and obtain a straight line result. So in order to get "a," various values of "a" are assumed until values of Q against $(H + a)$ gives a straight line on log-log paper.

When a quantity has to be added to the gauge heights in order to obtain a straight line, "a" is taken as positive, that means, the zero of the gauge is positioned at a level above the point of zero heights. Conversely, when a quantity has to be subtracted from the gauge heights, "a" is taken as negative and, that means the zero of the gauge is positioned at a level below the point of zero heights (Herschy 1985).

In order to get the Stage–Discharge Curves for each basin, the discharge and water level data (2002–2005) are extracted for each gauging station, these measurements are plotted on a long-log paper, and then various values of "a" are

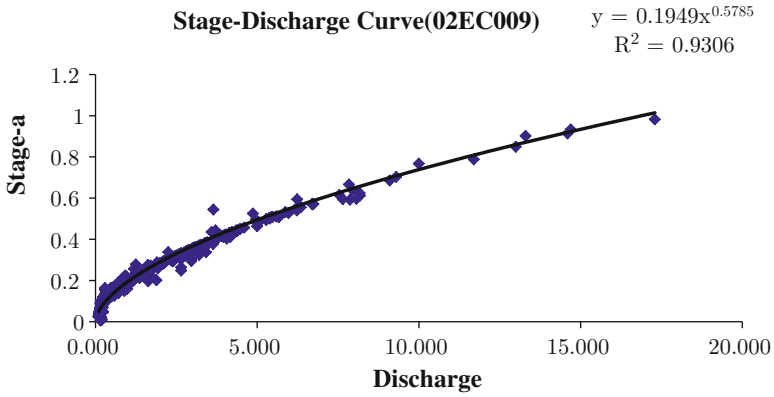


Fig. 4.6 Stage-discharge curve at O2EC009 station

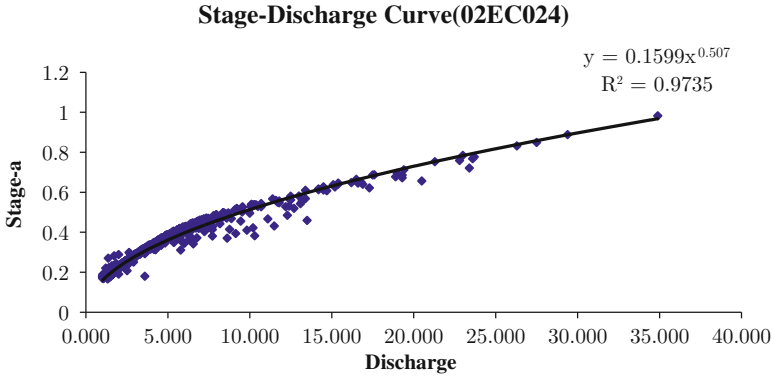


Fig. 4.7 Stage-discharge curve at 02HC024 station

assumed until values of Q against (H + a) give a straight line on log-log paper to determine the best fitting curve to model the Stage–Discharge relationship.

Figures 4.6 and 4.7 show the Stage-Discharge curves at 02EC009 and 02HC024 gauging stations by using the multi-regression method. Using Stage–Discharge Curves, the discharge can be interpolated on the Stage–Discharge Curves for water level prediction.

Rating curve need to be extrapolated when a water level is recorded below the lowest or above the highest gauged level (Maidment 1993). Table 4.12 shows the extrapolation of current rating curves used by WSC at 02EC008 station. And Figs. 4.8 and 4.9 show the extrapolation of the rating curve at 02ED008 and 02EC009 stations separately.

Table 4.12 Working table of current rating curves for 02EC008 station

Gauge height (m)	Discharge (m^3/s)	Gauge height (m)	Discharge (m^3/s)	Gauge height (m)	Discharge (m^3/s)	Gauge height (m)	Discharge (m^3/s)
8.1	0	8.25	0.484	8.5	1.7	9.6	11.2
8.11	0.022	8.26	0.522	8.55	1.98	9.7	12.4
8.12	0.048	8.27	0.561	8.6	2.28	9.8	13.6
8.13	0.075	8.28	0.6	8.65	2.58	9.9	14.8
8.14	0.103	8.29	0.64	8.7	2.89	10	16.1
8.15	0.132	8.3	0.682	8.75	3.2	10.1	17.4
8.16	0.163	8.32	0.77	8.8	3.52	10.2	18.8
8.17	0.196	8.34	0.862	8.85	3.85	10.3	20.3
8.18	0.23	8.36	0.96	8.9	4.2	10.4	21.9
8.19	0.265	8.38	1.06	8.95	4.58	10.5	23.5
8.2	0.3	8.4	1.16	9	4.99	10.6	25.2
8.21	0.336	8.42	1.26	9.1	5.85	10.7	26.9
8.22	0.372	8.44	1.37	9.2	6.82	10.8	28.7
8.23	0.409	8.46	1.48	9.3	7.8	10.9	30.6
8.24	0.446	8.48	1.59	9.5	10	11	32.7

Source: WSC

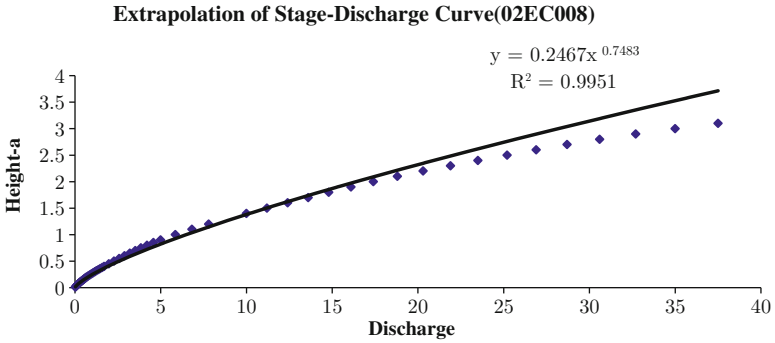


Fig. 4.8 Extrapolation of stage-discharge curve at 02EC008 station

4.7 Summary

The historical data for the past several decades (river gauging, precipitation, ground water, census and land use) are used to model the relationship among the stream runoff, precipitation and hydrological-geographical features to develop a hydrological model for river runoff prediction. The whole process includes three steps. First, base flow separation, since the SCS model is designed to model the relationship between direct flow and precipitation, so the first step is to study the method to

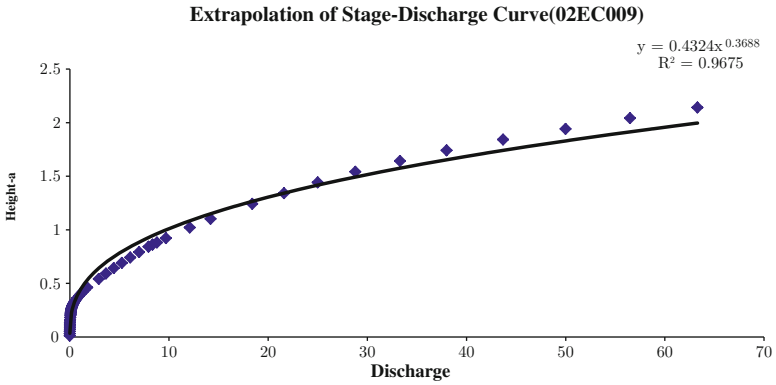


Fig. 4.9 Extrapolation of stage-discharge curve at 02EC009 station

separate base flow from runoff to get direct flow, and therefore each watershed has pairs of rainfall and direct flow depth, and these pairs of rainfall and direct flow depth can be used to calculate pairs of composite CN value. Second, the land cover classification map is interacted with the watershed boundary map to get the land cover percentage in each watershed. Then multiple regression is applied to those pairs of composite CN values and land cover percentage in each watershed to get the CN values for different land covers in the ORM area. Finally, these CN values are used to calculate the composite CN values for each watershed, and these new composite CN values are used for river runoff prediction.

Acknowledgments The author expresses the appreciation of funds received from Hainan peovince major science and technology projects (#ZDKJ2016015-1), the Hainan Province Natural Science Foundation (#20164178), the Sanya City Key Laboratory projects(#L1404) and the Sanya City science and technology cooperation projects (#2015YD19 and #2014YD08).

References

Arnold JG, Engel BA, Srinivasan R (1993) Continuous time, grid cell watershed model, application of advanced information technologies. Effective management of natural resources. ASAE Publication, 04–93. American Society of Agricultural Engineers, 267–278

Choi JY, Engel BA, Chung HW (2002) Daily streamflow modeling and assessment based on the curve number technique. Hydrol Process 16:3131–3150

DeGagne MPJ, Douglas GG, Hudson HR, Simonovic SP (1996) A decision support system for the analysis and use of stage-discharge rating curves. J Hydrol 184:225–241

Environment Canada (1999) Stage discharge relation. http://www.wsc.ec.gc.ca/CDP/Lesson18/index_ie_e.htm. Accessed on 01 Dec 2008

Geetha K, Mishra SK, Eldho TI, Rastogi AK, Pandey RP (2008) SCS-CN-based continuous simulation model for hydrologic forecasting. Water Resour Manag 22:165–190

Grove M, Harbor J, Engle B (1998) Composite vs. distributed curve numbers: effects on estimates of storm runoff depths. J Am Water Resour Assoc 34(5):1015–1023

Hawkins RH (1993) Asymptotic determination of runoff curve number from data. J Irrig Drain Eng ASCE 119(2):334–345
Herschy RW (1985) Streamflow measurement. Elsevier Applied Science Publishers, London
Ko C (2004) Storm runoff volume estimation in the Oak Ridges Moraine area, using GIS and remote sensing techniques. Unpublished M.Sc. thesis, York University
Liu XZ, Li JZ (2008) Application of SCS model in estimation of runoff from small watershed in Loess Plateau of China. Chin Geogr Sci 18(3):235–241
Maidment DR (1993) Handbook of hydrology. McGraw-Hill, New York
Mishra SK, Singh VP (2003) Soil conservation service curve number (SCS-CN) methodology. Kluwer Academic Publishers, Dordrecht/Boston
Nageshwar RB, Wesley PJ, Ravikumar SD (1992) Hydrologic parameter estimation using geographic information system. J Water Resour Plan Manag 118(5):492–512
Ontario Ministry of Natural Resources (1999) National-scale Ontario land cover map. http://open.canada.ca/data/en/dataset/fb44cd63-deb3-5efb-a07c-818e36db4c89. Accessed on July 2016
Ponce VM, Hawkins RH (1996) Runoff curve number: has it reached maturity? J Hydrol Eng 1(1):11–19
Shi PJ, Yuan Y, Zheng J, Wang JA, Ge Y, Qiu GY (2007) The effect of land use/cover change on surface runoff in Shenzhen region, China. Catena 69:31–35
Steenhuis TS, Winchell M, Rossing J, Zollweg J, Walter MF (1995) SCS runoff equation revisited for variable-source runoff area. J Irrig Drain Eng ASCE 121(3):234–238
USDA-SCS (1968) National engineering handbook. U.S. Department of Agriculture, Washington, DC
USDA-SCS (1972) National engineering handbook. Washington, DC, USDA-SCS
USDA-SCS (1985) National engineering handbook, section 4 – hydrology. USDA-SCS, Washington DC
Viessman W, Lewis GL (2003) Introduction to hydrology. Pearson Education Inc., New York

Chapter 5
Coastal Flood Frequency Modeling

Lei Wang and Xin Zhang

Abstract Because of the complexity of flood forecasting, hydrological models and statistical models need to be developed for flood frequency analysis, and flood forecasting. In this chapter, the performance of LP3, GEV, PL and GP models was investigated for flood frequency analysis to find the most suitable flood frequency analysis method in the Oak Ridges Moraine (ORM) area, southern Ontario. Historical data of river runoff peak discharge for the past several decades are used to model the relationship between flood discharge and return period by fitting a theoretical statistical distribution for each gauging station. The correlation coefficients of observations and model fitted lines are compared to evaluate the performance of each flood frequency analysis method. The comparison of different flood frequency analysis methods suggest that there are no significant differences of the LP3, GEV, PL and GP models for less than approximate 10-year return period flood frequency analysis in the ORM area. And there are no significant differences of the LP3, GEV, PL and GP models for greater than approximate 10-year return period flood frequency analysis in the ORM area.

Keywords Flood frequency • Oak Ridges Moraine (ORM) area • Concentration-area fractal model

L. Wang (✉)
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

Hainan Province Key Laboratory of Earth Observation, Sanya Institute of Remote Sensing, Sanya 572029, Hainan Province, China
e-mail: wanglei98@radi.ac.cn

X. Zhang
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zhangxin@radi.ac.cn

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_5

5.1 Flood Frequency and Return Period

Flood frequency (return period) is very important for flood control, land-used regulation, emergency planning, and insurance considerations (Malamud and Turcotte 2006). The purpose of flood frequency analysis is to relate the discharge of extreme events to their frequency of occurrence through the use of probability distribution (Chow et al. 1988). The primary methodology for estimating flood frequency is by fitting a theoretical statistical distribution to measurements of flood peak discharges and using the distribution to estimate the frequencies of different flood discharge (Beven 2001).

The frequency of a hydrological event is the probability that discharge of extreme events will occur in any given year (Viessman and Lewis 2003). The Return period is a statistical measure of how often an event of a certain size is likely to occur. In hydrology, return period is normally calculated in the unit of year. The definition of return period is the reciprocal of exceedance probability of the event,

$$T = \frac{1}{P(X \geq x)} \tag{5.1}$$

where T is the return period in years, $P(X >= x)$ is the probability of the event which has a value of $X >= x$.

In this study, both annual extreme series and partial-duration series are used for flood frequency analysis. Annual extreme series use the LP3 and GEV distribution for flood frequency analysis. In the next section, Concentration-Area (CA) fractal model is used for flood threshold selection, and this threshold is used as a partial-duration series flood threshold for flood frequency analysis. The PL and GP distribution are used for partial-duration series flood frequency analysis. These four flood frequency analysis methods are then evaluated and compared from which the most suitable method for flood frequency analysis and prediction in the ORM area can be determined. This effort includes tests of descriptive ability and tests of predictive ability.

5.2 Application of Log-Pearson III (LP3) Model for Flood Frequency Analysis

5.2.1 Introduction to LP3

The LP3 distribution Probability Density Function (PDF) is defined as (Maidment 1993):

$$f(x) = \frac{1}{\mathrm{x}|\beta|\Gamma(\alpha)} \left(\frac{\ln(\mathrm{x}) - \gamma}{\beta}\right)^{\alpha-1} \exp\left(-\frac{\ln(\mathrm{x}) - \gamma}{\beta}\right) \tag{5.2}$$

Where γ is the lower bound of the distribution, β is a scale parameter, α a shape parameter, and Γ () the gamma function. These three parameters are related to the mean $= \gamma + \alpha\beta$, variance $= \alpha/\beta^2$ and skewness $= 2/\sqrt{\alpha}$ of the distribution.

The equation used for coefficient of skewness calculation is:

$$\frac{n}{(n-1)(n-2)}\sum\left(\frac{x_i - \overline{x}}{s}\right) \tag{5.3}$$

And the Cumulative Distribution Function (CDF): F (x) is:

$$F(x) = \frac{\Gamma_{(\ln(x)-\gamma)/\beta}(\alpha)}{\Gamma(\alpha)} \tag{5.4}$$

5.2.2 Frequency Analysis by Frequency Factors Method

USGS Bulletin #17B recommends the technique for computing flood frequency curves using annual flood series. The recommended technique fits a LP3 distribution to base 10 logarithms of the discharge at selected exceedance probability by the equation:

$$\text{Log x} = \mu + \text{K} \bullet \text{s} \tag{5.5}$$

Where x — flood discharge
μ — sample mean of the base 10 logarithms
s — standard deviation of the base 10 logarithms
K — a frequency factor, which depends on the skew coefficient and selected exceedance probability.

Since the K value (Viessman and Lewis 2003) changes with probability, the flood discharge can be calculated for different flood probabilities by using Eq. (5.5).

5.2.3 Application of LP3 Model for Flood Frequency Analysis in the ORM Area

In order to apply the LP3 model for flood frequency analysis, the annual maximum flood peak at each gauging station is extracted as the study datasets. A log transform is then applied on these annual maximum flood peaks, and the mean, variability and skewness of log transform of the annual maximum flood peaks are calculated. A frequency factor analysis method is then applied to fit a LP3 distribution for each gauging station. Figure 5.1 shows the fitted LP3 line at 02HB001 station. And Table 5.1 shows the parameters of fitted LP3 line for each gauging station.

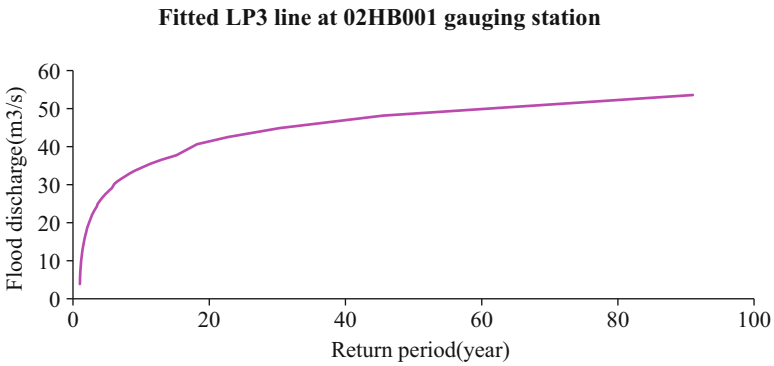


Fig. 5.1 Fitted LP3 line at 02HB001 station

After estimating the parameters of the LP3 distribution for each station, we can obtain a CDF F(x) for each station for a given flood peak x, and the corresponding return period T from the LP3 model fitted line can be calculated as given below (Chow et al. 1988):

$$T = 1/(1 - F(x)) \tag{5.6}$$

In this way, for a given return period, the flood peak from the LP3 model fitted lines can be compared with the observation data set to evaluate the performance of the LP3 model for flood frequency analysis in the ORM area.

In this study, the Weibull plotting position method (Weibull 1939) is used for calculating the probability of exceedance for an unknown probability distribution. The process of calculating probability of exceedance involves two steps. First, all annual observed flood peaks are ranked as descending, creating series of Q1, Q2, …, and Qk as the annual flood peak series, and these data show descending order. Then, the Weibull plotting position method is applied to get the return period in years, as $T = (K + 1)/M$. K is the number of all annual series flood peaks and M is the ordered sequence number of flood event Q. For example, for 02HB001 gauging station, the total number of annual series flood peaks is 90, and they are ranked as descending, so the return period in years for the largest flood event is $T = (90 + 1)/1 = 91$ years, and the return period in years for the second largest flood event is $T = (90 + 1)/2 = 45.5$ years. The return period in years for other flood events can be calculated in the same way. So for a given return period, the flood peak from the LP3 model fitted lines and the observation data set are plotted against the return period, and the correlation coefficients between model fitted lines and the observation data set are calculated to evaluate the performance of the LP3 for flood frequency analysis in the ORM area. Figure 5.2 show the LP3 model fitted lines and the observation data set plotted against return period at O2HB001 gauging station.

The correlation coefficients between model fitted lines and observations are shown in Table 5.1: the range distributes from 0.94 to 0.99 for all gauging stations.

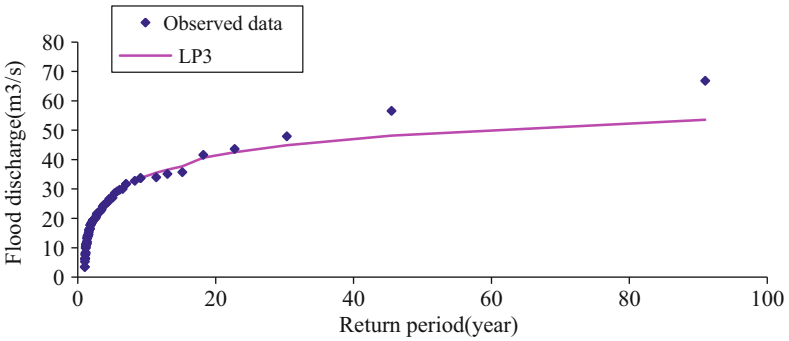


Fig. 5.2 Fitted LP3 line and observations at 02HB001station

Table 5.1 Model parameters of fitted LP3 curve

Gauging station	Number of years	α	β	γ	Correlation coefficients
02EC008	19	632.57	0.02048	−9.8853	0.9438
02EC009	40	11.778	0.12388	1.5049	0.9820
02EC010	33	164.91	0.03067	−3.3235	0.9916
02ED003	55	320.02	0.02482	−3.3387	0.9756
02ED100	25	17.626	0.10758	4.0204	0.9869
02HB001	90	20.66	−0.12451	5.422	0.9766
02HB018	22	17421.0	−0.00286	53.08	0.9801
02HB025	16	22.214	−0.09301	5.7255	0.9789
02HC003	57	32.595	0.10224	1.204	0.9411
02HC009	51	39.794	−0.0945	6.6976	0.9774
02HC018	40	16.39	−0.13166	4.7644	0.9705
02HC019	43	24.05	−0.11968	5.6011	0.9719
02HC022	43	73.685	−0.06135	7.6732	0.9846
02HC024	43	77.408	0.04735	0.05512	0.9946
02HC027	38	41.627	−0.0574	2.417	0.9742
02HC028	41	12.398	−0.10404	3.9415	0.9885
02HC030	39	33.787	0.07619	6.2301	0.9730
02HC031	35	114.81	0.03997	−1.4302	0.9732
02HC032	30	85.85	−0.04726	6.3448	0.9799
02HC033	39	20.565	0.08617	0.85833	0.9706
02HD003	44	449.21	0.02388	−8.5242	0.9544
02HD008	46	13.839	−0.18602	5.2794	0.9839
02HD009	35	12.455	0.16012	0.42052	0.9432
02HD013	23	2091.1	0.00978	−18.242	0.9702
02HG002	11	12.421	0.08172	−0.12474	0.9509

Table 5.2 Comparison of observations and LP3 model fitted line of flood discharge for approximate 10-year return period

Gauging station	Observed flood discharge (m^3/s)	LP3 prediction (m^3/s)	Error	Relative error
02EC008	38.5	41.95	3.45	0.09
02EC009	38.5	34.07	−4.43	−0.12
02EC010	11	9.68	−1.32	−0.12
02ED003	175	182.21	7.21	0.04
02ED100	15.1	15.17	0.07	0.00
02HB001	33.7	33.70	0.00	0.00
02HB018	52.6	47.81	−4.79	−0.09
02HB025	69	64.0	−4.99	−0.07
02HC003	163	200.97	37.97	0.23
02HC009	38.5	39.567	1.07	0.03
02HC018	25.5	25.93	0.43	0.02
02HC019	29.1	32.24	3.14	0.11
02HC022	48	46.58	−1.43	−0.03
02HC024	76	72.99	−3.02	−0.04
02HC027	21.7	20.27	−1.43	−0.07
02HC028	22	22.05	0.05	0.00
02HC030	72.6	66.99	−5.61	−0.08
02HC031	39.2	40.04	0.84	0.02
02HC032	19.2	17.07	−2.13	−0.11
02HC033	23	22.13	−0.87	−0.04
02HD003	19.8	16.92	−2.88	−0.15
02HD008	34.3	33.55	−0.75	−0.02
02HD009	25	22.79	−2.21	−0.09
02HD013	19.5	17.04	−2.47	−0.13
02HG002	3.89	3.73	−0.16	−0.04

For most gauging stations, the values range from 0.97 to 0.99, indicating generally good agreement between the model fitted lines and observations.

A further evaluation is applied to investigate the performance of the LP3 model in estimating the flood peak discharge for different return periods. Tables 5.2 and 5.3 compare observations and LP3 model fitted lines of flood discharge for approximately 10-year and 40-year return periods.

For approximately 10-year return period, for most stations except 02HC003 station, the LP3 fitted lines generally match the observation flood discharge very well, showing that the LP3 model works well in estimating 10-year return period flood events.

For approximately 40-year return period, for most stations except 02HC003, 02HC018, 02HC031, 02HD003 and 02HD009, the LP3 fitted lines generally match well the observation flood discharge, showing that the LP3 model works well in estimating 40-year return period flood events.

Table 5.3 Comparison of observations and LP3 model fitted line of flood discharge for approximate 40-year return period

Gauging station	Observed flood discharge (m^3/s)	LP3 prediction (m^3/s)	Error	Relative error
02EC008	–	–	–	–
02EC009	53.4	52.64	−0.76	−0.01
02EC010	12.8	12.19	−0.62	−0.05
02ED003	254	220.89	−33.11	−0.13
02ED100	–	–	–	–
02HB001	56.6	48.159	−8.44	−0.15
02HB018	–	–	–	–
02HB025	–	–	–	–
02HC003	218	304.16	86.16	0.40
02HC009	46.2	50.045	3.85	0.08
02HC018	30	35.01	5.01	0.17
02HC019	49.5	44.54	−4.96	−0.10
02HC022	61.4	64.55	3.15	0.05
02HC024	108	102.06	−5.94	−0.06
02HC027	24.6	26.39	1.79	0.07
02HC028	26	26.80	0.80	0.03
02HC030	94.9	88.39	−6.51	−0.07
02HC031	65	54.45	−10.56	−0.16
02HC032	21.1	21.54	0.44	0.02
02HC033	38.5	33.86	−4.64	−0.12
02HD003	33	26.28	−6.72	−0.20
02HD008	59.9	50.40	−9.50	−0.16
02HD009	52.9	36.75	−16.15	−0.31
02HD013	–	–	–	–
02HG002	–	–	–	–

5.2.4 *Apply LP3 Model for Flood Frequency Prediction in the ORM Area*

This fitted LP3 distribution can also be used to estimate flood discharge for different return periods. Table 5.4 lists the flood discharge calculated for different return period flood events.

Furthermore, the fitted LP3 curves can also be used for flood frequency (return period) prediction. The return period of a given discharge can be obtained by interpolated on the fitted LP3 curve.

Table 5.5 shows the return period interpolation results at 02HC033 gauging station for selected flood events. The return period is the reciprocal of exceedance probability of the event.

Table 5.4 Flood discharge calculated for different return periods using LP3 distribution

Gauging station	5-year return period (m^3/s)	10-year return period (m^3/s)	50-year return period (m^3/s)	100-year return period (m^3/s)
02EC008	33.229	41.947	63.55	73.734
02EC009	27.234	34.068	52.638	62.189
02EC010	7.8624	9.4384	13.135	14.805
02ED003	144.7	177.26	255.14	290.84
02ED100	12.318	14.515	18.784	20.391
02HB001	28.046	34.569	48.159	53.578
02HB018	38.485	45.385	60.568	67.046
02HB025	56.506	66.517	86.239	93.781
02HC003	150.71	200.97	344.48	421.03
02HC009	31.367	39.567	57.851	65.611
02HC018	21.394	25.928	35.011	38.508
02HC019	25.148	31.335	44.541	49.921
02HC022	36.628	45.29	64.545	72.745
02HC024	58.288	71.034	102.06	116.54
02HC027	16.85	19.831	26.391	29.19
02HC028	19.404	22.046	26.801	28.476
02HC030	56.487	66.993	88.391	96.838
02HC031	33.602	41.07	59.17	67.576
02HC032	14.291	17.071	22.993	25.433
02HC033	19.072	23.244	33.856	39.01
02HD003	13.836	17.415	26.277	30.455
02HD008	27.057	34.555	50.403	56.73
02HD009	17.608	23.68	42.056	52.391
02HD013	13.187	16.102	22.934	26.008
02HG002	3.0689	3.5692	4.7836	5.3506

Table 5.5 Return periods for flood events at O2HC033 gauging station

Year	Month	Day	Discharge (m^3/s)	Log discharge	Return period (year)
1968	2	2	24.7	1.3927	13
1984	2	14	23	1.3617	10
1985	2	24	23	1.3617	10
1980	3	21	19.2	1.2833	5
1974	5	17	18.9	1.2765	5
1992	8	28	18.1	1.2577	4
1979	3	4	17	1.2304	4
2000	4	21	12.7	1.1038	2
1977	3	13	11.4	1.0569	2
2004	5	24	9.61	0.9827	2

5.3 Application of GEV Model for Flood Frequency Analysis and Prediction

5.3.1 Introduction to GEV Model

The Generalized Extreme Value (GEV) (Jenkinson 1955) distribution is a flexible three-parameter model that combines the Gumbel, Fréchet, and Weibull maximum extreme value distributions. It has the following PDF:

$$f(x)\begin{cases}\frac{1}{\sigma}\exp\left(-(1+kz)^{-1/k}\right)(1+kz)^{-1-1/k} & k\neq 0\\ \frac{1}{\sigma}\exp(-z-\exp(-z)) & k=0\end{cases} \tag{5.7}$$

where $z = (x\text{-}\mu)/\sigma$, and k, σ, μ are the shape, scale, and location parameters respectively.

Various values of the shape parameter yield the extreme value type I, II, and III distributions. Specifically, the three cases $k = 0, k > 0$, and $k < 0$ correspond to the Gumbel, Fréchet, and Weibull distributions.

And the Cumulative Distribution Function (CDF) is:

$$F(x) = \begin{cases}\exp\left(-(1+kz)^{-1/k}\right) & k\neq 0\\ \exp(-\exp(-z)) & k=0\end{cases} \tag{5.8}$$

where $z = (x\text{-}\mu)/\sigma$.

5.3.2 GEV Parameter Determination by Maximum Likelihood Method (MLM)

In order to apply the GEV distribution for flood frequency analysis, the annual maximum flood peak is extracted as the study datasets. EasyFit software (MathWave Technologies 2009) and MLM are used to fit the GEV distribution to the extracted study datasets. Table 5.6 shows the parameters of fitted GEV distribution for each gauging station.

5.3.3 Evaluation of GEV Model for Flood Frequency Analysis in the ORM Area

After estimating the parameters of the GEV distribution for each station, we can obtain CDF F(x) for each station for a given flood peak x, and the corresponding

Table 5.6 Model parameters of fitted GEV distribution curve

Gauging station	Number of years	Shape parameter (k)	Scale parameter (σ)	Location parameter (μ)	GEV type	Correlation coefficients
02EC008	19	0.21823	7.8501	17.928	II	0.9388
02EC009	40	0.21744	6.0817	16.125	II	0.9788
02EC010	33	0.08033	1.8482	4.8851	II	0.9914
02ED003	55	0.09731	36.046	85.493	II	0.9763
02ED100	25	−0.13413	3.5465	7.5361	III	0.9878
02HB001	90	0.06601	7.9258	14.902	II	0.9864
02HB018	22	−0.01844	9.4486	24.675	III	0.9806
02HB025	16	−0.06415	15.338	34.347	III	0.9801
02HC003	57	0.28778	38.209	76.363	II	0.9342
02HC009	51	0.0381	9.5532	16.343	II	0.9827
02HC018	40	−0.12705	6.761	12.229	III	0.9789
02HC019	43	−0.07119	8.159	13.581	III	0.9709
02HC022	43	0.01306	10.744	20.271	II	0.9838
02HC024	43	0.11794	13.825	35.217	II	0.9844
02HC027	38	−0.01272	4.0319	10.903	III	0.9759
02HC028	41	−0.2014	4.9442	13.051	III	0.9896
02HC030	39	0.02495	14.107	33.901	II	0.9788
02HC031	35	0.1096	8.1503	20.091	II	0.9750
02HC032	30	−0.04692	3.8914	8.6933	III	0.9800
02HC033	39	0.15914	4.1635	11.834	II	0.9654
02HD003	44	0.16118	3.5154	7.604	II	0.9630
02HD008	46	0.15689	7.9703	12.441	II	0.9925
02HD009	35	0.34821	4.0977	8.8083	II	0.9576
02HD013	23	0.10847	3.2603	7.6972	II	0.9700
02HG002	11	0.1558	0.53965	2.1238	II	0.9490

return period T from the GEV model fitted line can be calculated using Eq. (5.6). For a given return period, the flood peak from the GEV model fitted lines can be compared with the observation data set to evaluate the GEV model for flood frequency analysis in the ORM area.

For a given return period, the flood peak from the GEV model fitted lines and the observation data set are plotted against return period, and the correlation coefficients between model fitted lines and the observation data set are calculated to evaluate the performance of GEV model for flood frequency analysis in the ORM area. Figures 5.3 and 5.4 show the GEV model fitted lines and the observation data set plotted against the return period at 02HB001 and 02HC033 gauging stations separately. The correlation coefficients between model fitted lines and the observation data set are listed in Table 5.7.

From Table 5.7, we note that the GEV fitted lines generally match well the observation flood discharge. Correlation coefficients are 0.99, 0.98, 0.97 or 0.96, except for a few stations such as 02EC008, 02HC003 and 02HG002. These data

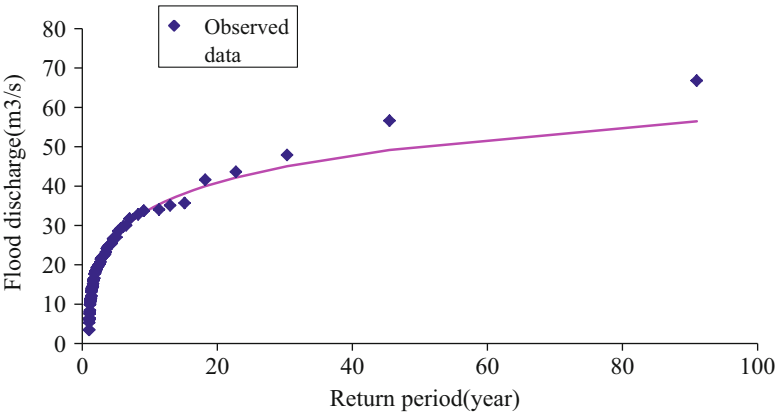


Fig. 5.3 Fitted GEV line and observations at 02HB001gauging station

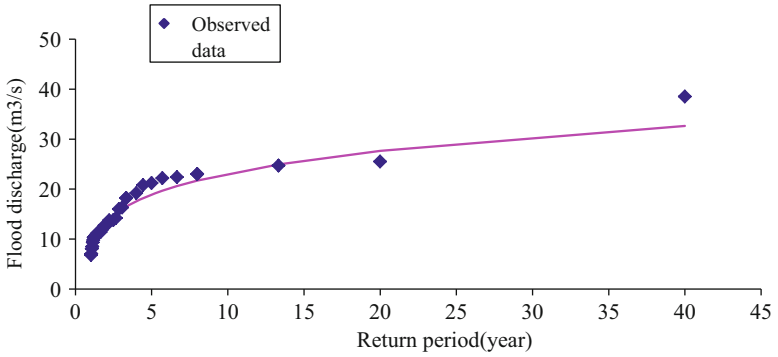


Fig. 5.4 Fitted GEV line and observations at 02HC033 gauging station

prove that it is effective and applicable to use the GEV model for flood frequency analysis in the ORM area.

A further evaluation is applied to investigate the performance of the GEV model in estimating flood peak discharge for different return periods. Tables 5.7 and 5.8 compare observations with GEV model fitted lines of flood discharge for approximately 10-year and 40-year return periods.

For approximately 10-year return period, for most stations except 02HC003 and 02HD003, the GEV fitted lines generally match well the observation flood discharge, showing that the GEV model works well in estimating 10-year return period flood events.

For approximately 40-year return period, for most stations except 02HC003, 02HC019, 02HC031, 02HD003 and 02HD009, the GEV fitted lines generally match well the observation flood discharge, showing that the GEV model works well in estimating 40-year return period flood events.

Table 5.7 Comparison of observations and GEV model fitted lines of flood discharge for approximate 10-year return period

Gauging station	Observed flood discharge (m^3/s)	GEV prediction (m^3/s)	Error	Error (%)
02EC008	38.5	40.74	2.238	0.06
02EC009	38.5	33.78	−4.72	−0.12
02EC010	11	9.69	−1.31	−0.12
02ED003	175	181.2	6.17	0.04
02ED100	15.1	15.03	−0.07	−0.01
02HB001	33.7	33.21	−0.49	−0.01
02HB018	52.6	46.5	−6.09	−0.12
02HB025	69	63.9	−5.1	−0.07
02HC003	163	197.31	34.31	0.21
02HC009	38.5	38.79	0.29	0.01
02HC018	25.5	25.46	−0.04	0.01
02HC019	29.1	31.31	2.21	0.08
02HC022	48	46.03	−1.97	−0.04
02HC024	76	72.86	−3.14	−0.04
02HC027	21.7	20.28	−1.42	−0.07
02HC028	22	22.00	0.00	0.00
02HC030	72.6	66.56	−6.04	−0.08
02HC031	39.2	39.85	0.65	0.02
02HC032	19.2	17.00	−2.20	−0.11
02HC033	23	21.71	−1.29	0.06
02HD003	19.8	16.64	−3.17	−0.16
02HD008	34.3	32.82	−1.48	−0.04
02HD009	25	21.92	−3.08	−0.12
02HD013	19.5	16.99	−2.51	−0.13
02HG002	3.89	3.76	−0.13	−0.03

Table 5.8 Comparison of observations and GEV model fitted lines of flood discharge for approximate 40-year return period

Gauging station	Observed flood discharge (m^3/s)	GEV prediction (m^3/s)	Error	Error (%)
02EC008	–	–	–	–
02EC009	53.4	53.49	0.09	0.01
02EC010	12.8	12.33	−0.47	−0.04
02ED003	254	220.8	−33.2	−0.13
02ED100	–	–	–	–
02HB001	56.6	49.19	−7.41	0.13
02HB018	–	–	–	–
02HB025	–	–	–	–
02HC003	218	306.22	88.22	0.40
02HC009	46.2	48.84	2.64	0.06
02HC018	30	33.03	3.03	0.10
02HC019	49.5	41.38	−8.12	−0.16

(continued)

Table 5.8 (continued)

Gauging station	Observed flood discharge (m^3/s)	GEV prediction (m^3/s)	Error	Error (%)
02HC022	61.4	63.28	1.88	0.03
02HC024	108	103.72	−4.28	−0.04
02HC027	24.6	26.25	1.65	0.07
02HC028	26	26.41	0.41	0.02
02HC030	94.9	91.72	−3.18	−0.03
02HC031	65	54.76	−10.24	−0.16
02HC032	21.1	21.23	0.13	0.01
02HC033	38.5	32.63	−5.87	−0.15
02HD003	33	26.70	−6.30	−0.19
02HD008	59.9	55.34	−4.56	−0.08
02HD009	52.9	36.73	−16.17	−0.31
02HD013	–	–	–	–
02HG002	–	–	–	–

5.3.4 Application of the GEV Model for Flood Frequency Prediction in the ORM Area

The fitted GEV distribution can also be used to estimate flood discharge for different return periods. Table 5.9 lists the flood discharge calculated for different return period flood events, including 5-year, 10-year, 50-year, and 100-year events.

A further evaluation is applied to compare the GEV distribution with the widely used LP3 distribution for different return periods to compare the flood frequency predictive ability (Fig. 5.5).

For short return period predictions (5-year and 10-year), the GEV distribution prediction result is less than that of the LP3 distribution; however, for long return period predictions (50-year and 100-year), the GEV distribution prediction result is larger than that of the LP3 distribution. Further investigation finds that the relative differences between these two distributions for different return periods are relatively small: most of them are less than 5 %, and only a few are larger than 10 % (Table 5.10).

02HB001 gauging station is selected as an example to compare the GEV distribution with the LP3 distribution to compare the flood frequency predictive ability (Fig. 5.5).

Figure 5.5 shows the fitted GEV and LP3 line at 02HB001 gauging station, and it is observed that for less than 40-year return periods, there is not too much difference for different distribution predictions. However, for long return periods, for example, longer than 40-year return periods, the prediction of GEV distribution is larger than prediction of LP3 distribution.

Table 5.9 Flood discharge calculated for different return period using GEV distribution

Gauging station	5-year return period (m^3/s)	10-year return period (m^3/s)	50-year return period (m^3/s)	100-year return period (m^3/s)
02EC008	31.858	40.738	66.245	80.119
02EC009	26.91	33.779	53.492	64.204
02EC010	7.8313	9.444	13.355	15.171
02ED003	143.7	176.18	256.57	294.64
02ED100	12.355	14.425	18.31	19.711
02HB001	27.399	34.131	50.176	57.503
02HB018	38.653	45.502	60.248	66.347
02HB025	56.282	66.488	87.294	95.448
02HC003	148.03	197.31	351.7	442.52
02HC009	31.09	38.79	56.532	64.375
02HC018	21.463	25.462	33.03	35.781
02HC019	25.188	30.546	41.378	45.587
02HC022	36.545	44.808	63.28	71.21
02HC024	57.901	70.847	103.72	119.66
02HC027	16.893	19.847	26.251	28.918
02HC028	19.451	21.997	26.412	27.88
02HC030	55.462	66.556	91.716	102.67
02HC031	33.379	40.892	59.776	68.846
02HC032	14.33	17.004	22.568	24.794
02HC033	18.887	23.1	34.351	40.072
02HD003	13.569	17.14	26.7	31.573
02HD008	25.92	33.952	55.34	66.188
02HD009	16.88	22.804	42.83	55.432
02HD013	13.008	16.007	23.535	27.145
02HG002	3.0357	3.5783	5.0216	5.7527

Table 5.10 Relative difference of flood discharge calculated for different return periods between GEV and LP3 distributions

Gauging station	5-year return period	10-year return period	50-year return period	100-year return period
02EC008	−0.04	−0.03	0.04	0.09
02EC009	−0.01	−0.01	0.02	0.03
02EC010	0.00	0.00	0.02	0.02
02ED003	−0.01	−0.01	0.01	0.01
02ED100	0.00	−0.01	−0.03	−0.03
02HB001	−0.02	−0.01	0.04	0.07
02HB018	0.00	0.00	−0.01	−0.01
02HB025	0.00	0.00	0.01	0.02
02HC003	−0.02	−0.02	0.02	0.05
02HC009	−0.01	−0.02	−0.02	−0.02
02HC018	0.00	−0.02	−0.06	−0.07
02HC019	0.00	−0.03	−0.07	−0.09
02HC022	0.00	−0.01	−0.02	−0.02

(continued)

Table 5.10 (continued)

Gauging station	5-year return period	10-year return period	50-year return period	100-year return period
02HC024	−0.01	0.00	0.02	0.03
02HC027	0.00	0.00	−0.01	−0.01
02HC028	0.00	0.00	−0.01	−0.02
02HC030	−0.02	−0.01	0.04	0.06
02HC031	−0.01	0.00	0.01	0.02
02HC032	0.00	0.00	−0.02	−0.03
02HC033	−0.01	−0.01	0.01	0.03
02HD003	−0.02	−0.02	0.02	0.04
02HD008	−0.04	−0.02	0.10	0.17
02HD009	−0.04	−0.04	0.02	0.06
02HD013	−0.01	−0.01	0.03	0.04
02HG002	−0.01	0.00	0.05	0.08

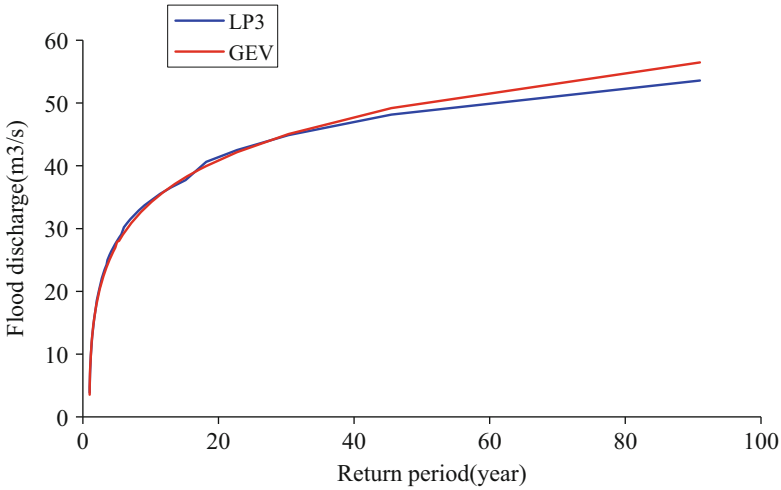


Fig. 5.5 Fitted LP3 and GEV lines at 02HB001gauging station

5.4 Concentration-Area Fractal Model for Flood Threshold Selection

5.4.1 Threshold Selection for Partial-Duration Series Flood Frequency Analysis

Although POT modeling is theoretically easy and is well adapted to the analysis of extreme values (Lang et al. 1999), the problems of choice of threshold represent difficulties associated with the POT approach. PDS flood frequency usually

involves subjectivity for selection of the threshold, and less arbitrary methods for selecting thresholds need to be developed in applications of the PDS model. The threshold selection is crucial, but by no means straightforward, so the issue of threshold requires more research (Adamowsk et al. 1998).

Normally, there are two approaches for flood threshold selection: (1) physically based approaches, and (2) statistically based approaches. The first approaches select the flood threshold according to some physically relevant criterion, such as the overflowing of a river; the second approaches use some statistical analysis to determine "flood threshold" (Birikundavyi and Rousselle 1997a). For example, Birikundavyi and Rousselle (1997a) used a PP plot to identify the best value for the threshold level when modeling POT with the Poisson-exponential model.

5.4.2 Introduction to Fractal Models

Fractal models have provided effective tools to characterize complex physical processes and their end products (Cheng et.al. 2009). Natural patterns normally appear very complex. However, they may display an underlying simplicity through scale-invariance. Scale-invariance means that the pattern appears unchanged under magnification or contraction. The key idea in fractal geometry is self-similarity. An object is self-similar if it can be decomposed into small copies of itself. Fractals are a natural consequence of self-similarity resulting from scale-independent processes (Evertsz and Mandelbrot 1992). It can be said that fractals apply to sets and that multifractals apply to measures from which spatially associated fractal sets can be derived. Both fractals and multifractals are characterized by their fractal dimension and singularity intensity of the measure (Evertsz and Mandelbrot 1992; Cheng 1997). Fractal models can be used to characterize complex physical processes and dynamic systems (Kigami 2001).

The Power-law distribution is the only type of distribution that shows scale invariance and the self-similarity property. Many singular processes have shown that the end products of singular processes can be modeled by power-law models (fractal models) (Cheng et.al. 2009). One way to determine whether data fit a power law relationship is to plot the log(y) versus the log(x) and check whether there exists a linear relation between log(y) and log(x). The Power law relation is a common tool used in fractal analysis.

5.4.3 Concentration-Area Fractal Method

Fractal models have been applied to separate populations based on distinctive self-similarities observed in populations; for example, a Concentration-Area (CA) fractal model was proposed for separating geochemical anomalies from background values on the basis of distinct power-law relations (Cheng

et al. 1994), since regular flow events can be considered as background, and flood events can be considered as anomalies. This CA fractal model can be adopted to identify the best value of flood threshold (Cheng et.al. 2009).

This Concentration-Area method associates element concentration value (v) and the areas within which element concentration values are above a threshold ρ, (A ($v \geq \rho$)) as

$$A(v \geq \rho) = C\rho^{-\alpha} \tag{5.9}$$

where C and α are two parameters determining the power-law relation. C is constant and α is the exponent of the power-law function (Cheng et.al. 2009).

5.4.4 *Application of Concentration-Area Fractal Method for Flood Threshold Selection*

This CA fractal model can also be used for flood "threshold selection." Using the historical discharge data and variable threshold for each gauging station to calculate the number of days with discharge above the threshold, the log-log plot is applied to model the relationship between discharge and number of days with flow above the threshold. Since each straight line segment represents a power law relationship means the values in each group are self-similar. This provides a basis for separating river runoff into different populations. The group showing different properties from those of normal flow belongs to a separate population as flood events (Li 2005; Cheng et al. 2009).

The 02HB001 station is used as a case study for flood threshold selection. The flow discharge threshold values are set as 51 groups of an interval of 1.33 m^3/s, and the minimum and maximum values are 0.17–66.8 m^3/s, respectively. The number of days with flow discharge in each group is calculated, and shown in Table 5.11 and plotted in Fig. 5.6.

Table 5.11 shows that the number of days against flow data generally show linear trend in the log-log plot. We can draw two straight lines on the log-log plot. The pink straight line shows a linear relation between log-transformed number of days and the log-transformed flow data, with a linear correlation coefficient R2 = 0.999. The yellow straight line shows a linear relation between log-transformed number of days and the log-transformed flow data, with a linear correlation coefficient R2 = 0.990. Both of these straight lines ensure that the number of days and flow data follow a power-law relation, implying that the values in each group are self-similar. The group showing a different property from that of normal flow belongs to a separate population as flood events. The intersection of these two lines can be considered as the breakdown discharge as the "flood threshold."

Table 5.11 Accumulative days and discharge at 02HB001 Station

Threshold discharge (m^3/s)	Number of days	Log threshold	Log days	Threshold discharge (m^3/s)	Number of days	Log threshold	Log days
0.17	33,112	−1.77	10.41	34.82	9	3.55	2.20
1.50	11,776	0.41	9.37	36.15	6	3.59	1.79
2.84	3646	1.04	8.20	37.48	6	3.62	1.79
4.17	1891	1.43	7.54	38.82	6	3.66	1.79
5.50	1188	1.70	7.08	40.15	6	3.69	1.79
6.83	810	1.92	6.70	41.48	6	3.73	1.79
8.17	609	2.10	6.41	42.81	5	3.76	1.61
9.50	475	2.25	6.16	44.15	4	3.79	1.39
10.83	375	2.38	5.93	45.48	4	3.82	1.39
12.16	310	2.50	5.74	46.81	4	3.85	1.39
13.50	257	2.60	5.55	48.14	3	3.87	1.10
14.83	196	2.70	5.28	49.48	3	3.90	1.10
16.16	166	2.78	5.11	50.81	2	3.93	0.69
17.49	138	2.86	4.93	52.14	2	3.95	0.69
18.83	111	2.94	4.71	53.47	2	3.98	0.69
20.16	88	3.00	4.48	54.81	2	4.00	0.69
21.49	72	3.07	4.28	56.14	2	4.03	0.69
22.82	58	3.13	4.06	57.47	1	4.05	0.00
24.16	49	3.18	3.89	58.80	1	4.07	0.00
25.49	36	3.24	3.58	60.14	1	4.10	0.00
26.82	31	3.29	3.43	61.47	1	4.12	0.00
28.15	25	3.34	3.22	62.80	1	4.14	0.00
29.49	22	3.38	3.09	64.13	1	4.16	0.00
30.82	20	3.43	3.00	65.47	1	4.18	0.00
32.15	17	3.47	2.83	66.80	1	4.20	0.00
33.49	13	3.51	2.56				

For 02HB001 station, the breakdown discharge is 14.83 m^3/s. This breakdown can be used as the "flood threshold". If the flow discharge is greater than the "flood threshold," it can be considered as flood event.

Similarly, for other stations, we can calculate the number of days in which the flow discharge is greater than the threshold value, then we can make the log-log plot of number of days against flow data. The intersection of these two lines can be considered as "flood threshold." For example, for 02ED003 gauging station, the Log-log plot of number of days against the flow discharge at 02ED003 gauging station is shown in Fig. 5.7. The breakdown discharge is 76.59 m^3/s.

Similarly, we can log-log plot of number of days against the flow discharge and get the breakdown value for all other stations in the ORM area. Table 5.12 shows the break down value for all gauging stations in the study area. If the flow discharge is greater than the "flood threshold," it can be considered as flood event.

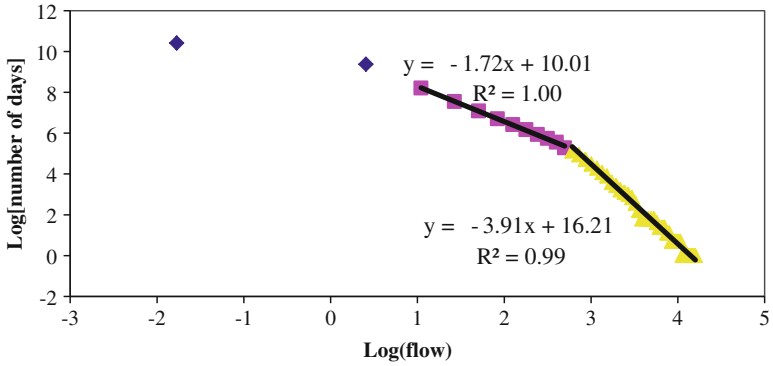


Fig. 5.6 Log-log plot of number of days against the flow discharge at 02HB001 Station. Logarithmic transformation is nature based.

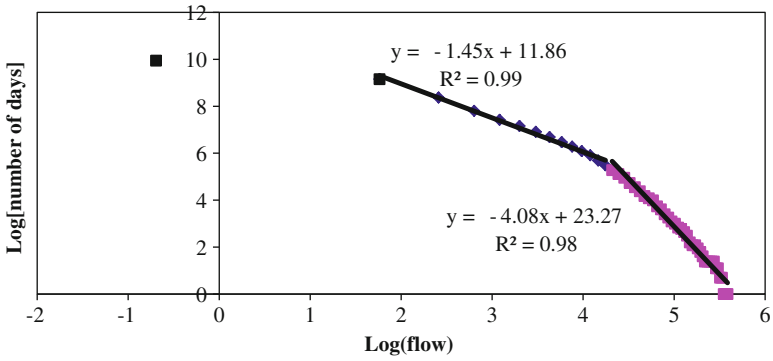


Fig. 5.7 Log-log plot of number of days against the flow discharge at 02ED003 Station

We can also map the spatial distribution of the flood thresholds and find the spatial distribution pattern of flood thresholds in the ORM area (Fig. 5.8). It is observed that flood threshold is related to watershed area: Large area watersheds have high flood thresholds and small area watersheds have small flood thresholds. This is expected because large area watersheds have bigger total volume of water that appears downstream and this increases watershed flood threshold.

5.4.5 Discussion

There are a number of peaks over threshold methods (Lang et al. 1999) in which the peaks over a chosen threshold, typically one to five per year, are used to define the partial-duration flood series. Yevjevich and Taesombut (1979) found that the value of peaks over a chosen threshold per year should be greater than 1.8. Ashkar et al. (1987) found that the value could be less than 1.65, while Birikundavyi and

Table 5.12 Flood thresholds in the ORM area

Gauge station	Min discharge (m^3/s)	Max discharge (m^3/s)	Flood threshold (m^3/s)
02EC008	0.062	77.5	12.45
02EC009	0.062	53.4	13.93
02EC010	0.001	12.8	4.10
02ED003	0.5	267	76.59
02ED100	0.014	17	6.81
02HB001	0.17	66.8	14.83
02HB018	1.2	54.1	22.36
02HB025	1.29	83.9	32.68
02HC003	0.255	838	84.03
02HC009	0.006	83.3	18.32
02HC018	0.023	30	11.41
02HC019	0.099	49.5	13.93
02HC022	0.051	61.4	18.46
02HC024	0.985	108	30.95
02HC027	0.042	24.6	9.37
02HC028	0.02	26.6	12.78
02HC030	0.108	94.9	28.55
02HC031	0.001	65	19.50
02HC032	0.026	21.1	7.61
02HC033	0.002	38.5	9.24
02HD003	0.221	33	8.09
02HD008	0.201	59.9	13.33
02HD009	0.17	52.9	9.66
02HD013	0.011	23.9	7.66
02HG002	0.006	5	2.20

Rousselle (1997b) found the most appropriate threshold level to be 1.7. In the CA method, the number of flood events per year changes from 1 to 2, proving the CA method effective and applicable for flood threshold selection.

5.5 Application of Power Law (PL) model for partial-duration series flood frequency analysis

5.5.1 Introduction to Power Law Model for Flood Frequency Analysis

Turcotte and Greene (1993), Malamud et al. (1996), Kidson and Richards (2005) proposed the application of the power-law model for flood frequency analysis. According to Malamud and Turcotte (2006), several authors have argued that a power-law distribution, which is scale-invariant over the range considered and consists of just two parameters, is preferable to estimate the flood hazard. Turcotte

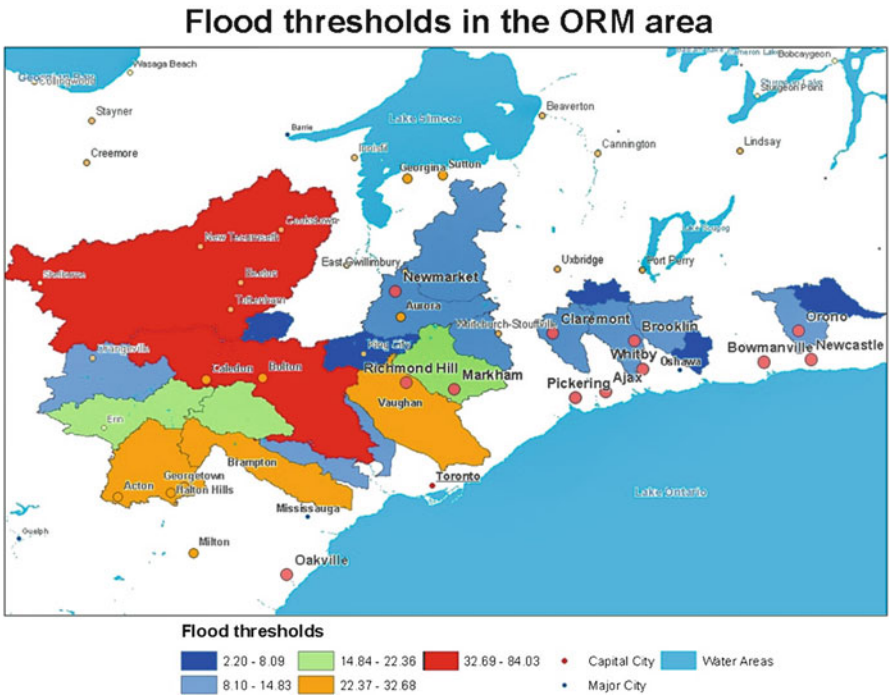


Fig. 5.8 Flood thresholds distribution in the ORM area

and Greene (1993) found that the power-law and log-Gumbel distributions were in good agreement with historical flood data and gave the most conservative predictions of large future floods. Kidson and Richards (2005) found that power law statistics provide a much better estimate of the most extreme historical flood event and found that the power-law model can be best used for large events above a fairly high threshold.

The power law model for flood frequency takes the form (Malamud et al. 2006)

$$Q\,(T) = C\,T\sigma \tag{5.10}$$

where Q(T) is the maximum flood discharge associated with a return period of T years, C and σ are regression coefficients.

5.5.2 *Application of Power Law Model for Flood Frequency Analysis in the ORM Area*

Sections 5.1.2 and 5.1.3 describe annual maximum flood peak for flood frequency analysis. Previous research studies have shown partial duration series for flood

frequency analysis (Choulakian 1990; Birikundavyi and Rousselle 1997a). A partial-duration series is obtained by taking all flood peaks equal to or greater than a pre-defined base flood threshold. The flood threshold derived from the CA model (Table 5.12) can be used as the flood threshold for the partial-duration series of a flood frequency analysis.

For the partial duration series flood frequency analysis, flood events must be independent. Several independence criteria are found in the literature (Lang et al. 1999). According to Lang et al. (1999), one criterion to define independent flood is "the relative minimum value between two consecutive peaks should be less than 50 % of the previous peak value." And Mathier et al. (1993) considered two neighboring flood peaks are independent if they are separated by at least 7 days and if the flow between them falls below 50 % of the smaller peak. In this study, this criterion (Mathier et al. 1993) is used to select independent flood events for partial duration series flood frequency analysis because this approach helps to discard those flood events coming from the same flood generating process, and only the peak discharge of those flood events is used for flood frequency analysis.

Similarly, the Weibull plotting position method (Weibull 1939) is used to calculate the probability of exceedance for observed flood discharge. For example, for the 02EC010 gauging station, the total number of partial-duration series flood peaks is 49, and the peaks are ranked as descending; so the return period in years for the largest flood event is $T = (49 + 1)/1 = 50$ years, and the return period in years for the second largest flood event is $T = (49 + 1)/2 = 25$ years. The return period in years for other flood events can be calculated using the same method.

Figure 5.9 shows the fitted power law model line at 02EC010 station. The x-axis stands for log-transformed return period and the y-axis stands for log-transformed partial-duration flood discharge. The partial-duration flood series generally correlate reasonably well with the power-law relation. The good power-law correlations obtained for the partial-duration series indicate that the statistical distribution of

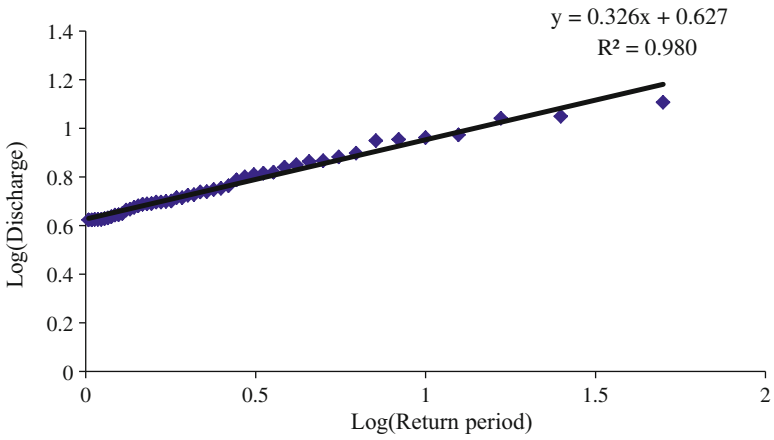


Fig. 5.9 Fitted power law model line at 02EC010 station

flood might be fractal. Similarly, we can get the fitted power law model curve for other gauging stations. Table 5.13 shows the parameters of the fitted power law model curve.

A further evaluation is applied to investigate the performance of the power law model in estimating the flood peak discharge for different return periods. Tables 5.14 and 5.15 compare the observations and power law model fitted lines of flood discharge for approximate 10-year and 40-year return periods.

For approximately a 10-year return period, for most stations except 02EC008, 02HD009 and 02HC032, the power law model fitted lines generally show good agreement with the observation flood discharge, showing that the power law model works well in estimating 10-year return period flood events.

For the approximately 40-year return year, for most stations except 02ED003, 02HB001, 02HC019, 02HC031, 02HC033 and 02HD003, the power law model fitted lines generally do not match well the observation flood discharge, normally

Table 5.13 Model parameters of fitted power law model curves

Gauging station	Number of years	σ	Correlation coefficients
02EC008	19	0.37	0.965
02EC009	40	0.37	0.983
02EC010	33	0.33	0.980
02ED003	55	0.34	0.982
02ED100	25	0.28	0.908
02HB001	90	0.34	0.985
02HB018	22	0.36	0.954
02HB025	16	0.34	0.973
02HC003	57	0.30	0.951
02HC009	51	0.31	0.964
02HC018	40	0.28	0.873
02HC019	43	0.31	0.949
02HC022	43	0.37	0.982
02HC024	43	0.30	0.992
02HC027	38	0.25	0.959
02HC028	41	0.25	0.937
02HC030	39	0.31	0.985
02HC031	35	0.30	0.978
02HC032	30	0.31	0.963
02HC033	39	0.37	0.983
02HD003	44	0.38	0.986
02HD008	46	0.44	0.975
02HD009	35	0.51	0.979
02HD013	23	0.38	0.972
02HG002	11	0.30	0.941

Table 5.14 Comparison of observations and power law model fitted lines of flood discharge for approximate 10-year return period

Gauging station	Observed flood discharge (m^3/s)	Power law prediction (m^3/s)	Error	Relative error
02EC008	30.30	33.60	3.30	0.11
02EC009	34.60	33.71	−0.89	−0.03
02EC010	9.15	8.97	−0.18	−0.02
02ED003	177.00	181.21	4.21	0.02
02ED100	14	14.65	0.65	0.05
02HB001	34	35.84	1.84	0.05
02HB018	50.00	53.31	3.31	0.07
02HB025	69	67.46	−1.54	−0.02
02HC003	171.00	180.91	9.91	0.06
02HC009	38.50	39.24	0.74	0.02
02HC018	25.20	25.29	0.09	0.00
02HC019	29.10	31.65	2.55	0.09
02HC022	43.90	42.85	−1.05	−0.02
02HC024	66.10	62.87	−3.23	−0.05
02HC027	19.30	18.39	−0.91	−0.05
02HC028	22.10	24.22	2.12	0.10
02HC030	62.30	58.14	−4.16	−0.07
02HC031	40.30	40.36	0.06	0.00
02HC032	18.10	16.14	−1.96	−0.11
02HC033	22.20	21.75	−0.45	−0.02
02HD003	20.00	19.22	−0.78	−0.04
02HD008	34.30	34.67	0.37	0.01
02HD009	31.50	27.32	−4.18	−0.13
02HD013	19.50	18.10	−1.40	−0.07
02HG002	5.00	5.34	0.34	0.07

Table 5.15 Comparison of observations and power law model fitted lines of flood discharge for approximate 40-year return period

Gauging station	Observed flood discharge (m^3/s)	Power law prediction (m^3/s)	Error	Relative error
02EC008	77.50	59.63	−17.87	−0.23
02EC009	53.40	61.35	7.95	0.15
02EC010	12.80	15.17	2.37	0.19
02ED003	254.00	264.40	10.40	0.04
02ED100	–	–	–	–
02HB001	56.60	57.73	1.13	0.02
02HB018	54.10	79.15	25.05	0.46
02HB025	–	–	–	–
02HC003	218.00	250.77	32.77	0.15
02HC009	52.21	46.2	−6.01	−0.13

(continued)

Table 5.15 (continued)

Gauging station	Observed flood discharge (m^3/s)	Power law prediction (m^3/s)	Error	Relative error
02HC018	30.00	39.99	9.99	0.33
02HC019	49.50	48.56	−0.94	−0.02
02HC022	60.30	68.18	7.88	0.13
02HC024	86.70	95.90	9.20	0.11
02HC027	22.20	24.30	2.10	0.09
02HC028	26.60	34.10	7.50	0.28
02HC030	80.60	88.94	8.34	0.10
02HC031	54.40	56.17	1.77	0.03
02HC032	21.10	24.93	3.83	0.18
02HC033	35.30	34.38	−0.92	−0.03
02HD003	33.00	32.57	−0.43	−0.01
02HD008	59.90	70.06	10.16	0.17
02HD009	52.90	47.73	−5.17	−0.10
02HD013	–	–	–	–

the relative error is greater than 0.10, showing that the power law model does not work well in estimating 40-year return period flood events.

5.5.3 *Application of the Power Law Model for Flood Frequency Prediction in the ORM Area*

This fitted power law line can also be used to estimate flood discharge for different return periods. Table 5.16 lists the flood discharge calculated for different return period flood events, including 5-year, 10-year, 50-year, and 100-year events.

A further evaluation is applied to compare the PL distribution with the widely used LP3 distribution for different return periods, including 5-year, 10-year, 50-year, and 100-year flood events, to compare the flood frequency predictive ability (Table 5.17).

For short return period predictions (5 year and 10 year), the power law model prediction result is less than that of the LP3 distribution; however, for long return period predictions (50 year and 100 year), the power law model prediction result is larger than that of the LP3 distribution. Further investigation shows that the relative difference between these two distributions for different return periods is relatively much larger than the relative difference between GEV distribution and the LP3 distribution. Most of them are greater than 10 % and only a few are less than 5 %, mostly for 5-year and 10-year return periods.

Table 5.16 Flood discharge calculated for different flood return periods using power law model

Gauging station	5-year return period (m^3/s)	10-year return period (m^3/s)	50-year return period (m^3/s)	100-year return period (m^3/s)
02EC008	26.35	35.1	68.32	91.01
02EC009	25.49	32.98	60.03	77.7
02EC010	7.16	8.97	15.17	19.02
02ED003	140.4	178.2	309.94	393.37
02ED100	12.17	14.78	23.24	28.24
02HB001	27.89	35.39	61.55	78.11
02HB018	41.14	52.75	93.96	120.48
02HB025	55.89	70.83	122.73	155.52
02HC003	144.41	177.46	286.35	351.88
02HC009	31.43	39	64.39	79.91
02HC018	20.76	25.29	39.99	48.71
02HC019	25.36	31.41	51.63	63.94
02HC022	33.68	43.55	79.1	102.28
02HC024	51.3	63.36	103.44	127.76
02HC027	15.24	18.17	27.32	32.58
02HC028	19.93	23.65	35.19	41.75
02HC030	47.75	59.06	96.76	119.67
02HC031	32.76	40.36	65.5	80.69
02HC032	13.31	16.54	27.4	34.05
02HC033	16.88	21.75	39.16	50.45
02HD003	14.9	19.4	35.8	46.61
02HD008	26.81	36.3	73.37	99.33
02HD009	19.95	28.34	64.02	90.93
02HD013	14.9	19.44	36.08	47.09
02HG002	4.63	5.71	9.3	11.47

Table 5.17 Relative difference of flood discharge calculated for different flood return period between power law model and LP3 distribution

Gauging station	5-year return period	10-year return period	50-year return period	100-year return period
02EC008	−0.21	−0.16	0.08	0.23
02EC009	−0.06	−0.03	0.14	0.25
02EC010	−0.09	−0.05	0.15	0.28
02ED003	−0.03	0.01	0.21	0.35
02ED100	−0.01	0.02	0.24	0.38
02HB001	−0.01	0.02	0.28	0.46
02HB018	0.07	0.16	0.55	0.80
02HB025	−0.01	0.06	0.42	0.66
02HC003	−0.04	−0.12	−0.17	−0.16
02HC009	0.00	−0.01	0.11	0.22
02HC018	−0.03	−0.02	0.14	0.26

(continued)

Table 5.17 (continued)

Gauging station	5-year return period	10-year return period	50-year return period	100-year return period
02HC019	0.01	0.00	0.16	0.28
02HC022	−0.08	−0.04	0.23	0.41
02HC024	−0.12	−0.11	0.01	0.10
02HC027	−0.10	−0.08	0.04	0.12
02HC028	0.03	0.07	0.31	0.47
02HC030	−0.15	−0.12	0.09	0.24
02HC031	−0.03	−0.02	0.11	0.19
02HC032	−0.07	−0.03	0.19	0.34
02HC033	−0.11	−0.06	0.16	0.29
02HD003	0.08	0.11	0.36	0.53
02HD008	−0.01	0.05	0.46	0.75
02HD009	0.13	0.20	0.52	0.74
02HD013	0.13	0.21	0.57	0.81
02HG002	0.51	0.60	0.94	1.14

5.6 Apply Generalized Pareto (GP) Distribution for Partial-Duration Series Flood Frequency Analysis

5.6.1 Introduction to GP Distribution for Flood Frequency Analysis

According to Pickands (1975) the partial-duration series over a high threshold follow the Generalized Pareto (GP) distribution:

$$f(x) = \begin{cases} \frac{1}{\sigma}\left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-1-1/k} & k \neq 0 \\ \frac{1}{\sigma}\exp\left(-\frac{(x-\mu)}{\sigma}\right) & k = 0 \end{cases} \tag{5.11}$$

where k, σ, μ are the shape, scale, and location parameters respectively.

And the Cumulative Distribution Function (CDF) is:

$$F(x) = \begin{cases} 1 - \left(1 + k\frac{(x-\mu)}{\sigma}\right)^{-1/k} & k \neq 0 \\ 1 - \exp\left(-\frac{(x-\mu)}{\sigma}\right) & k = 0 \end{cases} \tag{5.12}$$

5.6.2 GP Distribution Parameter Determination by Maximum Likelihood Estimation (MLE)

In order to apply the GP distribution for flood frequency analysis, the independent partial duration series flood event is extracted as the study datasets using the method described in the previous section. EasyFit software (MathWave Technologies 2009) and MLM were used to fit the GP distribution to the extracted study datasets. Table 5.18 shows the parameters of fitted GP distribution for each gauging station. The correlation coefficients between model fitted lines and observations are also shown in Table 5.18. The data range from 0.94 to 0.99 for all gauging stations; for most gauging stations, the range is from 0.97 to 0.99, indicating very good agreement between model fitted lines and observations.

Table 5.18 Model parameters of fitted GP distribution curve

Gauging station	Number of years	k	Σ	μ	GP type	Correlation coefficients
02EC008	19	0.21374	6.7533	13.131	II	0.9805
02EC009	40	0.17546	6.4076	13.54	II	0.9893
02EC010	33	0.07339	1.8798	4.0384	II	0.9916
02ED003	55	0.14774	34.92	77.883	II	0.9859
02ED100	25	−0.39227	5.0498	6.773	III	0.9787
02HB001	90	0.0825	7.901	15.103	II	0.9884
02HB018	22	−0.28249	16.029	20.765	III	0.9882
02HB025	16	0.02421	15.426	30.58	III	0.9811
02HC003	57	0.36071	29.572	85.841	II	0.9445
02HC009	51	0.17028	6.7542	18.746	II	0.9856
02HC018	40	−0.50897	10.61	10.871	III	0.9932
02HC019	43	−0.14015	8.5562	14.06	III	0.9447
02HC022	43	0.05016	9.8844	17.524	II	0.9803
02HC024	43	0.16839	11.05	30.846	II	0.9902
02HC027	38	−0.13865	4.4357	9.4072	III	0.9805
02HC028	41	−0.30746	6.2767	12.413	III	0.9817
02HC030	39	0.09711	11.724	28.041	II	0.9890
02HC031	35	0.0688	8.1479	19.362	II	0.9792
02HC032	30	−0.03574	3.8805	7.5234	III	0.9738
02HC033	39	0.14568	4.5123	8.9412	II	0.9813
02HD003	44	0.20727	3.5022	7.9421	II	0.9694
02HD008	46	0.1607	7.9405	12.46	II	0.9887
02HD009	35	0.43483	3.7366	9.2737	II	0.9615
02HD013	23	0.01853	4.5868	7.4646	II	0.9863
02HG002	11	−0.06584	0.96982	2.0756	II	0.9428

5.6.3 Evaluation of GP Distribution for Flood Frequency Analysis in the ORM Area

A further evaluation was applied to investigate the performance of the GP distribution in estimating flood peak discharge for different return periods. Tables 5.19 and 5.20 compare the observations and GP model fitted lines of flood discharge for approximately 10-year and 40-year return periods.

For the 10-year return period, for most stations except 02HC003, 02HD009 and 02HD013, the GP fitted lines generally match well with the observation flood discharge.

For the 40-year return period, for most stations except 02EC008, 02HC003, 02HC019, 02HD003 and 02HD009, the GP model fitted lines generally match the observation flood discharge.

Table 5.19 Comparison of observations and GP model fitted lines of flood discharge for approximate 10-year return period

Gauging station	Observed flood discharge (m^3/s)	GP prediction (m^3/s)	Error	Relative error
02EC008	30.3	32.178	1.88	0.06
02EC009	34.6	32.74	−1.86	−0.05
02EC010	9.15	8.75	−0.40	−0.04
02ED003	177	173.66	−3.34	−0.02
02ED100	14	14.43	0.43	0.03
02HB001	34	35.14	1.14	0.03
02HB018	50	47.90	−2.10	−0.04
02HB025	69	64.14	−4.86	−0.07
02HC003	171	199.26	28.26	0.17
02HC009	38.5	37.79	−0.71	−0.02
02HC018	25.2	25.26	0.06	0.00
02HC019	29.1	30.90	1.80	0.06
02HC022	43.9	41.65	−2.25	−0.05
02HC024	66.1	61.93	−4.17	−0.06
02HC027	19.3	18.15	−1.15	−0.06
02HC028	22.1	23.09	0.99	0.04
02HC030	62.3	56.90	−5.40	−0.09
02HC031	40.3	39.69	−0.61	−0.02
02HC032	18.1	15.76	−2.34	−0.13
02HC033	22.2	21.29	−0.91	−0.04
02HD003	20	18.28	−1.72	−0.09
02HD008	34.3	33.50	−0.80	−0.02
02HD009	31.5	23.12	−8.38	−0.27
02HD013	19.5	17.38	−2.12	−0.11
02HG002	4	4.07	0.07	0.02

Table 5.20 Comparison of observations and GP model fitted lines of flood discharge for approximate 40-year return period

Gauging station	Observed flood discharge (m^3/s)	GP prediction (m^3/s)	Error	Relative error
02EC008	77.5	48.389	−29.11	−0.38
02EC009	53.4	49.57	−3.83	−0.07
02EC010	12.8	12.56	−0.24	−0.02
02ED003	254	238.32	−15.68	−0.06
02ED100	–	–	–	–
02HB001	56.6	51.58	−5.02	−0.09
02HB018	54.1	56.44	2.34	0.04
02HB025	–	–	–	–
02HC003	218	294.29	76.29	0.35
02HC009	46.2	47.70	1.50	0.03
02HC018	30	28.87	−1.13	−0.04
02HC019	49.5	39.83	−9.67	−0.20
02HC022	60.3	55.419	−4.88	−0.08
02HC024	86.7	83.66	−3.04	−0.04
02HC027	22.2	21.73	−0.47	−0.02
02HC028	26.6	26.70	0.10	0.00
02HC030	80.6	77.02	−3.58	−0.04
02HC031	54.4	51.67	−2.73	−0.05
02HC032	21.1	20.31	−0.79	−0.04
02HC033	35.3	29.59	−5.71	−0.16
02HD003	33	26.00	−7.00	−0.21
02HD008	59.9	55.70	−4.20	−0.07
02HD009	52.9	35.52	−17.38	−0.33
02HD013	–	–	–	–
02HG002	–	–	–	–

5.6.4 Application of the GP Model for Flood Frequency Prediction in the ORM Area

The fitted GP distribution can also be used to estimate flood discharge for different return periods. Table 5.21 lists the flood discharge calculated for different return period flood events, including 5-year, 10-year, 50-year, and 100-year events.

A further evaluation was applied to compare the GP distribution with the LP3 distribution for different return periods, including 5-year, 10-year, 50-year, and 100-year events, to compare the flood frequency predictive ability (Table 5.22).

For short return period predictions (5-year and 10-year), the GP distribution prediction results are generally less than that of the LP3 distribution; and for long return period predictions (50-year and 100-year), the GP distribution prediction

Table 5.21 Flood discharge calculated for different return periods using GP distribution

Gauging station	5-year return period (m^3/s)	10-year return period (m^3/s)	50-year return period (m^3/s)	100-year return period (m^3/s)
02EC008	26.103	33.221	54.441	66.084
02EC009	25.456	31.72	49.568	58.95
02EC010	7.2498	8.7541	12.557	14.338
02ED003	141.33	173.66	262.81	308.24
02ED100	12.799	14.429	16.871	17.532
02HB001	28.702	35.139	51.583	59.366
02HB018	41.494	47.899	58.715	62.057
02HB025	55.896	67.107	93.875	105.73
02HC003	150.36	191.98	340.03	435.52
02HC009	31.252	37.787	56.297	65.97
02HC018	22.528	25.259	28.87	29.717
02HC019	26.389	30.899	39.827	43.094
02HC022	34.092	41.65	60.246	68.73
02HC024	51.274	61.927	92.03	107.73
02HC027	15.806	18.151	22.801	24.505
02HC028	20.382	22.771	26.696	27.873
02HC030	48.464	58.291	83.831	96.122
02HC031	33.229	39.691	55.938	63.509
02HC032	13.593	16.101	21.691	24.001
02HC033	17.125	21.286	32.732	38.551
02HD003	14.633	18.277	29.06	34.934
02HD008	27.044	34.585	55.7	66.617
02HD009	17.982	24.068	47.77	64.333
02HD013	14.958	18.255	26.074	29.515
02HG002	3.5567	4.1477	5.4204	5.9283

Table 5.22 Relative difference of flood discharge calculated for different flood return periods between GP distribution and LP3 distribution

Gauging station	5-year return period	10-year return period	50-year return period	100-year return period
02EC008	−0.21	−0.21	−0.14	−0.10
02EC009	−0.07	−0.07	−0.06	−0.05
02EC010	−0.08	−0.07	−0.04	−0.03
02ED003	−0.02	−0.02	0.03	0.06
02ED100	0.04	−0.01	−0.10	−0.14
02HB001	0.02	0.02	0.07	0.11
02HB018	0.08	0.06	−0.03	−0.07
02HB025	−0.01	0.01	0.09	0.13
02HC003	0.00	−0.04	−0.01	0.03
02HC009	0.00	−0.04	−0.03	0.01
02HC018	0.05	−0.03	−0.18	−0.23

(continued)

Table 5.22 (continued)

Gauging station	5-year return period	10-year return period	50-year return period	100-year return period
02HC019	0.05	−0.01	−0.11	−0.14
02HC022	−0.07	−0.08	−0.07	−0.06
02HC024	−0.12	−0.13	−0.10	−0.08
02HC027	−0.06	−0.08	−0.14	−0.16
02HC028	0.05	0.03	0.00	−0.02
02HC030	−0.14	−0.13	−0.05	−0.01
02HC031	−0.01	−0.03	−0.05	−0.06
02HC032	−0.05	−0.06	−0.06	−0.06
02HC033	−0.10	−0.08	−0.03	−0.01
02HD003	0.06	0.05	0.11	0.15
02HD008	0.00	0.00	0.11	0.17
02HD009	0.02	0.02	0.14	0.23
02HD013	0.13	0.13	0.14	0.13
02HG002	0.16	0.16	0.13	0.11

results are less than that of LP3 distribution. A further investigation finds that the relative difference between these two distributions for different return periods is larger than the relative difference between the GEV distribution and the LP3 distribution. Some of them are greater than 10 %, and only a few are less than 5 %, mostly for 5-year and 10-year return periods.

A further evaluation is applied to compare the GP distribution with the PL distribution for different return periods, including 5-year, 10-year, 50-year and 100-year, to compare the flood frequency predictive abilities. Generally, for short return period predictions (5-year and 10-year), the relative differences between the GP distribution and the PL distribution are relatively small. The GP distribution prediction results are generally greater than that of the PL distribution for 5-year return period, but generally less than that of the PL distribution for 10-year return period (Table 5.23). However, for long return period predictions (50-year and 100-year), GP distribution prediction results are generally less than that of PL distribution. A further investigation finds the relative differences between these two distributions for long return periods are relatively large, most of them are greater than 20 % (Table 5.24).

Figure 5.10 shows the fitted GP and PL line at 02HB001 gauging station, and it is observed that for less than 10-year return periods, there is not too much difference of model prediction for different distributions. But for greater than 10-year return periods, the prediction of PL model is larger than the prediction of GP distribution.

Table 5.23 Relative difference of flood discharge calculated for different flood return periods between GP distribution and PL distribution

Gauging station	5-year return period	10-year return period	50-year return period	100-year return period
02EC008	−0.01	−0.05	−0.20	−0.27
02EC009	0.00	−0.04	−0.17	−0.24
02EC010	0.01	−0.02	−0.17	−0.25
02ED003	0.01	−0.03	−0.15	−0.22
02ED100	0.05	−0.02	−0.27	−0.38
02HB001	0.03	−0.01	−0.16	−0.24
02HB018	0.01	−0.09	−0.38	−0.48
02HB025	0.00	−0.05	−0.24	−0.32
02HC003	0.04	0.08	0.19	0.24
02HC009	−0.01	−0.03	−0.13	−0.17
02HC018	0.09	0.00	−0.28	−0.39
02HC019	0.04	−0.02	−0.23	−0.33
02HC022	0.01	−0.04	−0.24	−0.33
02HC024	0.00	−0.02	−0.11	−0.16
02HC027	0.04	0.00	−0.17	−0.25
02HC028	0.02	−0.04	−0.24	−0.33
02HC030	0.01	−0.01	−0.13	−0.20
02HC031	0.01	−0.02	−0.15	−0.21
02HC032	0.02	−0.03	−0.21	−0.30
02HC033	0.01	−0.02	−0.16	−0.24
02HD003	−0.02	−0.06	−0.19	−0.25
02HD008	0.01	−0.05	−0.24	−0.33
02HD009	−0.10	−0.15	−0.25	−0.29
02HD013	0.00	−0.06	−0.28	−0.37
02HG002	−0.23	−0.27	−0.42	−0.48

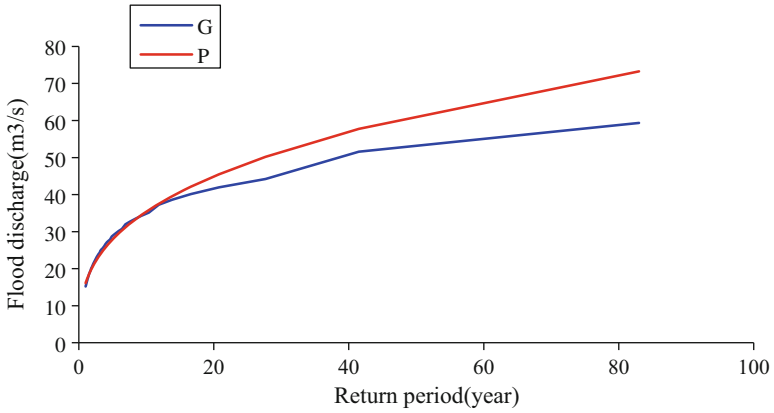


Fig. 5.10 Fitted GP and PL line at 02HB001 gauging station

Table 5.24 Comparison of correlation coefficients of different flood frequency analysis methods

Gauging station	Correlation coefficients of LP3	Correlation coefficients of GEV	Correlation coefficients of PL	Correlation coefficients of GP
02EC008	0.9438	0.9388	0.9646	0.9805
02EC009	0.9820	0.9788	0.9832	0.9893
02EC010	0.9916	0.9914	0.9801	0.9916
02ED003	0.9756	0.9763	0.9822	0.9859
02ED100	0.9869	0.9878	0.9078	0.9787
02HB001	0.9766	0.9864	0.9851	0.9884
02HB018	0.9801	0.9806	0.9535	0.9882
02HB025	0.9789	0.9801	0.9726	0.9811
02HC003	0.9411	0.9342	0.9511	0.9445
02HC009	0.9774	0.9827	0.9643	0.9856
02HC018	0.9705	0.9789	0.9732	0.9932
02HC019	0.9719	0.9709	0.9497	0.9447
02HC022	0.9846	0.9838	0.9817	0.9803
02HC024	0.9946	0.9844	0.9922	0.9902
02HC027	0.9742	0.9759	0.9586	0.9805
02HC028	0.9885	0.9896	0.9371	0.9817
02HC030	0.9730	0.9788	0.9849	0.9890
02HC031	0.9732	0.9750	0.9781	0.9792
02HC032	0.9799	0.9800	0.9626	0.9738
02HC033	0.9706	0.9654	0.9833	0.9813
02HD003	0.9544	0.9630	0.9860	0.9694
02HD008	0.9839	0.9925	0.9748	0.9887
02HD009	0.9432	0.9576	0.9798	0.9615
02HD013	0.9702	0.9700	0.9722	0.9863
02HG002	0.9509	0.949	0.9408	0.9428

5.7 Comparison of Flood Frequency Analysis Using LP3, GEV, PL and GP Methods

In this study, both annual extreme series and partial-duration series are used for flood frequency analysis. For annual extreme series, the LP3 and GEV methods are used for flood frequency analysis, and the PL and GP methods are used for partial-duration series flood frequency analysis. Reliability and predictive ability are compared for these four flood frequency analysis methods to find which method performs better for flood frequency analysis in the ORM area.

Table 5.24 compares correlation coefficients of different flood frequency analysis methods. It is observed that, for most gauging stations, there is not too much difference of correlation coefficients for different flood frequency analysis methods in the ORM area.

Table 5.25 Correlation coefficients of observations and model fitted lines for less than approximate 10-year and greater than approximate 10-year return period

	Correlation coefficients for less than approximate 10 year return period				Correlation coefficients for greater than approximate 40 year return period			
Gauging station	LP3	GEV	PL	GP	LP3	GEV	PL	GP
02EC008	0.9438	0.9388	0.9587	0.9724	–	–	–	–
02EC009	0.9691	0.9671	0.9880	0.9764	0.9740	0.9772	0.9649	0.9589
02EC010	0.9907	0.9905	0.9942	0.9864	0.9297	0.9318	0.9267	0.9387
02ED003	0.9883	0.989	0.9921	0.9925	0.6938	0.6885	0.7516	0.8390
02ED100	0.9894	0.9852	0.9285	0.9683	–	–	–	–
02HB001	0.9937	0.9954	0.9878	0.9955	0.9874	0.9913	0.9828	0.9852
02HB018	0.9779	0.9812	0.9736	0.9885	–	–	–	–
02HB025	0.9759	0.9771	0.9859	0.9866	–	–	–	–
02HC003	0.9417	0.9417	0.9488	0.9432	0.7269	0.7182	0.8707	0.8405
02HC009	0.9662	0.9733	0.9612	0.9695	0.9874	0.9871	0.8854	0.9846
02HC018	0.9909	0.9811	0.9229	0.9930	0.8671	0.8740	0.9244	0.9603
02HC019	0.9802	0.9809	0.9521	0.9863	0.9722	0.9694	0.9521	0.9432
02HC022	0.9848	0.982	0.9865	0.9883	0.7781	0.7769	0.9206	0.8261
02HC024	0.9909	0.9909	0.9964	0.9928	0.9672	0.9673	0.9480	0.9695
02HC027	0.9747	0.9756	0.9505	0.9829	0.7278	0.7304	0.9264	0.9240
02HC028	0.9911	0.9891	0.9469	0.9760	0.8792	0.8806	0.9083	0.7425
02HC030	0.9593	0.9622	0.9871	0.9799	0.9740	0.9756	0.9653	0.9674
02HC031	0.9885	0.9886	0.9907	0.9898	0.9146	0.9127	0.9532	0.9341
02HC032	0.9749	0.9742	0.9630	0.9562	0.9644	0.9631	0.9886	0.9397
02HC033	0.9714	0.9708	0.9966	0.9913	0.9614	0.9115	0.9234	0.8709
02HD003	0.9416	0.9437	0.9750	0.9619	0.9579	0.9631	0.9604	0.9501
02HD008	0.9848	0.9845	0.9683	0.9624	0.9689	0.9815	0.9851	0.9893
02HD009	0.9716	0.9295	0.9635	0.9327	0.9549	0.9503	0.9076	0.8819
02HD013	0.9580	0.9510	0.9895	0.9753	–	–	–	–
02HG002	0.9402	0.9390	0.9258	0.9374	–	–	–	–

The correlation coefficients of observations and model fitted lines for less than approximately 10-year and greater than approximately 10-year return periods for different flood frequency analysis methods are listed in Table 5.25. The goodness-of-fit tests are applied for these four models to find which model performs better than other models for less than approximately 10-year and greater than approximately 10-year return period flood frequency analysis in the ORM area. The criteria of goodness are both the smallness in median and the narrowness (max-min ranges) of the relative error (Haktanir and Horlacher 1993).

The comparison of different flood frequency analysis methods suggests that there are no significant differences of the LP3, GEV, PL and GP models for less than approximate 10-year return period flood frequency analysis in the ORM area. The median of the LP3, GEV, PL and GP models are all around 0.97, there are

Table 5.26 Comparison of correlation coefficients of observations and model fitted lines for less than approximate 10-year and greater than 10-year return period

	Correlation coefficients for less than approximate 10 year return period				Correlation coefficients for greater than approximate 10 year return period			
Gauging station	LP3	GEV	PL	GP	LP3	GEV	PL	GP
02EC008	0.9438	0.9388	0.9587	0.9724	–	–	–	–
02EC009	0.9691	0.9671	0.9880	0.9764	0.9740	0.9772	0.9649	0.9589
02EC010	0.9907	0.9905	0.9942	0.9864	0.9297	0.9318	0.9267	0.9387
02ED003	0.9883	0.989	0.9921	0.9925	0.6938	0.6885	0.7516	0.8390
02ED100	0.9894	0.9852	0.9285	0.9683	–	–	–	–
02HB001	0.9937	0.9954	0.9878	0.9955	0.9874	0.9913	0.9828	0.9852
02HB018	0.9779	0.9812	0.9736	0.9885	–	–	–	–
02HB025	0.9759	0.9771	0.9859	0.9866	–	–	–	–
02HC003	0.9417	0.9417	0.9488	0.9432	0.7269	0.7182	0.8707	0.8405
02HC009	0.9662	0.9733	0.9612	0.9695	0.9874	0.9871	0.8854	0.9846
02HC018	0.9909	0.9811	0.9229	0.9930	0.8671	0.8740	0.9244	0.9603
02HC019	0.9802	0.9809	0.9521	0.9863	0.9722	0.9694	0.9521	0.9432
02HC022	0.9848	0.982	0.9865	0.9883	0.7781	0.7769	0.9206	0.8261
02HC024	0.9909	0.9909	0.9964	0.9928	0.9672	0.9673	0.9480	0.9695
02HC027	0.9747	0.9756	0.9505	0.9829	0.7278	0.7304	0.9264	0.9240
02HC028	0.9911	0.9891	0.9469	0.9760	0.8792	0.8806	0.9083	0.7425
02HC030	0.9593	0.9622	0.9871	0.9799	0.9740	0.9756	0.9653	0.9674
02HC031	0.9885	0.9886	0.9907	0.9898	0.9146	0.9127	0.9532	0.9341
02HC032	0.9749	0.9742	0.9630	0.9562	0.9644	0.9631	0.9886	0.9397
02HC033	0.9714	0.9708	0.9966	0.9913	0.9614	0.9115	0.9234	0.8709
02HD003	0.9416	0.9437	0.9750	0.9619	0.9579	0.9631	0.9604	0.9501
02HD008	0.9848	0.9845	0.9683	0.9624	0.9689	0.9815	0.9851	0.9893
02HD009	0.9716	0.9295	0.9635	0.9327	0.9549	0.9503	0.9076	0.8819
02HD013	0.9580	0.9510	0.9895	0.9753	–	–	–	–
02HG002	0.9402	0.9390	0.9258	0.9374	–	–	–	–
MEDAIN	**0.9759**	**0.9771**	**0.9736**	**0.9799**	**0.9579**	**0.9503**	**0.9267**	**0.9397**
RANGE	**0.0535**	**0.0659**	**0.0737**	**0.0628**	**0.2936**	**0.3028**	**0.237**	**0.2468**

almost no difference. This suggests that the LP3, GEV, PL and GP models have almost same performance for less than approximate 10-year return period flood frequency analysis (Table 5.26).

Similarly, the comparison of different flood frequency analysis methods also suggests that there are no significant differences of the LP3, GEV, PL and GP models for greater than approximate 10-year return period flood frequency analysis in the ORM area. The medians of the LP3, GEV, PL and GP models are all greater than 0.92, which shows there are good correlation between observations and model fitted lines (Table 5.26).

Furthermore, the discharge calculated for different flood return periods, including 5-year, 10-year, 50-year and 100-year, is compared. Table 5.27 compares

Table 5.27 Comparison of flood discharge calculated for 5-year and 10-year return periods

Gauging station	Flood discharge calculated for 5-year return period (m^3/s)				Flood discharge calculated for 10-year return period (m^3/s)			
	LP3	GEV	PL	GP	LP3	GEV	PL	GP
02EC008	33.2	31.86	26.35	26.1	41.95	40.74	35.1	33.22
02EC009	27.2	26.91	25.49	25.46	34.07	33.78	33	31.72
02EC010	7.86	7.831	7.16	7.25	9.438	9.444	8.97	8.754
02ED003	145	143.7	140.4	141.3	177.3	176.2	178	173.7
02ED100	12.3	12.36	12.17	12.8	14.52	14.43	14.8	14.43
02HB001	28	27.4	27.89	28.7	34.57	34.13	35.4	35.14
02HB018	38.5	38.65	41.14	41.49	45.39	45.5	52.8	47.9
02HB025	56.5	56.28	55.89	55.9	66.52	66.49	70.8	67.11
02HC003	151	148	144.4	150.4	201	197.3	177	192
02HC009	31.4	31.09	31.43	31.25	39.57	38.79	39	37.79
02HC018	21.4	21.46	20.76	22.53	25.93	25.46	25.3	25.26
02HC019	25.1	25.19	25.36	26.39	31.34	30.55	31.4	30.9
02HC022	36.6	36.55	33.68	34.09	45.29	44.81	43.6	41.65
02HC024	58.3	57.9	51.3	51.27	71.03	70.85	63.4	61.93
02HC027	16.9	16.89	15.24	15.81	19.83	19.85	18.2	18.15
02HC028	19.4	19.45	19.93	20.38	22.05	22	23.7	22.77
02HC030	56.5	55.46	47.75	48.46	66.99	66.56	59.1	58.29
02HC031	33.6	33.38	32.76	33.23	41.07	40.89	40.4	39.69
02HC032	14.3	14.33	13.31	13.59	17.07	17	16.5	16.1
02HC033	19.1	18.89	16.88	17.13	23.24	23.1	21.8	21.29
02HD003	13.8	13.57	14.9	14.63	17.42	17.14	19.4	18.28
02HD008	27.1	25.92	26.81	27.04	34.56	33.95	36.3	34.59
02HD009	17.6	16.88	19.95	17.98	23.68	22.8	28.3	24.07
02HD013	13.2	13.01	14.9	14.96	16.1	16.01	19.4	18.26
02HG002	3.07	3.036	4.63	3.557	3.569	3.578	5.71	4.148
02EC008	33.2	31.86	26.35	26.1	41.95	40.74	35.1	33.22

different discharges calculated for 5-year and 10-year return periods. And Table 5.28 compares different discharges calculated for 50-year and 100-year return periods.

Generally, the LP3 model prediction results are the highest among these four models, and the GEV model prediction results are the second; the PL model prediction results are the third, and the GP distribution prediction results are the fourth. We also find that annual extreme series prediction results are generally greater than that of part-duration series.

Long return period predictions (50-year and 100-year) are different from short return period predictions; generally, the PL model prediction results are the highest among these four distributions, the GEV distribution prediction results are the

Table 5.28 Comparison of flood discharge calculated for 50-year and 100-year flood return periods

Gauging station	Flood discharge calculated for 50-year return period (m^3/s)				Flood discharge calculated for 100-year return period (m^3/s)			
	LP3	GEV	PL	GP	LP3	GEV	PL	GP
02EC008	63.55	66.25	68.32	54.44	73.73	80.119	91	66.1
02EC009	52.64	53.49	60.03	49.57	62.19	64.204	77.7	59
02EC010	13.14	13.36	15.17	12.56	14.81	15.171	19	14.3
02ED003	255.1	256.6	309.9	262.8	290.8	294.64	393	308
02ED100	18.78	18.31	23.24	16.87	20.39	19.711	28.2	17.5
02HB001	48.16	50.18	61.55	51.58	53.58	57.503	78.1	59.4
02HB018	60.57	60.25	93.96	58.72	67.05	66.347	120	62.1
02HB025	86.24	87.29	122.7	93.88	93.78	95.448	156	106
02HC003	344.5	351.7	286.4	340	421	442.52	352	436
02HC009	57.85	56.53	64.39	56.3	65.61	64.375	79.9	66
02HC018	35.01	33.03	39.99	28.87	38.51	35.781	48.7	29.7
02HC019	44.54	41.38	51.63	39.83	49.92	45.587	63.9	43.1
02HC022	64.55	63.28	79.1	60.25	72.75	71.21	102	68.7
02HC024	102.1	103.7	103.4	92.03	116.5	119.66	128	108
02HC027	26.39	26.25	27.32	22.8	29.19	28.918	32.6	24.5
02HC028	26.8	26.41	35.19	26.7	28.48	27.88	41.8	27.9
02HC030	88.39	91.72	96.76	83.83	96.84	102.67	120	96.1
02HC031	59.17	59.78	65.5	55.94	67.58	68.846	80.7	63.5
02HC032	22.99	22.57	27.4	21.69	25.43	24.794	34.1	24
02HC033	33.86	34.35	39.16	32.73	39.01	40.072	50.5	38.6
02HD003	26.28	26.7	35.8	29.06	30.46	31.573	46.6	34.9
02HD008	50.4	55.34	73.37	55.7	56.73	66.188	99.3	66.6
02HD009	42.06	42.83	64.02	47.77	52.39	55.432	90.9	64.3
02HD013	22.93	23.54	36.08	26.07	26.01	27.145	47.1	29.5
02HG002	4.784	5.022	9.3	5.42	5.351	5.7527	11.5	5.93
02EC008	63.55	66.25	68.32	54.44	73.73	80.119	91	66.1

second, the LP3 distribution prediction results are the third, and GP distribution prediction results are the last.

Figure 5.11 shows the fitted GP, PL, GEV and LP3 line at 02HB001 gauging station, it is observed that for less than 10-year return periods, there is not too much difference of model prediction for different distributions. But for long return periods, for example, larger than 40-year return period, the prediction of the LP3 distribution is the least, then the GEV distribution, then the GP distribution and the PL model prediction is the largest at 02HB001 gauging station.

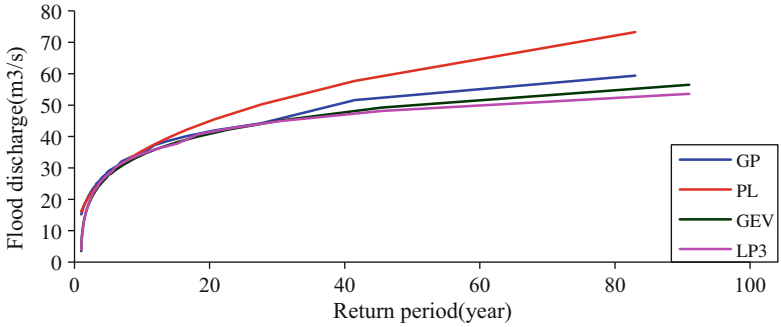


Fig. 5.11 Fitted GP,PL,GEV and LP3 line at 02HB001 gauging station

Acknowledgments The author expresses the appreciation of funds received from Hainan peovince major science and technology projects (#ZDKJ2016015-1), the Hainan Province Natural Science Foundation (#20164178), the Sanya City Key Laboratory projects (#L1404) and the Sanya City science and technology cooperation projects (#2015YD19 and #2014YD08).

References

Adamowski K, Liang GC, Patry GG (1998) Annual maxima and partial duration flood series analysis by parametric and non-parametric methods. Hydrol Process 12:1685–1699

Ashkar F, El-Jabi N, Bobee B (1987) On the choice between annual flood series and peaks over threshold series in flood frequency analysis. In: Feldman AD (ed) Proceeding Engineering Hydrology Virginia, ASCE, New York, 276–280, 3–7 August 1987

Beven KJ (2001) Rainfall-runoff modelling: the primer. Wiley, Chichester/New York

Birikundavyi S, Rousselle J (1997a) A technique for selecting the threshold in applications of the poisson-exponential flood model. Can J Civ Eng 24:12–19

Birikundavyi S, Rousselle J (1997b) Use of partial duration series for single-station and regional analysis of floods. J Hydrol Eng ASCE 2:68–75

Cheng Q (1997) Fractal/multifractal and spatial analysis. In: Vare Pawlowsky Glahn (ed) Proceeding International Association for Mathematical Geology Meeting, the International Centre for Numeric Methods in Engineering (CIMNE), Barcelona, I:57–72

Cheng Q, Li L, Wang L (2009) Characterization of peak flow events with local singularity model. Nonlinear Process Geophys 16:503–513

Cheng Q, Agterberg FP, Ballantyne SB (1994) The separation of geochemical anomalies from background by fractal methods. J Geochem Explor 51:109–130

Choulakian V (1990) On the distribution of flood volume in partial duration series analysis of flood phenomena. Stoch Hydrol Hydraul 4:217–226

Chow VT, Maidment DR, Mays LR (1988) Applied hydrology. McGraw-Hill Book Company, New York

Evertsz CJG, Mandelbrot BB (1992) Multifractal measures. In: Peitgen H-O, Jurgens H, Saupe D (eds) Chaos and Fractals. Springer, New York, pp 922–953

Haktanir T, Horlacher HB (1993) Evaluation of various distributions for flood frequency analysis. Hydrol Sci J 38:15–32

Jenkinson AF (1955) The frequency distribution of the annual maximum (or minimum) value of meteorological elements. Q J R Meteorol Soc 81(348):158–171

Kidson R, Richards KS (2005) Flood frequency analysis: assumptions and alternatives. Prog Phys Geogr 29(3):392–410
Kigami J (2001) Analysis on fractals. Cambridge University Press, Cambridge
Lang M, Ouarda TBMJ, Bobee B (1999) Towards operational guidelines for over-threshold modeling. J Hydrol 225:103–117
Li LY (2005) Analysis on extreme hydrological events in the Oak Ridges Moraine area. Unpublished M.Sc. thesis, York University
Malamud BD, Turcotte DL, Barton C (1996) The 1993 Mississippi river flood: a one hundred or a one thousand year event? Environ Eng Geosci 2:479–486
Malamud BD, Turcotte DL (2006) The applicability of power-law frequency statistics to floods. J Hydrol 322:168–180
Maidment DR (1993) Handbook of hydrology. McGraw-Hill, New York
Mathier L, Roy R, Bobée B, Perron H, Fortin V (1993) Estimation régionale des crues: description des banques de donnnes hydrométriques, météorologiques et physiographiques pour le Québec et l'Ontario. INRS-Eau, Sainte-Foy, Que., Internal Report I-123
MathWave Technologies (2009) http://www.mathwave.com/products/easyfit.html
Pickands J (1975) Statistical inference using extreme order statistics. Ann Stat 3(1):119–131
Turcotte DL, Greene L (1993) A scale-invariant approach to flood-frequency analysis. Stoch Hydrol Hydraul 7:33–40
Viessman W, Lewis GL (2003) Introduction to hydrology. Pearson Education Inc., New York
Weibull W (1939) A statistical theory of strength, Ingeniors Vetenskaps Akademien (The Royal Swedish Institute for Engineering Research), proceedings no. 51, pp 5–45
Yevjevich V, Taesombut V (1979) Information on flood peaks in daily flow series. In: McMean EA, Hipel KW, Unny TE (eds) Input for risk analysis in water systems. Water Resources Publications, Fort Collins, pp 171–192

Chapter 6
Spatial Decision Making and Analysis for Flood Forecasting

Lei Wang and Xin Zhang

Abstract The application of flood forecasting models requires the efficient management of large spatial and temporal datasets, involving data acquisition, storage, processing, analysis and display of model results. Difficulty in linking data, analysis tools, and models is one of the barriers to be overcome in developing an integrated flood forecasting system. The current revolution in technology and the online availability of spatial data facilitate Canadians' need for information sharing in support of decision making. This need has resulted in studies demonstrating the suitability of the web as a medium for implementation of flood forecasting. Web-based Spatial Decision Support Services (WSDSS) provides comprehensive support for information retrieval, model analysis and extensive visualization functions for decision-making support and information services. This chapter develops a prototype WSDSS that integrates models, analytical tools, databases, graphical user interfaces, and spatial decision support services to help the public and decision makers to easily access flood and flood-threatened information. Flood WSDSS helps to mitigate flood disasters through river runoff prediction, flood forecasting, and flood information (flood discharge, water level and flood frequency) dissemination. The ultimate aim of this system is to improve access to flood model results by the public and decision makers.

Keywords Flood forecasting • GIS • Web • Spatial decision support system

L. Wang (✉)
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China

Hainan Province Key Laboratory of Earth Observation, Sanya Institute of Remote Sensing, Sanya 572029, Hainan Province, China
e-mail: wanglei98@radi.ac.cn

X. Zhang
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zhangxin@radi.ac.cn

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_6

6.1 Introduction

Chapters 4 and 5 have applied SCS CN model for runoff prediction, LP3, GEV, PL and GP models for flood frequency analysis and prediction, and the CA and singularity fractal model for flood threshold selection and characteristics description. However, these flood forecasting models are not accessible to the public and decision-makers. Due to the background difference, it is very difficult for the public and decision-makers to use these models directly for flood forecasting and management as they require familiarity with flood forecasting models and analysis tools. Difficulty in linking data and analysis tools and models is one of the barriers to be overcome in developing an integrated flood forecasting system. The main reason for developing a new system and tool is that the public and decision makers cannot effectively utilize available models and data unless data, analysis tools, and models are already integrated.

Meanwhile, the current revolution in technology and online availability of spatial data, such as the National Water Data Archive and National Climate Data and Information Archive, has enabled Canadians to access a wide variety of spatial data through the Internet. This has resulted in research studies demonstrating the suitability of the web as a medium for the implementation of flood forecasting.

This chapter describes the architecture and components as well as the functionality of a WFSDSS that provides a comprehensive environment for flood forecasting in the ORM area, integrating information retrieval, analysis and model analysis for information sharing and decision-making support services. It is capable of non-expert implementation and permits users for web-based river runoff prediction, water level prediction and flood frequency prediction. The ultimate aim of this system is to improve access to flood model results to the public and decision makers. This WFSDSS will improve the understanding of environmental, planning, management issues, and emergency management and response associated with the ORM's water environment, and allow development of sustainable solutions.

6.2 Problems of Real-Time Flood Forecasting

The ORM area has faced the impacts of extreme hydrological events. Water-related disasters, i.e. floods, have been more devastating as far as deaths, suffering, and economical damages are concerned, than other natural hazards (earthquakes, volcanoes, etc.). Floods not only have an impact on the ORM's economic, environmental, and social well-being, and particularly public safety, but also exacerbate major environmental problems. Despite the progress in science and technology, human settlements are still vulnerable to floods. The losses increase due to the continuing development of costly infrastructure, rise in population density, and decrease of the buffering capacities because of deforestation, urbanization, draining wetlands, etc.

The mechanism of flood forecasting is complex, involving precipitation, drainage basin characterizes, land use/cover types and runoff discharge. The application of flood forecasting models requires the efficient management of large spatial and temporal datasets, which involves data acquisition, storage, processing, analysis and display of model results. The extensive datasets usually involve multiple organizations, but no single organization can collect and maintain all the multidisciplinary data. The possible usage of the available datasets remains limited, primarily because of the difficulty associated with the combining of data from diverse and distributed data sources. The current revolution in technology and online availability of spatial data has enabled Canadians to access a wide variety of spatial data through the Internet. Internet technology has been widely used for application development because of advantages such as platform independency, reductions in distribution costs and maintenance problems, ease of use, ubiquitous access and sharing of information by the worldwide user community (Peng and Tsou 2003). This revolution has resulted in research studies demonstrating the suitability of the web as a medium for implementation of flood forecasting.

With the construction of the CGDI, much spatial data and information can be accessed from distributed sources over the Internet to facilitate Canadians' need for information sharing in support of decision-making. The CGDI can provide the user with access to integrated and timely data and services, such as Web Map Service (WMS), Web Feature Service (WFS), Filter Encoding, Geodata Discovery Service, etc for decision-making support. Using metadata services, users can know what data are available, how they have been collected, and who is responsible for managing and distributing the data. Using geodata services, users can access and query the data from distributed sources. However, WMS, WFS and Filter Encoding cannot meet the requirements for on-line flood forecasting, new application is still required to integrate WMS, WFS, statistical analysis, modeling and spatial analysis for decision making support.

Advances in flood DSS can improve flood forecasting. Flood DSS improves the decision-making process by providing data display, analytical results, and model output to summarize critical flood information. However, flood DSS are often difficult to access for non-expert public and decision-makers because they cannot effectively utilize available models and data, unless data, analysis tools and models are already integrated. There are still some challenges for the public and decision-makers to use flood DSS for flood forecasting and flood management. These challenges include (1) friendly interface development, (2) real-time data access and flood map dissemination, (3) distributed data sharing and services, (4) model sharing for decision-making support, and (5) integrating data and modeling for decision-making support.

(1) Friendly interface development

A Graphical User Interface (GUI) is the interface for WSDSS to interact with users. A GUI should be designed to be user-friendly, allow people to easy access

and interact with this system and visualize flood and flood-threat information. A multiple-level web-based interface require development, through which the distributed database, the model base, and the spatial decision services can be integrated.

(2) Real-time data access and flood map dissemination

Flood forecasting and flood management require real-time data access for model input and development of decision-making services systems for real-time runoff prediction, water level prediction, and flood frequency prediction This helps to mitigate flood disasters and minimize the potential flood hazard.

(3) Distributed data sharing and services

The third challenge lies in that flood forecasting decision-making support require collaboration, because the data are distributed and managed by different organizations. For example, river runoff prediction requires diverse information that includes current and historical in-situ measurements, such as river runoff data from gauging stations and precipitation data from weather stations, which usually involves multiple organizations, so collaboration is required to facilitate an interactive decision-making process. However, it is not always possible to find all these datasets in one place or on one server. Sometimes, it is difficult to obtain the required model input data. In addition, a variety of data sources and multiple data formats of the available data present a formidable challenge to users who wish to compile information into a form that is usable for their applications. The possible usage of the available datasets remains limited primarily because of the difficulty associated with the combining of data from diverse and distributed data sources. Assuming that all the information is at the same scale and has been formatted according to the same standards, users can potentially overlay spatial information, which is originally collected and maintained by a separate organization to examine how the layers interrelate. Analyzing this layered information as an integrated whole can significantly aid decision makers in considering complex choices. The extensive data requirements have long been an obstacle to the timely and cost-effective use of a complex decision-making model, which is facilitating an increasing realism to distributed geodata sharing, supports the information exchange and knowledge sharing among different organizations through web.

(4) Model sharing for decision-making support

Flood forecasting includes rainfall-runoff model development, flood frequency prediction model development, and flood "threshold" selection to decide whether flood warnings should be issued. The complexity of flood forecasting requires decision makers to access a range of models, for example, the SCS model for river runoff prediction, LP3, GEV, PL and GP models for flood frequency prediction, CA and singularity fractal models for flood threshold selection and characteristics description. A Web-based system can be ideal for sharing these models for public information services and decision-making support.

(5) Integrating data and model for decision-making support

The last challenge is integrating data and models for decision-making support. Some hydrological models for flood forecasting have been developed and

successfully applied in the ORM area. However, most of these hydrological models are inaccessible to the public and decision makers. Users need to download and install models on their local desktop computers and learn how to operate these models. Sometimes, it is difficult to obtain the required model input data. Few of these models are well integrated with GIS and are capable of non-expert implementation. Difficulty in linking data and analysis tools and models is one of the barriers to be overcome in developing integrated spatial decision-making techniques. An integrated approach is still needed to integrate models, data, decision-making support, and interface to help decision makers and the public to easily access flood and Flood-threatened information. The integrated system with a common database and a common web-based interface would make flood management easier.

Thus, further research for web-based flood forecasting is still required to integrate information retrieval, analysis, model analysis, and extensive visualization functions for information sharing and decision-making support in the ORM area. WFSDSS provides the framework within which spatially distributed data accessed remotely to prepare model input and model calculation and evaluate model results for runoff prediction, water level prediction, and flood frequency prediction.

6.3 System Overview

The chapter represents an effort to develop a spatial decision support service for flood forecasting. Obviously, flood forecasting is not a new topic, one innovation of this chapter is to develop a web-based decision support services approach, concept model, and prototype system that integrate hydrological models, analytical tools, databases, interface, and decision-making support services to help decision makers and the public to easily access flood and flood-threaten information. This system integrates information retrieval, model analysis, and extensive visualization functions for information sharing and decision-making support. The ultimate aim of this system is to improve access to flood model results to the public and decision makers.

WFSDSS is an integrated system that integrates hydrological models, analytical tools, databases, graphical user interfaces, and spatial decision-support services to help the public and decision makers for flood forecasting. WFSDSS helps to mitigate flood disasters through river runoff prediction, flood forecasting, and disaster information (flood discharge, water level and flood frequency (return period)) dissemination.

Based on the challenges in analysis of a web-based hydrological model for flood management, the key advantages of this system should be:

1. Integrated system: This system consists of different components. The runoff prediction model, water level prediction model and flood frequency prediction

model, the databases, decision support services and the user interfaces are separate components, but are integrated.
2. Web based interface, through which the distributed database, the model base and decision support services can be integrated.
3. Containing different forecasting models and decision-support services, including river runoff prediction model, water level prediction model, and flood frequency prediction model.
4. It is capable of non-expert implementation and provides flood forecasting decision support services for decision makers and the public. This system integrates information retrieval, analysis, model analysis and extensive visualization functions for information sharing and decision-making support.
5. Extensive visualization capabilities: All flood model results are visualized in maps, table or graph on the web, providing rapid dissemination of flood forecasting and flood threaten information.

6.4 Architecture of WSDSS

In this section, the architecture of flood WSDSS is proposed. This system has a multi-tier architecture consisting of presentation, business logic tier, and data tiers. Figure 6.1 provides an overview of system architecture. The presentation tier is the interface for users to interact with system, users can submit requests from the

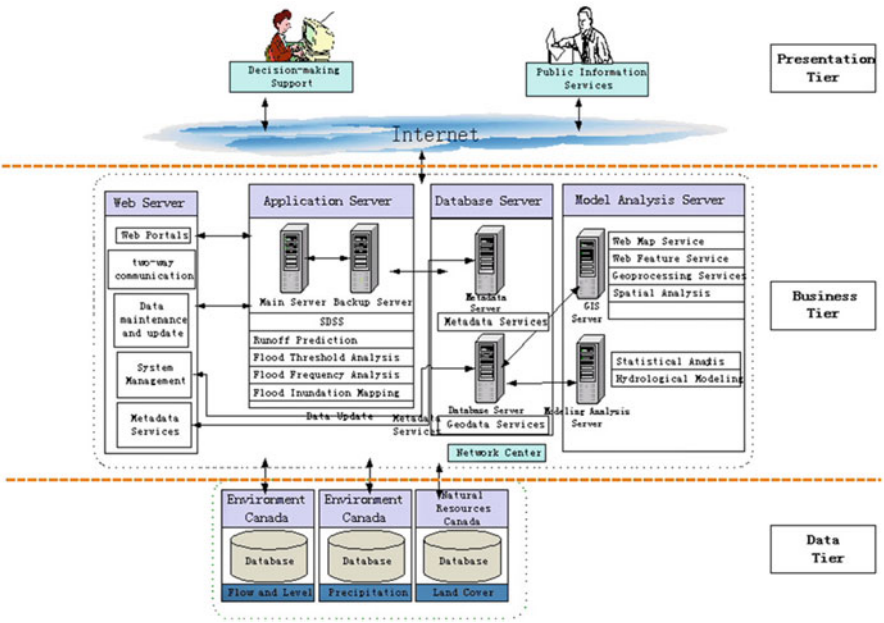


Fig. 6.1 Architecture of WSDSS

presentation tier, and it also can be used as the system client viewer for accessing geographic data and analysis results. The business tier copes with the requests from the presentation tier. The components in the business tier, including web server, application server, metadata server, geodata server and spatial analysis and model analysis server, are used for handling requests and modelling analysis. The web server is the information exchange center between users and application server, supplies two-way communication between client and application server, and is used for receiving user requests, retrieving information, transferring requests to application server, and returning the results in the proper form to the explorer at HTML/HTTP standard. The application server is the core of this system: it responds with the request from users, and transfers the requests to a database server or model analysis server according to the requirements. The metadata server supplies metadata services for users. The geodata server accesses data remotely for model analysis, statistical analysis and spatial analysis, automatically coping with problems in data heterogeneities. The spatial analysis and model analysis server supplies spatial analysis, hydrological model and statistical model analysis and geoprocessing services for river runoff prediction and flood forecasting. The data tier includes all available distributed data sources from different organizations, including river runoff, water level and precipitation data from Environment Canada, and Land Use data from Natural Resources Canada etc. The Database server is used to manage those sharing database in this system, and it also supplies data services for the application server and spatial analysis and model analysis server.

This system has two main sub-systems: a hydrological model for river runoff prediction sub-system, and hydrological models for flood frequency prediction sub-system. First, the hydrological model for river runoff prediction sub-system applies SCS CN model for river runoff prediction. Then, hydrological models for the flood frequency prediction sub-system apply LP3, GEV, PL and GP models for flood frequency prediction. Finally, river runoff, water level, and flood frequency (return period) prediction maps are published on the web for the public and decision maker to use for flood management.

6.5 Components of WSDSS

The integrated components of this system include database, model base, interface, and spatial and graphic data display system Fig. 6.2.

1. Distributed and central management database: all spatial data are stored in GIS databases and all attribute data are stored in the Relational database management system (RDBMS). There are two sets of databases in this system: distributed database for model input and central management database for model input, running and result display. The distributed database allows for data acquisition from various agencies, including river runoff, water level and precipitation data

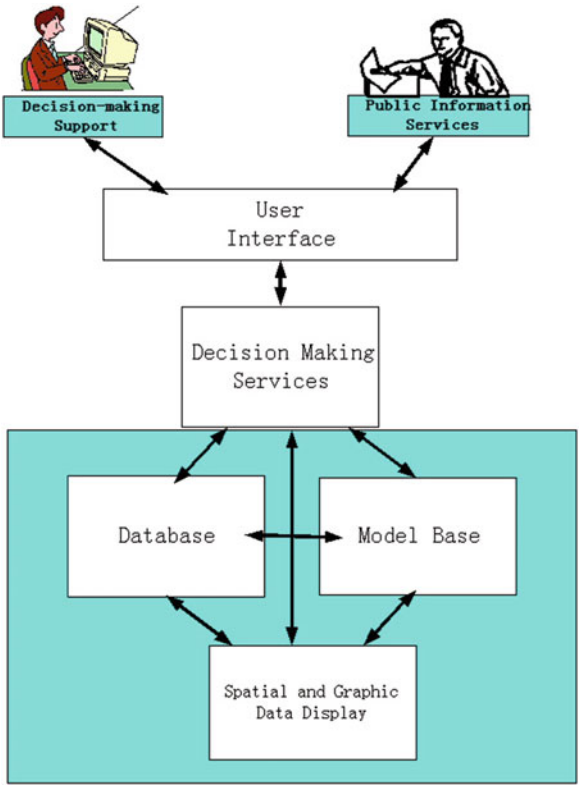


Fig. 6.2 WSDSS components

from Environment Canada. The central management database stores data for model input, running and result display. Spatial data management is either as file based or spatial databases.

2. Model base: several models are integrated to support decision making, including hydrological models for runoff prediction, water level prediction, and flood frequency prediction.
3. Interface: a multiple-level web-based interface needs to be developed, through which the distributed database, the model base and the spatial decision services can be integrated. Users can submit their requests through this web interface and FWSDSS will publish result map to users through this interface.
4. Spatial and graphic data display system: WebGIS supplies users the functions to be able to easily visualize model output results in map such as runoff prediction, water level prediction and flood frequency prediction. The implementation of the web visualization system is supported by ESRI ArcIMS.

6.6 System Functionality

From a technical perspective, WSDSS supplies three levels of decision support and services: metadata services, geodata services, and geoprocessing services. The first level is metadata services. Metadata, which gives a detailed description of the data, provides the following information: identification information, data quality information, spatial data organization information, spatial reference information, entity and attribute information, distribution information, and metadata reference information – as well as other information. Using metadata services, users can know what data are available in this system, how the data have been collected, and who is responsible for managing and distributing the data. This system supplies metadata services for National Water Data Archive, National Climate Data and Information Archive, and National-Scale Ontario Land Cover data set. The metadata protocol used in this system is the Federal Geographic Data Committee (FGDC) metadata standard.

The second level of decision support and services is geodata services. Geodata services include Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Services (WCS) and filter Encoding etc al, users are shielded from the inconvenience of worrying about a variety of data sources, multiple data formats and data maintenance and update of the available distributed data because WSDSS detects and automatically resolves data heterogeneities in the underlying datasets, supplying data services and web mapping services for the public. The decisions can be made better when the decision makers are provided with the most up-to-date and complete information. This system supplies geodata services for the National Water Data Archive and the National Climate Data and Information Archive.

The third level of decision support is geoprocessing services. Distributed geoprocessing services is very critical to help users in geodata processing, modeling and analysis. Effective sharing of methods and tools has a potentially much higher return than sharing of data. GIServices can generate new information values by sharing geographic information, spatial analysis methods, and users' experiences and knowledge. Distributed GIServices encourages sharing of analysis methods, spatial models, and geographic knowledge (Tsou 2001). In spatial decision-making, traditional SDSS and GIS software, analysis methods and models are difficult for the untrained professional to use. An easy-to-use, inexpensive set of analytical tool still need to be developed for distributed spatial decision-making support, improving accessing data, analysis tools and models. WSDSS can integrate information management, information retrieval, spatial analysis and model analysis, supporting information exchange and knowledge, analysis tools and model sharing through web. With such a system, users require only a simple web browser to access the data, the model and perform spatial analyses without the requirements or costs of installing GIS and modeling processing software packages.

6.6.1 Decision Support Services for River Rainfall-Runoff Prediction

The US SCS has developed a widely used lumped, empirical and mathematical model, SCS CN method, for estimating runoff (USDA-SCS 1968). The underlying theory of the SCS-CN procedure is that runoff can be related to soil cover complexes and rainfall through a parameter known as a CN. So river runoff can be predicted using the CN value and precipitation. The SCS model has been widely used for river runoff prediction and flood forecasting, but users need to be trained on how to use it. Difficulty in linking data, analysis tools and models is one of the barriers to be overcome in developing a river rainfall-runoff prediction system. Thus, WSDSS need to be developed to integrate models, data, web interface, and decision-making support to help decision makers and the public to easily access river runoff and water level prediction results. The most important benefits to users are that this system provides comprehensive support for integrating information retrieval, analysis, and model analysis for decision-making support.

First, the river runoff prediction sub-system uses the historical data for the past several decades (river gauging, precipitation, ground water, census and land use) to model the relationship among the stream runoff, precipitation and hydrological-geographical features to get the CN value for each watershed. The CN value is

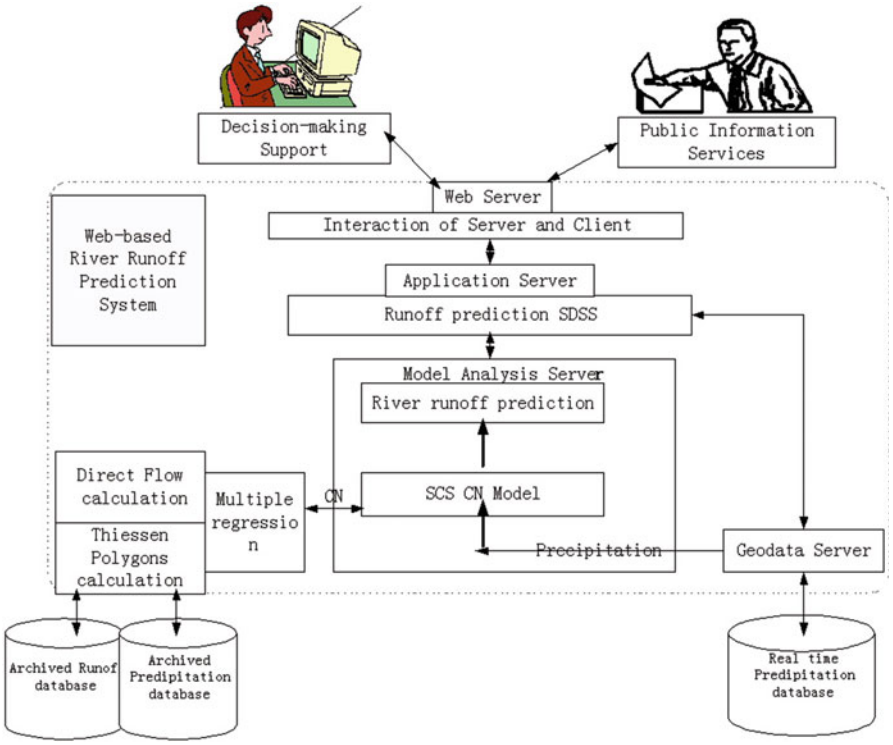


Fig. 6.3 River runoff prediction system

saved in database, and the SCS model uses this CN value and precipitation prediction results for river runoff prediction. When users submit the request from a client, the web server transfers this request to an application server. Then Geodata services access precipitation prediction data remotely from Environment Canada (precipitation prediction fall outside the scope of this study and will use the precipitation prediction result directly from Environment Canada). Next, Geoprocessing services runs the SCS model and necessary data processing for river runoff prediction and water level prediction. Web mapping services publishes the runoff prediction and water level prediction result maps on the web through ArcIMS. Figure 6.3 shows the architecture of the river runoff prediction sub-system.

6.6.2 Decision Support Services for Flood Frequency Prediction

The flood frequency prediction sub-system applies LP3, GEV, PL and GP models for flood frequency prediction.

The LP3, GEV, PL and GP models are applied for flood frequency prediction in the ORM area. The past several decades of historical data of river runoff peak discharge are used to model the relationship between flood discharge and return period by fitting theoretical statistical distributions for each gauging station. The parameters of these fitted flood frequency curves for each gauging station are saved in the database, and these fitted flood frequency curves can be used for flood frequency prediction.

The system supplies a two-way of flood frequency prediction. Users can provide flood peak discharge to predict flood frequency (return period), and the flood peak discharge is interpolated on the fitted flood frequency curve for flood frequency prediction. Or users can provide flood frequency (return period) to predict the required flood peak discharge, and the flood return period is also interpolated on the fitted flood frequency curve for required flood peak discharge prediction. The web mapping services component publishes the flood frequency prediction result maps on the web through ArcIMS. Figure 6.4 shows the architecture of flood frequency prediction sub-system.

6.7 System Development Approach: An Integrated Approach

Although many hydrological models have been developed for river runoff prediction and flood forecasting, most of these models are inaccessible to the public and decision makers, so it is very difficult for the public and decision makers to use

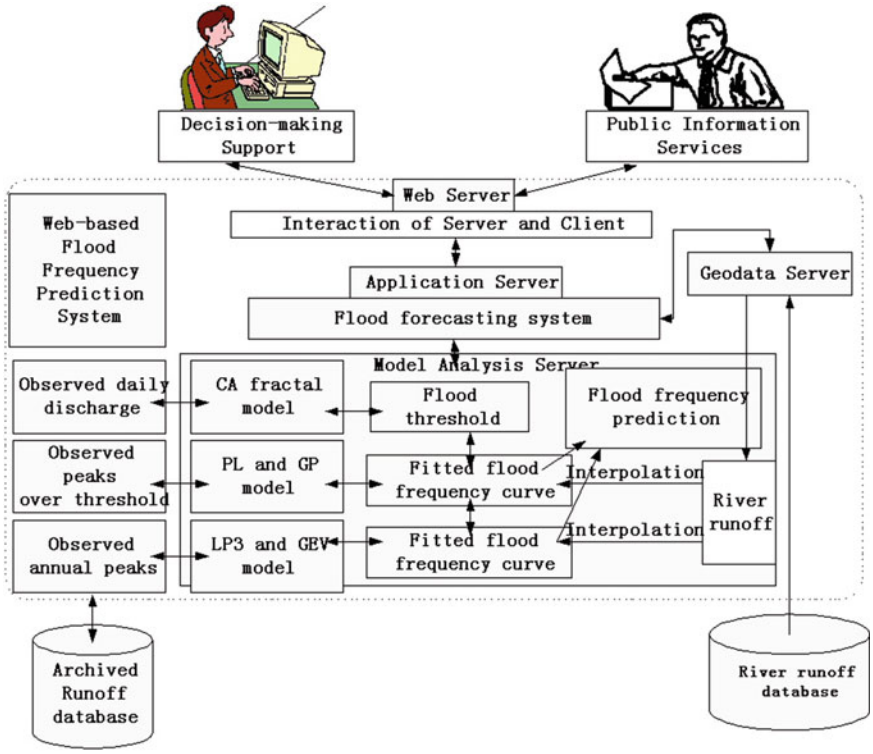


Fig. 6.4 Flood frequency prediction system

these models directly for flood forecasting and flood management. For this reason, a new system and tool are needed by the public and decision makers to enable them to utilize available models and data. An integrated approach is still needed to integrate models, data, decision making support, and interface to help decision makers and the public to easily and fast access flood and flood threaded information.

A flood WSDSS integrates models, data, interface, and decision-making support services to help decision makers and the public to easily access flood and flood-threatened information. Integrated hydrological models, statistical models and fractal models can meet model requirements for flood forecasting. The integrated system has a common database and a common web-based interface. The integration is performed using GIS as a common platform for database, model base and interface management. The SCS CN model is used for river runoff prediction, and the LP3, GEV, PL and GP models are used for flood frequency prediction. Extensive visualization capabilities are implemented, and end users can visualize the model generating results as maps on the web. The design of this system is based on database, web technology, GIS and spatial decision support services. The implementation of this system is supported by ESRI ArcIMS. The theory and

knowledge required to develop this flood WSDSS include hydrology, GIS, database management system, and SDSS.

Full implementation of WSDSS is not easy. Rinner (2003) advocates for Web-based decision services as the building blocks for future WebSDSS. He believes that complex geo-processing services, such as distributed spatial decision-support services, are the highest level of SDSS. Web-based decision services can support information exchange and knowledge, and analysis tools and model sharing from different organizations through the web. The development of WSDSS is based on the concept of web services. Advanced network techniques such as Java language, ActiveX controls, Common Object Request Broker Architecture (CORBA), Distributed Component Object Model (DCOM), and Java Beans, which focus on the development of distributed component technology, have substantially enhanced the potential value of GIS because now it is possible to locate and harness data from many disparate GIS databases to develop very rich analytical information on almost any topic associated with physical locations.

Full implementation of the functionality of WSDSS requires advanced computer sciences knowledge and much programming work. In this paper, instead of developing a system to fully implement WSDSS, a prototype system is developed to demonstrate the basic functionality of flood WSDSS.

6.8 Summary

This chapter introduces the architecture, components, functionality and an integrated system development approach of flood WSDSS. This system has a multi-tier architecture consisting of presentation, business logic, and data tiers. The presentation tier allows users to interact with the system, submit requests from presentation tier, and access geographic data and analysis results. The business tier copes with the requests from the presentation tier. The components in the business tier, including web server, application server, metadata server, geodata server and spatial analysis and model analysis server, are used for handling requests and modelling analysis. The web server is the information exchange center between users and application server, supplying two-way communication between client and application server. Then the application server transfers the requests to the database server or model analysis server according to the requirements. The metadata server supplies metadata services for users. The geodata server accesses data remotely for model analysis, statistical analysis, and spatial analysis, automatically coping with problems in data heterogeneities. The spatial analysis and model analysis server supplies spatial analysis, hydrological model, and statistical model analysis and geoprocessing services for river runoff prediction and flood forecasting. The data tier includes all available distributed data sources from different organizations.

The integrated components of this system include the database, model base, interface, and spatial and graphic data display system.

This system has two main sub-systems: the hydrological model for river runoff prediction sub-system, and the hydrological model for flood frequency prediction

sub-system. First, the river runoff prediction sub-system applies the SCS CN model for river runoff prediction using precipitation data and the CN value of each watershed. Then, the flood frequency prediction sub-system applies fitted flood frequency curves for flood frequency prediction.

Acknowledgments The author expresses the appreciation of funds received from Hainan peovince major science and technology projects (#ZDKJ2016015-1), the Hainan Province Natural Science Foundation (#20164178), the Sanya City Key Laboratory projects (#L1404) and the Sanya City science and technology cooperation projects (#2015YD19 and #2014YD08).

References

Peng ZR, Tsou MH (2003) Internet GIS: distributed geographic information services for the internet and wireless networks. Wiley, Hoboken

Rinner C (2003) Web-based spatial decision support: status and research directions. J Geogr Inf Decis Anal 7(1):14–31

Tsou MH (2001) A dynamic architecture for distributing geographic information services on the internet. Ph.D dissertation, University of Colorado at Boulder, Boulder

USDA-SCS (1968) National engineering handbook. U.S. Department of Agriculture, Washington, DC

Chapter 7
Ocean and Coast Disaster Data Modeling

Xin Zhang

Abstract This chapter focuses on data and system modeling for ocean and coast disasters. At first, the ocean and coast disaster monitoring data are introduced. According to its characteristics, we analyse how to organize data in the data warehouse. Then, we discuss how to use an ocean stereo monitoring data management system for data management. Next, according to the characteristics of ocean disaster factors and the software architecture, this chapter introduces the elements of ocean disasters and a variety of expressions and implementations. The chapter ends with the introduction of the storm surge disaster analysis model system and the sea level rising disaster analysis model system, including the data needed by the system, the software structure design and the function realization process.

Keywords Digital earth • Digital ocean • Ocean disaster

7.1 Multidimensional Data Organization of Ocean Disasters

7.1.1 Thematic Structure of an Ocean Disaster Spatio-temporal Data Warehouse

There are many common factors related to ocean disasters, such as water temperature, salinity, density, tides, tidal waves, currents and meteorological elements, water quality, sea ice, typhoons, and storm surge. Ocean disaster data are typically characterized by multidimensionality. The organization of ocean disaster multidimensional data is based on the concept of a data warehouse.

In research related to the oceans (Shen et al. 2007.), not every factor is isolated. In the process of solving problems, we must not only consider the characteristics of a certain single factor or phenomenon, but we must also analyze the relationships and effects of different factors.

X. Zhang (✉)
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zhangxin@radi.ac.cn

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_7

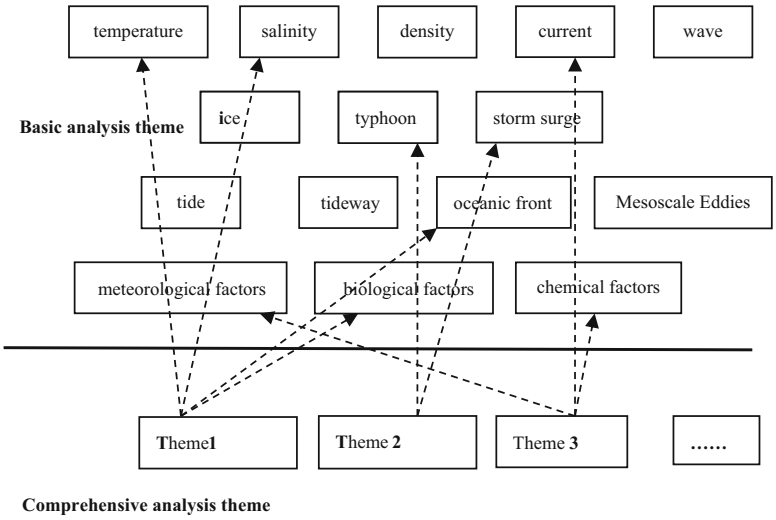


Fig. 7.1 Thematic structure of the ocean disaster spatio-temporal data warehouse

Figure 7.1 shows the thematic structure of the ocean disaster spatio-temporal data warehouse.

A thematic generalization is proposed for the application of data integration, which is the advanced application of the data warehouse. According to the customer's specific demands, thematic generalization joins various tables by a base factor. By sharing and integrating these tables, we could create a multidimensional data cubic. The cubic provides many topics aimed at different areas of ocean research and forms a flexible and efficient ocean disaster thematic structure.

7.1.2 Modeling the Dimensions of the Ocean Disaster in Spatio-temporal Data Warehouse

Dimension refers to the organized and classified hierarchy structure, and it describes the table of the factor in a data warehouse (Su et al. 2006a, b, c). Dimension is the basic component of the multidimensional data cubic, and it generally describes a member's similar set on which the user can base an analysis. The dimension of the cubic often has many attributes that can be organized into a hierarchical structure according to their details. The hierarchical structure includes the member set in the dimension and the relative position between these sets. For example, the time dimension can be divided into day, month, season, and year, where the "day" is the most detailed level. Dimension hierarchies can clearly reflect drill-down and roll-up operations. Drill-down means to observe data along some dimension more specifically, and roll-up

means to observe data along some dimension more systematically. Dimension hierarchies can be divided into generalized relationships, aggregation relationships and member relationships. A dimension's value is referred to as one of its members. If the dimension is multi-level, the dimension member is one value in the combination of different dimensional levels. For example, the time dimension has a month/day/year level. A dimension member is combined the value obtained from each level.

The ocean disaster spatio-temporal data warehouse adopts a metadata-driven method of sharing the metadata dimension to create the database. The Shared dimension refers to the dimension of the scope created in the database warehouse. It can be used by the data warehouse in an arbitrary cube. By creating the Shared dimensions and by being used in more than one cube, much time can be saved. Standardization of Shared dimensions, such as the standardization of time and geographical space Shared dimensions, will ensure that different data sets are used in a similar manner. This integration of data sharing is very important. Therefore, in the ocean disasters space-time dimension data warehouse, modeling of dimension tables can limit the maximum shared by the fact table. Regarding time and space in the data warehouse, the ocean disaster time and spatial dimensions have an obvious hierarchy, but there are also other dimensions that have no obvious hierarchy or for which the operations of hierarchy (roll-up and drill-down) do not make sense (Thomas, 2003). For these, a description of the theme data is needed, such as data type, data structure, and spatial reference information. This section will focus on the logical structure of the shared dimensions of ocean disaster data in the spatio-temporal data warehouse.

7.1.2.1 Time Dimension

The time data related to ocean disasters include observation time, analysis time, statistics time and forecast time. There are many levels of precision of the time. For example, many statistics record data in a period of 10 days or of a year or month. Data analysis and observations usually record time in hours. Valid ocean disaster data use hours to days, month, season and year as a time hierarchy, which will ensure the integrity of the data granularity level. The time dimension table structure is shown in Fig. 7.2.

In the data warehouse, most of the multidimensional model needs to build a relationship between every fact and the lowest level of a dimension value for each dimension. In the data warehouse, for time and space related to the ocean disasters, regardless of the particle size, the fact table must have the same level of measurement as the rules. If those multi-level granularity facts are mixed or are stored as a mixture of multiple-granularity data, the user will feel confused and it will be easy to make errors in data analysis and applications. Thus, the time dimension table contains interrelated chronology tables, such as season, month, 10-day, day and hour tables. Each time level can be aggregated into any upward level (not drawing connecting lines for clarity above). Each time a new level of granularity of data at a

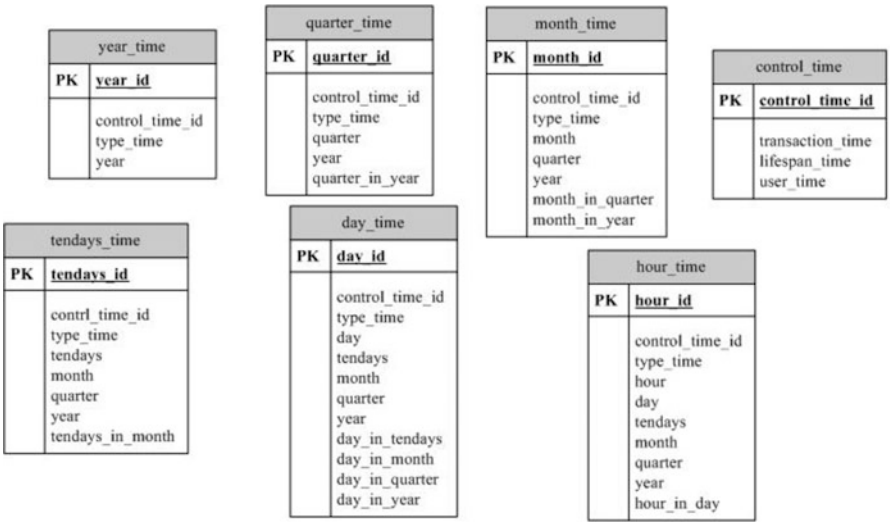


Fig. 7.2 The time dimension table structure of the ocean disaster spatio-temporal data warehouse

time is attained, the fact table produces a corresponding level of granularity. The meanings of the fact table fields are shown in Table 7.1. Each record in the fact table also corresponds to the time, life cycle and user-customized time (according to the need to add), which is collectively known as time control.

7.1.2.2 Space Dimension

It is observed through the space of knowledge in the field of oceans that the sea is a continuous 4D field. However, in the computer world, continuous space is discretized into latitude and longitude and depth of the point (X, Y, Z) and then into the point values of the record, such as station, line, navigation data, remote sensing information product, and 3D data. In the ocean disasters data warehouse, time and space data with a few points, such as station, line, walk navigation data (and the point, line class data), still adopt the method of recording a scatter of latitude and longitude. Three-dimensional multi-layer data, other than the analysis of remote sensing information products, includes adopted names of bedding face products, coordinates, and the scope and the depth of expressions to be recorded. The content of the spatial dimension table is derived from the space of metadata information and browsed diagrams. The space dimension table structure is shown in Fig. 7.3.

In the time and space data of the ocean disaster data warehouse, the hierarchy of the spatial area dimension table was identified through the data grid resolution. Regarding the spatial dimension, the roll-up and drill-down operations include the aggregation and further subdivision of the spatial data grid resolution.

Table 7.1 Field descriptions of the time dimension table in the ocean disaster spatio-temporal data warehouse

No.	Fields name	Data type	Description
1	year	char(4)	calendar year, such as“2010”
2	quarter	char(8)	Calendar Quarter, such as 2010–001. In order to distinguish the calendar month, use three digits
3	month	char(7)	calendar month:“2010-01”
4	tendays	char(9)	Ten days in a month: 2010-01-001,001, 002, 003 represents 3 time slots in a month
5	day	date	Day: 2010-3-1
6	hour	date	Hour: 2010-3-1 08:00:00
7	type_time	varchar	The type of time,including:observation, forecast, analysis, and statistics, etc.
8	quarter_in_year	nunber (1)	which quarter in the current year, such as 1
9	month_in_quarter	number (1)	Which month in a quarter of a year
10	month_in_year	number (2)	Which month in current yeah, such as 10
11	tendays_in_month	number (1)	Which tendays in a month, such as 1
12	day_in_tendays	number (2)	Which day in a ten-day, such as 1
13	day_in_month	number (2)	Which day in a month
14	day_in_quarter	number (2)	Which day in a quarter of a year
15	day_in_year	number (3)	Which day in a year
16	hour_in_day	number (2)	Which hour in a day
17	control_time_id	number (8)	The primary key of controlling time dimension table
18	transaction_time	date	Time that data load into the database
19	lifespan_time	date	End of life cycle time of the data
20	user_time	varchar	Custom date formats. Forecast data, or reanalysis data analysis time, etc.

7.1.2.3 Dimensions of the Source Project, Contact Information, and Investigation Information

The data source and the contact information of the project are important. They contain project entities from the source of the metadata and the data source. The investigation information dimension records the data observation instruments, platforms and the investigation information dimension table from the metadata. The unit table records all of the data related to the unit information. The dimension

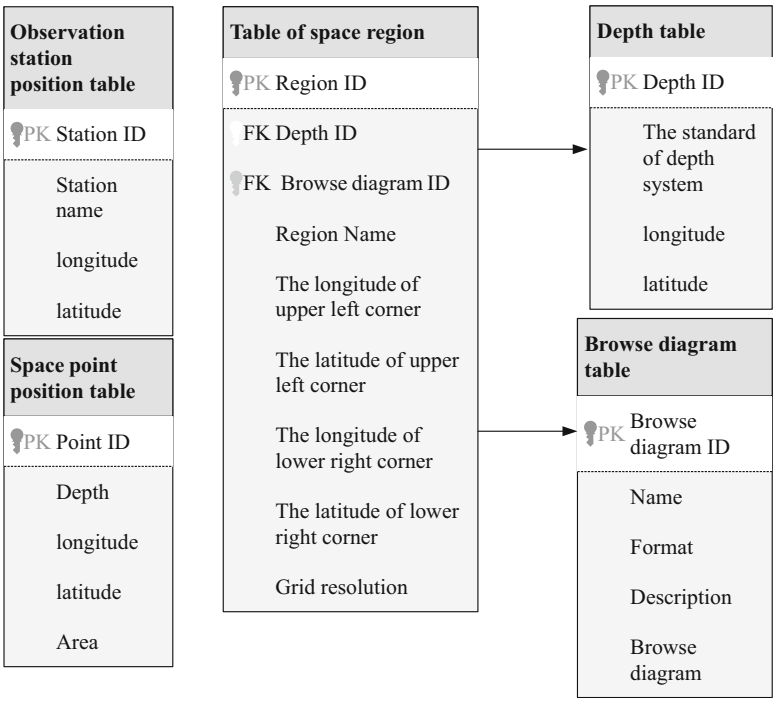


Fig. 7.3 The structure of a space dimension table in the ocean disaster spatio-temporal data warehouse

table is shared for the source project dimension table, the contact information dimension table and the survey information dimension table.

7.1.2.4 Category, Data Format, and the Spatial Reference Dimensions

The element category dimension records the classification of data information. The elements category is the primary key of the dimension table, and it is also the identity of the dimension hierarchy data category. Data on the category dimension can start roll-up or drill down operations according to the level of the category code to obtain the categories of data or information on the next level of classification information. The element category dimension is derived from the data classification of the metadata entity. Because of the different data sources and access, data on the same topic in the fact table may have a variety of data formats. The data structure dimension records the format of the data and uses the information. It is easy for a program to automatically identify and handle this information. The data structure dimension is derived from the data format of a metadata entity. The spatial reference dimension records the data of spatial reference information and reference

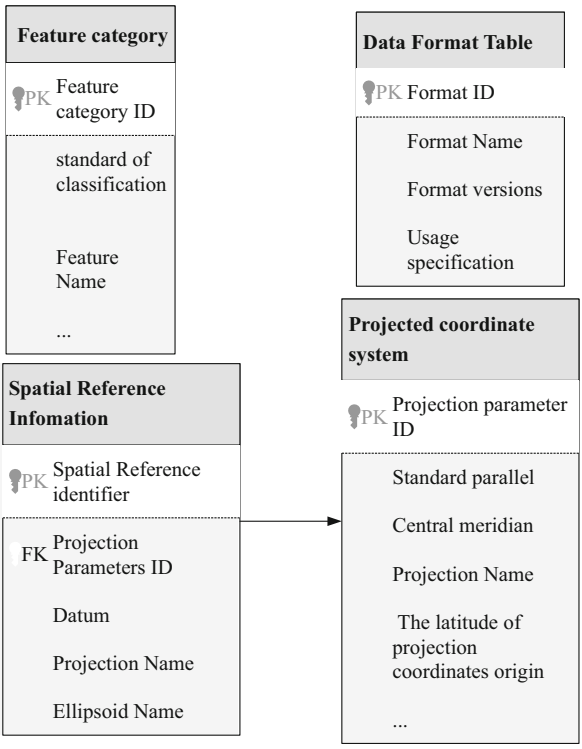


Fig. 7.4 The structures of the time element category, the data format and the spatial reference table in the ocean disaster spatio-temporal data warehouse

entities from the metadata of spatial data. The data type, data structure and spatial reference dimensions are shown in Fig. 7.4.

7.1.2.5 Warehouse Model of the Ocean Disaster Multi-dimensional Spatio-temporal Data

The ultimate goal of the design of the data warehouse model is to form the multidimensional data structure of joining the fact table with the dimension table, and to provide the multidimensional data for OLAP analysis. Considering the time and space complexity of the data, we adopt star and snowflake structures to organize the ocean disaster spatio-temporal data warehouse, and we adopt the star schema structure to organize the comprehensive analysis of multidimensional data. After the definition of data standards (such as data types, constraint condition and index), the definition of primary keys, foreign keys, properties and so on, we translate the model structure into the physical structure of the warehouse data. In the process of physical implementation, when the dimension hierarchies of different fact tables of some dimension of size are consistent, we can make the dimension table pointer point to the Shared dimension.

7.1.2.6 The Measurement of the Fact Table

A fact is the intersection of each dimension. It records the measurement of a particular subject. In the ocean disasters spatio-temporal data warehouse fact table, its value could be the temperature, salinity, flow velocity or flow direction. When building the model of basic multidimensional analysis theme data, the columns of the fact table are all the primary key of each dimension table, except for the measured variables (e.g., hour ID, area ID, point ID, contact information ID, item ID, ID, survey information data category code, data format ID, spatial reference identifier). These columns are the foreign keys of the fact table and cannot be empty. The benefit of this structure is that the fact table could join the dimension table through a foreign key of a few bytes, and much storage space could be saved. This strategy is particularly effective when the theme involves multiple dimensions or if one dimension has multiple layers. An analysis theme may correspond to several tables, such as temperature, salinity, density, statistical fact (similar to the season fact table), and monthly statistical fact tables. The tidal theme corresponds to the tidal site observation fact table and the tide and bedding face forecast data table. A fact table may belong to multiple analysis themes. For example, the themes of temperature, salinity and density can be used for fishery resources analysis, analysis of red tides and various other topics.

7.1.2.7 Multidimensional Basic Analysis Theme Data Model

By joining the fact table and the dimension table, we can get the multidimensional data structure of each analysis theme. For example, with the temperature theme, we will study the multidimensional model of the research analysis theme as follows. The fact table of the temperature theme includes tables of temperatures at different levels of time for specific areas. Table 7.2 illustrates that the measurement for temperature is the value of temperature. The values are stored in the "temperature" field. In the table of area temperature, the type of the temperature field is the BFile type in ORACLE. In the record of the area temperature field data, the graphics file or connection information of the data node is used to release the data (such as WMS service address, parameter information). We obtain the data format through data format tables. We record temperature directly in the table of the observation temperature.

7.1.2.8 Modeling for Comprehensive Analysis

The comprehensive analysis theme uses multiple basic analysis data as the data source for comprehensive analysis according to user needs. The comprehensive analysis theme adopts a star schema structure to construct a multidimensional data model in the ocean disasters spatio-temporal data warehouse. The table of the comprehensive analysis theme is necessary for joining various basic analysis

Table 7.2 Measurement of the ocean disaster theme in the spatio-temporal data warehouse

Theme	Attribute set (Measurement)
Temperature	The temperature of the seawater
salinity	The salinity of the seawater
Density	The density of the seawater
Current	The flow velocity, flow direction of the current
Wave	Wave height, period, direction, the frequency of a wave height in along a direction
Ice	Ice range, thickness, ice temperature, density, stack height
Typhoon	Wind velocity, direction, atmospheric pressure, radius of influence
Strom surge	The rise of water level, water level
Tide	Time of tide, tide height, high and low tide, etc.
Tideway	Flow velocity, flow direction
Oceanic front	Frontal line, width, intension
Mesoscale Eddies	Vortex position, average sea altitude intercept, the average radius of whirlpool
Meteorological factors	Air temperature, humidity, air pressure, precipitation, visibility the sea fog, etc.
Biological factors	Primary productivity, viruses, microbial, slightly biological, slightly photosynthetic eukaryotes, radiolarian, foraminifera, zooplankton biomass, etc.
Chemical factors	Dissolved oxygen, pH, total alkalinity, suspended solids, total organic carbon, total nitrogen, total phosphorus, nitrate nitrogen, nitrite nitrogen, mental elements, etc.
.....	.....

theme tables. A multidimensional data cube is formed by sharing the dimension table. The comprehensive analysis of a typhoon storm surge and the comprehensive analysis of a red tide environment are used as examples for studying the multidimensional data model of the comprehensive analysis theme in the ocean disaster spatio-temporal data warehouse.

A comprehensive analysis of the multidimensional structure of the typhoon storm surge theme is shown in Fig. 7.5. A foreign key in the storm surge fact table "typhoon number" joins the primary key of the fact table. The fact table for the typhoon storm surge data has a time granularity of hours, and the two datasets have a Shared schedule in the time dimension.

Red tide is a type of ocean disaster affecting many coastal countries and it is a complicated ecological anomaly. The causes of red tide are very complex. Water eutrophication is the material basis and the primary condition of red tides. The water detection results associated with red tides show that the sea water has been badly polluted and subjected to eutrophication, and that nitrogen and phosphorus nutrient substances greatly exceed regulatory limits. Second, some organic substances would also prompt the rapid proliferation of red tide organisms. The changes of the hydro meteorological and physical and chemical factors of sea water are the important reasons for the red tides. The temperature of the water is an important environmental factor. The optimum temperature for the red tides is 20–30 °C. Scientists have found that a sudden increase in water temperature of

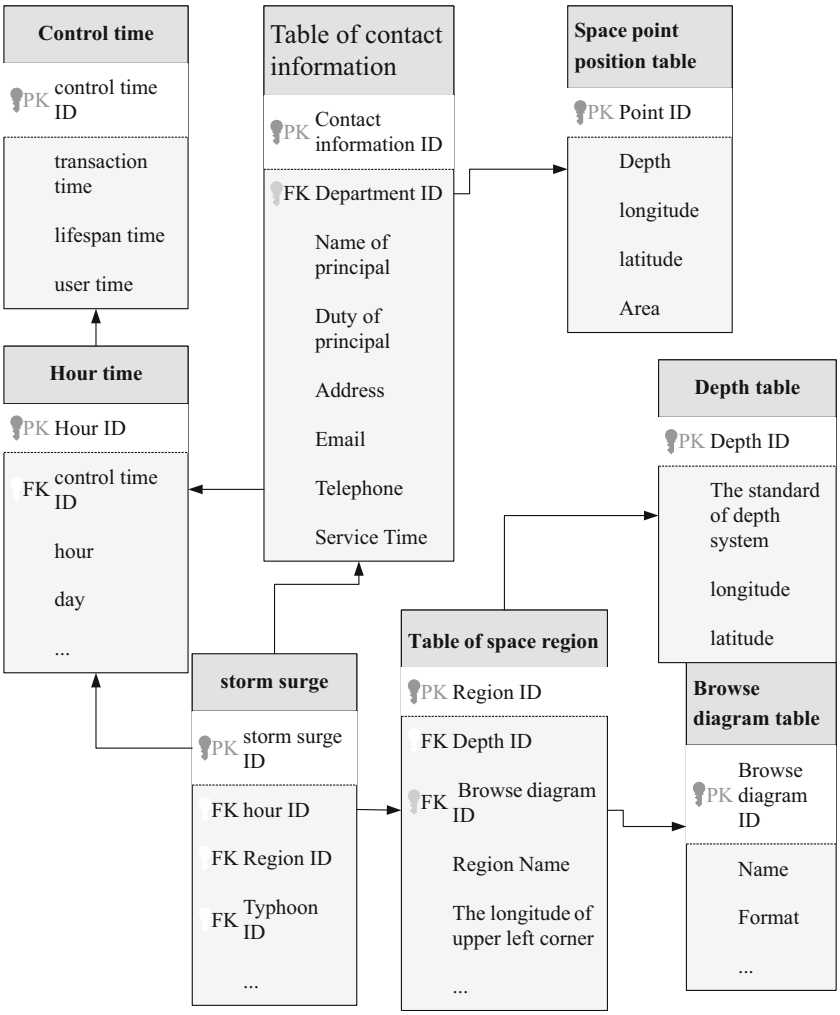


Fig. 7.5 The structure of storm surge table

more than 2 °C over the course of a week is the precursor for the occurrence of red tides. Chemical factors such as the salinity of the water are also contributing to biological factors, which is one of the reasons why the red tide organisms multiply. Salinity in the range of 26–37 is necessary for a red tide to occur. When salinity is in the range of 15–21.6, it is easy to form a thermocline and a halocline. The existence of the thermocline and the halocline provides conditions for the gathering of red tide organisms. Accordingly, monitoring data shows that when red tides occur, the area is always lacking in rainfall, the weather is hot, the water temperature is high, the wind is weak, and the conditions of the water environment change slowly.

When analyzing the occurrence of red tides in an environment, the time level of temperature, salinity, tide and tidal current measurements is "hour" in regional data.

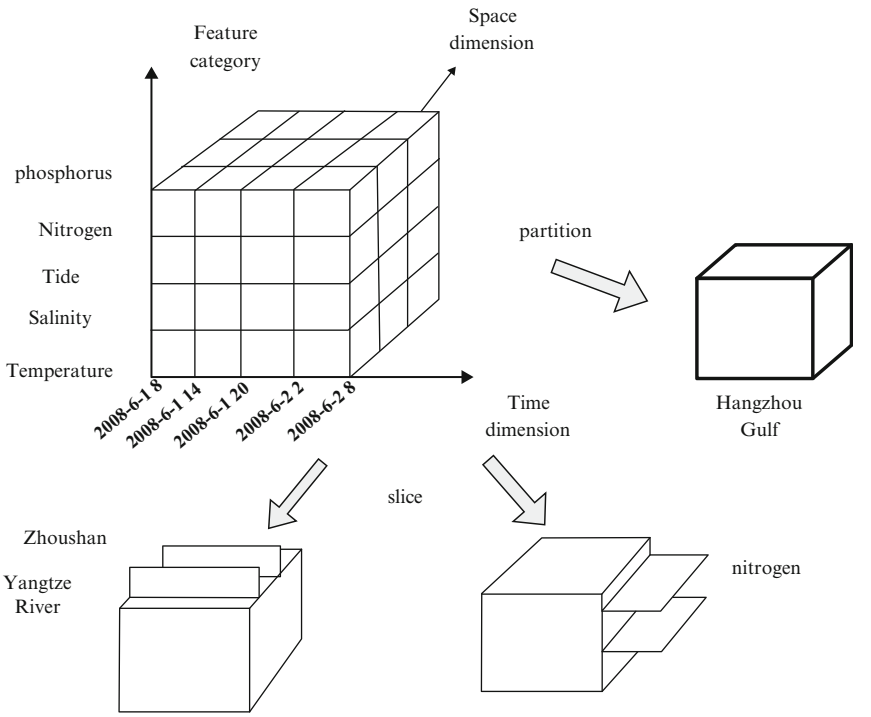


Fig. 7.6 The multidimensional data cube of the "red tide environment comprehensive analysis theme"

Combined with biological facts and chemical data (biological and chemical element analysis types through the filter element category table), the multidimensional data model of the theme associated with comprehensive analysis of the environment of red tide occurrence is formed. The procedure is shown in Fig. 7.6.

In multidimensional analysis, the analysis could involve temperature, salinity, tide, tidal current, the biological and chemical fact and dimension tables, and the space between regional dimension tables.

7.2 Three-dimensional Ocean Disaster Monitoring Data Management

7.2.1 Software Structure Design

The three-dimensional ocean monitoring data management system adopts database technology, web service technology and visualization technology (Wright, et al. 2010, 2007). The system can store, query and visualize monitoring data. It mainly

includes 4 modules: device management, data management, data query and task management.

1. The function of the device management module is mainly to manage the ground-wave radar, shore based device, buoy and subsurface buoy data. It includes accessing the device, device modification, visualizing the monitoring devices' data, and operating conditions.
2. The data management module provides the function of managing monitoring data, which supports input data manually or by importing files. The manual input function means that all types of monitoring data are entered manually. The import file function means that the data are input by importing the files, including wind speed, temperature, flow velocity, flow and wave height, dissolved oxygen, and PH.
3. The function of the data query module is to query monitoring data according to the conditions and visualize the result in many ways. This module can realize a simple query, a joint query, or display the query results in the form of a chart or graph.

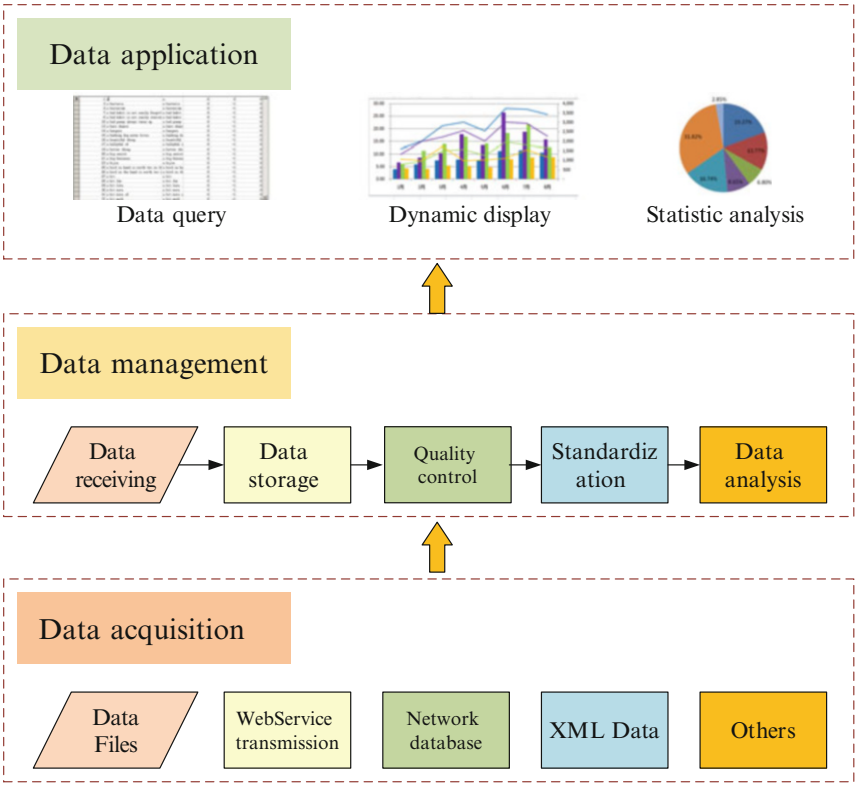


Fig. 7.7 The structure of the management software for three-dimensional ocean monitoring data

4. The function of the task management module is to manage and monitor the data quality through workflow technology. In particular, the module can define the task, create the workflow and execute the task, check the task, etc (Fig. 7.7).

To improve the visualization quality of the ocean monitoring data, this subject takes SKYLINE as the basic development platform. The data management system uses the C/S structure, which is based on Microsoft Visual Studio and Skyline. The skyline extensions and 3D display plugin should be installed on the client. The function of 3D ocean monitoring data management is to take a scientific and effective approach to manage and query the monitoring data. This system aims at effectively managing and querying the monitoring data of shore-based, ground wave radar, buoys and subsurface buoys. The system will play a role in the realm of ocean information management and application.

7.2.2 Software Development

The system development mainly aims at the application of the network environment, and the data access interface can be obtained by web service, as shown in Fig. 7.8.

7.3 Multi-dimensional Expression of the Process of Ocean Disaster Factors

7.3.1 Software Structure Design

From the perspective of technical implementation, the software of the multi-dimensional expression of ocean disaster factors from the bottom-up can be divided into the data access layer, the business logic layer and the presentation layer.

1. Data access layer

The data access layer is used to access the data warehouse and to realize the operation of SELECT, INSERT and UPDATE on the data tables, especially the SELECT operation. All of these operations are used for obtaining the data for the use of the business logic layer. The core of this layer is the data access class GetData. According to the data type, the class's public interfaces include: IGetPoint (access to a single point of value, a time series, or a depth sequence value), IGetCurve (line method to get a line of a point set, or a time series line trajectory); IGetSurface (obtains a plane, or a set of time series of a plane); and IGetVolume (three-dimensional data to obtain a certain hour, or a three-dimensional data time series).

Fig. 7.8 The development of the data acquisition interface based on web service

2. Business logic layer

The business logic layer is used to realize the visualization of the ocean environmental data. The layer, which is between the data access layer and the presentation layer, plays an essential role in the process of data exchange. The software is designed with the interface-oriented method in the hierarchical design process. The dependent relationship between the upper layer and bottom layer is a type of weak dependence. The upper layer is independent of the lower layer. Changing the upper layer's design has no effect on the lower layer.

3. Presentation layer

The presentation layer is used to display data and receive input data, and it is used to provide users with an interactive operation interface. The user is the core of all processing. The process should not be determined by the application. Thus, the user interface should be designed for the user to control the application work and it should respond conveniently. The design should follow the graphical user interface (GUI) design principles, which means an intuitive and consistent interface that is transparent to users

This study adopts SKYLINE as the basic research and development platform, which integrated with visualization development based on DirectX. The study realizes many expression methods, such as 2D graphics of single points and single elements associated with an ocean disaster changing over time, 2D graphics of a single point and multilayer elements varying with depth, graphics of a single element field changing over time, and simulation of multilayer dynamic processes.

Direct extension (DirectX, DX), which is developed by Microsoft, is a multimedia programming interface. Implemented in the C++ programming language and following the standard of COM, DirectX can let the Windows platform execute more efficiently for games or multimedia applications. It can strengthen the 3D graphics and sound effects, and it provides the designer with a common standard for hardware drivers. Thus, the developers do not have to write a different driver for each brand of hardware. At the same time, this strategy also reduces the work required for the user to install and set the hardware.

DirectX is composed of many APIs. According to the property, these APIs can be divided into four categories: display, voice, input and network. The display part is the key of the graphics processing. This part can be divided into DirectDraw (DDraw) and Direct3D (D3D). The former is mainly responsible for 2D graphics acceleration, and the latter is mainly responsible for the display of 3D effects and for providing an interface for the built-in 3d color modulation of most new video adapters. Direct3D provides powerful and efficient communication methods between the software programs and the hardware accelerator. It contains a special CPU instruction set that can provide further support for new computers. The core API of the voice part is DirectSound. This part is responsible for playing sound, audio mixing, strengthening the 3D sound, and recording. Sound card compatibility problems are solved by DirectSound. The input part, API DirectInput, supports a variety of input devices and makes these devices work at optimal levels. These

devices include not only the keyboard and the mouse but also the handle, rocker, simulator and so on. The network part, API DirectPlay, is mainly responsible for network development. It offers a variety of connections, such as TCP/IP, IPX, modem, and a serial port. In addition, it also provides a network dialogue function and security measures.

The DirectX Utility Library is a program framework established on the Direct3D 9 and Direct3D 10 APIs. It is used to create a series of powerful and convenient Direct3D development samples, prototypes and tools. DXUT works on Direct3D 9 and Direct3D 10. A program building on DXUT can easily use the API. DXUT simplifies the application of a typical Windows and Direct3D API.

The system develops a 3D visualization engine SeaGeo for ocean environmental data based on DXUT. Developed with the base of the MFC ActiveX framework, SeaGeo supports Direct3D 9.0 and its packaging component DXUT. The spatial coordinate transformation adopts DXUT camera components. From top to bottom, the engine can be divided into the control interface, 3D realization and data management, which are respectively implemented by the classes CseaGeoCtrl, CSeaGeo3D, and CdataFile.

Field data are stored in the BFILE format in the data warehouse. According to the user's query, we can retrieve the data obtained from the access layer of the visualization software in the data warehouse and form the 3D data field. Class CDataFile can read this data field into the computer memory and assign values to the 3D structure array (struct DATAGEOS for 3D scalar fields and DATAGEOV for 3D vector fields as well as build the relationship between each grid point's attribute and color (or the direction of the arrow and its color). These structures are used to establish a MESH on each section for Class CSeaGeo3D and start rendering it. The functions of Class CDataFile are shown in Table 7.3.

Class CSeaGeo3D is mainly used for 3D rendering. Its main features include 3D scene initialization, building the profile of the MESH, rendering of the MESH, and picking up attribute values of the node in the scene. Creating a MESH means assigning the color corresponding to the values of every grid calculated in the Class CdataFile to the color attribute of the vertex object. The function of CSeaGeo3D is shown in Table 7.4.

Table 7.3 The main functions of Class CDataFile

Function name	Functional description
ReadData()	Read the 3D field data into the memory, and assign the data to the array DATAGEO[i*j*k]
SetTemperatureColor()	Calculate the color of each dot in the 3D temperature field
SetSaltColor()	Calculate the color of each dot in the 3D salinity field
SetDensityColor ()	Calculate the color of each dot in the 3D density field
SetSoundspeedColor ()	Calculate the color of each dot in the 3D sound velocity field
SetFlowColor ()	Calculate the arrow color of a three dimensional flow field data vector (the color represents the flow velocity)

Table 7.4 The main functions of Class CSeaGeo3D

Function name	Functional description
InitApp()	Scene initialization, loading images, mark and vector layer
RenderOutline()	Draw the coordinate system guides in the 3D scene
LoadModis()	Load land image as background in the scene
LoadNotice()	Load the standard information
LoadShp()	Load feature layer as background
CreatDepthMesh()	Create 3D field data profile in the direction of the depth
CreatLonMesh ()	Create 3D field data profile along the longitude
CreatLatMesh ()	Create 3D field data profile along the latitude
CreatAnySectionMesh()	Create 3D field data profile along any angle
CreatTimingDepthMesh()	Create timing 3D field data profile in the direction of the depth
CreatTimingLonMesh ()	Create timing 3D field data profile along the longitude
CreatTimingLatMesh ()	Create timing 3D field data profile along the latitude
CreatTimingAnySectionMesh()	Create timing 3D field data profile along any angle
Render()	Load the MESH object in the 3D scene, and render the data
RenderFlowDir()	Called in the Render function. Draw the vector arrow if the field is a vector field
FreeDataMesh ()	Release the MESH object
SelectPoint()	Capture a vertex in the 3D scene
RenderText()	Shows point coordinates and attributes

7.3.2 Software Development

7.3.2.1 Single-Point and Single-Layer Element Changes Over Time

The main function of this module is to display the data of coordinates in the element field changing over time based on the earth sphere model. We express this change through the curve of the single-point elements of changes in two-dimensional field data over time. The visualization method of the point process aims at space point objects. Time series data for each station can be inversed into a point process for this station. The spatial locations of these point data are relatively constant. They change as the time attribute information changes. For example, the seasonal changes of water temperature can be expressed by a point process. Ocean elements change over time. This change could be expressed by a continuous curve of 2D geometric coordinates, with the horizontal axis as time and the vertical axis as property values. The interface is shown below (Fig. 7.9, Tables 7.5 and 7.6).

1. Input and output items

When accessing the sub module, users can choose the element type in the element type drop-down list box (including four options: temperature, salinity, density and sound velocity). An operation type for “single point of time changes” is selected in the operation type drop-down list box. The data type of forecast data is

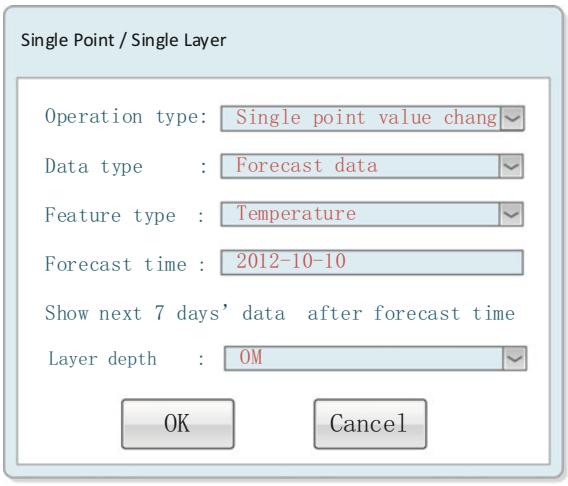


Fig. 7.9 The function interface of single-point element changes over time

Table 7.5 List of the input items of the "single-point elements changing over time" module

Name	ID	Type	Description
Feature Type	DDSJ_Yslx	string	Input optional element types, including temperature, salinity, density and sound velocity
Data sources	DDSJ_Sjly	string	Input the data sources: forecast data
Operation type	DDSJ_Czlx	string	Input operation type
Layer depth	DDSJ_Cs	string	The depth from the surface to the bottom of the sea
The starting date	DDSJ_Qbrq	DateTime	The starting date of the 7-days forecast data

Table 7.6 List of the output items of the "single-point elements changing over time" module

Name	ID	Type	Description
The curve of the single-point changing over time	DDSJ_Ddsdbh	Object	When the operation type is "single-point data changing over depth", query the data, call the graphics interface, and draw the curves of single-point data changing with depth

selected in the data type drop-down list box. The time and deep layer in time are selected in the deep layer drop-down list box, and OK is clicked. Users set up a start time to show a week's forecast result, and the interval time is 24 h.

After clicking the OK button, the user can see the data area on the sphere model. The cursor turns into an arrow. After the user clicks on any position on the data area, a new page will show the data curve of the position. After clicking the cancel

button, the temporarily-created position label and the regional box are removed and returned to the initial state.

7.3.2.2 Single-Point and Single-Layer Element Changes Over Sea Depth

The main function of this module is to display the data at a set of coordinates in the elements field that are changing over sea depth based on the earth sphere model. We express this change through the curve of the single-point elements of two-dimensional field data changes over sea depth (Tables 7.7 and 7.8).

1. Input item
2. Output items

When accessing the sub module, users can choose the element type in the element type drop-down list box (including four options: temperature, salinity, density and sound velocity). The depth change of a single-point change is selected as the operation type in the operation type drop-down list box. The forecast data is selected in the data type drop-down list box. The time is selected in the time drop-down list box, and OK is clicked.

After clicking the OK button, the user can see the data area on the sphere model. The cursor turns into an arrow. After the user clicks on any position of the data area, a new page will show the data curve of the position. After clicking the cancel button, the temporarily-created position label and regional box are removed and returned to the initial state.

Table 7.7 List of the input items of the "single-point elements changing over depth" module

Name	ID	Type	Description
Feature Type	DDSD_Yslx	string	Input optional element types, including temperature, salinity, density and sound velocity
Data sources	DDSD_Sjly	string	Input the data sources: forecast data
Operation type	DDSD_Czlx	string	Input the operation type
Forecast date	DDSD_Ybrq	DateTime	Visual

Table 7.8 List of the output items of the "single-point elements changing over depth" module

The curve of the multi-layers of single-point changing over depth	DDSD_Ddsdbh	Object	When the operation type is "single-point data changing over depth", query the data, call the graphics interface, and draw the curves of single-point data changing with depth

7.3.2.3 Single Element Field Dynamic Changing Over Time

The visualization of the element field is very intuitive. The user can easily visualize the changing process of the physical value through the visualization of element field process. The element field object can be a flat or an arbitrary angle profile.

The visualization of a single-element field process changing over time can be realized in two ways: animation and rendering directly. Animation generates animation files by the program or animation generation tools (such as Moviemaker). In one time sequence, the images of different discrete times can be generated through an image generation interface. These images are used as consecutive frames to generate an animation. The similarities of the two methods are that they are both based on the idea of snapshots and that they convert an element field to a sequence of snapshots. However, the direct rendering method does not generate the image, but it renders the data directly after the user sends a request and renders the data of the next moment after a certain time interval. Compared with the animation, rendering is a real-time method (it does not need to generate frame by frame images and animations). Thus, rendering is easier for human-computer interaction. Users can pause at any time in the rendering process and use the mouse to pick up points for queries.

The main function of this module is to display the tidal data field of the water level changing with time based on the spherical earth model. The data are derived from the Taiwan Strait. To show the tidal bedding face data, different colors represent different water levels. This module also provides the function of querying of a single point of data and the function of numerical mapping of regional data. The visualization of surface processes is very intuitive. We can easily recreate the process of changing physical values through the visualization of the surface process and by finding the rule of this change. The interface is shown below (Tables 7.9 and 7.10).

1. Input item
2. Output items

After accessing the sub module, select the tidal bedding face data in the user data type drop-down list box. Tidal bedding face data include: dynamic change, information query and numerical mapping. Time is selected in the time drop-down list

Table 7.9 List of input items of the "single-factor data changes over time" submodule

Name	ID	Type	Description
Data sources	CXCL_Sjly	string	Input the data sources: Tidal bedding data
Operation type	CXCL_Czlx	string	User input operation types, including: dynamic change and single-point information and regional numerical mapping
The starting time	CXCL_Qsrq	DateTime	Input the start time as a condition. The data interval is 1 hour. Load 24 frame of data at a time by default

Table 7.10 List of output items of the "single-factor data changes over time" submodule

Name	ID	Type	Description
Tidal bedding face data of a single layer in a period	CXCL_Dccx	Object	When the operation type is "dynamic expression of single tidal data", display the tidal cloud image generated in the process of data preprocessing on the 3D sphere
The water level of tide at a specific time	CXCL_Ddxx	Float	When the operation type is "information query of the single point", click in the data region in the user data area and the water level will be displayed on the left side of the page
The mapping of tidal bedding face value	CXCL_Sztt	Object	When the operation type is "region mapping of numerical value", click twice on the data area and form a rectangular query. Query data and display water level values of the selected area in a 3D scene

box. Tide forecast data time resolution is selected as 1 h, namely 24 frames of data per day: 24 data in 1 day is chosen for dynamic display (the default). The "data load" button is clicked, then 24 data frames are loaded.

After data loading is completed, if the operation type is "dynamic", the start button is clicked. The dynamic change process of the tidal bedding face data starts to show in the scene of the three-dimensional sphere. After clicking the cancel button, the location label is removed. A regional box is temporarily created, and the 3D sphere window returns to its initial state Figs. 7.10 and 7.11.

After data loading is completed, if the operation type is "information query of a single point", the cursor will turn into an arrow. When the user clicks the data area of the sphere model, the water level of the location that the user clicked on will display on the left side of the page. After clicking the cancel button, the location label and the temporarily created regional box are removed, and the 3D sphere window returns to its initial state.

After data loading is completed, if the operation type is "regional information query", the cursor will turn into an arrow. When the user clicks the data area of the sphere model twice, a data rectangle is generated. After the query of the data files, the result is shown in the form of numerical mapping on the sphere model. After clicking the cancel button, the location label and the temporarily created regional box are removed, and the 3D sphere window returns to its initial state.

7.3.2.4 Dynamic Changes in the Multilayer Elements Field Over Time

The main function of this module is to visualize and display the ocean environmental factors in 3D. The feature types include sea temperature, salinity, density, sound velocity, and current. Through the 3D visualization plugin, we can query and display the elements of a single point of information. This system adopts the

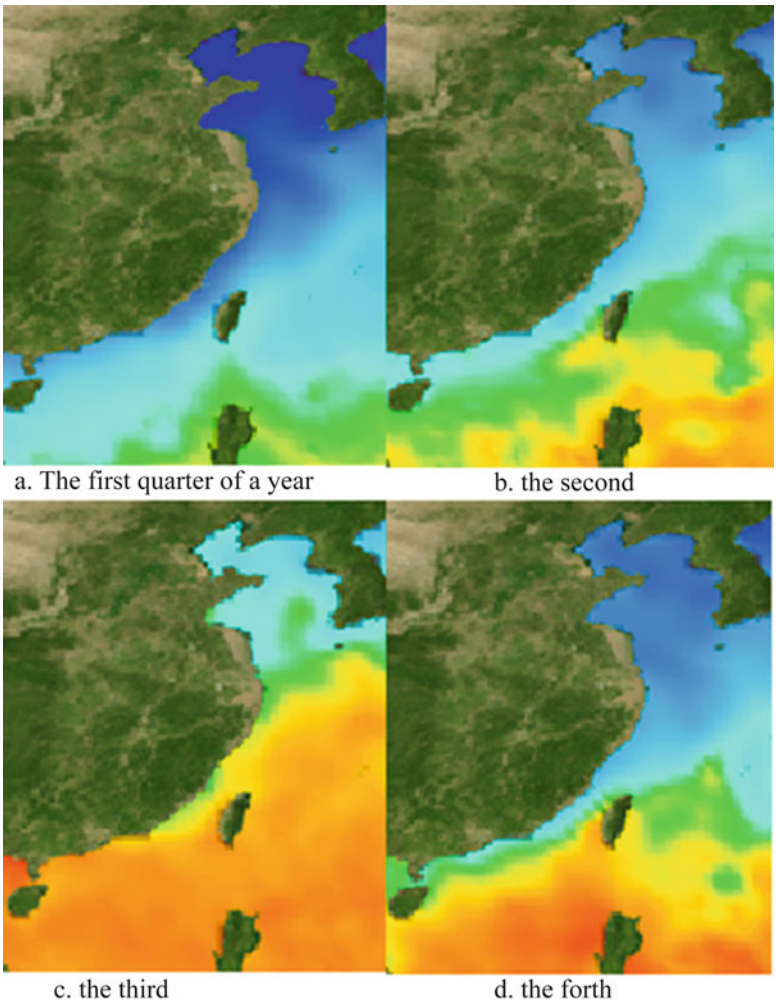


Fig. 7.10 The visualization of sea surface temperature field

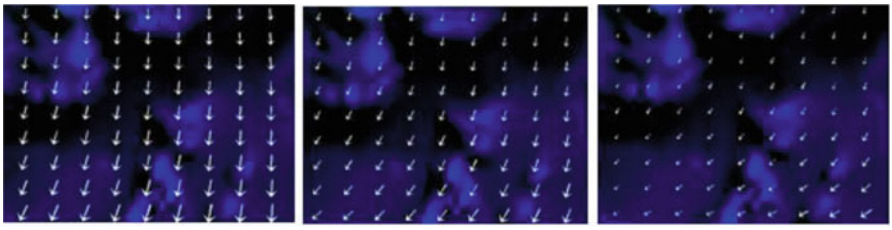


Fig. 7.11 The visualization of sea surface wind field

method of profile reconstruction of a static 3D visualization data field. Through the 3D visualization plugin, the 3D profile data structure is constructed and the data is then rendered frame by frame (Tables 7.11 and 7.12).

1. Input item
2. Output items

After accessing the sub module "operation page load" at the same time, the main window is redirected to the 3D display page.

The operation type is selected as "3D display" in the user operation type drop-down list box. The data type is selected as "forecast data" in the data type drop-down list box. The time is selected in the time drop-down list box. The OK button is clicked and the 3D data is loaded.

The 3D data of the element selected by the user will show in the 3D window. The left side of the function page shows the legend and operations, including the "mark", "land template", and "boundary". The check box corresponds to the contents shown in the window. The check box options such as "depth", "longitude", "dimension" and "arbitrary" are used to control the angle.

After data loading is completed, the data would display in the 3D window. The space key is clicked and then the 3D data grid begins to be displayed. When selecting a profile type as "depth", the "play" button is activated. The "play" button is clicked, and the changing process of 3D field data from the ocean surface to the underlying water is displayed.

In the 3D window, the user can control the point of view and the section position through the keyboard. The "depth", "longitude" or "dimension" checkboxes are checked in the 3D control panel. The keyboard "+" and "−" (for notebooks, use CTRL + numlk to open the keypad) are used to control different layers or different longitude or latitudes. The "arbitrary" checkbox is checked, and the t, g, f, and h

Table 7.11 List of input items of the "multilayer elements field dynamic process simulation" module

Name	ID	Type	Description
Feature type	DCDT_Yslx	string	Input optional element types, including temperature, salinity, density and sound velocity
Data sources	DCDT_Sjly	string	Input data source: forecast data
Operation type	DCDT_Czlx	string	Operation type: the visualization of temperature, salinity and density

Table 7.12 List of output items of the "multilayer elements field dynamic process simulation" module

Name	ID	Type	Description
3D scene	DCDT_Dcjtzs	Object	When the operation type is "3D visualization expression", display the 3D model in a new page, and profile the model along the axis or at an arbitrary angle. Show the elements changing with depth

keys are used to control the first point of latitude and longitude and the i, j, k, l keys are used to control the second point of latitude and longitude. When checking the "depth" checkbox, you can click on the "play" button. The 3D window will automatically display the changing process of a single element along with the layer depth. The "mark", "land", and "boundary" checkboxes are checked, and the corresponding template is then loaded into the 3D window.

When the element type is salinity, density, sound velocity or flow field, the OK button is clicked. The corresponding data will be loaded into the 3D window. The other operation is the same as with the operation of the temperature field described above. When the element type is "flow", "show vector arrow" is checked in the 3D control panel. The arrows indicate the flow field of the vector in the 3D window. After clicking the cancel button, the location label and the regional box that was temporarily created are removed, and the 3D sphere window returns to its initial state.

The dynamic process simulation of the elements field of a multilayer can be realized in two ways. The first method is to display the data from the surface to the underlying water to indicate the numerical change in the vertical direction. The second method is to visualize the dynamic process of the multilayer elements field through a vertical section of the dynamic process. With the aid of the 3D visualization engine, the latter constructed a 3D profile data structure at first, then it rendered the data frame by frame. By controlling the interval time, the method can achieve the purpose of expressing a body process. The process of a 3D salinity field data profile is shown in Figs. 7.12 and 7.13.

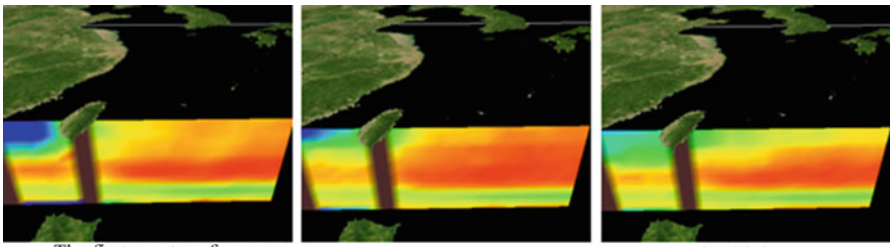


Fig. 7.12 The North latitude 23.5° profile of 3D salinity field visualization process

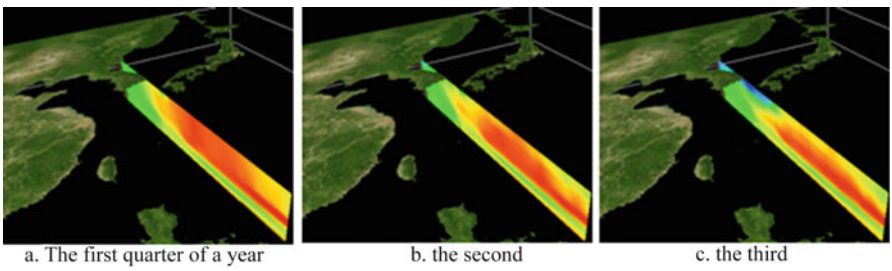


Fig. 7.13 The east longitude 130° profile of 3D salinity field visualization process

To improve the speed of real-time rendering, the system refers to a spatio-temporal data model for the ground state in which to render the data. In this model, the ground state generally refers to the data of the system last updated, namely a complete rendering of data for a moment (all grid points are rendering). The ground state distance means the number of adjacent two-ground-state data frames. The sea flow field data are different from other spatio-temporal data, because the land (continental shelf, continental shelf template) is constant, so no object will be added or deleted. Change domain means that the values of the domain attribute are different between the two adjacent data sets (including vector data of two attributes: the size and direction; any changes in the two attributes are considered as object state changes).

The ground state distance is mainly influenced by product data interval, area size, and grid density. The combination of single moment data volume and time sequence influences the distance. The density of a data grid will cause the total amount of data to grow geometrically. The ocean changes constantly. The change domain is determined mainly by the data reflecting the sea conditions at that moment. When the selected parameters change distinctly, the domain can be relatively large. When it starts to render, the system will determine the ground state, and the correct domain is dynamically selected according to the length of the selected time slot and the sea area. The grid point data values are compared. The grid is skipped when the data value is in the domain. The grid is rendered only when its value is out of the range of the domain. This strategy can save some rendering time to some extent.

7.4 The Prototype System of Analysis of the Storm Surge Disaster Process

7.4.1 Data Structure

The numerical forecast and data processing of storm surge is shown in Fig. 7.21. At first, according to the typhoon track data from Japan, South Korea, the United States and the national weather forecast, as well as years of experience, we forecast a relatively accurate typhoon path. Then, parameters were input into the system in the standard format. Numerical calculations were conducted and the storm surge forecast data were generated, including site forecast data and field data. Finally, the process in time and space was expressed and released online Fig. 7.14.

The existing numerical calculation model generally adopts the method of segmenting the research object to simplify the boundary problem in the calculation of the model. The segmentation in the storm surge forecast model is mainly manifested in two aspects. The first one is the grid segmentation of the study space. The other one is discretization processing of the storm surge process in

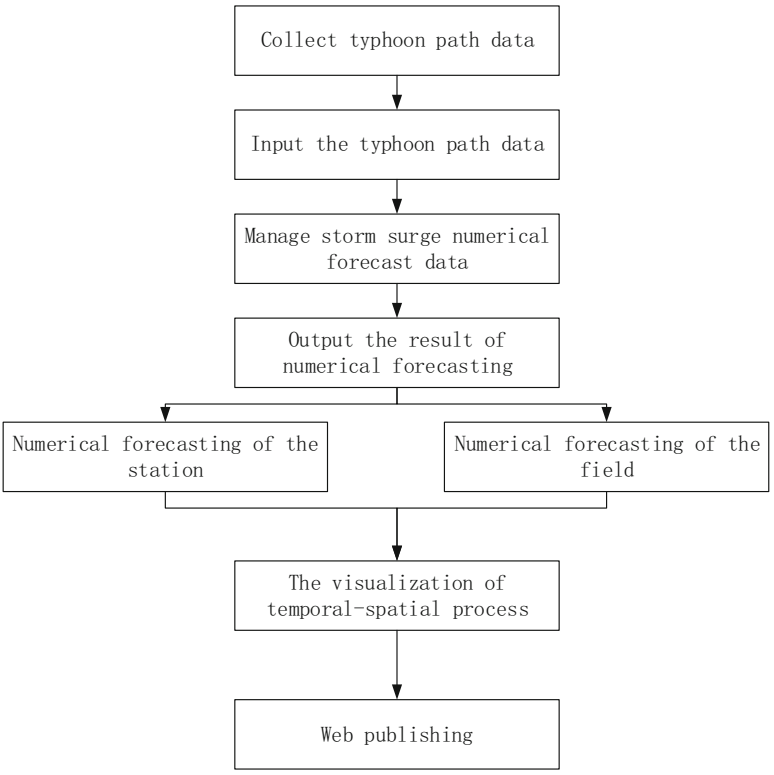


Fig. 7.14 Flow chart of the storm surge forecast and data processing

time. The output information of the model is a collection of discrete information in time and space.

The output data of the storm surge numerical forecast has the following characteristics: (1) Regarding a calculation model, the study area is usually fixed, especially those numerical models aimed at local characteristics. The model parameters always contain some regional experience parameters. Thus, the output data of the model is often recorded for the value of the element only. The relative positions are represented by the sequence of data, and the areal coverage is reflected in some parameters of the model. (2) Considering the structure of the initial input data and the model calculation problem, we use the regular shaped grid to divide the whole study area into a series of smallest uniform cells. In this way, the output values are the values of the corresponding elements of each cell. (3) For the predictive calculation, the time step is constant in this ocean disaster forecasting model. (4) Forecast data are divided into two parts: site forecast data and field data.

Therefore, the output data of the ocean storm surge disaster forecast model is a group of sequences repeated in a particular time interval and arranged in an orderly way. Each of these numerical sequences corresponds to a series of regular grids with small areas. To consider the model in the view of GIS, the forecast data at one

single time is very similar to the raster image data in terms of form. The differences are just in the method of recording and in the meaning of the data.

Based on the analysis of the characteristics of storm surge forecast data, the system needs to solve the standardization of the storm surge forecast data in a distributed network environment and for the network transport problems, so that the forecast data can be displayed and released on a unified platform. A typical application is the "digital ocean" project in China. In this project, to avoid repeated construction, each node at the provincial level is mainly responsible for production and release of the forecasting data. The National Center for Ocean Information is responsible for the operation and maintenance of the whole system to achieve the goal of the storage, distribution, and release of the data. Thus, to transfer the data through the Data Service Bus, it is necessary to standardize the storm surge forecast data.

Considering the requirements of multi-type data and scalability, we standardize the data by XML. The features of XML are that it separates the structured data from the other described information and that it allows the structured data from different sources to merge (integration). The client can easily extend the XML data to adapt to various data application requirements. From the perspective of the data description, XML is flexible, scalable, and it has good structure and constraints. From the perspective of data processing, its format is simple and easy to read, and it is easy to be processed by the program.

The process of transmission, display and release of storm surge forecast data is as follows. At first, the storm surge forecast site data and field data are produced by the node storm surge forecasting system. Second, the data are transformed to standard XML and are transferred to the data server through the data bus. The server will designate storage space for the data received. When the transmission is over, the data sever will add new records in the log. The monitoring service will read the log records to find the data directory and start to process these data. The service will generate visual raster data and update the process flag. The network publishing web foreground program can find that the new data has been processed at this time. "Loading data" is selected in the interface. The raster data, which has been processed, will be downloaded from the web server to the client. After matching the projection and optimizing the detail, storm surge data are loaded into the sphere. Now, the users can browse the dynamic space-time process information of storm surge, the tide field of each time and the water level of each point.

The numerical forecast of storm surge generates site data. These data contain a series of site latitude and longitude coordinates, adding water data, astronomical tide level and total water forecast values. These site data are discrete and are represented by multiple time sequences. Based on this characteristic, we can design a data storage structure of time series points combining elements of the ocean data model in terms of time sequence point storage structure.

The time sequence type description table is the description of specific forecast parameter value types, such as the name and units of the specific parameter, whether the time step is regular (this attribute field of the irregular time series

could not be defined), data type (mean, maximum, minimum, instantaneous value, etc.) and specific source (including the observation record and numerical model for calculating). The table can also include other description attribute information. The user can extend the table in a flexible way. For example, the storm surge water level, the total water level value and astronomical tide are important for analyzing the storm surge.

The time series storage table mainly stores the specific parameter values and timetable (Yasuko et al. 2010). The table is associated with a time series description table by the field of time sequence and number. This field is defined for elements for the situation where a point element has several time sequences.

The transition table is the connection table between the time sequence elements and the time series numerical storage table. The time sequence of one single parameter type may correspond to several elements of time elements. One element of single parameter types may have multiple time sequence values. To address this problem, a transition table is designed.

Using the storage structure of the time sequence of point elements above, we could store the storm surge forecast time series data efficiently. The advantage of this method is that it is concerned with multiple time sequences of storm surge forecast data. It is advantageous to query, retrieve and visualize the storm surge forecast data.

The storm surge numerical forecast data belongs to the scalar field data. It is unusual at present in the research of the organization and storage of scalar field data. Most existing methods are based on file management. We design a suitable storage structure for storm surge forecast field data, combining the field data storage structure with data in the grid model.

In the grid model, we take the area of some field data in a different dimension as a grid. This grid constitutes the basic structure of the whole grid network (Ying jie and ZhenhuaLv 2015). This method can effectively store and manage the regular vector field data and the scalar field data. In this article, the water level data of the 2D scalar field is the primary data.

As shown above, the scalar field numerical storage table stores the specific time and the numerical data associated with the field data. The grid elements represent the center of the field data in the grid unit. Its property fields represent the location information a grid point, including the grid number, grid column number, etc. The grid object table describes the entire field area, including the longitude and latitude of the starting position, the resolution, the total number of grid points, and so on. The time index table is designed to speed up the query of time elements. It assigns different times to different identities. This method could avoid low efficiency when querying data in a single table.

Based on this storage structure of the grid model, we can effectively manage storm surge numerical forecast field data. When the region of ocean elements field data is very large, the data have many spatial resolutions or the resolution in terms of latitude and longitude are different, we should also consider the problem of space region partitioning to efficiently access and query data.

7.4.2 *Software Structure Design*

The user inputs related parameters in the system interface. The prediction and analysis of the model needs necessary parameters. These parameters are used to trigger sub-functions. Combined with the relevant corresponding data from the data layer, the calculation of the model is completed. The result will be stored in the corresponding dataset in the data layer Fig. 7.15.

7.4.2.1 The Design of the Data Layer

The data layer is the foundation of the whole system. It mainly includes basic geographic information, auxiliary information and process information. From the perspective of the data set, it is divided into image data sets, the terrain data set, the social and economic data set, the data set of storm surge disaster prediction, the data set of analysis of the disaster result, etc. The layer is based on the Oracle database management platform. Oracle can manage the spatial data and attribute data efficiently. The content is as follows:

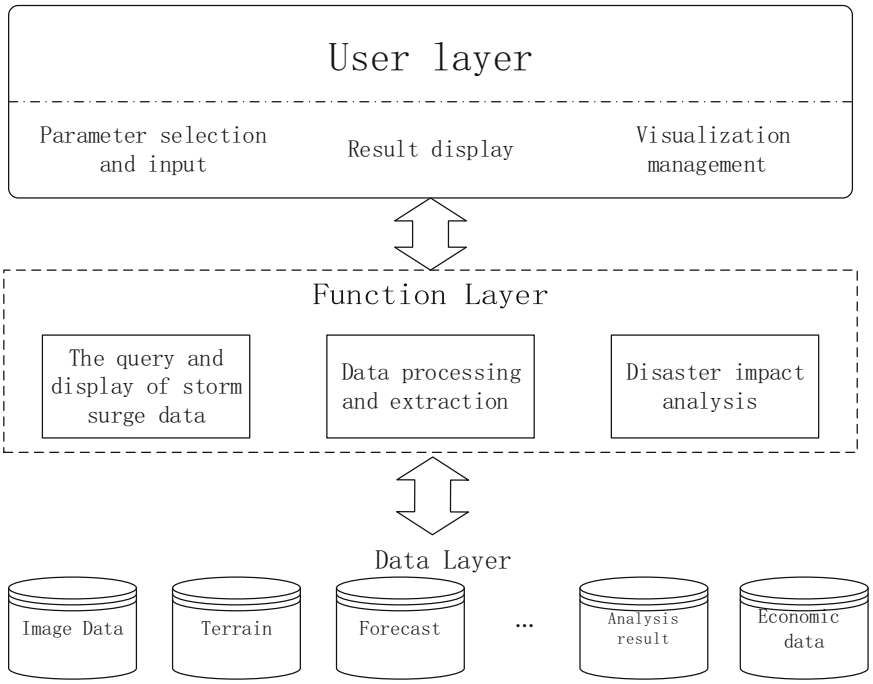


Fig. 7.15 The software framework of the typhoon storm surge process analysis prototype

- Basic geographic information

Basic geographic information is the basis for simulating a real 3D scenario and storm surge disaster forecast and analysis. It includes the global basic terrain and 1 km resolution remote sensing image data, coastal key research area high resolution terrain and image data, 1:250000 administrative division place names, terrain, roads, rivers, map data of key research areas, etc.

- Auxiliary information

Auxiliary information is necessary for disaster impact analysis. Different requirements need different degrees of detailed data. Auxiliary information can be divided into observation data, social and economic data and multiple model data. The content includes coastal tidal stations, the measured water level, station location information, population density with administrative region as the statistical unit, population, economy, land, lake area, other regional coastal buildings, offshore facility models, etc.

- Process information

Process information is the intermediate results of information produced with operation of the system. The information mainly includes prediction information of sea level, scope of flood information, analysis information of the effects of disasters, etc. The output will not change if the basic parameters are the same between different processes. Thus, the storage of the process information can raise the efficiency of the system.

7.4.2.2 The Design of the Function Layer

The functional layer is a collection of functions of the server. It is the core part of the whole system. It can be divided into three major sub functions: data query and expression, storm surge disaster prediction, and disaster impact analysis.

- Data query and expression

This module integrates the function in the software of multi-dimensional expression of ocean disaster elements in space and time.

- Storm surge disaster prediction

The storm surge disaster prediction sub-function is realized by the integration of the storm surge disaster prediction model. It can be divided into sea level forecast and the submerged range estimates.

- Disaster impact analysis

Disaster impact analysis is used for the statistics of the disaster area and the social economic index of this area. This function uses the prediction results in the storm surge disaster prediction and spatial analysis function in GIS.

7.4.2.3 The Design of the User Layer

The user layer is the interaction window between the user and the system. This layer realizes the client operations through the browser and plugin. It can be divided into three parts: parameter control, visual control and results display.

Parameter control determines the function and the parameters according to the input parameter and menu that the user chooses, and it triggers the corresponding sub-function needed. Visual control can realize the function of zooming and positioning on the sphere by the control panel. The system determines relevant parameters as part of the data extraction and optimization of parameters based on the state of the user perspective. The results are shown in the final results window. This part renders and displays the data back to the client's scene. These data are processed through a 3D plugin on the server.

The analysis prototype system of storm surge disaster in time and space adopts a three-layer structure. It includes the data layer (data supporting), the application layer (application service supporting) and the user layer. The configuration of the software environment of the data layer varies with the different data from different servers. The software needed mainly includes the operating system on the server, the 3D sphere scene software, and the database and spatial database engine. The application layer is a program of the deployment management prototype system. It provides browser-based access services for end users. The software includes the operating system on the server, the dynamic link library components that are dispensable for the application, etc. The configuration of the user layer refers to the configuration of the machine that needs to access the system and use all of the functions. In addition to the basic operating system and the browser, this machine still needs to download and install the corresponding plug-in. The list of the modules of this system is shown in Table 7.13.

7.4.3 Software Development

The typhoon spatio-temporal data query sub-module manages the historical typhoon data. It provides 3 query methods: query by the typhoon number, query

Table 7.13 List of modules of the analysis prototype system of the process of storm surge disaster in time and space

No.	Module name	ID	The parent module ID
1	Main module	MAIN_M	–
2	Data management	SJGL_M	MAIN_M
3	Data query	SJCX_M	MAIN_M
4	The expression of time and space disaster process	GCBD_M	MAIN_M
5	The expression of disaster forecast results	JGYC_M	MAIN_M
6	User-privilege management	YHGL_M	MAIN_M

by the time period and query by the typhoon level. This module returned typhoon information (one or more) to the user in the form of a data table, including the coordinate information of the typhoon path, wind speed, and wind direction. Its interface is shown below.

The site data generated by the storm surge numerical forecast contains a series of latitude and longitude coordinates of the sites, adding water, astronomical tide levels and total water forecast values. The system could also mark the site of the warning level on the sequence diagram with a red line. It shows the relationship between the numerical forecast of the total level of storm surge and the warning level.

Combining the storm surge forecast sites with the query of the forecast time sequence, we could realize the expression of adding water, astronomical tide, and total water level of storm surge forecast sites.

The field data of a storm surge forecast is raster data in a time sequence. The value of the raster data is the value of the storm surge water level on this grid. The grid data of storm surge forecast is converted into a sequence of texture feature information. The mapping relationship is kept between texture feature information and the original data. The texture information could transfer to the original forecast data through the inverse process of the mapping relationship. Thus, the visualization and query of storm surge tide field data can be realized. The principle is shown as follows. The sea surface height of storm surge information is expressed by color. A corresponding criterion is established between the sea surface height value and the color texture. A color sequence can be generated by this relationship. The color sequence is filled in the corresponding grid. We can obtain the corresponding texture map data. The texture mapping relationship can be expressed by the following equation:

$$H(i) = R(i)G(i)B(i)$$

H represents the relative height of the storm tide. R, G, and B represent the three primary colors, and 'i' is the value of the corresponding ordinal. One sequence (e.g., 123321123) will yield a corresponding color value mapping sequence. The structure is shown in Fig. 7.16.

By the same token, texture information can be inverted to the original information according to the corresponding standards. Texture data could be generated through the former, whereas the latter is used to query the results of the original visual information. We can display the change in the whole storm surge water level dynamically and query the value of any point or any time. We can display the multilevel resolution information after establishing the image pyramid. The water level information of the powerful typhoon "SEN" is shown as Fig. 7.17.

The field data of the numerical prediction of storm surge is commonly a grid dataset as a time sequence. Each moment corresponds to a tidal field datum. These data are rendered into raster data. The raster data unit value is the value of storm surge water level in this space grid. The field data of storm surge forecast can be

1	2	3
3	2	1
1	2	3

R(1)G(1)B(1)	R(2)G(2)B(2)	R(3)G(3)B(3)
R(3)G(3)B(3)	R(2)G(2)B(2)	R(1)G(1)B(1)
R(1)G(1)B(1)	R(2)G(2)B(2)	R(3)G(3)B(3)

Fig. 7.16 The mapping of storm surge texture field data

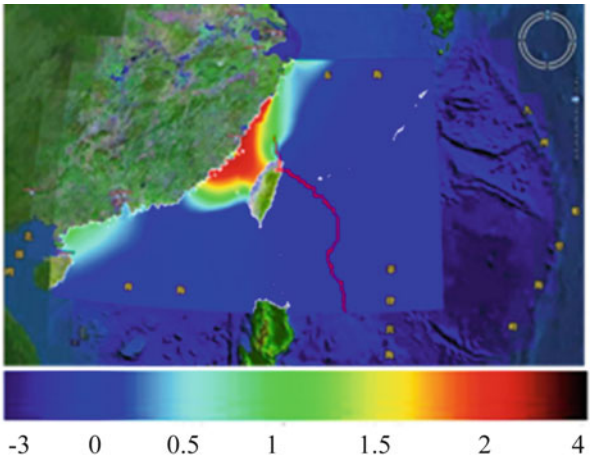


Fig. 7.17 Tidal information on 2008-09-09 14:00

processed as a sequence of texture feature information. The mapping relationship is kept between the texture feature information and the original data. The texture information could transfer to the original forecast data through the inverse process of the mapping relationship. Thus, the visualization and query of the storm surge tide field data can be realized.

The query and expression of the storm surge tide field data can be divided into three modes: (1) the query is based on a single position at some storm surge water level value; (2) the entire field of storm surge water from a static query display is based on a certain point in time; and (3) visual expression of spatial and temporal variation processes of water is based on a certain period of time in the process of a storm surge.

Figure 7.17 is a static display of storm surge water level at some point based on the field. The left and right sides show the storm surge water level information on September 9, 18 and 21 of 2009.

Figure 7.18 shows the dynamic change process as a function of time based on the field of the storm surge water level. When the system operates in actual use, the user needs to choose the beginning and ending times. The image below corresponds to the dynamic change of storm surge water level information in one moment.

Various methods are synthesized in a scenario. Multiple factors are shown in different ways by calling the different visual expression methods of a space object. The pic shows the comprehensive expression of the No. 0813 typhoon storm surge process. The typhoon path and the storm surge water field are displayed in the same

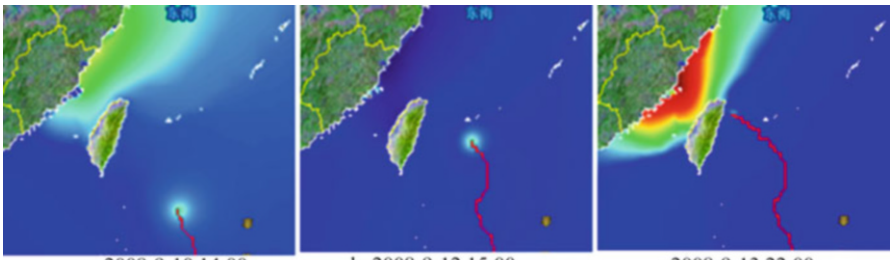


Fig. 7.18 The comprehensive visual expression of typhoon storm surge process

scene. We could see that the water level increases distinctly in the typhoon eye. At the same time, this system supports the comprehensive visual expression of cloud and numerical mapping in the same scenario or when using two-dimensional scalar field data, two dimensional scalar field data cloud or two dimensional vector field diagram methods.

The storm surge of a powerful typhoon storm (0608) during August of 2006 is taken as an example for analysis. The site data includes Shacheng, Sansha, Meihua, Pingtan, Wungang, Chongwu, Shenhu, Xiamen and Dongshan. The three-dimensional flow field data adopted the calculation results by a numerical model developed by Xiamen University. The disaster field three-dimensional grid interval is 1/32 degrees; it has 25 layers, and it has a vertical space range of 110 to 130 degrees east longitude and 18 to 30° north latitude. Prediction and simulation experiments use the data of Tianjin Binhai, a new area combined with the national ocean information center forecast data.

The hazard prediction results expression module is responsible for the prediction of storm surge disasters on the basis of the analysis of storm surge data. The hazard prediction includes GDP, population, transportation facilities, water area, and so on.

7.5 The Analysis of Sea Level Rising

7.5.1 The System Structure

This system builds on the LAN environment. It can be divided into the user layer, the function layer and the data layer. The system framework is shown in Fig. 7.19. The data layer is mainly used for the organization and management of the system data, and it provides data support for the realization of system functions. The data mainly includes basic data and process data. Function layers mainly implement data processing and extraction, analysis of a range of sea level rising forecasts, flood and disaster impact analyses, and other functions. Finally, this layer provides forecast and analysis results for the users. The user layer provides users with a browser-

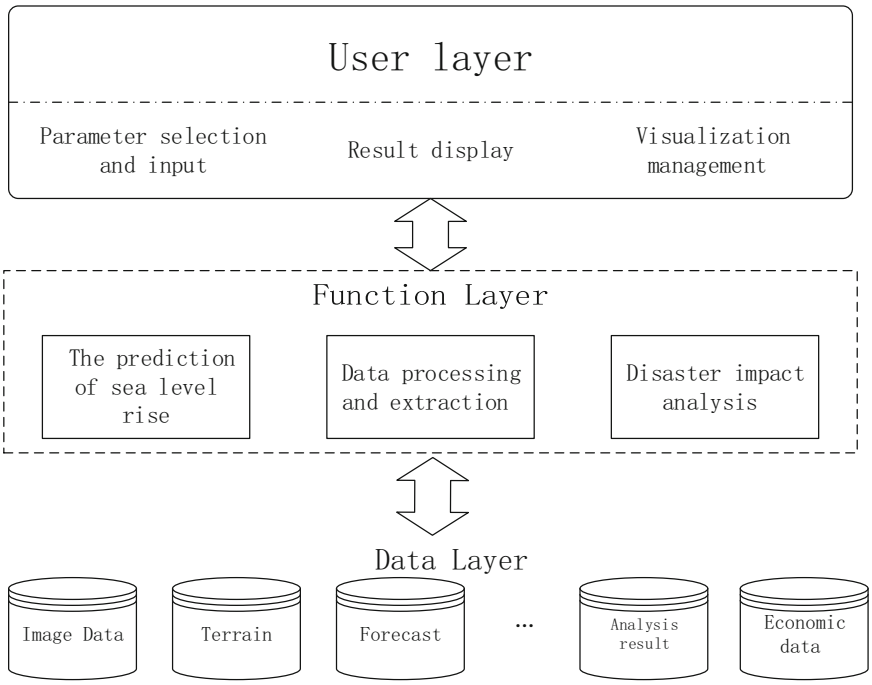


Fig. 7.19 The software framework of the sea level forecast system

based parameter control interface and information display window based on a three-dimensional sphere.

User input parameters are related through the system interface. The prediction and analysis of the model need necessary parameters. These parameters are used to trigger relative sub-functions. Combining with the relevant corresponding data from the data layer, the calculation of the model is completed. The result will be stored in the corresponding dataset in the data layer (Zhang Xin et al. 2011, 2013).

According to user's choice of scenarios, the required data is extracted and the data is optimized. After receiving this data, the client starts to render and display them.

The data layer is the foundation of the whole system. It mainly includes basic geographic information, auxiliary information and process information. From the perspective of the data set, it is divided into image data sets, the terrain data set, the social and economic data set, the data set of storm surge disaster prediction, the data set of analysis of the disaster result, etc. The layer is based on the Oracle database management platform. Oracle can manage the spatial data and attribute data efficiently. The content is as follows.

1. Basic geographic information

Basic geographic information is the basis of simulating a real 3D scene and storm surge disaster forecast and analysis. It includes the global basic terrain and 1 km resolution remote sensing image data, coastal key research area high resolution terrain and image data, 1:250000 administrative division place names, terrain, roads, rivers, map data of key research areas, etc.

2. Auxiliary information

Auxiliary information is necessary for disaster impact analysis and the display of results. Different requirement need different degrees of detail in the data. Auxiliary information can be divided into observation data, social and economic data and multiple model data. The content includes coastal tidal stations, the measured water level, station location information, population density with the administrative region as the statistical unit, population, economy, land, lake area, other regional coastal buildings, offshore facility models, etc.

3. Process information

Process information is the intermediate result of information produced by the operating of the system. The information mainly includes predictions of sea level, scope of flood information, information related to the analysis of the effects of disasters, etc. The output will not change if the basic parameters are the same between different processes. Thus, the storage of the process information can raise the efficiency of the system.

Functional layers are a collection of functions of the server. They are the core part of the whole system. They can be divided into three major sub functions: sea level rising forecast, disaster impact analysis and data extraction and optimization.

1. Sea level rising forecast

The sea level rising forecast sub-function integrates the forecasting model. The forecast can be divided into sea level forecast and submerged range estimates.

2. Disaster impact analysis

The disaster impact analysis function uses the prediction results in the rise in sea level forecast sub-function. It analyzes the disaster area and the related social economic index of statistics through spatial analysis in GIS.

3. Data extraction and optimization

Data extraction and optimization improve the effectiveness and efficiency when showing the result. According to the parameters that the user chooses and the angle of the view, the system determines the type of data extraction and the degree of optimization, and it provides the required data for the client.

The user layer is the interaction window between the user and the system. This layer realizes the client operations through the browser and plugin. It can be divided into three parts: parameter control, visual control and result display.

Parameter control determines the function and the parameters according to the input parameter and menu that the user chooses, and it triggers the corresponding sub-function needed. Visual control can realize the function of zooming and positioning on the sphere by the control panel. The system determines relevant parameters as the data extraction and optimization of parameters based on the state of the user perspective. The results display is the final window for displaying results. This part renders and displays the data back to the client's window. These data are processed through a 3D plugin on the server.

7.5.2 System Implementation

7.5.2.1 Model Integration Based on Service

The sea level rising forecast model is mainly used for sea level change predictions and the estimation of corresponding flooded areas. Currently, there are two methods to forecast sea level change: the observation method and the numerical simulation method. The observation method forecasts by changing trends based on years of sea surface height observation data. The numerical simulation method is based on a certain prediction model. This method predicts sea surface height through the adjustment of the parameters under different conditions. The estimation method of the flooded area can be divided into non-source flood and source flood.

The system adopts the numerical simulation method to predict sea levels in the forecasting model of sea level rising in the Tianjin area. The estimates of the flooded area are improved at the base of the source flood algorithm. By adding water bursts into the model, the method is used to calculate the evolution of the water. The specific idea is as follows. The user sets the tide level, risk prediction and embankment parameters to calculate the largest rise in sea level. A sea surface of 0 m is taken as the original surface. The range of the submerged area when the sea level risings to a certain height through a flooded area model is estimated. The data set predicting changes in height from 0 m in the same interval is produced.

The sea level rising forecast model varies with the prediction area. In this system, we adopt the web service technology to encapsulate the existing model uniformly. The system calls the service dynamically according to the region that the user specifies. In the process of service encapsulation, the system follows unified rules of the service description so that the system can realize the integration of the different models and calls through a unified interface.

We use standard XML files to store the description of model service according to the following rules:

1. The file must contain three level 1 labels: forecast area, controllable parameters and input data.
2. The forecast regional tag must contain the name and area. Text is used to describe the model forecast area in the label name tags, such as “Tianjin”, etc. A coordinate (X0, Y0, X1, Y1) is used to determine the precise location of an area in the area label. The four values represent the longitude and latitude of the upper left corner and the longitude and latitude of the bottom right hand corner, respectively. At least to two decimal places are kept.
3. Controllable parameters must contain the number of labels and the total number of parameters, which is same as the label number. Parameter labels include three level 3 tags: parameter descriptions, parameter types and parameter values.
4. The input data label must also contain a number of data labels that shows the total number of input data. This number is consistent with the number of data labels. Each data tag contains three levels 3 tags: data description, data type, and file location.

The above label uses a unified identifier corresponding to the system to ensure the standardization of different service descriptions and the identification of the system.

After the encapsulation of the model, the service description file and the service are released and registered at the same time in the system. The system provides users with an optional forecast area through the area list. The user needs to input the related parameters so that the system extracts the relevant data from the data layer based on the description of the area that the user selected. The forecasting model starts to predict and store the result into the corresponding data set.

7.5.2.2 Data Extraction and Optimization Based on Scenarios

The simulation of a real 3D scenario requires terrain data. The integration of image and other types of data leads to the expansion of data. The volume of a range area of a 3D scene is often several times or more the volume of the two-dimensional map data. For the system based on a network, the network transmission and the machine itself has many pressures to process mass data. To display the 3D scene more smoothly and to simulate rising sea levels based on the three-dimensional sphere, the amount of data must be reduced as much as possible, and the display effect should not be affected. Thus, it is necessary to extract and optimize the original data.

The area that the computer screen displays is visible to users. Under the premise of a constant window, the size of the object that can be distinguished increases with the increase of the visible range. This principle determines the extraction area and the optimization degree based on the user’s perspective.

1. The current user's visible range and the surrounding related scene data and predicted results are extracted.
2. According to the current user's visual angle height, the current optimization level is determined. The higher the angle, the higher the optimization level and the lower the actual resolution of that scene.
3. Based on a dynamic simulation rate, the data time optimization level is determined. The higher the rate, the higher the optimization levels.
4. The user perspective is kept unchanged when dynamically simulating the process of sea level rising.

In the Tianjin coastal area, which has the largest risk of a once-in-a-century tide, the predicted rises in sea level are taken as an example. The specific extraction and optimization method and processes include:

1. Extract: the current visual area (X0, Y0) – (X1, Y1) is calculated from the user perspective parameters and the screen size. The original data space is cut by the visual area. The data after the extraction are within the scope of this data.
2. Space optimization: The current user perspective height is obtained from the user parameters. The space optimize level and the optimization parameters (a, N) are confirmed. When the change of the slope of the adjacent nodes is greater than a, this node is retained. When continuous, the slope change of multiple nodes is less than a. When the node number is larger than N, 1 of every N is kept, and the data are optimized on the basis of the extraction.
3. Time optimization: the time optimization level and parameter h are determined by the rate identified by the user. The submerged range data are optimized according to interval h.

Table 7.14 shows the situation from different perspectives, including the optimization ratio in the same data simulation rate. The values in the table refer to the proportion of the current operating data and the data at the next higher level. The Tanggu area selected is approximately 1/100 of the area of Tianjin selected. Figures 7.20 and 7.21 show the sea level rising simulation.

Table 7.14 The contrast table of optimized data in different areas

Area	Tianjin	Tanggu
Original data	1	1
Scope of extraction	1	1/10
Spatial optimization	1/10	1
Time optimization	1/10	1/10
Total	1/100	1/100

Fig. 7.20 The simulation picture of Binhai district in Tianjin

Fig. 7.21 The simulation picture of Tanggu district in Tianjin

Acknowledgments The study is funded by the Free Exploring Program of the State Key Laboratory of Remote Sensing Science of China (No. 14ZY-03), the Special Research Project for the Commonweal of the Ministry of Water Resources of the People's Republic of China (grant no. 201201092), the 908 Project of the State Oceanic Administration, China (No. 908-03-03-02), the National Natural Science Foundation of China (grant no. 61473286 and 61375002), the National Science & Technology Pillar Program (2015BAJ02B01, 2015BAJ02B02). All members

of the Research Team are gratefully acknowledged for their contributions to the work carried out in this chapter in recent years. In particular, special thanks are given to Dr. Jian Liu, Miss Qiong Zheng, Mr. Yongxin Chen and Mr. Yuqi Liu for their contributions in the document revision.

References

Shen YZ, Austin JA, Crouch JR, Dinniman MS (2007) Interactive visualization of regional ocean modeling system. In: Proceedings of the IASTED international conference on graphics and visualization in engineering, pp 74–82

Su FZ, Du YY, Pei XB (2006a) Constructing digital sea of China with the Datum of coastal line. Geo-Inf Sci 8(1):12–15

Su FZ, Yang XM, Xu J (2006b) Basic theory and key technologies for ocean geographic information system. Acta Oceanol Sin 25(2):80–86

Su FZ, Zhou CH, Zhang TY (2006c) Constructing a raster-based spatial-temporal hierarchical data model for ocean fisheries application. Acta Oceanol Sin 25(1):57–63

Thomas CM (2003) The coastal module of the Global Ocean Observing System (GOOS): an assessment of current capabilities to detect change. Ocean Policy 27(3):295–302

Wright DJ, Blongewicz MJ, Halpin PN, Breman J (2007) Arc ocean: GIS for a Blue planet. ESRI Press, Redlands

Wright TE, Burton M, Pyle DM, Caltabiano T (2010) Visualising volcanic gas plumes with virtual globes. Comput Geosci 35(3):1837–1842

Yasuko Y, Yanaka H, Suzuki K, Tsuboi S, Isse T, Obayashi M, Tamura H, Nagao H (2010) Visualization of geoscience data on Google Earth: development of a data converter system for seismic tomographic models. Comput Geosci 36(2):373–382

Ying jieHu, Zhenhua LA (2015) Multistage collaborative 3D GIS to support public participation. Int J Digital Earth 8(3):212–234

Zhang Xin, Dong Wen, Li Sihai, Luo Jiancheng, Chi Tianhe (2011) China digital ocean prototype system. Int J Digital Earth 4(3):211–222

Zhang Xin, Dong Wen, Xiaoyi Jiang et al (2013) “Digital earth” in support of an online oceanic educational public service and popularization. Acta Oceanol Sin 32(5):82–86

Chapter 8
Coastal Remote Sensing

Changming Zhu and Xin Zhang

Abstract Coastal zones are a junction between the land and the ocean. The dynamic nature of the coast makes it difficult to clearly define the borders of coastal zones. Sometimes these zones are referred to as tidewater areas, extending from the coast to approximately 10 miles inland. Ketchum defined a coastal zone as the band of dry land and adjacent ocean space (water and submerged land) in which terrestrial processes and land uses directly affect oceanic processes and uses and vice versa (Ketchum BH, The water's edge: critical problems of the Coastal zone. MIT Press, Cambridge, MA, 1972). Coastal zones are one of the most ecologically valuable and biodiverse regions in the world. They provide food, freshwater and other marine products for human beings, as well as play an important role in the local environment. However, human activities and global climate change have given rise to natural disasters that threaten coastal zones environmental security. In particular, the effect of industrial projects on coastal zones is worsening. Thus, it is important to study these areas to provide guidance for coastal resource management and environmental protection. Numerous researchers have attempted to monitor coastal zone dynamics based on remote sensing data. In this chapter, we summarize various topics concerning the application of remote sensing in coastal zones. We focus on describing several new methods, including coastline automatic extraction, intertidal zone identification, coastal wetland classification and coastal invasive plant detection, using remote sensing.

Keywords Coastal zone • Remote sensing • Digital coast • Coastline • Coastal wetlands

C. Zhu
Department of Geography and Environment, Jiangsu Normal University, Xuzhou 221116, Jiangsu, China
e-mail: ablezhu@163.com

X. Zhang (✉)
State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China
e-mail: zhangxin@radi.ac.cn

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_8

8.1 Coastline Automatic Extraction with Remote Sensing Data

8.1.1 Introduction of Coastline Extraction Technology

Coastlines are the baseline dividing sea and land in ocean management. The ability to rapidly and accurately retrieve the position of coastlines and detect their variation is very important in coastal zone management and planning. Traditionally, coastline acquisition and updating has been accomplished through field surveys, which are time consuming and labor intensive. Remote sensing techniques have been used to address these challenges. The development of remote sensing technology has provided a new way to collect coastline information quickly and accurately. Coastline extraction and delineation is a major application filed of remote sensing technology. Many studies have used remote sensing data to automatically delineate coastlines (Chen et al. 2014; Ekercin 2007; Lee and Jurkevich 1990; Mason and Davenport 1996; Sun et al. 2014; White and El Asmar 1999; Yao et al. 2013).

Essentially, coastline extraction method is to separate water bodies and land on remotely sensed image. There are many methods used for water extraction, including water index segmentation and spectral classification (Frazier and Page 2000; Liu and Jezek 2004; Luo et al. 2009; Ouma and Tateishi 2006; Ryu et al. 2002). These methods can be divided into 3 categories—edge detection, threshold segmentation and classification (Table 8.1)—each of which has advantages and

Table 8.1 A summary of the major algorithms used for coastline extraction

Algorithm	Description	Advantages	Disadvantages	References
Edge detection methods	Using an edge detector (such as Canny, Soble and etc.) to detect water boundary according to gradient	Suitable for coastline extractiononSAR or radar images	The calculation process is complicated and the continuity of result is poor	Lee and Jurkevich (1990), Han and Jing (2005), Niedermeier et al. (2001), Du Tao (1999), Mason and Davenport (1996)
Supervised classification methods	Using supervised classification algorithm to discriminatewater bodies from the wholeimage	High accuracy and stability	Sample selection need human interference	Ryan et al. (1991), Zhu (2002), Zhu et al. (2013)
Threshold segmentation methods	According to the difference between water spectra and others via a threshold to separate sea and land. It is referred to as density slice method	Simple and easy toimplement	The optimal threshold is hard to determine	Sohn and Jezek (1999); Huo et al. (2003), Wang et al. (2005), Ryu et al. (2008), Cui et al. (2007)

disadvantages. Edge detection is suitable for coastline extraction from SAR; however, the extracted line is not community. In turn, classification yields the most accurate results; however, human interaction is required to select the classification samples. Threshold segmentation is the simplest method, but its weakness lies in determining the optimal threshold determination parameter.

Because the spectral characteristics of coastal waters depend on the regional environment, the traditional normalized difference water index (NDWI) threshold segmentation method easily misclassify water pixels into land. This error will seriously affect the accuracy of shoreline delineation. Thus, an automatic coastline extraction method that uses SVM classification and sample auto-selection has been proposed based on the NDWI segmentation model (Zhu et al. 2013).

8.1.2 New Methods for Coastline Extraction

The sample auto-selection and SVM classification method for automatic extraction of coastlines consists of five steps (Fig. 8.1). First, the initial water distribution information is obtained using the NDWI calculation and global threshold segmentation. Second, the classification samples are automatically selected based on NDWI information. Third, water and land are separated with an SVM classifier. Finally, small terrestrial water bodies are filled and coastlines are automatically tracked. Experimental results show that this method can enhance the ability to identify coastal waters and improve the accuracy and automation of coastline extraction from remote sensing imagery.

8.1.2.1 NDWI Threshold Segmentation

The NDWI histogram has a bimodal distribution, with a background peak on the left and a water peak on the right (Li and Sheng 2012; Luo et al. 2009; Ryu et al. 2002). This feature is the theory basis of NDWI segmentation algorithm

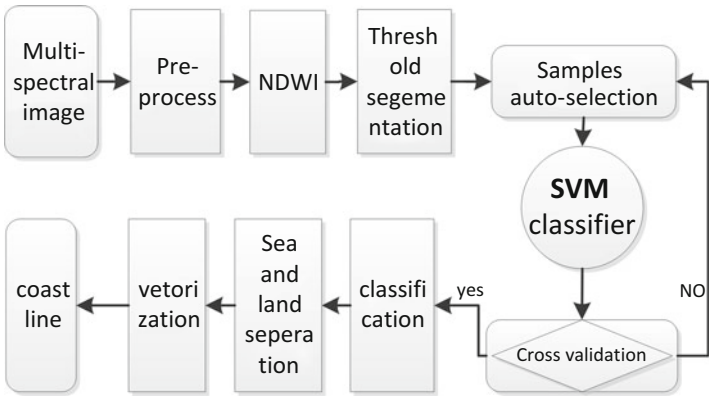


Fig. 8.1 The overall technical flow chart of coastline extraction

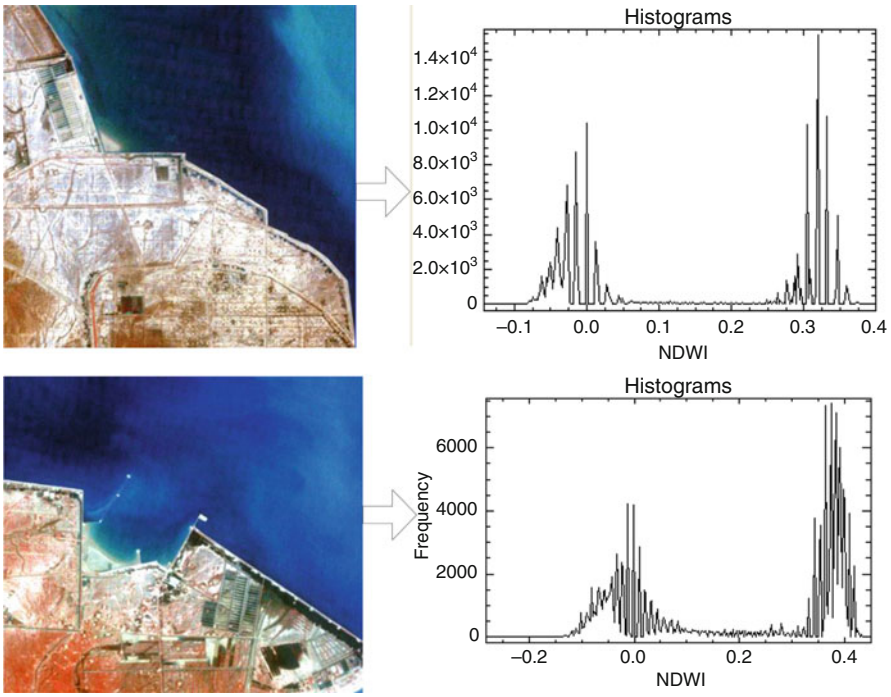


Fig. 8.2 The NDWI histogram in the coastline bufferzone

used for water body extraction, as shown in Fig. 8.2, where the X-axis is the NDWI value and the Y-axis is the frequency.

In theory, the histogram's bimodal distribution can be described as a Gaussian bimodal distribution (Fig. 8.3). The first wave peak of the NDWI histogram represents the pixels that are most similar to water. Next, the threshold segmentation method is used to find the first trough (T_0) of the histogram waveform, which is taken as the threshold used to segment the index image.

The segmentation threshold is determined by the formula (Li and Sheng 2012) (Otsu 1975) Eq. (8.1):

$$T_0 = \frac{\mu_{Water} \times \sigma_{Land} + \mu_{Land} \times \sigma_{Water}}{\mu_{Water} + \mu_{Land}} \tag{8.1}$$

Where μ_{water} and μ_{Land} are the mean values of the water peak and the land peak, respectively, and σ_{water} and σ_{Land} are the corresponding standard deviations.

The formula of NDWI calculation is as follows Eq. (8.2) McFeeters (1996):

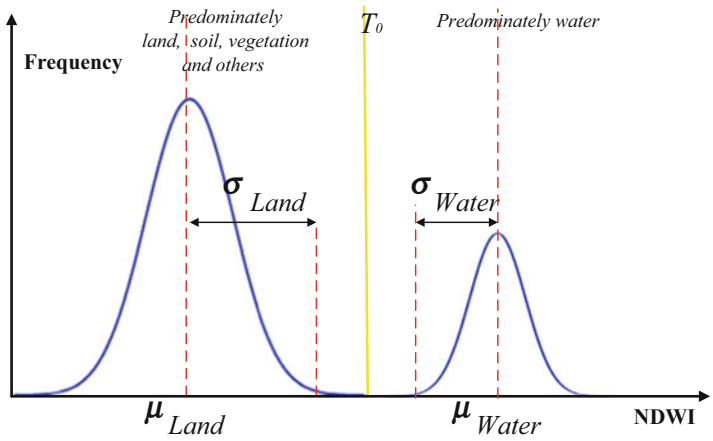


Fig. 8.3 Sketch of the NDWI histogram distribution for water bodies

$$\mathrm{NDWI} = \frac{\rho_{\mathrm{Green}} - \rho_{\mathrm{NIR}}}{\rho_{\mathrm{Green}} + \rho_{\mathrm{NIR}}} \tag{8.2}$$

Where ρ_{Green} is the green band reflectance and ρ_{NIR} is the near infrared band reflectance. It should be noted that, the NDWI model has its physical basis, which is according to the spectral characteristics of water bodies. Thus, prior to calculation of the NDWI, remote sensing imagery preprocessing is essential to convert image DN values to radiance (Chander et al. 2009; McFeeters 1996).

8.1.2.2 Automatic Selection of Samples

After NDWI threshold segmentation, we determined the primary water distribution. Classification samples were randomly and automatically selected from land and water areas at the basis of NDWI preliminary segmentation. Next, the samples were divided into three sample datasets, of which two were used for testing and one was used for validation.

8.1.2.3 SVM Classification Model

The support vector machine (SVM) supervised machine-learning algorithm was used for water reclassification. The SVM algorithm was proposed and used to find the optimal hyperplane of two classifiable samples (Mountrakis et al. 2011; Vapnik et al. 1996; Vapnik and Kotz 1982). The SVM model aims to construct a function following form given in Eq. (8.3):

$$y = f(x, w) = wT\varphi(x) + b, \tag{8.3}$$

where the non-linear mapping function $\varphi(x)$ maps the input data into a higher dimensional feature space, wT is the regression coefficient and b is the regression constant.

The kernel is related to the transform function $\phi(x)$ by the equation $k(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$. The kernel function $k(x_i, x_j)$ can be chosen as any one of the following:

(a) Polynomial (homogeneous): $k(x_i, x_j) = (x_i \cdot x_j)^d$

(b) Polynomial (inhomogeneous): $k(x_i, x_j) = (x_i \cdot x_j + 1)^d$

(c) Gaussian radial basis function: $k(x_i, x_j) = \exp(-\gamma ||x_i - x_j||^2), \quad for \quad \gamma > 0$

(d) Radial basis function: $k(x_i, x_j) = \exp\left(-\frac{||x_i - x_j||^2}{2\delta^2}\right)$

(e) Hyperbolic tangent: $k(x_i, x_j) = \tanh(kx_i \cdot x_j + c)$

In this article, we used the LIBSVM 2.9 algorithm software package, which can be downloaded for free (http://wenku.baidu.com/view/d32b9a40be1e650e52ea992f.html/). We selected the radial basis function (RBF) as our kernel function. More detailed descriptions of the SVR method have been given by previous authors (Chang and Lin 2011; Xi et al. 2011; Zhang et al. 2005; Zhu et al. 2012).

Based on the sample dataset, SVM model training was initiated (Fig. 8.4). To improve the classification accuracy, the training samples were divided into 3 subsets. Each time, 2 sample subsets were selected; one was used for training and the other for validation. Samples with large error samples were rejected. The classification model used threefold validation to determine the optimal values for the parameters g and C.

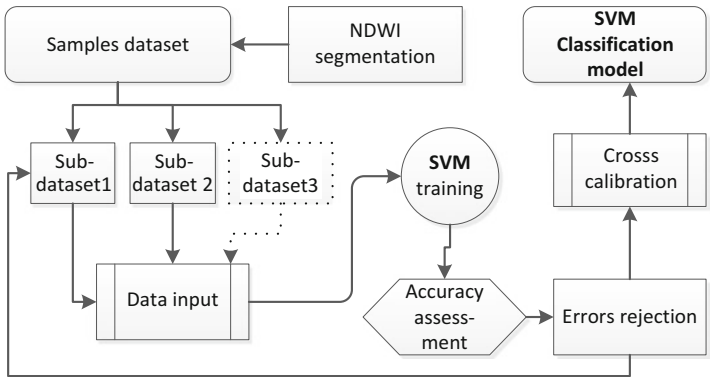


Fig. 8.4 The process of SVM model training

8.1.2.4 Post-processing

1. Land water body removal

Using a priori spatial knowledge, all lakes in the land area were removed from the classification image. The algorithm used in this paper aimed to divide sea and land through spatial analysis, so water bodies that are not adjacent to the ocean were removed.

2. Raster to vector conversion

In binary classification image, 0 indicates a land pixel and 1 indicates a water pixel. After removal of the lakes, all pixels labeled as water pixels could be regarded as sea areas. Next, vectorization of the boundary between sea and land got coastline distribution information.

3. Estuary and lagoon revision

For large rivers, the characteristics of estuaries should be maintained. For small rivers, the coastline is considered to parallel the land. For lagoons, the environmental conditions should be considered. When there were wide passages (in this study, areas greater than 3 pixels) linking the sea with a lagoon, the inside of the lagoon was used as the coastline. Otherwise, the coastline was defined as outside of the lagoon.

8.1.3 Experiment and Analysis

Enhanced thematic mapper plus (ETM+) images were selected as test data. The images were acquired on May 13, 2015 (path row) and the study area is located on the Liaodong peninsula. The process of coastline extraction is illustrated in Fig. 8.5. Figure 8.5a shows the ETM+ 4/3/2 false color composite image after radiometric correction. Figure 8.5b shows the NDWI image. Figure 8.5c shows the NDWI threshold segmentation. Figure 8.5d shows the automatic selection of classification samples based on the NDWI. Figure 8.5e shows the location of the sample distribution. Figure 8.5f shows the classification result based on the SVM classifier. Figure 8.5g shows the results of land and sea separation after removal of lakes and water bodies on land. Figure 8.5h shows the results of vectorization and post-processing. Figure 8.5i shows the final coastline distribution results.

Figure 8.6 compares the accuracy of the SVM classification and the threshold method. Figure 8.6a shows the coastline overlaid on the NDWI threshold segmentation image. Figure 8.6b shows the coastline overlaid on the SVM classification image. Figure 8.6c, d show zoomed out images corresponding to Fig. 8.6a, b, respectively. Figure 8.6a, c show that a significant amount of water near the shoreline was incorrectly classified as land. However, Fig. 8.6a, c also show that the extracted coastline fits the real shoreline fairly well. These results suggest that

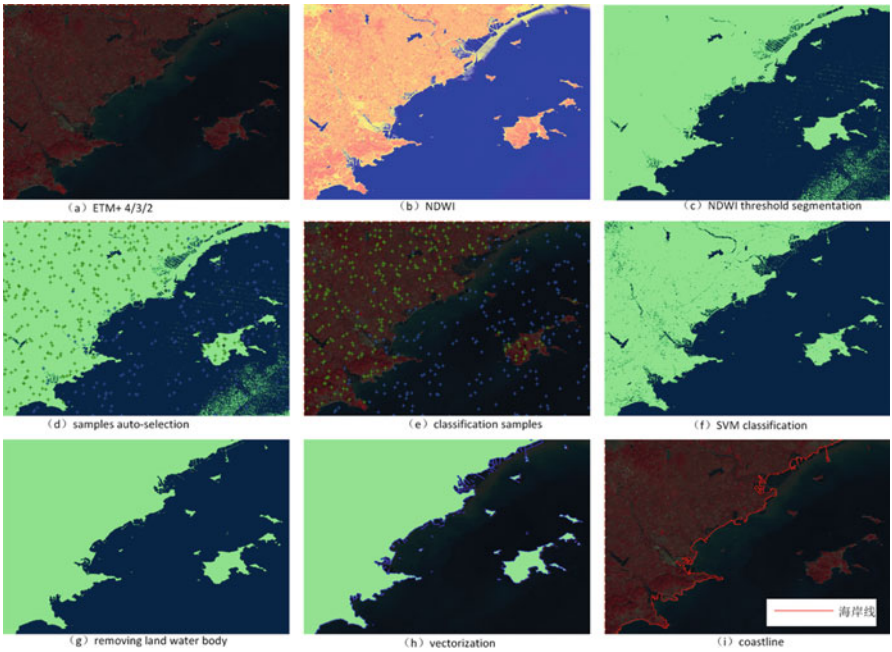


Fig. 8.5 The process of coastline extraction based on sample auto-selection and the SVM model

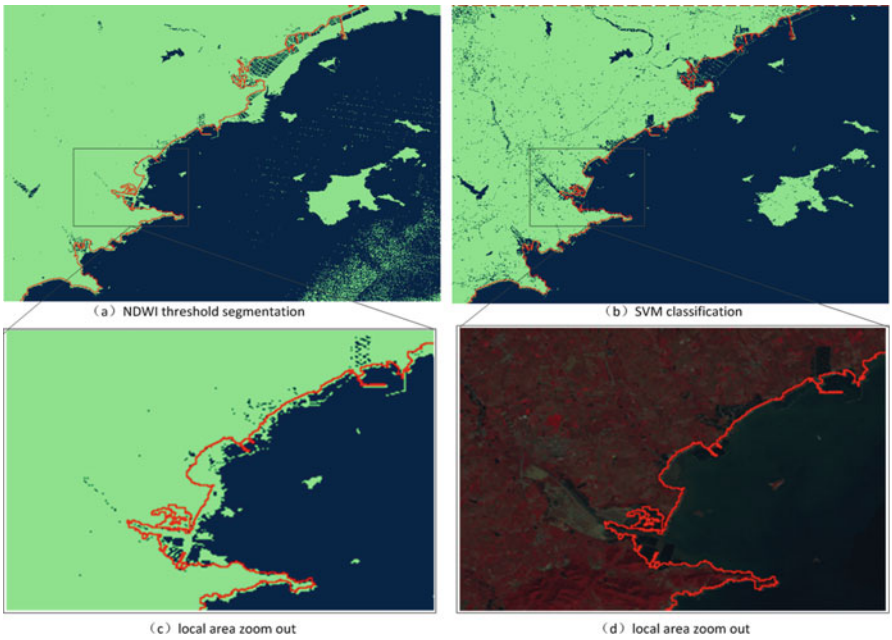


Fig. 8.6 Comparison of accuracy between the SVM model and the NDWI threshold segmentation

the SVM classification method significantlyimproved the accuracy of coastline retrieval.

8.1.4 Summary

The procedure of coastlines accurate extraction from remote sensing images is quite complex. The spectral characteristics of near shore water make it susceptible to misclassification; for example, red band shifts induced by river sand and augmentation of NIR bands due to red tides. Otherwise, the NDWI uses the two bands (Green and NIR) of the multi-spectral image. The variation in the water spectrum results in distortion of the NDWI near the shoreline areas. It is difficult to find an optimal segmentation threshold. Thus, the unique threshold segmentation method will incorrectly classify some near shoreline water as land.

To improve the accuracy of coastline delineation, sample auto-selection and the SVM classification method were used for coastline extraction on the basis of the NDWI. This method integrates the advantages of the NDWI threshold and supervised classification, and the NDWI threshold segmentation solved the problem of supervised classification in sample auto-selection. We conclude that (1) NDWI threshold segmentation can achieve rapid water body extraction. However, the bands used by NDWI are limited, and the extraction accuracy is affected by spectral diversification. Through supervised classification by all bands, the extraction accuracy can be significantly improved. (2) Sample auto-selection avoids human interference during classification samples selection, which improves computational efficiency and the degree of automation.

8.2 Intertidal Zone Identification by Remote Sensing Image

8.2.1 Introduction

The intertidal zone is a arrow fringe area of the land between the land and sea, where is above water at low tide and under water at high tide. So it is exposed to air or covered by water alternately (Feder and Bryson-Schwafel 1988; Vernberg and Vernberg 2012). It lacks a universal definition or classification system, because of the heterogeneity of structure and biotic characteristics in intertidal zone. Many studies have classified the intertidal zone using remote sensing techniques; however, they have been largely unsuccessful because affected by tidal level, the zone is only exposed at low tide, which may not overlap with the time when satellites overpass (Murray et al. 2012; Ryu et al. 2008; Zhao et al. 2008). Thus, it is not easy to accurately identify the intertidal zone with remote sensing techniques.

Because the location of the high tide line is difficult to acquire. Even though many remote sensing projects have examined the coastal zone, there is limited information concerning the global or regional distribution of tidal flats. There are rare literature reported use Medium resolution image. Some specific applications of intertidal zone mapping used aircraft-borne LIDAR or SAR (Beijma et al.; Donoghue et al. 1994; Gade et al. 2008; Green et al. 1996). However, these methods are limited by the high cost of airborne data acquisition. What more, generally there are no history archived data can be used for dynamic study.

Therefore, to dig out history information from the archive moderate resolution image is quite important. In this chapter, based on previous studies, we explored a simple and practical method for intertidal zone identification by NDWI histogram characteristics analysis, which is completely based on optical remote sensing. In local windows, the NDWI histogram has a trimodal distribution, in which different peaks represent various landscapes (land、intertidal zone and sea). According to the histogram characteristic, we can achieve intertidal zone identification through NDWI threshold segmentation. A key component of this method involves finding the optimal segmentation thresholds.

8.2.2 Methodology

On NDWI image, there is significant difference between areas with water and no water. The NDWI frequency distribution shows a trimodal distribution, with the background land on the left peak, the intertidal zone in the middle peak and the water on the right peak (Fig. 8.7). The three peaks oftheNDWI histogram represent pixels most closely fitting land (soil, vegetation, etc.), water and the intertidal zone, respectively. Based on the histogram shape feature, NDWI threshold segmentation can be used for extracting intertidal zone in most cases.

However, threshold is not a fixed value and different areas need different segmentation thresholds. In order to deal with this problem, we used dynamic thresholds in the NDWI for intertidal zone automatic extraction by spatial adaptive iteration model (Li et al. 2011; Li and Sheng 2012; Luo et al. 2009). The procedure includes 4 steps: (1) Window operation. We open a window of approximately 200*200 pixels, where the center of the window is the shoreline. Next, we calculate the histogram of the window to find the first and the second trough value, and then take those values as initial thresholds (T_1, T_2). (2) Adjusting the segmentation threshold value. We use step-by-step adaptive iteration ($Max\_T_b = T_1 + 0.002$, $Min\_T_b = T_2 - 0.002$) to find the optimal threshold value (Max_T_b, Min_T_b) at which the object information became relatively stable. (3) Stability judgment. The criterion of stability is that less than 20 pixels increase after threshold adjustment. This prompts the procedure to break its iterative circle and move to the next window. Repeating cycle of the above three procedures until the entire image has been evaluated. (4) Output and post-processing. The results are output and the intertidal zones are evaluated in the post-processing based on expert knowledge.

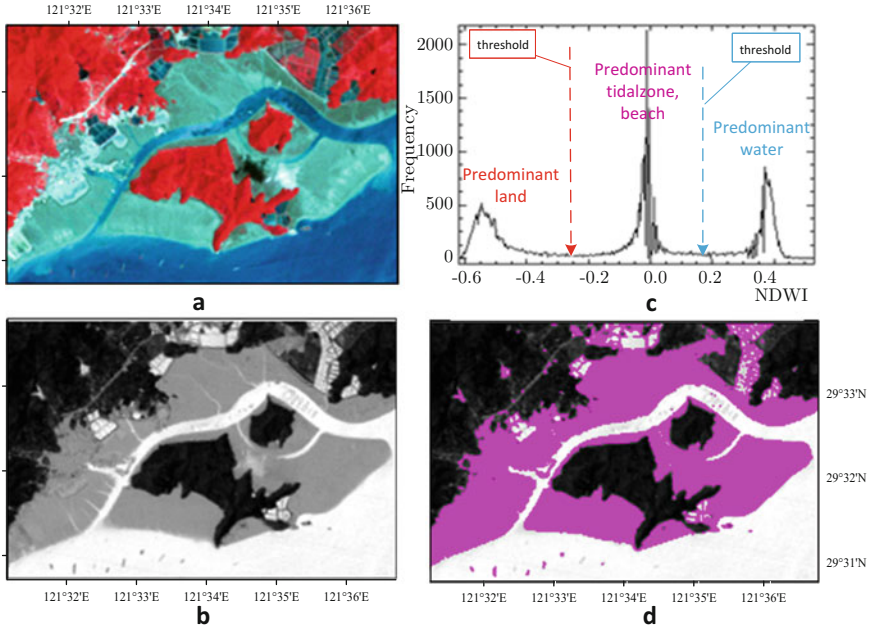


Fig. 8.7 Flow chart for intertidal zone extraction based on NDWI (Note: (**a**) ETM 4/3/2 band composite image; (**b**) NDWI; (**c**) histogram of coastal zone area; (**d**) intertidal zone area)

1. **Window operation and initial segmentation**

Based on the NDWI and coastline spatial distribution information, we open a window along the coastline. The width of the window is approximately 200*200 pixels and the window's central pixel is located on the coastline. The NDWI histogram is computed within the window. In the histogram, seawater, the intertidal zone and the land can ideally be described with three peaks. Figure 8.8 shows the three peaks of the histogram waveform, from left to right, that represent the pixels most closely fitting land (soil, vegetation, etc.), the intertidal zone and water. During initial segmentation, the first and second trough (T_1, T_2) of the histogram are identified and taken as the initial two thresholds. The mean values for water, the intertidal zone and land, respectively, are μ_{Water}, $\mu_{intertidal}$ and μ_{Land}, and σ_{Water} and σ_{Land} are their corresponding standard deviations.

2. **Step-by-step adaptive iteration**

In this study, we apply a local adaptive iteration algorithm based on the step-by-step iterative transformation that was first proposed by Luo in 2009 (Bai et al. 2011; Li et al. 2011; Li and Sheng 2012; Luo et al. 2009). The procedure includes the following steps:

- **Step 1:** Statistical local histogram. A local NDWI histogram was computed for each operational window zone.

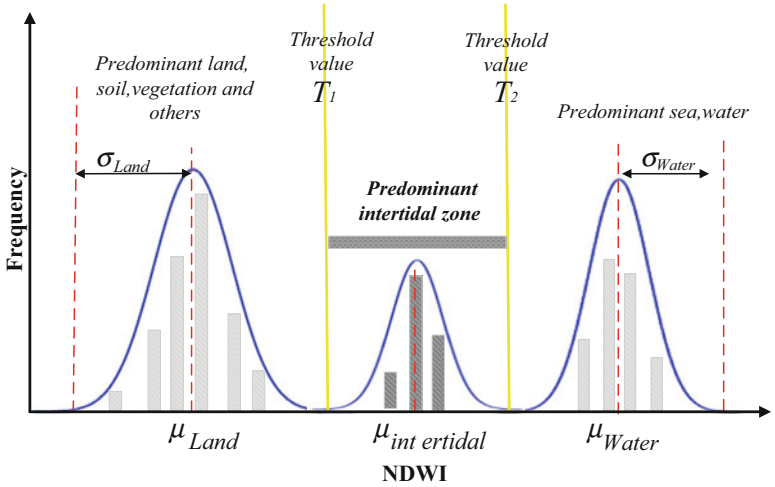


Fig. 8.8 Sketch of the NDWI histogram's trimodal distribution

- **Step 2:** Identification of the initial threshold. Based on minimum values, the first and second trough values (T_1, T_2) of the local window's NDWI histogram are found. T_1 and$T_2$are treated as the initial threshold segmentation values.
- **Step 3:** Step-by-step adaptive iteration. The segmentation thresholds (Max_T, Min_T_b) were auto-adjusted until the object informationbecame stable. The algorithm judgment criterion compares the new result with the previous one. Suppose that the previous time, the pixel area was N_0, and the new pixel area is N. If $N > N_0$, then $N = N_0$, $Max\_T_b = T_1 + 0.002$, and $Min\_T_b = T_2 - 0.002$; otherwise, the procedure will break its iterative loop and move to the next window.

8.2.3 Intertidal Zone Mapping

Using the above method, we extracted the intertidal zone at the mouth of the Yellow River in Shandong province, China, using Landsat time-series data, and we analyzed its dynamic changes. The intertidal zone distribution maps are shown in Fig. 8.9. Figure 8.9a shows the intertidal zone distribution in 1973. Figure 8.9b shows the intertidal zone distribution in 1984. Figure 8.9c shows the intertidal zone distribution in 2000. Figure 8.9d shows the intertidal zone distribution in 2010. Comparing Fig. 8.9a with d, it demonstrates that the intertidal zone has undergone significant changes in this region.

The intertidal zone area in 1973, 1984, 2000, and 2010 ware counted separately. In 1973, the total area was 732.9 km^2; in 1984, the area was 710.64 km^2; in 2000, the area dropped to 610.64 km^2; and in 2010, the area was 423.8 km^2. The statistical results show that the intertidal zone area in the Yellow River estuary decreased

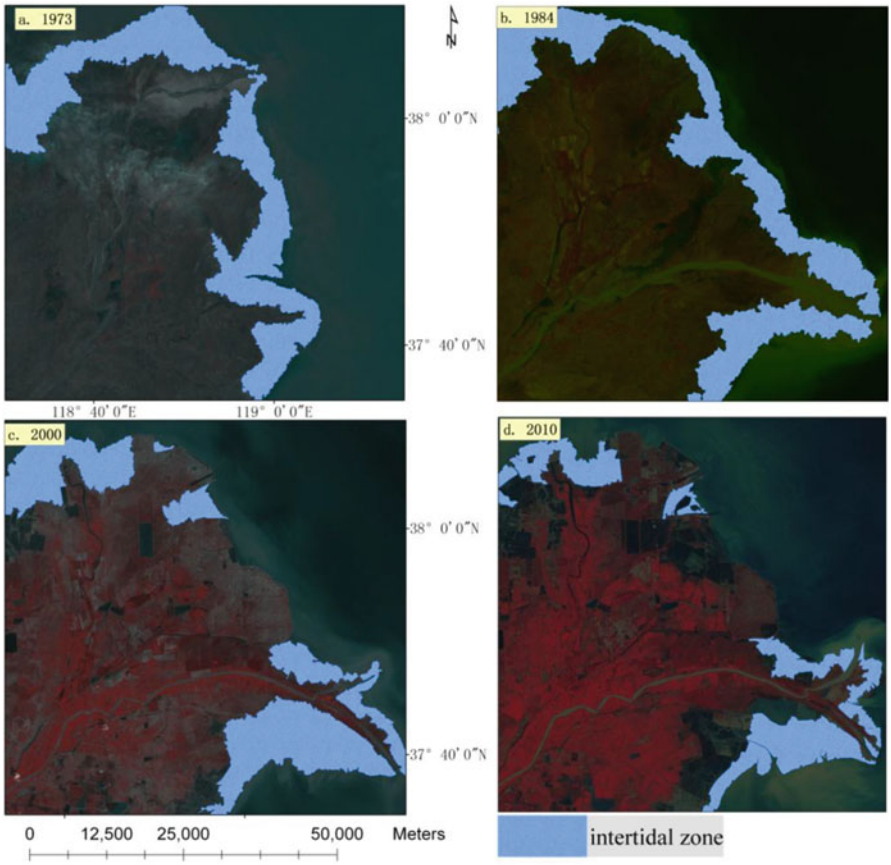


Fig. 8.9 Maps of the intertidal zone distribution in 1973, 1984, 2000, and 2010

significantly after 1973. Figure 8.10 shows the changes in the intertidal zone area from 1973 to 2010; the red line in the figure is the best fit curve, with the fitting function $y = -41.125*x^2 + 102.92*x + 670.66$.

8.2.4 Conclusions

This section characterizes the intertidal zone completely based on optical remote sensing data. Research indicates that the NDWI histogram shape features can provide useful information for the intertidal zone extraction. This method addressed the optimal threshold selection by step-by-step iteration model. Through a dynamic segmentation threshold rather than an arbitrary value to extract the intertidal zone adaptively. Compared with traditional methods, this method is simple and practical.

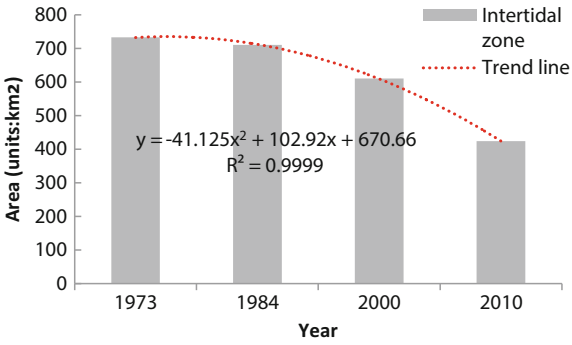


Fig. 8.10 Changes in the intertidal zone area

It constitutes a promising method for intertidal zone identification from multispectral images without using auxiliary data such as DEM, topographic maps or field survey data. However, the accuracy of this method has yet to be shown by further experimentation. Nevertheless, this method is a useful alternative approach at the rarity of data source. It provides a feasible way to rapidly retrieve intertidal zone information and has many domains of application.

8.3 Coastal Wetland Classification Using Remote Sensing

8.3.1 Introduction

Coastal wetlands are productive ecological systems that serve as buffer zones between the sea and land. They play an important role as carbon sinks, as well as absorbing nitrogen, regulating geochemical cycles, conserving water and causing sedimentation (Boesch et al. 1994). Meanwhile, coastal wetlands are also important areas for breeding and habitat for fish and other invertebrates (Sklar et al. 1985). However, with increases in population, industry development and overexploitation of wetland resources in recent decades, coastal wetland systems are now considered one of the most vulnerable ecosystems in the world (Boesch et al. 1994; Cahoon et al. 2006; Gedan et al. 2011; Lee et al. 2006; Li et al. 2007; Nicholls 2004; Simenstad et al. 2006). It is imperative to quantify the status and trends of regional wetlands resource for management and conservation.

Traditional wetland surveys in the field are time-consuming and difficult. The development of remote sensing technology and the availability of abundant multisource satellite data have provided ways to obtain wetland information quickly and accurately. Numerous researchers have attempted to extract wetlands using classification techniques for various remote sensing datasets. Augusteijn and Warrender (1998) established a method for automatic wetland classification using a neural network algorithm based on optical and radar data. Islam et al. (2008) used Landsat ETM+ and SRTM data to map wetlands by semi-automated methods. Bwangoy

et al. (2010) integrated optical and radar remote sensing data and topographical indices to delineate wetlands in the Congo basin.Davranche et al. (2010) used classification trees and SPOT-5 seasonal time series to monitor wetlands. Zhang et al. (2011) classified coastal wetland vegetation with a Landsat Thematic Mapper image.

Using Object-Oriented (per-parcel) image classification technology, some researchers have extracted wetland information (Dronova et al. 2011; Frohn et al. 2011; zhu et al. 2014). These studies have been moderately successful in wetland extraction, but automation remains a challenge. Wetland ecosystems are complex, and the shape, size and form of wetlands vary. In addition, wetlands do not have unique spectral characteristics (Schmidt and Skidmore 2003). Thus, the level of automatic classification is not high when remote sensing is used for wetland mapping (Mwita et al. 2012). Therefore, we propose a new hybrid approach for the automatic extraction of coastal wetlands based on "Tu-pu" (spatial and spectral features) coupled information.

8.3.2 Coastal Wetland Classification Schemes

The definitions of wetlands vary widely, and various existing classification systems for wetlands are on a regional scale. Several classification schemes have been developed to describe coastal wetlands. Before coastal wetlands remote sensing survey and mapping, it is necessary to unify the classification schemes. Considering the divisibility of spectral features and the wetland classification scheme of the Ramsar Convention (Convention on Wetlands of Importance Especially as Waterfowl Habitat), we propose a coastal wetland classification system based on remote sensing (Table 8.2).

8.3.3 "Tu-Pu" Coupled Coastal Wetland Hierarchical Classification

The distribution of coastal wetlands depends on water and has spatial relationships with natural features such as coastlines, lakes, and rivers. We analyze the spectral and spatial characteristics of wetlands and propose a "Tu-Pu" Coupled Coastal Wetland Hierarchical Classification approach for coastal wetland extraction (Fig. 8.11).

The entire procedure can be described as follows. First, we extract estuaries, coastlines, lakes and other features based on their spectral characteristics. Second, we retrieve ground moisture with the NDMI algorithm. Third, using a global spectral clustering algorithm, we convert the image pixels to objects. We classify the potential coastal wetland distribution based on the ground humidity index.

Table 8.2 Wetland classification system based on remote sensing

Code1	Type	Code2	Subtype
1	Coastal wetlands	11	Rivers
		12	Lakes and lagoons
		13	Marshes and swamps
		14	Cultivated areas
		15	Salt pans
		16	Intertidal zones
		17	Shallow seas

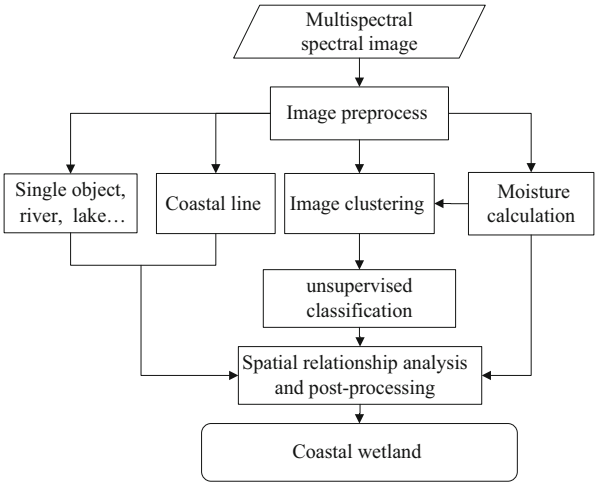


Fig. 8.11 Flow chart for coastal wetland extraction

Finally, through analysis of the spatial dependence relationship with features such as coastlines and lakes, we revise the classification result for coastal wetlands.

1. **Image clustering and wetland classification**

Amongclustering algorithms, the Iterative Self-Organizing Data Analysis Technology Algorithm (ISODATA) is one of the most popular. It has been widely used in unsupervised classification of remote sensing images. For example, unsupervised clustering is often used to obtain land use and cover change (LUCC) maps. This approach is useful when relatively little priori information for the data is available. In view of this, we use an ISODATA-based unsupervised clustering method for image clustering and wetland classification. Obtained through soil moisture retrieval, the ground soil moisture image is an important input parameter for clustering and classification. The property value directly determines the possibility that a pixel represents a wetland area.

2. **Spatial relationship analysis and result optimization**

Spatial distance and the spatial dependence relationship were used to adjust categories and optimize the classification accuracy. The procedure included

reference object extraction, target object pixel searching and reference adjacent object identification. The detailed steps are:

- (1) Through extraction of water body and shoreline surface features, the reference object (target object) was determined for wetland extraction.
- (2) The lake and river classification was derived from knowledge of the spatial patterns in the water layers. In other layers, we recognized swamp areas according to soil moisture and spatial-relationship analysis.
- (3) We selected uncertain target objects in a range (buffer). The vegetation coverage, soil moisture and other information determined whether the object belonged to the wetland category.

To aided wetlands judgment, some empirical spectral-index models were used as input parameters:

Normalized Difference Snow/Ice Index (NDSII) Xiao et al (2001):

$$\mathrm{NDSII} = \frac{\rho_{\mathrm{Green}} - \rho_{\mathrm{infraRed}}}{\rho_{\mathrm{Green}} + \rho_{\mathrm{infraRed}}} \tag{8.4}$$

Normalized Difference Moisture Index (NDMI) Wang et al 2012:

$$\mathrm{NDMI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{infraRed}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{infraRed}}} \tag{8.5}$$

Normalized Difference Impervious Surface Index (NDISI) Xu H. 2008:

$$\mathrm{NDISI} = \frac{\rho_{\mathrm{TIR}} - \left(\rho_{\mathrm{MIR}} + \rho_{\mathrm{NIR+MNDWI}}\right)/3}{\rho_{\mathrm{TIR}} + \left(\rho_{\mathrm{MIR}} + \rho_{\mathrm{NIR+MNDWI}}\right)/3} \tag{8.6}$$

$$\mathrm{MNDWI} = \frac{\rho_{\mathrm{Green}} - \rho_{\mathrm{MIR}}}{\rho_{\mathrm{Green}} + \rho_{\mathrm{MIR}}} \tag{8.7}$$

where ρ_{Green} is the green-band reflectance, $\rho_{\mathrm{NIR}}$is the near-infrared-band reflectance, ρ_{infraRed} is the infrared-band reflectance, ρ_{TIR} is the thermal-infrared-band reflectance, and $\rho_{\mathrm{MIR}}$is the mid-infrared-band reflectance. For the ETM+ imagery,ρ_{Green} corresponds to band 2,ρ_{NIR} to band 4, ρ_{infraRed} to band 5,ρ_{MIR} to band 7, and ρ_{TIR} to band 6. The original data storage was an 8-bit digital number (DN) image. All imagery should convert the DN to reflectance (Chander et al. 2009; McFeeters 1996). A Lambert Azimuthal Equal Area Projection was adopted as the distribution area statistic at different times.

8.3.4 Experiment and Validation

The test data used in the experiments is an ETM+ image, which covers part of a coastal city in eastern China (Fig. 8.12a). The classification result is shown in Fig. 8.12b). To check the validity of the proposed method, the accuracy of wetland maps was assessed by comparison with reference data (the result from visual interpretation). We also compare the proposed method and the maximum likelihood classification. The quantitative accuracy evaluation results are listed in Table 8.3.

Table 8.3 shows that maximum likelihood classification, which is the most basic and widely used method, can extract relatively pure water and other major categories. However, it misclassifies many areas, and its overall accuracy is only 89.2 %. Comparatively, the TUPU hybrid method proposed in this study, which is based on spectral and spatial characteristics, has a total accuracy of 92.6 %. In addition, the former method mainly concentrates on spectral information, and does not sufficiently use spatial information. Our proposed method considers both spectral and spatial information in classification. The proposed method achieves a satisfactory classification result, which verifies its feasibility and validity.

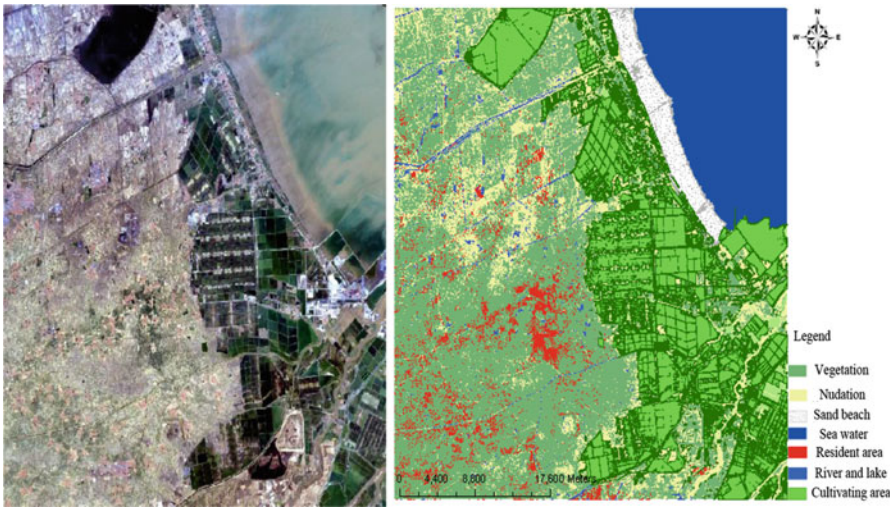


Fig. 8.12 Coastal wetland classification map

Table 8.3 Accuracies for wetland classification

Classes / Method	Marshes and swamps	Rivers, lakes and lagoons	Intertidal zones	Shallow seas	Cultivated areas	Total accuracy
TUPU hybrid	95.6 %	88.9 %	96.0 %	95.6 %	90.4 %	92.6 %
Maximum likelihood	94.2 %	88.1 %	87.3 %	91.2 %	85.5 %	89.2 %

8.3.5 *Regional Application: Analysis of Dynamic Changes in the Yellow River Estuary Wetland*

8.3.5.1 Background of the Study Area

The Yellow River estuary is located in Dongying city, in Shandong province. The study areas is an inlet of the modern Yellow River, at a longitude from 118.30 to 119.30 °E and a latitude of 37.05 to 38.20 °N (Fig. 8.13). North of the Yellow River estuary is Bohai and to the east is Laizhou Bay. Generally, the modern Yellow River estuary coastline is from Wu Hao Zhuang (in the north) to Song Chun Rong Gou (in the south) (Yang et al. 2001).

Because of the Yellow River sedimentation, the estuary is continuously expanding to the sea. So, here is one of the youngest land areas in the world. Wetland area accounted for 80 % of total area, and it is considered one of the six most beautiful wetlands in China. Most of the wetland is natural and concentrated in the eastern zone from the mouth of the Xiaodao River in the south to the mouth of the Tuhai River in the north. Artificial wetland areas, including paddy fields, reservoirs and salty fields are widely distributed in the central to western part of the estuary. The Yellow River estuary is an internationally important area for protection of new wetland areas and endangered birds. But, according to a wetland survey report from Shandong province, the wetland ecosystem in the this area has

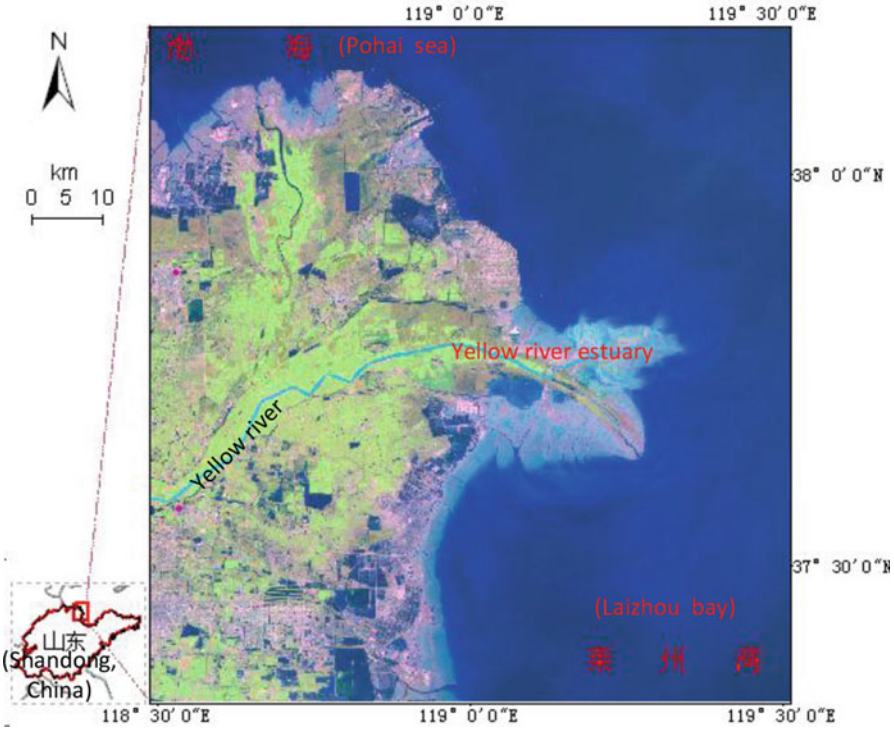


Fig. 8.13 Location of the study area

suffered tremendous degradation over the last 20 years, due to decreasing river discharge, increasing human activities and lack of scientifically based management.

8.3.5.2 Yellow River Estuary Wetland Mapping

Using the "Tu-Pu" Coupled Coastal Wetland Hierarchical Classification approach, we mapped the Yellow River estuary and analyzed time series change. The Yellow River estuary coastal wetland remote sensing classification maps are shown in Fig. 8.14. Figure 8.14a shows the Yellow River wetland distribution map in

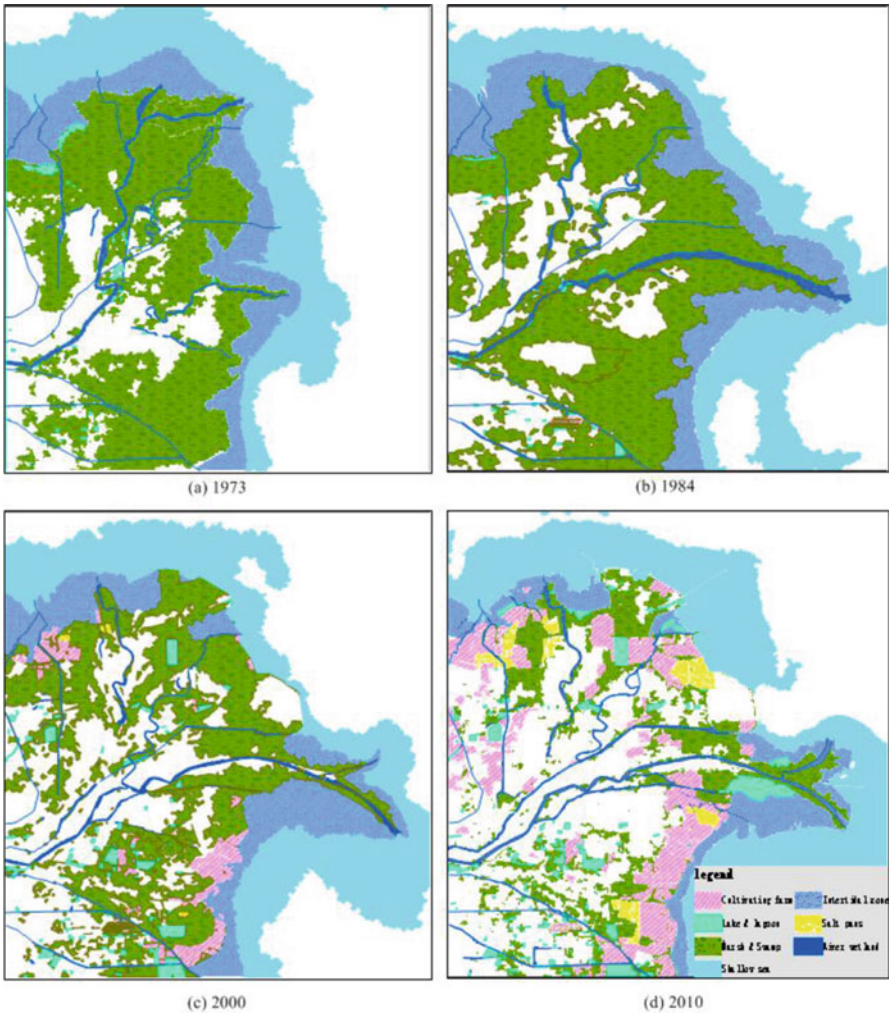


Fig. 8.14 Yellow River estuary coastal wetland dynamics from 1973 to 2010 (Note: (**a**) wetland distribution map for 1973, (**b**) wetland distribution map for 1984, (**c**) wetland distribution map for 2000, (**d**) wetland distribution map for 2010)

1973; Fig. 8.14b shows the wetland distribution map in 1984; Fig. 8.14c shows the wetland distribution map in 2000; and Fig. 8.14d shows the wetland distribution map in 2010. The wetland distribution maps demonstrate that the coastal wetlands at the mouth of the Yellow River have undergone significant change in the past decades.

Figure 8.15 shows an area histogram of various types of wetlands; it is apparent that swamp wetlands and intertidal zones are the dominant wetland types in this region. The different types of wetlands show different dynamics. Swamp wetlands and intertidal zones decreased sharply from 1973 to 2010. In contrast, river wetlands were relatively stable, lake areas increased slightly, and cultivated areas and salt pans increased significantly.

8.3.5.3 Change Analysis of Yellow River Estuary Wetlands

Detailed statistics of the changes in wetland area in the Yellow River estuary are shown in Table 8.4. Table 8.4 indicates that the total wetland area in this region decreased dramatically over the past 40 years. From 1973 to 2010, the total wetland area decreased by approximately 570 km^2. However, the changes in various wetland types are not the same. The most obvious change is in swamps and artificial

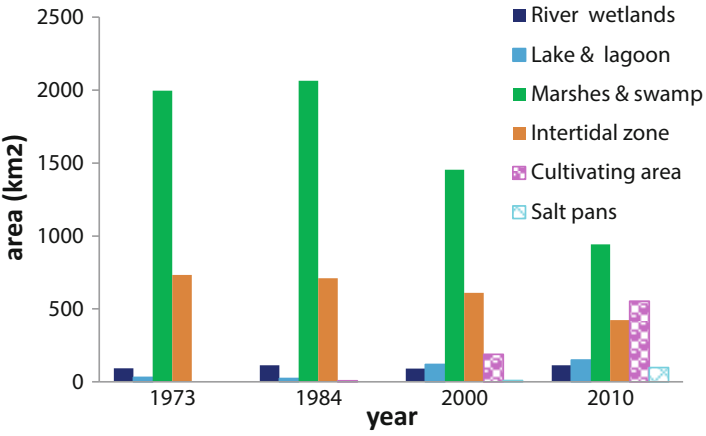


Fig. 8.15 Accumulation histogram of area change of all wetland types

Table 8.4 Area statistics for wetlands in the Yellow River estuary at different periods (units:km^2)

	Marshes and swamps	River wetlands	Lakes and lagoons	Cultivating area	Salt pans	Intertidal zones	Total
1973	1996.44	92.18	30.23	0.00	0.00	732.90	2851.75
1984	2064.86	113.22	23.38	8.13	0.00	710.64	2920.23
2000	1453.79	91.12	119.86	188.20	9.68	610.64	2473.29
2010	942.42	112.68	150.18	552.67	98.77	423.88	2280.60

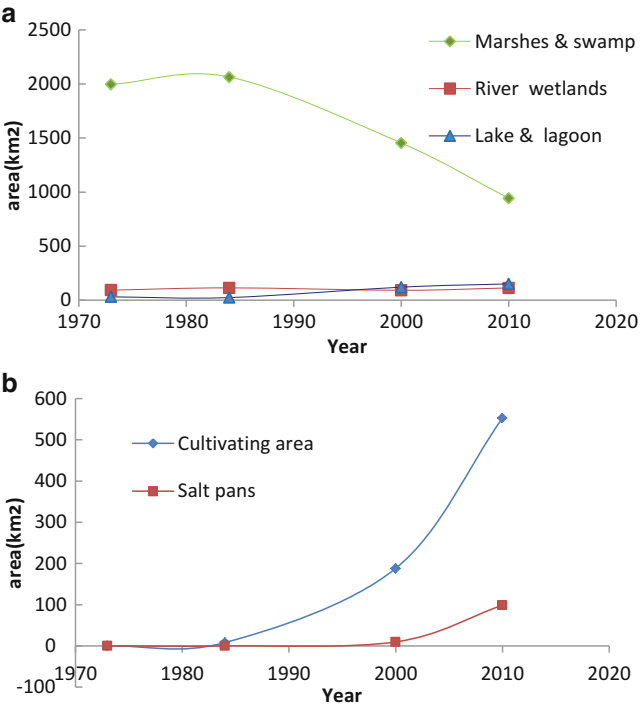


Fig. 8.16 Changes in wetland area (Note: (**a**) natural wetland area change; (**b**) artificial wetland area change)

wetlands (salt pans and cultivated areas). Swamp area decreased from 1996.44 km^2 in 1973 to 942.42 km^2 in 2010. Cultivated areas increased from 8.13 km^2 in 1984 to 522.67 km^2 in 2010. Salt pans area increased from 9.68 km^2 to 98.77 km^2 since 2000.

Figure 8.16 graphs the changes in wetland area over the past four decades. Figure 8.16a shows the change in the natural wetland area; Fig. 8.16b shows the change in artificial wetland area. The two types of wetlands show different variations. Artificial wetlands display similar changes, increasing exponentially. However, Swamp wetland area declines rapidly.

8.4 Coastal Invasive Species Detection with Remote Sensing: A Case Study of *Spartina alterniflora* in Xiangshan Bay, China

8.4.1 Introduction

Spartina alterniflora is a major exotic plant in China and originates from the Atlantic coast of North America (Daehler and Strong 1996; Zhu et al. 2016). It

usually grows in intertidal zones and estuaries between the upper level of the middle tide and the low level of high tide (Vinther et al. 2001). Because it plays an important role in protecting the beach erosion, accelerating sidimentation, and inceasing the land area, S. alterniflora has been introduced to many countries and regions including the west coast of North America, Europe, New Zealand and China (Ayers and Strong 2002). In China, it was introduced in 1979 by coastal line ecological protection engineering (Chung 2006; Xu and Zhou 1985). Several years later, it had been planted on intertidal zones from Dianbai County,the southern of Guangdong province, to Ye County, the northern of Shandong province, and multiplied rapidly (Chung 2006; Xu and Zhou 1985).

Because of its high reproduction rate, the invasive S. alterniflora soon occupied a bare ecological niche as a pioneer weed species (Zhang et al. 2004). At present, it is widespread from the southern part of Guangdong province to the northern part of Liaoning province (Li et al. 2005). For example, in the mouth of Minjiang mouth, the area of S. alterniflora was less than 667 m^2 in 2004, but exceeded 200 hm^2 in 2007. In Hangzhou Bay, S. alterniflora increased from 600 hm^2 to 1500 hm^2 in just 1 year, from 2002 to 2003 (Deng et al. 2006; Li et al. 2005). The growing space of native plants has been severely encroached, with adverse effects on birds foraging and fish survival in coastal zones (Liu et al. 2007). The Chinese government lists *S. alterniflora* among the first batch of 16 invasive species in China identified in 2003 (Zhang 2010).

Numerous researchers have attempted to use remote sensing technology to detect S. alterniflora and other invasive plants (Dennison and Roberts 2003; Rosso et al. 2006; Wan et al. 2014; Zhang et al. 1997; Zuo et al. 2012). However, the existed studies have mainly focused on high resolution satellite imagery or Unmanned Aerial Vehicle (UAV) images (Wan et al. 2014; Zhang 2010). Few studies have used medium resolution images, although we note that high spatial resolution images are able to detect S. alterniflora better. However, the cost of high spatial resolution imagery is expensive. Generally, there is rare historically archived data available. So, compeletely based on the high resolution imagery is not suitable for long-term or dynamic observation in large scale. It can be appropriate to use historical data to determine past distributions. Therefore, it is important to used archived moderate resolution imagery to determine history distribution information. This study integrates spatial and spectral information and proposes a hybrid method for automatic extraction of the S. alterniflora distribution in Xiangshan Bay based on medium resolution remote sensing imagery, as well as its long-term detection (Zhu et al. 2016).

8.4.2 Study Area

Xiangshan Bay is located in Zhejiang province in southeast China, at a longitude of 121° 20′−112° 16′ and a latitude of 29° 10′−30° 00′, as shown in Fig. 8.17. It is a northeast to southwest-trending inland strip-type semi-closed bay. There is a large amount of S. alterniflora in the area, and high intensity of land exploitation. The

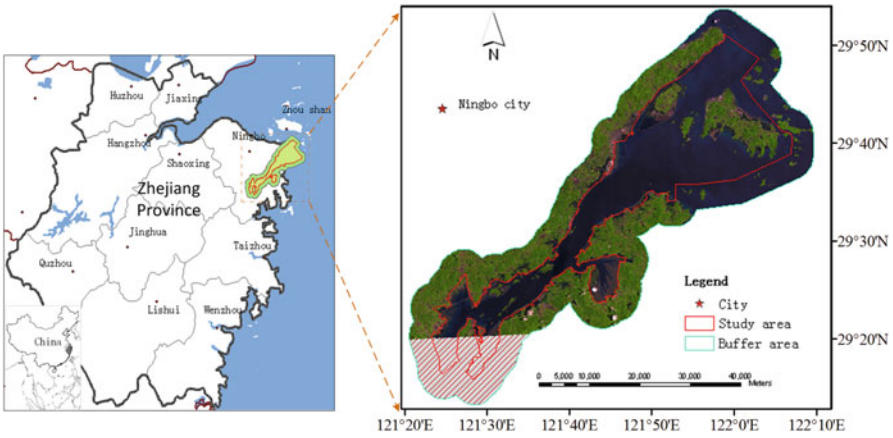


Fig. 8.17 Location of the study area

area has experienced high strength of human activities, including intertidal flat reclamation and marine resource development. S. alterniflora was first introduced and planted in Yumen County of Zhejiang province in 1983 (Zhang 2010), after which it rapidly spread to other beaches. At present, the major S. alterniflora distribution area is in the Zhejiang coastal zone.

8.4.3 Methodology

The patches of S. alterniflora are weak signal throughout the images. It is hard to extract S. alterniflora distribution information from the image solely based on spectral features. From previous studies and the characteristics of S. alterniflora habitat (its ecological niche), S. alterniflora is mainly distributed in the intertidal zone and estuary (Ayers and Strong 2002; Daehler and Strong 1996; Vinther et al. 2001; Yuan et al. 2009). The intertidal zone is a potential habitat area for S. alterniflora, where is the possibility geographic growth location of S. alterniflora. This geographic distribution (spatial information) provides an effective probability objective area, which can be used to assist S. alterniflora detection accurately.

The hybrid method consists of two phases: (1) delineation of the intertidal zone as the potential distribution area of S. alterniflora and (2) extraction of S. alterniflora distribution with a mixed pixel analysis. Details are illustrated in Fig. 8.18 (Zhu et al. 2016). During the first phase, we extract the coastline using threshold segmentation methods. Next, the intertidal zone is extracted using the "Moving Window Operation" NDWI histogram segmentation.The details were illustrated in section 2. For the mixed pixel decomposition, firsly, we obtain the end members in the intertidal zone and then use linear spectral mixture analysis (LSMA) for S. alterniflora extraction from the intertidal zone. Finally, we use the threshold segmentation technique to obtain distribution information from the fraction unmixing image.

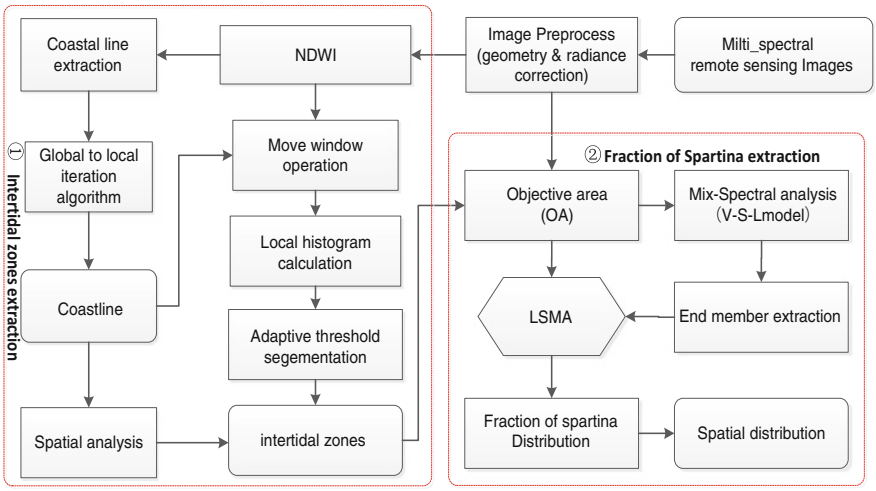


Fig. 8.18 Flow chart of *S. alterniflora* information extraction from remote sensing images

1. **Mixed spectral analysis and end member extraction**

On moderate resolution images, the phenomenon of mixed pixels is a widespread phenomenon that is an barrier in improving image classification accuracy and effective application of remotely sensed data (Cracknell 1998; Fisher 1997). Traditional per-pixel (hard) classifiers, such as the supervised and unsupervised classifier, the mixed-pixel problem was no take into consideration. To address the mixed pixel phenomenon, Many research has done a lot of researches and experiments. Ridd (1995) proposed a vegetation-impervious surface soil (V–I–S) concept model (Ridd 1995). It assumes that the spectral signature of land cover is a linear combination of three major components of vegetation, impervious surface, and soil. Using LSMA technique, the mixed pexel can be decomposed into different parts, and the fraction of each component represents the land scape area proportion (Adams et al. 1995; Aguiar et al. 1999; Cochrane and Souza 1998; DeFries et al. 2000; Smith et al. 1990). The concept of the V–I–S model and the LSMA algorithm provides a guideline for unmixing of urban landscapes and a link of these components to spectral signatures (Lu and Weng 2006).

Analyzing the *S. alterniflora* distribution and the characteristics of the study area showed that in the intertidal zones, land cover can be divided into three classes: water, bare land and vegetation (*S. alterniflora* and others). Thus, in the intertidal zone, the mixed pixel spectrum is composed of vegetation, bare land and dark objects. However, in Xiangshan Bay, the primary vegetation is *S. alterniflora*. Because of the extreme habitat conditions, there are little other plants in the intertidal zone (Zhang et al. 2004). To effectively identify high-quality image end members, minimum noise fraction (MNF) techniques are often used to transform multispectral images into a new dataset (Boardman and Kruse 1994; Green

et al. 1988). End members are then selected from the feature spaces of the transformed images (Cochrane and Souza 1998; GarciaHaro et al. 1996; Van der Meer and de Jong 2000).

The MNF rotation procedure was applied to convert the six reflective bands into a new coordinate dataset. After the MNF transform, the front component accounted for the main information and the first three components included more than 99 % of the information. The 2D scatterplots of MNF components between 1 and 2 are illustrated in Fig. 8.19. Analysis the link of the scanterplots and landscapes. Corner A represents bare land or sand with very high values in the MNF1 and MNF2 image; Corner B represents dark objects (shallow water and shadows) with very low values in the MNF1 and MNF2 image; Corner C represents vegetation (S. alterniflora) with high values in the MNF1 image and low values in MNF2 image. Based on 2D scatterplots feature space, three end members, S. alterniflora, high-albedo and low-albedo, were selected from the image by end member adaptive extraction methods (Zhu et al. 2011). We note that these may not be the true pixels (Lu et al. 2003; Theseira et al. 2003); however, it can represent the pure pixels in the local images (assumed image-based end members).

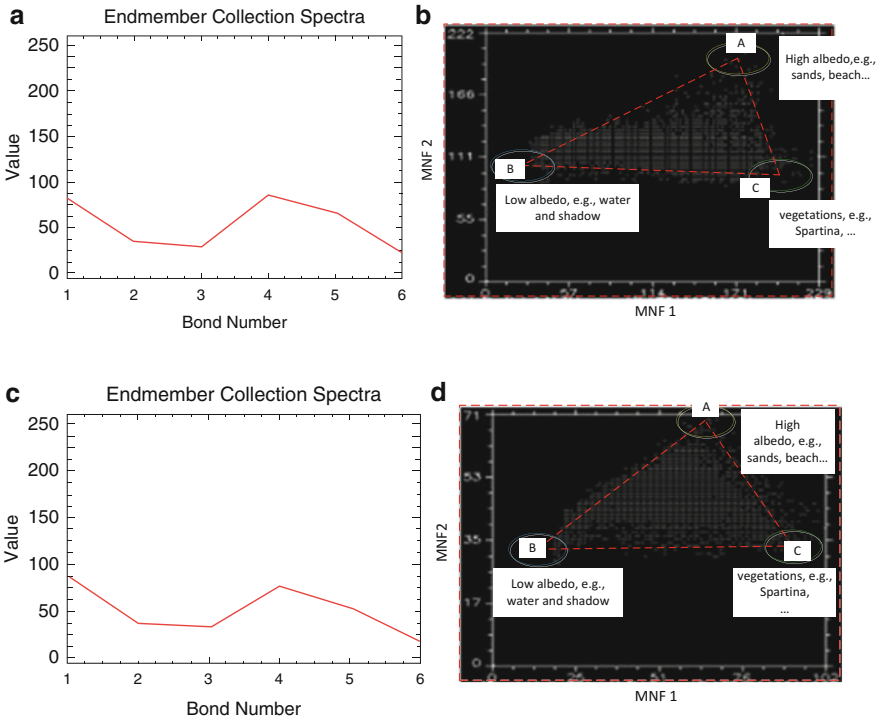


Fig. 8.19 End member spectrum extraction from the spectral feature space (**a**) End member of *S. alterniflora* from TM image, date: July 17, 2003. (**b**) End member of *S. alterniflora* from OLS image, date: June 13, 2014

2. **Fraction of *S. alterniflora* distribution estimated using LSMA**

The LSMA algorithm assumes that the spectrum measured by the sensor is a linear combination of the spectra of the end members. Within a pixel, the spectral proportions of the end members represent the ratio of the area covered by distinct landmarks on the ground. This algorithm has been regarded as a physically based image processing tool for accurate and quantitative extraction of sub-pixel information (Adams et al. 1995; Mustard and Sunshine 1999). Many studies have described this approach (Adams et al. 1995; Lu et al. 2003; Mustard and Sunshine 1999). In this study, a constrained least-squares algorithm was applied to decompose the 6 reflective bands into 3 fractional images and 1 error distribution image. The aim of the LSMA algorithm is to find suitable end members. Considering environmental effects and the scale conversion problem, image-based end members are better than pure spectra because they are measured at the same scale as the image data and in the same environmental conditions.

8.4.4 Results and Analysis

1. **Classification results**

Based on the image-based end members (S. alterniflora, high albedo and low albedo) a constrained least-squares solution was applied to decompose the six TM reflective bands into four fractional images (S. alterniflora, high albedo, low albedo and error). The fraction of S. alterniflora distribution is shown in Fig. 8.20. Figure 8.20a shows the fraction of S. alterniflora distribution in 2003, and Fig. 8.20b shows the fraction of S. alterniflora distribution in 2010. Comparing Fig. 8.20a with b demonstrates the spread of the invasive species S. alterniflora over the past 10 years in Xiangshan Bay. In 2003, S. alterniflora was sparsely distributed in the estuary and the intertidal zone. By 2014, both the area and the density of S. alterniflora in this region had increased significantly. Almost the entire muddy coastal beach had been occupied by S. alterniflora, indicating a significant threat to the regional environment.

2. **Spatiotemporal analysis**

Density slice methods were used to classify the fractional images, as shown in Fig. 8.21. Thus, a threshold of 0.6 was used for S. alterniflora extraction in this paper, meaning that where the probability exceeded 60 %, we considered a given area to be a S. alterniflora patch. The spatial expansion of S. alterniflora in Xiangshan Bay from 2003 to 2014 is apparent from the classification images (Fig. 8.21). The S. alterniflora distribution map (Fig.8.21a) indicates two characteristics. First, the distribution area has expanded significantly. In 2003, S. alterniflora only appeared in the freshwater area of the estuary. In 2009, S. alterniflora had expanded to the beach and some sea shoals. The second notable characteristic is the expansion of *S. alterniflora* towards the sea. The dynamic

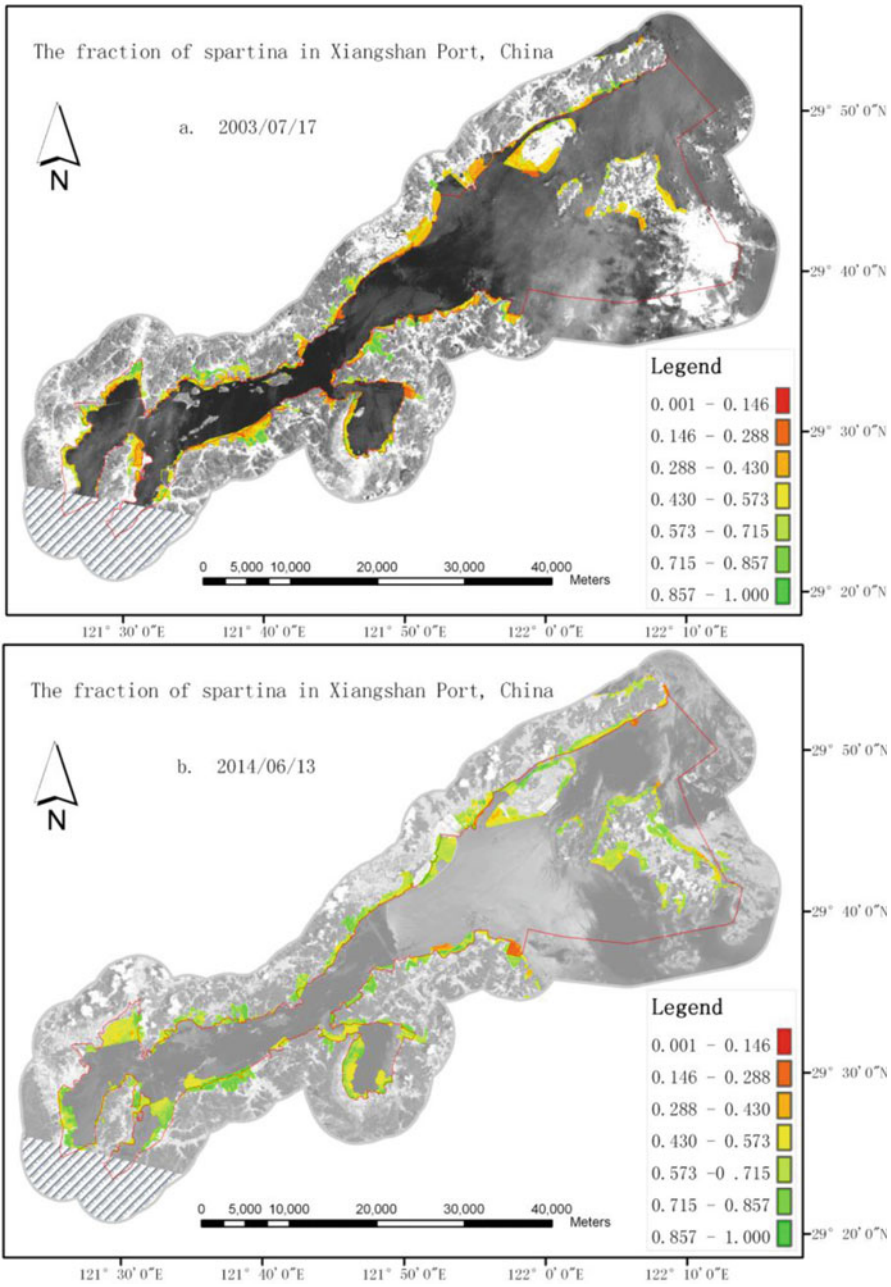


Fig. 8.20 Distribution maps of the fraction of *S. alterniflora* (**a**) fraction of *S. alterniflora* in 2003; (**b**) fraction of *S. alterniflora* in 2014 (The fraction values range from 0 to 1, with lowest values in *red* color and highest values in *green* on the fraction images)

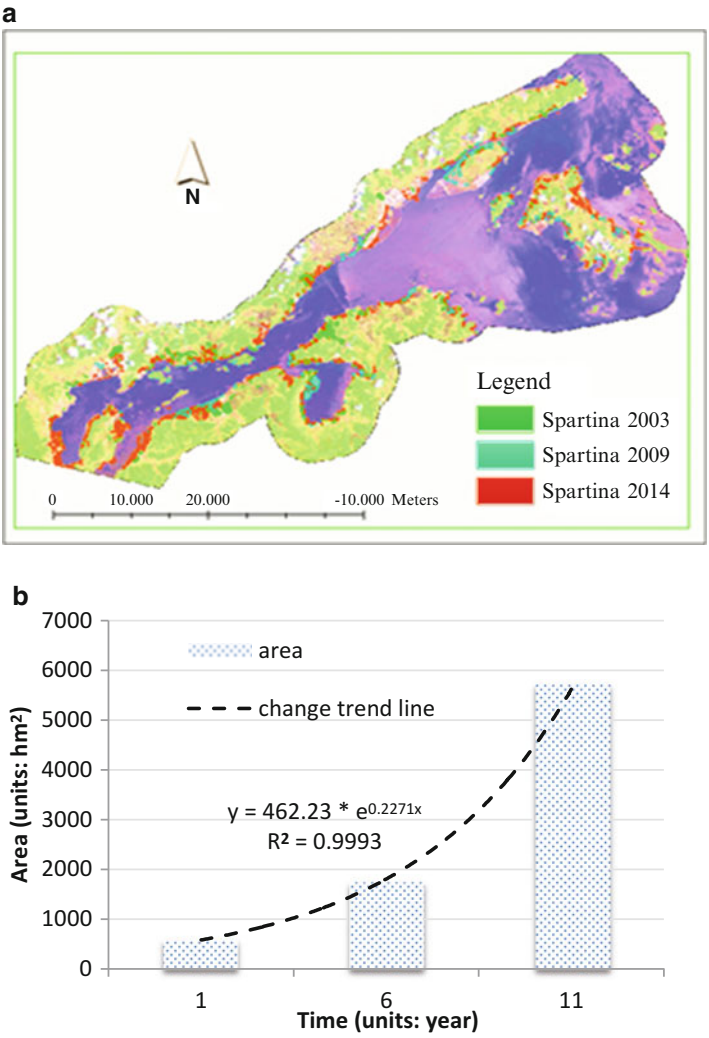


Fig. 8.21 Spatiotemporal change of *S. alterniflora* in Xiangshan bay (**a**) The Patches Distribution of *S. alterniflora* in the Xiangshan port, Zhejiang province, in 2003a, 2009a and 2014a; (**b**) Total area change of *S. alterniflora* in Xianshan port, in 2003a, 2009a and 2014a

change maps show that when S. alterniflora spreads to a new area, the original growth area may be altered to another landscape because of human activities. For example, there may be an S. alterniflora patch in 2003 but not in 2014. Thus, the increase in S. alterniflora area represents a dynamic balance between cultivation, proliferation, exploitation and spread to new areas. This is typical of sprawling in Xiangshang Bay. We use geospatial statistics to determine the exact distribution area of S. alterniflora. As shown in Fig. 8.21b, the spread of S. alterniflora in this region can be represented with an exponentially increasing function over the last

10 years. In 2003, the area of S. alterniflora is approximately 590 hm^2 and it is sparely distributed in the estuary; in 2009, the area is approximately 1745 hm^2; and in 2014, the area is approximately 5715 hm^2. In 2014, it occupies almost all muddy beaches. The area of S. alterniflora has increased 9.68-fold over the 10-year period. In the first 5 years, it increased 2.96-fold and it increased 3.27-fold in the last 5 years.

8.4.5 Summary

This paper describes a new hybrid method for using remote sensing technology to monitor the spread of the invasive species S. alterniflora in Xiangshan Bay, Zhejiang province during the period from 2003 to 2014. Our results demonstrate that the proposed algorithm is practical and effective. The overall accuracy reached 81.67 %, which shows that the method performs well. The approach provides a technological reference for S. alterniflora extraction based on medium resolution multispectral remotely sensed images. It can be applied in surveys of S. alterniflora macro survey, as well as dynamic detection over the long-term using 30 m per-pixel Landsat imagery.

Through the method of 3 time dynamic detection, we found that the distribution area of S. alterniflora enlarged ten folds between 2003 and 2014 in Xiangshan Bay. The change can be described by an exponentially increasing model. The area enlarged from 590 hm^2 to 5715 hm^2 in approximately 10 years. Within a decade, this invasive species had occupied almost all muddy beaches to become the most dominant coastal salt vegetation in the region. The spatial pattern of S. alterniflorashowed that it spread towards the sea, exploiting the intertidal zone. This trend is consistent with the results of some previous studies (Zhang 2010).

Acknowledgments The study is funded by the Free Exploring Program of the State Key Laboratory of Remote Sensing Science of China (No. 14ZY-03), the Special Research Project for the Commonweal of the Ministry of Water Resources of the People's Republic of China (grant no. 201201092), the 908 Project of the State Oceanic Administration, China (No. 908-03-03-02), the National Natural Science Foundation of China (grant no. 61473286 and 61375002), and the International Science & Technology Cooperation Program of China (grant no. 2010DFA92720).

References

Adams JB, Sabol DE, Kapos V, Filho RA, Roberts DA, Smith MO, Alan RG (1995) Classification of multispectral images based on fractions of endmembers: application to land cover change in the Brazilian Amazon. Remote Sens Environ 52:137–154

Aguiar APD, Shimabukuro YE, Mascarenhas NDA (1999) Use of synthetic bands derived from mixture models in the multispectral classification of remote sensing images. Int J Remote Sens 20:647–657

Augusteijn MF, Warrender CE (1998) Wetland classification using optical and radar data and neural network classification. Int J Remote Sens 19(8):1545–1560. doi:10.1080/014311698215342

Ayers DR, Strong DR (2002) The Spartina invation of Sanfrancisco bay. Aquat Nuis Species Dig 4:38–40

Bai J, Chen X, Li J, Yang L, Fang H (2011) Changes in the area of inland lakes in arid regions of central Asia during the past 30 years. Environ Monit Assess 178(1–4):247–256

Beijma S van, Comber A, Lamb A The use of high resulotion sat and optical remote sensing data for mapping and monitoring of intertidal salt marsh vegetation habitats

Boardman JW, Kruse FA (1994) Automated spectral analysis: a geological example using AVIRIS data, north Grapevine Mountains, Nevada. In: Proceedings, ERIM tenth thematic conference on geologic remote sensing, Ann Arbor, MI, pp 407–418

Boesch DF, Josselyn MN, Mehta AJ, Morris JT, Nuttle WK, Simenstad CA, Swift DJ (1994) Scientific assessment of coastal wetland loss, restoration and management in Louisiana. J Coast Res 20:1–103

Bwangoy J-RB, Hansen MC, Roy DP, Grandi GD, Justice CO (2010) Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. Remote Sens Environ 114(1):73–86. doi:10.1016/j.rse.2009.08.004

Cahoon DR, Hensel PF, Spencer T, Reed DJ, McKee KL, Saintilan N (2006) Coastal wetland vulnerability to relative sea-level rise: wetland elevation trends and process controls wetlands and natural resource management. Springer, Berlin, pp 271–292

Chander G, Markham BL, Helder DL (2009) Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sens Environ 113:893–903

Chang C-C, Lin C-J (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol (TIST) 2(3):27

Chen X, Ma J, Ma Y (2014) Analysis of the spatial and temporal changes of the coastline in the Haizhou Bay during the past 40 Years. Adv Mar Sci 32(3):324–333

Chung CH (2006) Forty years of ecological engineering with Spartina plantations in China. Ecol Eng 27:49–57

Cochrane MA, Souza CM Jr (1998) Linear mixture model classification of burned forests in the eastern Amazon. Int J Remote Sens 19:3433–3440

Cracknell AP (1998) Synergy in remote sensing–what's in a pixel. Int J Remote Sens 19:2025–2047

Cui B, Chang X, Chen Y et al (2007) Dynamic mornitoring of coastline in the Yellow river estuary by remote sensing. Sci Survey Map 32:108–109

Daehler CC, Strong DR (1996) Status, prediction and prevention of introduced cordgrass Spartinas invasions in pacific estuaries, USA. Biol Conserv 78:51–58

Davranche A, Lefebvre G, Poulin B (2010) Wetland monitoring using classification trees and SPOT-5 seasonal time series. Remote Sens Environ 114(3):552–562. doi:10.1016/j.rse.2009.10.009

DeFries RS, Hansen MC, Townshend JRG (2000) Global continuous fields of vegetation characteristics: a linear mixture model applied to multiyear 8 km AVHRR data. Int J Remote Sens 21:1389–1414

Deng ZF, An SQ, Zhi YB (2006) Preliminary studies on invative model and outbreak mechanism of exotic species, Spartina alterniflora Loisel (in chinese). Acta Ecol Sin 26(8):2678–2686

Dennison PE, Roberts DA (2003) The effects of vegetation phenology on endmember selection and species mapping in southern California chaparral. Remote Sens Environ 87:295–309

Donoghue D, Thomas DR, Zong Y (1994) Mapping and monitoring the intertidal zone of the east coast of England using remote sensing techniques and a coastal monitoring GIS. Mar Technol Soc J 28:19

Dronova I, Gong P, Wang L (2011) Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at Poyang Lake, China. Remote Sens Environ 115(12):3220–3236. doi:10.1016/j.rse.2011.07.006
Du Tao ZB (1999) A study of mapping coast by processing remote sensing image with wavelets. Mar Sci Bull 19–21
Ekercin S (2007) Coastline change assessment at the Aegean Sea coasts in Turkey using multitemporal Landsat imagery. J Coast Res 233:691–698
Feder HM, Bryson-Schwafel B (1988) The intertidal zone. Environmental Studies in Port Valdez, Alaska, pp 117–164
Fisher P (1997) The pixel: a snare and a delusion. Int J Remote Sens 18:679–685
Frazier SP, Page JK (2000) Water body detection and delineation with Landsat TM data. Photogramm Eng Remote Sens 66:1461–1467
Frohn RC, Autrey BC, Lane CR, Reif M (2011) Segmentation and object-oriented classification of wetlands in a karst Florida landscape using multi-season Landsat-7 ETM+ imagery. Int J Remote Sens 32(5):1471–1489. doi:10.1080/01431160903559762
Gade M, Alpers W, Melsheimer C, Tanck G (2008) Classification of sediments on exposed tidal flats in the German Bight using multi-frequency radar data. Remote Sens Environ 112(4):1603–1613
GarciaHaro FJ, Gilabert MA, Melia J (1996) Linear spectral mixture modeling to estimate vegetation amount from optical spectral data. Int J Remote Sens 17:3373–3400
Gedan KB, Kirwan ML, Wolanski E, Barbier EB, Silliman BR (2011) The present and future role of coastal wetland vegetation in protecting shorelines: answering recent challenges to the paradigm. Clim Chang 106(1):7–29
Green AA, Berman M, Switzer P, Craig MD (1988) A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Trans Geosci Remote Sens 26:65–74
Green E, Mumby P, Edwards A, Clark C (1996) A review of remote sensing for the assessment and management of tropical coastal resources. Coast Manag 24(1):1–40
Islam MA, Thenkabail PS, Kulawardhana RW, Alankara R, Gunasinghe S, Edussriya C, Gunawardana A (2008) Semi-automated methods for mapping wetlands using Landsat ETM + and SRTM data. Int J Remote Sens 29(24):7077–7106. doi:10.1080/01431160802235878
Han Z, Jing Y (2005) The Yangtze Delta muddy tidal flat shoreline extraction based on multi-source remote sensing data of Satellite infrared and microwave. Prog Nat Sci 18:1000–1006
Huo J, Wang C, Wang Z (2003) A multi-threshold based morphological approach for extracting coastal line feature from remote sensing images. J Image Graph 8:805–809
Ketchum BH (1972) The water's edge: critical problems of the Coastal Zone. MIT Press, Cambridge, MA
Lee J-S, Jurkevich I (1990) Coastline detection and tracing in SAR images. IEEE Trans Geosci Remote Sens 28(4):662–668
Lee SY, Dunn RJK, Young RA, Connolly RM, Dale P, Dehayr R,. . .Teasdale P (2006) Impact of urbanization on coastal wetland structure and function. Aust Ecol 31(2):149–163
Li J, Sheng Y (2012) An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: a case study in the Himalayas. Int J Remote Sens 33(16):5194–5213
Li J, Yang X, Tong Y, Zhang D, Sheng Y, Zhang R (2005) Influences of Spartina alterniflora invasion on ecosystem services of coastal wetland and its countermeasures. Mar Sci Bull 24(5):33–38
Li Q, Wu Z, Chu B, Zhang N, Cai S, Fang J (2007) Heavy metals in coastal wetland sediments of the Pearl River Estuary, China. Environ Pollut 149(2):158–164
Li J, Sheng Y-W, Luo J (2011) Automatic extraction of himalayan glacial lakes with remote sensing. Yaogan Xuebao- J Remote Sens 15(1):29–43

Liu H, Jezek KC (2004) Automated extraction of coastline from satellite imagery by integrating Canny edge detection and locally adaptive thresholding methods. Int J Remote Sens 25(5): 937–958
Liu JE, Zhou HX, Qin P, Zhou J (2007) Effects of Spartina alterniflora salt matshes on organic carbon acquisition in interfida; zones of Jiangsu province, China. Ecol Eng 7:240–249
Lu D, Weng Q (2006) Use of impervious surface in urban land-use classification. Remote Sens Environ 102:146–160
Lu D, Moran E, Batistella M (2003) Linear mixture model applied to Amazonian vegetation classification. Remote Sens Environ 87:456–469
Luo JC, Sheng YW, Shen ZF, LI JL (2009) Automatic and high-precise extraction for water information from multispectral images with the step-by-step iterative transformation mechanism. J Remote Sens 13(4):610–615
Mason DC, Davenport IJ (1996) Accurate and efficient determination of the shoreline in ERS-1 SAR images. IEEE Trans Geosci Remote Sens 34(5):1243–1253
McFeeters SK (1996) The use of normalized difference water index (NDWI) in the delineation of open water features. Int J Remote Sens 17(7):1425–1432
Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. ISPRS J Photogramm Remote Sens 66(3):247–259
Murray N, Phinn S, Clemens R, Roelfsema C, Fuller R (2012) Continental scale mapping of Tidal flats across East Asia using the Landsat archive. Remote Sens 4(12):3417–3426. doi:10.3390/rs4113417
Mustard JF, Sunshine JM (1999) Spectral analysis for earth science: investigations using remote sensing data. In: Remote sensing for the earth sciences: manual of remote sensing, vol 3. Wiley, New York, pp 251–307
Mwita E, Menz G, Misana S, Nienkemper P (2012) Detection of small wetlands with multi sensor data in East Africa. Adv Remote Sens 01(03):64–73. doi:10.4236/ars.2012.13007
Nicholls RJ (2004) Coastal flooding and wetland loss in the 21st century: changes under the SRES climate and socio-economic scenarios. Glob Environ Chang 14(1):69–86. doi:10.1016/j.gloenvcha.2003.10.007
Niedermeier A, Lehner S, Van der Sanden J (2001) Monitoring big river estuaries using SAR images. In: Geoscience and remote sensing symposium, 2001. IGARSS'01. IEEE 2001 International, IEEE, pp 1756–1758
Otsu N (1975) A threshold selection method from gray-level histograms. Automatica 11(285–296):23–27
Ouma YO, Tateishi R (2006) A water index for rapid mapping of shoreline changes of five East African Rift valley lakes: an empirical analysis using Landsat TM and ETM+ data. Int J Remote Sens 27(15):3153–3181
Ridd MK (1995) Exploring a V–I–S (Vegetation–Impervious Surface–Soil) model for urban ecosystem analysis through remote sensing: comparative anatomy for cities. Int J Remote Sens 16:2165–2185
Rosso PH, Ustin SL, Hastings A (2006) Use of lidar to study changes associated with Spartina invasion in San Francisco Bay marshes. Remote Sens Environ 100:295–306
Ryu J-H, Won J-S, Min KD (2002) Waterline extraction from Landsat TM data in a tidal flat, a case study in Gomso Bay, Korea. Remote Sens Environ 83:442–456
Ryu J-H, Kim C-H, Lee Y-K, Won J-S, Chun S-S, Lee S (2008) Detecting the intertidal morphologic change using satellite data. Estuar Coast Shelf Sci 78(4):623–632
Schmidt K, Skidmore A (2003) Spectral discrimination of vegetation types in a coastal wetland. Remote Sens Environ 85(1):92–108
Simenstad C, Reed D, Ford M (2006) When is restoration not?: incorporating landscape-scale processes to restore self-sustaining ecosystems in coastal wetland restoration. Ecol Eng 26(1):27–39
Sklar FH, Costanza R, Day JW (1985) Dynamic spatial simulation modeling of coastal wetland habitat succession. Ecol Model 29(1):261–281

Smith MO, Ustin SL, Adams JB, Gillespie AR (1990) Vegetation in deserts: I. A regional measure of abundance from multispectral images. Remote Sens Environ 31:1–26
Sohn H-G, Jezek K (1999) Mapping ice sheet margins from ERS-1 SAR and SPOT imagery. Int J Remote Sens 20:3201–3216
Sun X, Lv T, Gao Y (2014) Driving force analysis of Bohai bay coastaline change from 2000 to 2010. Res Sci 36(2):0413–0419
Theseira MA, Thomas G, Taylor JC, Gemmell F, Varjo J (2003) Sensitivity of mixture modeling to endmember selection. Int J Remote Sens 24:1559–1575
Van der Meer F, de Jong SM (2000) Improving the results of spectral unmixing of Landsat Thematic Mapper imagery by enhancing the orthogonality of end-members. Int J Remote Sens 21:2781–2797
Vapnik VN, Kotz S (1982) Estimation of dependences based on empirical data, vol 40. Springer, New York
Vapnik V, Golowich SE, Smola A (1996) Support vector method for function approximation, regression estimation, and signal processing. In: Advances in neural information processing systems, vol 9. Citeseer
Vernberg WB, Vernberg FJ (2012) Environmental physiology of marine animals. Springer Science & Business Media, New York
Vinther N, Christiansen C, Bartholdy J (2001) Colonisation of Spartina on a tidal water divide, Danish Wadden Sea. Dan J Geogr 101:11–20
Wan H, Wang Q, Jiang D, Fu J, Yang Y, Liu X (2014) Monitoring the invasion of Spartina alterniflora using very high resolution unmanned aerial vehicle imagery in Beihai, Guangxi (China). Sci World J 2014:1–8
Wang G, Guan D (2012) Effects of vegetation cover and normalized difference moisture index on thermal landscape pattern: a case study of Guangzhou, South China[J]. Yingyong Shengtai Xuebao 23(9):2429–2436
Wang L, Xu H, Pan S (2005) Dynamic monitoring of the shoreline changes in Xiamen Island with its surrounding areas using remote sensing technology. Rem Sens Technol Appl 20:404–410
White K, El Asmar HM (1999) Monitoring changing position of coastlines using Thematic Mapper imagery, an example from the Nile Delta. Geomorphology 29(1):93–105
Xi C, Jiancheng L, Zhanfeng S, Changming Z, Xin Z, Liegang X (2011) Estimation of impervious surface based on integrated analysis of classification and regression by using SVM. In: Geoscience and remote sensing symposium (IGARSS), 2011 I.E. International, pp 2809–2812: Ieee
Xu GW, Zhou RZ (1985) Preliminary studies of introduced Spartina alterniflora Loisel in China. J Nanjing Univ Res Adv Spartina 212–225
Xiao X, Shen Z, Qin X (2001) Assessing the potential of VEGETATION sensor data for mapping snow and ice cover: a normalized difference snow and ice index. Int J Rem Sens 22(13):2479–2487
Xu Hanqiu (2008) A new remote sensing index for fastly extracting impervious surface information. Geomat Inform Sci Wuhan Univ 11:1150–1155.
Yang H, Guo H, Wang C (2001) Coast line dynamic inspect and land cover classification at Yellow River Mouth using TM-SAR data fusion method. Geogr Territorial Res 17(4):185–191
Yao X, Gao Y, Du Y (2013) Spatial and temporal changes of Hainan Coast-line in the past 30 years based on RS. J Nat Res 28(1):114–125
Yuan H, Li S, Zheng H (2009) Evaluation of the influences of foreign Spartina alterniflora on ecosystem of Chinese coastal wetland and its countermeasures. Mar Sci Bull 28(6):122–128
Zhang Y (2010) The spatial distribution and bioenergy estimation of an invasive plant Spartina alterniflora in China. Master, Zhejiang University
Zhang M, Ustin SL, Rejmankova E, Sanderson EW (1997) Monitoring pacific coast maushes using remote sening. Ecol Appl 7(3):1039–1053
Zhang RS, Shen YM, Lu LY et al (2004) Formation of Spartina alterniflorasalt marshes on the coast of Jiangsu Province. China Ecol Eng 23:95–105

Zhang J, Pan Y, He C (2005) The high spatial resolution remote sensing image classification based on SVM with the multi-source data. In: Geoscience and remote sensing symposium, 2005. IGARSS'05, Proceedings. 2005, IEEE International (vol. 6, pp 3818–3821): IEEE

Zhang Y, Lu D, Yang B, Sun C, Sun M (2011) Coastal wetland vegetation classification with a Landsat Thematic Mapper image. Int J Remote Sens 32(2):545–561. doi:10.1080/01431160903475241

Zhao B, Guo H, Yan Y, Wang Q, Li B (2008) A simple waterline approach for tidelands using multi-temporal satellite images: a case study in the Yangtze Delta. Estuar Coast Shelf Sci 77(1):134–142

Zhu X (2002) Remote sensing monitoring of coastline changes in Pearl river estuary. Mar Environ Sci 21:19–22

Zhu C, Luo J, Shen Z, Li J (2011) A spatial adaptive algorithm for endmentber extraction on multispextral remote sensing image. Spectrosc Spectr Anal 31(10):2814–2818

Zhu C, Luo J, Ming D, Shen Z, Li J (2012) Method for generating SPOT natural-colour composite images based on spectrum machine learning. Int J Remote Sens 33(4):1309–1324

Zhu C, Zhang X, Luo J (2013) Automatic extraction of coastline by remote sensing technology based on SVM and auto-selection of training samples. Remote Sens Land Res 25(2):69–74

Zhu C, Zhang X, Jiaguo QI (2016) Detecting and assessing spartina invasion in coastal region of china: a case study in the xiangshan bay. Acta Oceanol Sin 35(4):35–43

Zhu C, Li J, Zhang X, Luo J (2014) Wetlands information automatic extraction from high resolution remote sensing imagery based on object-orient technology (in chinese with english abstract). Bull Surv Mapp 10:23–28.

Zuo P, Zhao S, Liu C a, Wang C, Liang Y (2012) Distribution of Spartina along China's coast. Ecol Eng 40:160–166

Chapter 9
Applications and Practice of Digital Ocean and Digital Coast

Xiaoyi Jiang

Abstract Firstly, this chapter introduces the digital ocean and digital coast applications from the standpoint of ocean ecological environmental monitoring, shoreline and island management, El Nio phenomenon and the Sea-level rising monitoring, fisheries and public service and so on. Then, the practice in the program of China offshore digital ocean information infrastructure is introduced. The achievements laid the technical foundation for the development of China's digital ocean, the sustainable development of China's ocean industry.

Keywords Digital Oceany system • Application • China offshore

9.1 Introduction

Currently, the most influential digital ocean system in the world is Google Ocean, launched by Google in 2009. It is a new functional module in Google Earth 5.0. By downloading and installing Google Earth software and connecting to the Internet, users can use Google Ocean through remote access to Google's database.

From the perspective of content and technology, Google Ocean is a tool for understanding and enjoying the ocean. Through a virtual depiction of earth, people can see the colorful ocean, mysterious trenches, and sub-ocean cracks and volcanoes. Through ocean observations, users can obtain data on the ocean environment, including, e.g., climate change and coastal zone changes. Users can find new attractions, such as surfing, diving and tourist areas, watch videos of exotic ocean life, read about shipwrecks in a certain ocean area, and explore under-ocean volcanoes and wrecks. Furthermore, users can view photos and videos of famous attractions such as the Bahamas, the Red Ocean and the Great Barrier Reef, track underwater creatures, which are located by satellite. For the deep application, users can develop their own software modules and integrate them into Google Earth

X. Jiang (✉)
National Marine Data and Information Service, State Oceanic Administration of China, Tianjin 300171, China
e-mail: andyjiangxy@126.com

X. Zhang et al. (eds.), *Modeling with Digital Ocean and Digital Coast*, Coastal Research Library 18, DOI 10.1007/978-3-319-42710-2_9

through the open API interface provided by Google Earth, but the number of open API interfaces is limited.

Except for the public information exploring applications stated above, there are much more applications and studies can be carried out by scientists and professional departments in the area of digital ocean and digital coast.

9.1.1 Ocean Ecological Environment Monitoring

The ocean ecosystem is a complex and multilateral system. It is increasingly influenced by human activities and natural factors, especially considering the current rapid development of an ocean economy and global changes (Alexandrakis et al. 2011; Brakenridge et al. 2012). On the one hand, there are major global environmental changes, such as environmental pollution, overfishing, habitat destruction, the invasion of alien species, global warming, ocean level rise, El Nino La Nina, and the Antarctic ozone hole expansion. These changes will allow marine ecosystems to make sensitive responses. On the other hand, many types of changes in the marine ecosystem will have a negative impact on human development and utilization of the ocean, which will cause adverse effects on ocean fisheries, marine aquaculture and coastal tourism. Therefore, understanding the evolutionary trend of the ocean ecosystem and the influence of human activities in a timely fashion and implementing ocean ecological environment monitoring are necessary for ocean ecological protection and management. In ocean ecological protection and management, we must understand the evolutionary trends of the ocean ecosystem and the influence of human activities, to implement ocean ecological environment monitoring.

Because ocean environmental monitoring gives priority to coastal zones and coastal waters, the coastal and near shore ocean ecosystems include the following: coastal wetlands, lagoons, mangroves, coral reefs, ocean grass beds, estuaries, bays and a variety of other ecological types (Chang 2001; Cutter 1996; Cutter et al. 2000; Cutter et al. 2008) . Therefore, when designing an ocean ecological environment monitoring project, we must consider how to choose the appropriate monitoring project according to the different ecological types, and these monitoring programs should be able to reflect the evolutionary trends of ocean ecosystems and the impact of human activities.

Although monitoring the ocean environment cannot provide a certain unity of projects and targets due to the complex and diverse ecological types compared with the monitoring of physical and chemical factors, some general monitoring projects can nevertheless be implemented. These projects involve water temperature, turbidity, salt degree value, dissolved oxygen content and distribution, nutrient content and distribution, richness and diversity of plankton, plants and animals, richness and diversity of benthos, the ocean chlorophyll content and primary productivity estimates. Except for the above, these projects also include benthic environmental conditions, suspended particulates, biological habitat status, the development and

utilization of land and ocean areas, and the impact of human activities. These monitoring projects can be applied to all types of marine ecological environments and supply the basic data used in mastering the state of the ocean environment. For specific ocean ecological types, we must select specific monitoring items according to their own ecological characteristics and laws. For example, the ecological monitoring of mangroves and coral reefs cannot adopt similar projects. However, some of the basic principles of ecological monitoring should be followed when selecting ecological monitoring projects. These basic principles include reflecting the ecological system's structure in a monitoring project, such as the population, community and ecological system level of a structure index that includes important fishing records of ocean animals due to the population dynamics of some top predators. For example, ocean birds can respond to an ocean environment's long-term changes and should thus be considered in a monitoring project. Such a monitoring project should reflect the ecological system function, such as an energy learning index through a nutrition structure determination, and several relatively easy determinations of pressure indexes of biology and ecology or monitoring projects that can evaluate the index of Pollutant discharge to the ocean environment load, and other indicators that reflect specific human activities. Examples include marine aquaculture in coastal wetland activities and recreational and tourist activities such as the distribution of coral reefs; in addition, some important indicators of ocean ecology evolution and ocean alien species are important aspects of a project on ocean ecological monitoring.

Ocean color remote sensing monitoring is mainly based on the study of spectral reflectance characteristics of different areas (Gao et al. 2014; Jing et al. 2012). Clean water has low reflectivity. The water color shows a strong absorption of light, and strong molecular scattering exists only in the shorter spectral region. Thus, in general, in remote sensing images, the water color is dark, especially in the infrared spectrum. For water quality monitoring, we can adopt remote sensing technology that regards water color spectral characteristics and water color as indicators. Marine oil pollution and the dumping of waste are important reasons for the deterioration of the ocean environment.

The chemical fertilizers of the farmland along the river along with wastewater from the city and industrial wastewater are dumped constantly into the ocean, which expands the scope of marine pollution and causes the ecological environment to deteriorate and the environmental quality to decline. Remote sensing satellites have been applied, especially remote sensing satellites that can be used to research for oil pollution and chemical pollution over a large range. It is also possible to estimate the scope of the pollution and its diffusion, thus providing the necessary data and information for the ocean environmental protection department.

Digital Ocean is a fundamental information platform for comprehensive integration and analysis of Earth observation data, model data, field monitoring data, remote sensing data and its products; it forms a mechanism for ecological environment evolution prediction research, prevention and emergency, and many other areas, to provide support (Kuo and Guo Sheng 2011, 2013; Liang and Zou 2005).

Fig. 9.1 The changes present of the coast line in a digital coast system

9.1.2 Management of Shoreline and Islands

The Digital Ocean fundamental platform can achieve three-dimensional visualization expression and a location query of the coast, important ocean areas, islands, high-precision images of coastal areas, vectors and terrain, integrated Ocean scene simulation functions of ocean water, landscape, submarine topography and coastal Marine life, which can reappear in the real-world islands. With respect to application of this project to islands, coastal zone investigation data can achieve island information (e.g., location, area, long coastline, population), including a variety of conditions for query and retrieval, positioning and island imaging and the visualization of a terrain display function. Fig. 9.1shows a coast line change.

9.1.3 The Application in El Nino and Sea-Level Rising Study

The El Nino phenomenon, which is also known simply as El Nino, is a descriptor of an unusual weather phenomenon in Peru and Ecuador. It mainly refers to the Eastern Pacific and Central tropical oceans' having abnormal ocean temperatures that continue to warm, which causes changes in the world's climate patterns and creates droughts in some of the areas, whereas other areas have excessive rainfall. Its frequency is not regular, but on average, it occurs approximately every 4 years. If the duration of the phenomenon is less than 5 months, then it is known as an El Nino condition; if the duration is 5 months or longer, then it is called an El Nino event (Lin et al. 2012; Lubna et al. 2010).

It is generally believed that ocean surface temperatures that are above average by 0.5° C for three consecutive months or more can be considered an El Nino phenomenon. Meteorology generally agrees that the occurrence of the El Nino phenomenon has an omen-like significance to global climate disasters in many areas at present, and thus, El Nino phenomenon monitoring has become one of the important contents in climate monitoring.

Sea-level rise is the phenomenon that is part of global Sea-level rise, which is caused by global warming, melting polar ice, upper ocean thermal expansion and other factors. Studies have shown that there has been a 10 ~ 20 cm rise in the global ocean level during the past century and that this rise will accelerate in the future. However, the actual ocean level changes in a certain region of the world are also influenced by the local land vertical movements – the slow crustal movements and local surface subsidence. Global Sea-level rise coupled with local land movements results in a relative Sea-level change in the area. Therefore, when research is performed on an area of Sea-level rise, it is only meaningful to study the relative ocean level rise (Nyaruhuma et al. 2012; Sheik Mujabar and Chandrasekar 2011; Shi et al. 2010).

The rise in the ocean level has a large influence on social and economic aspects and the natural environment and ecosystem of coastal areas. First, rising ocean levels could drown low-lying coastal regions, strengthen oceanic dynamic factors for progress toward the beach, and produce erosion of the coasts, for example, changing a mulberry field to the ocean. Second, a rise in the ocean level will increase the intensity of storm surges and increase the frequency, which would not only endanger the lives and property of the people in the coastal areas but also cause land salinization. In China, the impact of ocean level rise mainly affects the Bohai Bay area, the Yangtze River Delta region and the Pearl River Delta region.

The study of the El Nino phenomenon and the mechanisms of ocean level rise must be analyzed using forecasting based on the digital ocean platform. Temporal and spatial analysis and the expression of the El Nino phenomenon and ocean level rise are the key content of the "Digital Ocean" El Nino phenomenon and the rise in ocean level. In the space and time data process, when expressing the El Nino phenomenon and ocean level rise forecasting data, the most common method is to adopt a form of animation sequence, taking discrete time point images in a time series as consecutive frames. Through a single frame query or continuous data displays, disasters can be shown in the form of time and space processes; this method addresses problems in the data real-time display, numerical dynamic query and multilevel resolution. Based on the digital ocean, the El Nino phenomenon and the rise in ocean level, the prediction and forecasting results of the numerical model of the integration of the El Nino phenomenon and the ocean level rise provides a visualization of the results of the prediction in a three-dimensional scene, which is displayed and analyzed, and the relevant information is provided to the user on the related disasters, including answers to queries and statistics.

The core ideal of the temporal and spatial expression process of the phenomenon of El Nino and the Sea-level rise is the output of each type of prediction model

taken as a whole (Taramelli et al. 2010; Verhoeven et al. 2012). On the basis of the number of times, the whole data set is divided into a single process, and the data of each subset is a complete forecasting process. The single-process forecasting data set contains both the spatial distribution information and the time course information, and the information on the individual times is separated for each single time by the discrete marking of time. A Pyramid-based data structure is stored for each moment. Based on the basic idea of the Pyramid data structure for managing the single-time forecasting data, when a user views a certain moment, it is not necessary to obtain the most accurate and complete data; the result is based on the range of the region of interest and the current resolution.

In short, the information integration of the digital ocean's basic information platform is accomplished through comprehensive integration and analysis of the ground observation data, the model calculation data and field monitoring data. It uses disaster mechanism research, disaster forecasting, disaster prevention and emergency rescue and other links to provide support on the information integration of the digital ocean's basic information platform.

9.1.4 The Application in Fisheries

Ocean fisheries constitute one of the most important industries. The ecological environment of fishing waters is an important material basis for fisheries development. Starting from the 1990s, however, with new economic development, marine fishing waters are suffering from severe stress due to humans, including over fishing, environmental pollution, and habitat destruction. Ecological environment problems in the marine fisheries' areas are becoming increasingly prominent, such as red tide, acidification, ocean level rise, biological depletion, and ecosystem degradation. Therefore, it is urgent to strengthen the monitoring of ocean fishery resources and the ecological environment and to protect the fishery resources and the viability of the fisheries. For a long time, monitoring and research on fishery resources have become more dependent on conventional offshore field observations. This method is limited by time and space. It has a high cost, and it is slow and difficult to achieve synchronous sampling and measurement of large-scale amounts of water. Additionally, the acquired data are often not highly efficient, which cannot meet the needs of real-time management of the entire aquatic ecosystem fisheries resources.

With the increasing dependence on fishery resources and the rapid development of global environmental changes, this approach cannot meet the requirements of the traditional monitoring and research methods of fishery resources (Xu and Tang 1998; Yin et al. 2001). The digital ocean can provide powerful information support and an analysis function to the ocean fishery industry from different scales and different applications viewpoints.

The application of digital ocean to fishery resources and environment data acquisition mainly involves obtaining some important environmental parameters with respect to the fishery resources, which is based on the use of satellite images and other remote sensing data, such as ocean surface temperature, ocean color, ocean surface salinity, and water physical characteristics (height, tide, wind, waves). For the environmental monitoring of the offshore fishery resources, the satellite remote sensing images are more vulnerable to the continental aerosol and spectral characteristics of the surface features, while the airborne remote sensing images have more advantages in terms of spatial resolution, sampling time and free flight route. In recent years, new sensors have been dedicated to breaking through the limitations of the spatial and spectral resolutions; sensors with higher spatial and spectral resolution have been used in the ecological monitoring and evaluation of fishery resources. For example, the medium resolution imaging spectrometer (MODIS) is a new generation of the world's new generation of optical remote sensors; its wide spectrum range (400 ~ 1440 nm) and high spectral resolution (20 nm) can monitor the ocean water color and ocean surface temperature and is a powerful tool for monitoring the rising ocean.

The digital ocean can establish a broad database that includes fishery resources, the environment, and social and economic factors, forming the management module for the fishery resources and environment. This module is built under the premise of the integration of traditional methods and remote sensing to obtain a variety of environmental, biological and social and economic data.

Currently, the global network has many free and open source GIS software and data, which scientists can easily use. According to the moderate range of the marine fishery to the marine environment (the environmental parameters include water depth, salinity, temperature, primary productivity, ocean ice, coastal zone), the model of the distribution of ocean species is generated on a large scale. The China National Marine Information Center specially developed a marine information-sharing website to achieve the sharing of marine information services. This web site has many databases, which include the following: ocean information sharing website, website with ocean resources and space database, 1:1000000 ocean basic geographic information, ocean economic spatial database, ocean ecological spatial database, ocean environment, ocean disaster spatial database, and ocean remote sensing information products data. However, the existing distributed data resources cannot meet the requirements for the development of marine fisheries, which are based on many factors, different scales and high efficiency. However, the existing distributed data resources fail to meet the requirements of the development of Marine Fisheries based on multiple factors, different scales and high efficiency. The digital ocean can meet the requirements of the future development of marine fisheries from the aspects of integrated, comprehensive, multi-scale, and dynamic performance.

Based on the digital ocean platform, useful information about the species abundance in marine fisheries, the range of spawning grounds, and the location of the migrating channel in different time and space scales is obtained; remote sensing technology is used to provide the ocean environment parameters. It is the primary problem of the temporal and spatial variation of fishery resources to determine the

distribution of the fishery resources. The range of the fish spatial distribution can be obtained by observing the stomach content, fish type, developmental stage and sign data unit, fishing to catch and wild behavior. However, because these data are obtained at different sampling points, it can only represent the value of the point and the surrounding small range, and it is not possible to know the situation of the sample and the whole ocean or water body. Through the analysis of the spatial structure and the algorithm of the variation function, we can simulate the high precision of the fish abundance distribution data according to the limited sampling points.

9.2 Public Service Application of Digital Ocean

"Digital ocean" public service objectives is providing WEB-based ocean science knowledge and ocean forecasting information to the public, through the integration of ocean survey and evaluation data, historical data, ocean base geo-data, remote sensing data and product information that was data based on the earth's sphere. A typical public application system is iOcean, for which the site is www.iocean.cn. This site operates in the Internet environment, and the content can be divided into nine parts: the seabed, digital, virtual aquarium water, ocean resources, visit the polar ocean, Naval, ocean forecasting, oceanographic research and observation, and ocean science. The main functional interface is shown in Fig. 9.2.

Fig. 9.2 "Digital Ocean" public service function interface

9.3 China Digital Ocean Information Infrastructure

From 2006 to 2011, China's State Oceanic Administration led the implementation of the largest-ever ocean information project, "China offshore digital ocean information infrastructure". The purpose is to lay the technology foundation, the information foundation and the application foundation for the development of China's digital technology. The seven main achievements of the project include the establishment of the digital ocean standard system, the data acquisition, processing and updating system, the ocean data warehouse system, the network and exchange system, the security system, the application service system, and the operation and maintenance support system (Yin et al. 2006; Ying et al. 2013).

9.3.1 Architecture

Based on the international trends and the development of ocean technology, the digital ocean is an application process that is based on informational technology and guaranteed by ocean information infrastructure, in view of the specific requirements of ocean scientific research and ocean management applications and with the aim to fully tap and share massive, multi-source, long sequences of ocean data. China's ocean information work started late. The ocean infrastructure construction and the corresponding soft environment foundation are weak. The information of ocean management departments at all levels of the state and the study level different institutions are uneven, and ocean data and information resources are unevenly distributed. We have not yet built a global digital ocean hardware or software environment on the national level, so we must achieve a major breakthrough in many areas and face the many challenges to achieve independent innovation and development. Therefore, based on the principle of 'ramming foundation, service orientation', we should combine the reality of the ocean information construction of China, plan the content framework of China's digital ocean construction overall, carry out digital ocean construction planning based on basic ocean data acquisition, software and hardware environmental capability construction, ocean information resource integration and development, digital ocean visualization and system application and service. We must also focus on the limited material and technological advantages, establish our own core technology system, and implement digital ocean engineering in stages.

1. Digital ocean information infrastructure

The information infrastructure is the basis of the development of digital oceans. Its main contents include the establishment of an information exchange

network system used to connect with all levels of ocean authorities, the ocean operations center and ocean research institute, the establishment of a multi-level distributed ocean data center for the states and coastal provinces and cities, and the establishment and improvement of ocean information standardization system.

2. Ocean data acquisition and transmission

The acquisition and transmission of ocean data are the premise and foundation of building the digital ocean. Using ocean data acquisition means such as ocean surveys, observations, and monitoring, we can construct a three-dimensional ocean observation system formed by space-based, air-based, land-based, and ocean-based information. We can then determine whether the data acquisition and transmission ability of the distribution density is reasonable and the monitoring data are complete. We can finally fully obtain basic environmental data, such as ocean hydrology, meteorology, and ocean management data, including the use of ocean areas, island management, the ocean economy, law enforcement supervision, and the protection of rights, and basic geographic data, including ocean base geography, topography, and imaging. This is the main data source of the entire digital ocean framework.

3. The exploitation and utilization of digital ocean information resources

The system can realize the value added by the ocean information and product. Through the establishment of the unified framework for digital ocean spatial data, integrating and utilizing all types of data, constructing a basic ocean database, an ocean database and the corresponding metadata entity, establishing a metadata navigation and database research index, and developing ocean information integration, integration and analysis.

4. Digital ocean multidimensional visualization platform

The visualization platform of the digital ocean is a tool with which to carry out the exhibition of the ocean information visualization and spatial analysis. With the support of high-performance computer and advanced visualization equipment, and through the use of, e.g., scientific visualization technology, GIS, and virtual simulation, based on the framework of digital ocean spatial data, basic ocean data and powerful method model, we can build a digital ocean visualization platform. We can realize a multi-dimensional visual expression and types analysis of ocean information as well as provide platform support for all kinds of ocean information system research and development.

5. Digital ocean application system

Digital ocean application system is the highest level of the digital ocean frame system. It includes the digital ocean application service system, which is for private network services, the digital ocean public system, which is for public network

services, and the digital integrated ocean information system, which is for comprehensive ocean management services.

9.3.2 Construction and Development

China's digital ocean information infrastructure has been built, and its main work is as follows: standard system construction, network platform building, the warehouse construction of ocean data, application system development, all levels node construction of the digital ocean, safety-guarantee system construction, system integration and the research and development of operation control systems.

1. Standard specification system construction

Taking account of relevant international and domestic standards, and combining the needs of China's ocean information technology development, we develop and modify a number of standards, including ocean information resource management, data resource integration, application system construction and the operational management of a total of four runs in nine categories. The formulation of the standard specification primarily includes the following: the ocean environment database standards, comprehensive ocean management special database standards, the ocean information metadata standards, ocean special survey element classification codes and schema legend rules, comprehensive ocean information management system construction specifications, digital ocean node construction technical specifications, the special ocean data and result management approach, the special ocean data exchange format standard, and the digital ocean node deployment trial operation and management measures.

2. Network construction

Ocean digital network construction aims to connect the state oceanic administration of Juzhu units, coastal provinces, cities and counties of a four-level network platform to meet the needs of digital ocean levels between nodes of ocean information transmission, exchange and information services. Network construction includes a backbone network and a branch network. The backbone network is composed of national and coastal provinces, which is composed of the national unity. The branch network is composed of provincial and county nodes, and it is constructed by each province under the framework of national unity, according to the current situation of local network resources in coastal provinces and cities and relying on the electronic government affairs network. Currently, the digital ocean information network has connected 11 coastal provinces, five municipalities, four research centers, and 19 national ocean Jushu units in China, for a total of 39 nodes.

3. Ocean data warehouse construction

The goal of ocean data warehouse construction is to use the data warehouse technology and to provide support for the integrated management and application of ocean data through the unified and effective multi-source, heterogeneous, multi-level management of massive ocean data according to different themes, regions, and time dimensions and to provide different levels of data access interfaces. At present, the development of the ocean data warehouse system is composed of the ocean information database, the product database and the database of the ocean application service. In addition, it explores and completes the ocean data warehouse update and maintenance system, the data warehouse management system, which is used for the daily updating and management of data.

4. Application system research and development

Application system research and development include the digital ocean application service system, the digital ocean public service system and the comprehensive ocean information management system. The digital ocean application service system exists in a three-dimensional sphere. It integrates natural geography, natural environment and comprehensive information management, and it provides four large modules, including basic data query retrieval, the statistical analysis of the ocean environment, integrated ocean information query management and node interaction, which provides one-stop ocean data sharing and integrated information application service for digital Ocean network users. China's internet-based Digital Ocean public service system provides the public with ocean information services and ocean science service, promotes the popularization of ocean knowledge, and improves ocean awareness through intuitive and innovative forms and popular content. The exploitation of the integrated ocean information management system is data supported by an ocean data warehouse, including ocean management, island management, environmental protection, ocean economy, ocean law enforcement, disaster prevention and mitigation, ocean science and technology and eight other sub-systems. It provides software tools for all levels of ocean departments.

5. The construction of all nodes in Digital Ocean

China's Digital Ocean construction includes 11 coastal provinces (autonomous regions and municipalities directly under the central government), i.e., Liaoning, Hebei, Tianjin, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, Guangxi and Hainan. It also includes five cities specifically designated in the state plan, e.g., Dalian and Qingdao, four research institutes, e.g., Ocean University of China and Xiamen University, and 19 national ocean Jushu units, for a total of 39 nodes. Nodes at all levels are important components of the digital ocean information infrastructure, and they are the objects of the application and service of the digital ocean system. They are also an important source of digital ocean data

and products. Node construction includes the construction of a green network, hardware and software environments, node-based data processing, data product production, node characteristics, and system development.

6. Security system construction

The goal of the Digital Ocean security system is to provide a basic software and hardware environment that ensures that all types of ocean database systems and application systems can operate safely and stably. This requires a comprehensive consideration and design of the overall structure of the network and hardware environment and the design and construction from this aspect of physical security, network security, information security, information security, system security, and computer virus defense.

7. Digital ocean system total integration and operation control

Digital ocean information infrastructure work relates to many nodes and the participation unit. There is a complex relationship between the system and function modules through the development of system integration, the integration of results of the hardware and software of the task unit, the establishment of a complete structure, and the full-featured, unified user security management of ocean digital integrated systems. We can achieve collaboration between nodes at different levels of business and information sharing and ultimately create a digital ocean system for running the business. At the same time, the operation control system is developed, which realizes the real-time monitoring and management of hardware and software, data transmission and the network running status of nodes.

9.4 Digital Ocean Application Service System

The aim of the Digital Ocean application service system is to create a one-stop ocean data-sharing and information application service platform using technologies such as virtualization, visualization, and large data analysis. The system is based on three-dimensional sphere expression platform, integrated ocean-base geography, the ocean environment, ocean business management, and other data. It functions based on the research and development of basic ocean data query and retrieval, ocean environment statistical analysis, comprehensive information query management, and information exchange. It manages data collection, analyzes information integration displays, links node interactions and provides information support services for China's ocean development, management and maintenance and other fields.

9.4.1 System Architecture

The digital ocean application service system is composed of a data platform, a virtual environment platform, an application service platform and a management platform. The isolation between each layer effectively reduces coupling, which makes the system easy to expand and easy to maintain. The data platform is the basis of the system and is primarily based on the integration and refinement of the formation of the application-oriented theme of the data warehouse and the data content, including basic ocean geography, remote sensing ocean images, ocean-bed terrain data and various ocean business management data. On the basis of the data platform, the virtual environment platform is formed by two times the three-dimensional sphere software and data platform, which provides not only basic information query, expression and analysis functions but also the interface with the data platform. The application service platform is based on the visualization platform, which primarily realizes the interactive interface of the query, retrieval, analysis and sharing of various ocean thematic information. It can be said that the application service platform is a virtual environment platform and the ocean business process of packaging. The top layer is the digital ocean application service system, which is based on the application of the Web environment. Through the digital ocean application service system, the user accesses a variety of digital ocean data, comprehensive information access and interactions. Virtual reality and inter-operability technology, massive map database service technology (based on the XML of ocean information exchange technology), data platform middleware, integrated services technology, and a number of key technologies and research and development results protect the advancement, stability and scalability of the digital ocean application service system.

9.4.2 Main Function Module

The digital ocean application service system includes five main functional modules (which is the first-level menu): basic ocean data retrieval services, the statistical analysis of the ocean environment, the integrated management ocean information query, special ocean application services and node information exchange. Each main function module includes a number of sub-modules (which is the second-level menu). In addition, the system realizes the layer management, the management of users' login, the system announcement and the message board and other auxiliary functions.

9.4.3 Basic Ocean Data Retrieval and Services

The main goal of the basic ocean data retrieval and service module is to provide users with ocean data-sharing services through downloading and online use. The module is modeled in the way of online shopping, regarding the data as a "commodity", by allowing users to retrieve their own interest in the "data goods"and then generating data to submit orders. Finally, the background of the

audit system reviews the order. For a review of the order, according to the data confidentiality level and the user's network status and other conditions, the "data goods" provided for the user are downloaded or used online. For orders that are not reviewed, the reason for rejection will be stated. Regardless of whether the review is passed, the system will give feedback information to the user and record the user's data applications and usage.

Based on the ocean data classification and organization method, the basic ocean data retrieval service module provides a total of six types of data, such as special survey data, business data, business-monitoring data, international business data, international cooperation and exchange data, and product data. According to the characteristics of each category of ocean data, the corresponding data query and retrieval conditions will be provided. Users can query the data related to their assigned permissions.

1. Special survey data retrieval

Special survey data primarily refer to a variety of ocean data obtained by special ocean survey projects. Its retrieval conditions include time (optional), projects, disciplines, and elements, which can be seen in the following. The result of the query is shown in the bottom of the sphere.

The retrieval result shows all the voyages under the hydrology—CTD, including the survey project name, the investigated organization names, the investigated code blocks, the investigated area name, the survey vessel name, the voyage number, the start date of the observations, the end date of the observation, and the station number information. Users can also view the corresponding voyage of the station data and metadata information. In the station list information, users can view all fields. When users click on a voyage, the corresponding position information will be drawn and rendered on the sphere. When users click on a station, it can be positioned on the sphere. Users can select interesting stations to add to favorites and thus form data orders for the submission of applications.

2. Business data retrieval

Business data primarily refer to data consisting of, e.g., stations, buoys, and ground-wave radar, which can be acquired continuously. Taking the station data as an example, the data retrieval includes two types of real-time data and archival data. Users can query real-time data and archive data by selecting a specific station name.

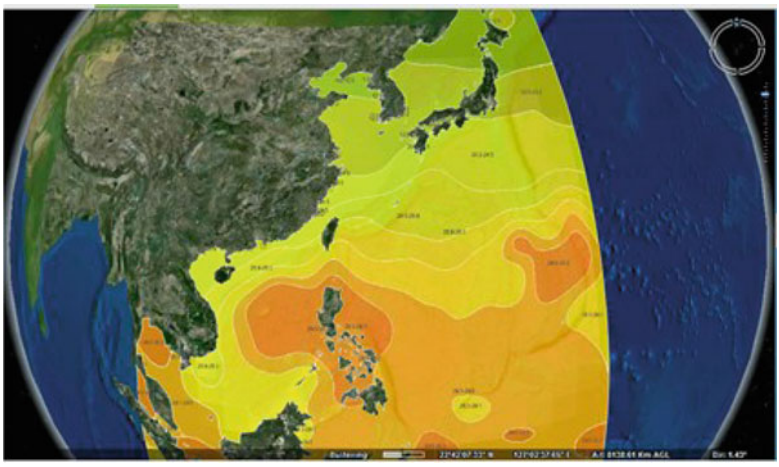
3. Data order processing and verification

In the favorite data list, the special basic data take the voyage number as collection granularity, the business data take the business means as collection granularity, and the international business and product data take the list of entries as collection granularity. After choosing the data of interest add to favorites, users can submit orders for the corresponding data, fill in the means of data use (including online and downloading), explain the demand and upload relevant certification

materials. Through ‘my data order’, users can view the status of the approval of the current application and can cancel orders that have been submitted but have not yet been processed. After the application is submitted, the platform administrator will review it. After the audit, if the order is for downloading data, then the download link is automatically downloaded for the user to download data. If the order is for online data use, users will be allocated to the virtual machine, and the application of the data will be stored in the virtual machine. Users log in to the virtual machine can directly process the data and create product in the virtual machine.

9.4.4 The Statistical Analysis Application

The goal of the statistical analysis module of the ocean environment is to draw the dynamic changes of the ocean temperature, salt, density, sound, flow, and higher ocean surface elements by means of analysis, the actual situation and the forecast data and in the form of contours, single points and single point curves and sections. The aim is to conduct an analysis to express the past, present, and future situation of the ocean environment. Users first select the modules, such as the ocean environment of conventional statistics, the reanalysis of the ocean environment, the live ocean environment, and the ocean environment forecasting, set parameters, such as the area, elements, time, visual expression, and layer depth, and then draw the corresponding graphics (A sample is shown in Fig. 9.3). In addition, the ocean environmental data product module provides users with reanalysis, live forecasting and other data retrieval and order services, the process of which is same as the ocean basic data research service.



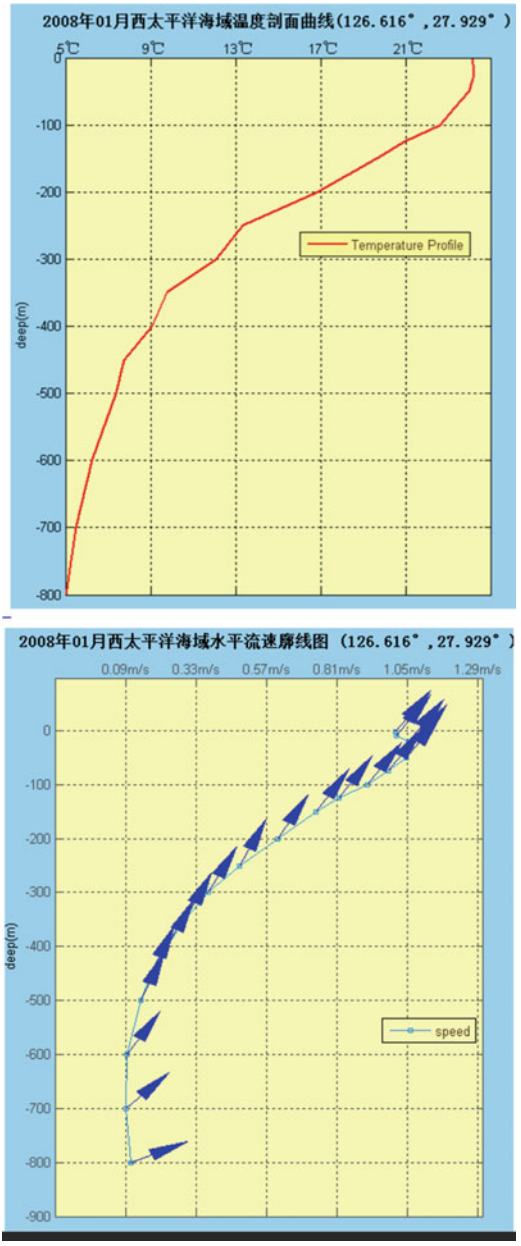


Fig. 9.3 The single point cross section mapping

9.4.5 Comprehensive Information Query

The comprehensive ocean information query management module is designed to provide the user with comprehensive management service support for existing integrated ocean management business requirements through bringing together a unified management of all types of ocean business outcome data, sorting, information displays and statistical analyses. The integrated management module primarily includes the use of ocean areas, island management, ocean economy, ocean environmental protection, ocean rights and interests, ocean science and technology, and ocean disaster reduction.

The ocean area project provides a query and statistical analysis of information about the ocean area project that is in use, including the name of the ocean project, the type of ocean, the user area, the ocean location, etc. The system provides the use of ocean projects and functional zoning and spatial positioning according to the time and the type of the ocean. At the same time, to gain an overall understanding of the use of ocean projects and functional zoning, the system provides a statistical analysis of the year and the coastal provinces and cities.

Island management provides census information on the comprehensive survey of the national island, the distribution of the island, and the name of the island. The island census includes the island name, the category, the province, the area, the offshore distance, the length of the coastline, the island classification, and the vegetation attribute information. The system provides information queries and spatial positioning functions according to provinces and cities. It also realized the function of statistical analysis for the island by the state and the province.

The module of maritime rights and interests provides an information query and spatial positioning function of the baseline point of the territorial ocean, the baseline of the territorial ocean, 12 nautical miles, 200 nautical miles, maritime borders, unilateral claim lines, islands and reefs of rights and interests and common

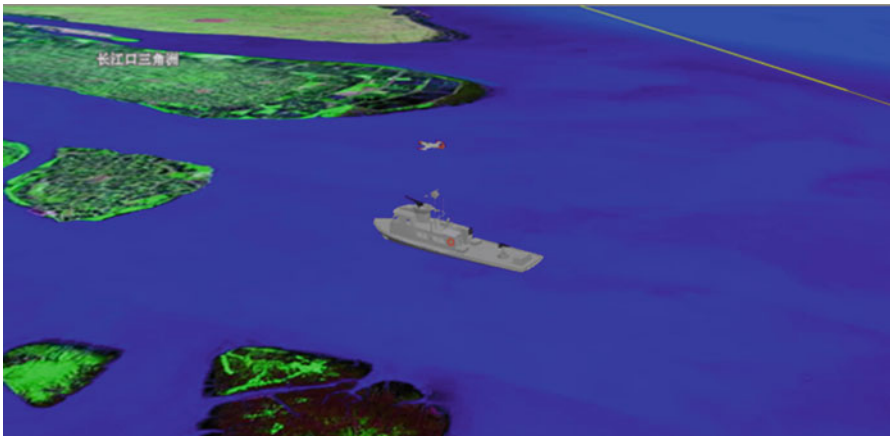
development areas, offshore blocks, military bases, agreement waters, air defense identification zones, the rights and interests of the event, foreign-related ocean surveys, oil and gas platforms, and dotted lines of the South China Ocean.

The module of ocean environmental protection provides an information query and spatial positioning function of environmental monitoring and carbon dioxide monitoring. The attribute information includes the monitoring name, the monitoring types, the monitoring tasks, regional monitoring, the monitoring mechanisms, the station numbers, and the monitoring year.

9.4.6 Thematic Application

Ocean thematic application services are designed to focus on hot spots and the specific needs of current ocean areas to provide in-depth information integration and analysis services for related topics by developing a dedicated application module through using massive amounts of data collected from the digital ocean, combined with spatial statistical analysis, visualization, integration and other means. At present, the system has completed the analysis of the interests of the surrounding the island of Diaoyu, the interests of the South China Ocean Islands and the integration of the scientific survey. The follow-up work will continue to carry out the special application of comprehensive assessments, such as rising ocean levels, typhoons, and tsunamis.

The ship-carrying scientific investigation, the application of the special subject, is the research and development of the ocean science investigation of the dynamics and operation of the ship. The system and digital ocean application service system share the same set of geographical background data. At present, the system has been installed and deployed in the ocean bureau system of 14 survey vessels. The main functions include ship monitoring, historical route mapping, task management, data services, sensitive areas, electronic charts and other functions.



9.4.7 Information Sharing and Exchange

China's digital ocean construction is a complex distributed system that covers the ocean bureau, coastal provinces and cities, the unit of the ocean bureau, and the scientific research institutes of the ocean. In addition to relying on the Digital Ocean master node, to provide users with data and information and to analyze application services, each node has its own characteristics and data. Determining how to develop a set of application service systems that support the management of unified linkage, the complementarity of macro and micro information, and the service focus on every aspect is important for the development of the digital ocean.

In this regard, the digital ocean application service system provides interactive services for nodes in three primary ways: data load released, interest points released, and the ability to display the results of the construction of the node web link service. Users can set access rights when publishing data services and interest points, including completely private node units and fully open and public nodes. The system determines the data services and interest points that are available to the user, based on the user's current permissions (super users, administrators, ordinary users) and registered units. Super users and administrators can view all publishing content, but only administrators can delete the published data service and interest points. Ordinary users can view their own publications or internal public and fully open data services and interest points but only to delete their own data services and interests. They have the authority to control the data and to ensure their security and reliability.

Data service publication provides users with a data interface, which supports 3D sphere integrated displays such as node vectors, terrain, remote sensing, 3D models, tilt camera and other data services. Users need to press the module's requirements for data production and services. The data need to be translated into *.fly format and to be released in IIS format. In the operation of this module, users are only required to publish an HTTP data service link. After the submission, the service will be added to the "layer management", which under the node directory corresponds to the user, and users can manage the service in the node data service list.

The interest point publishing service achieves the function of releasing interest points, which are defined by users. Node users can set the location of the interest points, the name of the interest points, and the access permissions of the interest points, among other properties, and then upload the information related to the points of interest (such as pictures and multimedia). The system will display the information on the interest points in the form of a list, which is released by users themselves, other users or all users. The system can be used to locate, edit, and delete interest points.

Acknowledgments The study is funded by the 908 Project of the State Oceanic Administration, China (No. 908-03-03-02). All members of the Research Team are gratefully acknowledged for their contributions to the work carried out in this chapter in recent years. In particular, special thanks are given to Linchong Kang, Chongjing Lv, Lili Song, Yi Wang for their contributions in the document.

References

Alexandrakis G, Poulos S, Petrakis S, Collins M (2011) The development of a Beach Vulnerability Index (BVI) for the assessment of erosion in the case of the North Cretan Coast (Aegean Ocean). Hell J Geosci 45:11–22

Brakenridge GR, Syvitski JPM, Overeem I, Higgins SA, Kettner AJ, Stewart-Moore JA, Westerhoff R (2012) Global mapping of storm surges and the assessment of coastal vulnerability. Nat Hazards (2013) 66:1295–1312. doi:10.1007/s11069-012-0317-z

Chang H (2001) Development of a Comprehensive Vulnerability Assessment model for local communities using GIS, in the department of city and regional planning. Cornell University, New York

Cutter SL (1996) Vulnerability to environmental hazards. Prog Hum Geogr 20:529–539

Cutter SL, Mitchell JT, Scott MS (2000) Revealing the vulnerability of people and places: a case study of Georgetown county, South Carolina. Ann Assoc Am Geogr 90(4):713–737

Cutter SL, Barnes L, Berry M et al (2008) A place-based model for understanding community resilience to natural disasters. Glob Environ Chang 18:598–606

Gao Y, Wang H, Liu GM, Sun XY, Fei XY, Wang PT, Lv TT, Xue ZS, He YW (2014) Risk assessment of tropical storm surges for coastal regions of China. J Geophys Res Atmos 119:5364–5374. doi:10.1002/2013JD021268

Jing Xiao, Markus Gerke, George Vosselman (2012) Building extraction from oblique airborne imagery based on robust facade detection. ISPRS J Photogramm Remote Sens 68(2012):56–68

Kuo Li, Guo Sheng Li (2011) Vulnerability assessment of storm surges in the coastal area of Guangdong province. Nat Hazards Earth Syst Sci 11:2003–2010

Kuo Li, Guo Sheng Li (2013) Risk assessment on storm surges in the coastal area of Guangdong Province. Nat Hazards (2013) 68:1129–1139. doi:10.1007/s11069-013-0682-2

Liang HY, Zou XQ (2005) Risk assessment on storm surge in the Haikou Bay. Acta Oceanol Sin 27(5):22–29

Lin N, Emanuel K, Oppenheimer M, Vanmarcke E (2012) Physically based assessment of hurricane surge threat under climate change. Nat Clim Chang 2:462–467. doi:10.1038/nclimate1389

Lubna R, Blaschke T, Zeil P (2010) Application of satelite derived information for disaster risk reduction: vulnerability assessment for southwest coast of Pakistan. In: Michel U, Civco DL (eds) Proceecling SPIE7831. Earth Resources and Environmental Remote Sensing/GIS Applications, Toulouse, doi:10.1117/12.864858

Nyaruhuma AP, Gerke M, Vosselman G, Mtalo EG (2012) Verification of 2D building outlines using oblique airborne images. ISPRS J Photogramm Remote Sens 71:62–75

Sheik Mujabar P, Chandrasekar N (2011) Coastal erosion hazard and vulnerability assessment for southern coastal Tamil Nadu of India by using remote sensing and GIS. Nat Hazards (2013) 69:1295–1314. doi:10.1007/s11069-011-9962-x

Shi PJ, Shuai JB, Chen WF et al (2010) Study on large-scale disaster risk assessment and risk transfer models. Int J Disaster Risk Sci 1(2):1–8

Taramelli A, Melelli L, Pasqui M, Sorichetta A (2010) Modelling risk hurricane elements in potentially affected areas by a GIS system, Geomatics. Nat Hazards Risk 1(4):349–373. doi:10.1080/19475705.2010.532972

UNDP (2004) Reducing disaster risk: a challenge for development. United Nations Development Programme, New York

Verhoeven G, Doneus M, Briese C, Vermeulen F (2012) Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. J Archaeol Sci 39(2012):2060–2070

Xu QW, Tang SD (1998) Research on the method of evaluating the economic loss caused by storm surge disaster. Mar Sci Bull 17(1):1–12

Yin BS, Wang T, Hou YJ et al (2001) Numerical study of the influence of waves and tide-surge interaction on waves in Bohai Ocean. Oceanologia et Limnologia Sinica 32(1):109–116

Yin BS, Xu YQ, Ren LC et al (2006) Risk assessment of overtopping dam under waves and surges in coastalareas of the Huanghe (Yellow) River delta. Oceanologia et Limnologia Sinica 37(5):457–463

Ying M, Zhang W, Yu H, Lu X, Feng J, Fan Y, Zhu Y, Chen D (2013) An overview of the China meteorological administration tropical cyclone database. J Atmos Ocean Technol 31:287–301. doi:10.1175/JTECH-D-12-00119.1