

Marine Big Data Science: Current Status and Prospects (Postprint)

Authors: Qian Chengcheng, Chen Ge

Date: 2023-03-19T00:00:00+00:00

Abstract

With the rapid advancement of marine observation and simulation capabilities, together with major breakthroughs in data science, modern marine science has experienced three major paradigm shifts: theory-driven, technology-driven, and data-driven. Marine big data has become an essential pathway for humanity's progression from understanding the ocean to governing it. This article analyzes the current development status across the entire upstream-midstream-downstream chain of marine big data acquisition, analysis, and application, and points out that tightly integrating the marine science and data science domains to effectively address challenges in scientific, technical, engineering, and humanities aspects of marine development is one of the core issues in marine scientific research. On this basis, it proposes the main directions and key technologies for the development of marine big data science in the next 5-10 years, which holds important practical significance and profound strategic importance for accurately grasping ocean conditions, enhancing marine forecasting service capabilities, and safeguarding national maritime rights and interests.

Full Text

Abstract

Modern oceanography has undergone three critical paradigm transformations—theory-driven, technology-driven, and data-dominant—propelled by rapid advancements in ocean observation and simulation capabilities alongside major breakthroughs in data science. Ocean big data has become an indispensable pathway for humanity's journey from understanding to managing the oceans, with its importance for societal development and national core competitiveness increasingly recognized worldwide. This paper analyzes the current development status across the entire chain of ocean big data, from acquisition and analysis to application. While big data can now be obtained from space-, air-, land-, and ocean-based platforms, breakthroughs are still needed for acquisition

in deep seas, extreme environments, and at high resolutions. In China, big data storage, management, and mining remain at a preliminary stage compared to other nations, although the country leads in data visualization. Many sectors are actively seeking to apply ocean big data and track the latest scientific discoveries; however, gaps sometimes persist between applications and upstream data acquisition and analysis.

Ocean big data exhibits two distinctive features: spatiotemporal coupling and geographic association. As multidimensional datasets with both temporal and spatial attributes, the increasing resolution and frequency of observations demand simultaneous analysis along both temporal and spatial axes, posing significant challenges for deep analysis given the diverse, high-dimensional factors involved. Unlike other big data that may be randomly distributed, ocean big data are geographically related—adjacent regions exhibit linear or nonlinear associations, forming modal characteristics across different spatiotemporal scales. Consequently, ocean big data acquisition, analysis, and application remain highly challenging. Effectively addressing scientific, technical, engineering, and humanities challenges in ocean development through the close integration of ocean science and data science has thus become a core issue in marine research.

Building upon this analysis, we propose five key research directions and enabling technologies for ocean big data science over the next 5–10 years: (1) investigating potential areas and methods for fusing ocean science with data science; (2) exploring observation and detection programs to support ocean big data development; (3) identifying oceanographic and information science trends suited to big data characteristics; (4) studying sharing mechanisms and collaborative innovation platforms for ocean big data; and (5) exploring emerging industry trends and technological requirements for ocean big data applications. Looking ahead, we aim to develop methods for acquiring high-resolution ocean big data under extreme conditions such as deep-sea trenches, create new analytical theories and techniques for multidimensional ocean big data analysis, establish open sharing platforms for ocean big data, and meet diverse needs in marine food safety, pollution and human health, maritime disasters, biodiversity, and the blue economy to facilitate the smooth establishment of new ocean big data industries. This paper makes a significant contribution to understanding the oceans and enhancing marine forecasting capabilities.

Keywords

Ocean big data, research status, future directions, key technologies

Author Biographies

QIAN Chengcheng is an engineer at the North China Sea Marine Forecasting Center of the State Oceanic Administration. She received her Ph.D. in Cartography and Geographic Information Systems from Ocean University of China (OUC) in 2015, completed a postdoctoral fellowship in Marine Science

at OUC, and spent one year as a visiting scholar at the University of California, Irvine. Her primary research focuses on ocean big data applications and marine remote sensing. She has led or participated in over ten projects, including the Qingdao Postdoctoral Research Fund, OUC Doctoral Innovation Research Fund, National Key R&D Program, and National Natural Science Foundation Key Special Projects. She has published 14 papers in major academic journals, including 12 SCI-indexed articles. E-mail: qianchengcheng@bhfj.gov.cn

CHEN Ge is a professor and doctoral supervisor at Ocean University of China, and Deputy Dean of the Institute for Advanced Marine Studies at OUC. He is a recipient of the Changjiang Scholar Award from the Chinese Ministry of Education and the National Science Fund for Distinguished Young Scholars. He has served on multiple national expert committees, including the Oceanography Discipline Evaluation Group of the State Council Academic Degrees Committee, the Earth Observation and Navigation Field Expert Group and Marine Environmental Monitoring Technology Theme Expert Group of the National “863” Program, the Secretariat of the Marine Science Teaching Guidance Committee of the Chinese Ministry of Education, the Information Science Committee of the Chinese Ministry of Education Science and Technology Committee, and is the Chief Scientist of the “Guanlan” marine science satellite. He has long been engaged in research and teaching on satellite ocean remote sensing and marine information technology, with recent research expanding into emerging interdisciplinary fields such as big data oceanography, achieving a series of original results with independent intellectual property rights. He has published over 220 papers and three monographs in academic journals spanning remote sensing, oceanography, atmospheric science, and information science. E-mail: gechen@ouc.edu.cn

References

15. Martin S. An Introduction to Ocean Remote Sensing. Cambridge: Cambridge University Press, 2004.
16. Cardenas A A, Manadhata P K, Rajan S P. Big data analytics for security. IEEE Security and Privacy Magazine, 2013, 11(11): 74-
17. Sue N. Big data: The harvard computers. Nature, 2008, 455: 36
18. Jiang X, Lin M, Zhang Y. Overview of the HY-2 satellite ground application system. China Engineering Science, 2014, 16(6): 4-12.
19. Dean J, Ghemawat S. MapReduce: simplified data processing on large clusters. Communications of the ACM—50th Anniversary, 2008, 51(1): 107-113.
20. Malewicz G, Austern M H, Bik A J C, et al. Pregel: a system for large-scale graph processing. SIGMOD’ 10 Proceedings of the 2010 ACM SIGMOD International Conference on Management of data. New York: ACM, 2010: 135-146.

21. Lorenz K Z. The comparative method in studying innate behavior pattern. In: *Physiological Mechanisms in Animal Behavior*. Oxford: Academic Press, 1950: 221-268.
22. Chen G. A novel scheme for identifying principal modes in geophysical variability with application to global precipitation. *Journal of Geophysical Research*, 2006, 111: D11103.
23. Chen G, Shen Z. Annual and semiannual cycles with oceanographic implications. *Journal of Climate*, 2016, 29: 37-59.
24. Chen G, Shen Z. Spatial-temporal seasonality: Extratropical spiral versus tropical phase lock. *Journal of Climate*, 2016, 29: 4021-4030.
25. Li D, Wang S, Li D. *Spatial Data Mining Theory and Application*. Beijing: Science Press, 2013: 10-13.
26. McLoughlin T, Laramee R S, Peikert R, et al. Over two decades of integration-based, geometric flow visualization. *Computer Graphics Forum*, 2010, 29: 1807-1829.
27. Weiskopf D. GPU-Based Interactive Visualization Techniques. [2018-07-20]. <http://www.scopus.com/inward/record.url?eid=2-s2.0-84892054283&partnerID=tZOtx3y1>.
28. Huang J, Pan Z, Chen G, et al. Image-space texture-based output-coherent surface flow visualization. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19: 476-1487.
29. Hansen C D, Johnson C R. *The Visualization Handbook*. Oxford: Academic Press, 2011.
30. Chen C K, Wang C, Ma K-L, et al. Static correlation visualization for large time-varying volume data. Hong Kong: 2011 IEEE Pacific Visualization Symposium, 2011: 27-34. DOI: 10.1109/PACIFICVIS.2011.5742369.
31. Wang C, Yu H, Ma K L. Application-driven compression for visualizing large-scale time-varying data. *IEEE Computer Graphics and Applications*, 2010, 30: 59-69.
32. Xie J. Fast uncertainty-driven large-scale volume feature extraction on desktop PCs. Chicago: 2015 IEEE 5th Symposium on Large Data Analysis and Visualization (LDAV), 2015: 17-24. DOI: 10.1109/LDAV.2015.7348067.
33. He Y, Su F, Du Y, et al. Web-based spatiotemporal visualization of marine environment data. *Chinese Journal of Oceanology and Limnology*, 2010, 28(5):1086-1094.
34. Zhang Y, Dong Z, Ma K L. Real-time volume rendering in dynamic lighting environments using precomputed photon mapping. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19: 1317-1330.

35. Wang C, Xie J, Li J, et al. Visualizing 3D atmospheric data massively distributed GPUs. Paris: 2014 IEEE 4th Symposium on Large Data Analysis and Visualization (LDAV), 2014: 3-10. DOI: 10.1109/LDAV.2014.7013198.
36. Yawen H, Fenzhen S, Yun D, et al. Web-based visualization of marine environment data. Beijing: 2010 18th International Conference on Geoinformatics, 2010, 2(1): 1-6. DOI: 10.1109/GEOINFORMATICS.2010.5567751.
37. Liu X, Shen H-W. Association Analysis for Visual Exploration of Multivariate Scientific Data Sets. IEEE Transactions on Visualization and Computer Graphics, 2016, 22: 955-964.
38. Woodring J, Petersen M, Schmeißer A, et al. In situ Eddy analysis in a high-resolution ocean climate model. IEEE Transactions on Visualization and Computer Graphics, 2016, 22: 857-866.
39. Wang B. A typical type of high-performance computation: earth system modeling. Physics, 2009, 38: 569-574.

Notes

Spatiotemporal database. [2018-07-20]. http://en.wikipedia.org/wiki/Spatiaotemporal_{database}.

AUV. [2018-07-20]. https://en.wikipedia.org/wiki/Autonomous_{underwater}_{vehicle}.

National Deep Sea Center. [2018-07-20]. <http://www.ndsc.org.cn/>.

Argo. [2018-07-20]. <http://www.argo.ucsd.edu/>.

Twitter Storm. [2018-07-20]. <http://www.oschina.net/p/twitter-storm>.

Digital Ocean. [2018-07-20]. <https://en.wikipedia.org/wiki/DigitalOcean>.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.