

August 2018

Big Data Science for Ocean: Present and Future

QIAN Chengcheng

*North China Sea Marine Forecasting Center of State Oceanic Administration, Qingdao 266061, China;
Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine
Science and Technology, Qingdao 266000, China*

See next page for additional authors

Recommended Citation

Chengcheng, QIAN and Ge, CHEN (2018) "Big Data Science for Ocean: Present and Future," *Bulletin of Chinese Academy of Sciences (Chinese Version)*: Vol. 33 : Iss. 8 , Article 18.

DOI: <https://doi.org/10.16418/j.issn.1000-3045.2018.08.018>

Available at: <https://bulletinofcas.researchcommons.org/journal/vol33/iss8/18>

This Article is brought to you for free and open access by Bulletin of Chinese Academy of Sciences (Chinese Version). It has been accepted for inclusion in Bulletin of Chinese Academy of Sciences (Chinese Version) by an authorized editor of Bulletin of Chinese Academy of Sciences (Chinese Version). For more information, please contact lcyang@cashq.ac.cn, yjwen@cashq.ac.cn.



Big Data Science for Ocean: Present and Future

Authors

QIAN Chengcheng and CHEN Ge

海洋大数据科学发展现状与展望

钱程程^{1,3} 陈戈^{2,3*}

1 国家海洋局北海预报中心 青岛 266061

2 中国海洋大学 信息科学与工程学院 青岛 266000

3 青岛海洋科学与技术国家实验室 区域海洋动力学与数值模拟功能实验室 青岛 266000

摘要 随着海洋观测、模拟手段的快速提升和数据科学的重大突破,现代海洋科学经历了理论牵引、技术驱动与数据主导三大范式变革,海洋大数据已成为人类从认识海洋到经略海洋的必经之路。文章从海洋大数据的获取、分析及应用,上、中、下游全链条分析了目前的发展现状,指出现今将海洋科学领域与数据科学领域紧密结合,有效应对海洋发展中科学、技术、工程、人文等方面的挑战,是海洋科研领域的核心问题之一。在此基础上,提出了未来5—10年海洋大数据科学发展的主要方向和关键技术,对于准确把握海洋状况、提升海洋预测服务能力、维护国家海洋权益具有重要的现实意义和深远的战略意义。

关键词 海洋大数据, 研究现状, 未来方向, 关键技术

DOI 10.16418/j.issn.1000-3045.2018.08.018

现代海洋科学的发展古往今来经历了3个阶段:①**理论牵引阶段**。该阶段的许多重大理论圆满地解释了海洋中的物理现象,长足推进了人们对海洋的认识。虽然该阶段缺少观测资料,但是发展的理论基本都具有里程碑的意义。②**观测牵引阶段**。该阶段各类观测手段逐渐发展起来,主要包括海洋调查船、浮标、潜水器、遥感以及Argo监测网等。③**数据牵引阶段**。随着观测技术手段的不断丰富,也带来了数据量的不断攀升,海洋科学迎来了第三阶段。从2008年开始,*Nature*、*Science*、*Economist*等杂志及Computing Community Consortium

(计算社区联盟)等组织将“大数据”引入到各个领域^[1-6]。“大数据”被定义为数据量增长速度快,用常规的数据工具无法在一定时间内进行采集、处理、存储和计算的数据集合,拥有数据量大(volume)、类型繁多(variety)、价值密度低(value)、速度快时效高(velocity)和在线式(online)五大特征^[7,8]。由于海洋数据的数据来源广泛、种类繁多,数据量已增至PB量级,时间分辨率跨越不同尺度,同时需要及时处理分析用于各类决策支撑,因此海洋数据已然成为“大数据”的典范。

*通讯作者

资助项目:青岛海洋科学与技术国家实验室“鳌山科技创新计划”(2018ASKJ01),中央高校基本科研基金(201762005)

修改稿收到日期:2018年8月20日

但是海洋大数据的独特性质,使得传统的理论基础、技术手段已逐渐暴露其弊端。海洋大数据有两个区别于其他数据的典型特征——时空耦合和地理关联。

(1) **时空耦合**。海洋大数据为同时拥有时间与空间属性的数据,即多维度数据。尤其随着观测技术的进一步发展,数据维度的采集分辨率与频率都越来越高。因此,数据分析过程需要同时从时间轴和空间轴两个维度进行分析,而在时间轴和空间轴上分析的因素又是多样的、高维的,这给大数据的分析带来了更大的挑战^[9]①。

(2) **地理关联**。海洋大数据不同于其他大数据的随机性与偶然性,由于其地理属性有着近邻效应,相邻区域空间位置关系存在线性或非线性的关联,从而组成了不同时空尺度的模态特征。

因此,在海洋大数据科学的发展过程中存在着诸多挑战。本文将从海洋大数据上、中、下游全链条论述海洋大数据科学的发展现状,并在此基础上提出未来5—10年海洋大数据科学发展的主要方向和关键技术。

1 海洋大数据获取现状

海洋大数据的获取手段主要得益于海洋观测技术的发展。目前海洋观测呈现出多元化、立体化、实时化的特征。从早期利用海洋调查船到浮标、潜水器、遥感的使用,再到海洋观测网Argo等的实现,无一不展现了人类不断探索海洋的决心与智慧。

(1) **海洋调查船**。这是一类专门从事海洋科学调查研究的船只,是搭载海洋仪器设备直接观测海洋、采集样品和研究海洋的工具。海洋调查船按其调查任务可分为综合调查船、专业调查船以及特种调查船^[10]。从世界上第一艘海洋调查船——“挑战者”号开始,已有100多年的历史。中国第一艘海洋调查船——“金

星”号始于1956年,目前中国已有共近50艘海洋调查船。目前,全球超过40个国家拥有海洋科考船,总数量超过500艘。

(2) **海洋浮标**。这是一类用于承载各类探测海洋和大气传感器的海上平台,是海洋立体监测系统的重要组成部分^[10]。根据浮标在海上所处位置不同,可分为锚定浮标、潜标、漂流浮标等。海洋锚定浮标最早出现于二战期间;20世纪70年代后期,随着计算机技术和卫星通信技术在浮标应用中的出现,使得浮标技术发展进入了飞跃期。海洋浮标在中国的开发研制始于20世纪60年代中期,90年代开始正式投入使用。目前,中国已经进入了海洋浮标监测的大国俱乐部^[11]。

(3) **潜水器**。又称为深潜器,是一种自带推动力的海洋考察设备——既能在水面行驶,又能在水下独立开展工作^[10]。1554年意大利人塔尔奇利亚发明的木质球形潜水器,对后来潜水器的研制产生了巨大影响。1717年英国人哈雷设计了第一个有实用价值的潜水器,此后直到20世纪60年代^②,人类对潜水器的研制主要致力于下潜深度的突破。“蛟龙”号是中国自行设计、自主集成研制的深海载人潜水器,目前下潜深度7062 m,成为世界上下潜能力最深的作业型载人潜水器^③。

(4) **海洋遥感**。该方法是利用传感器对海洋进行远距离非接触观测,以获取海洋景观和海洋要素的图像或数据资料;其发展历程大致可分为起步期(1939—1969年)、试验期(1970—1977年)、研究期(1978—1991年)、应用期(1992年至今)^[12]。20世纪90年代,遥感卫星开始大量发射;截至2012年底,在轨卫星数量为115颗,涉及超过30个空间机构^[13,14]。预测到2030年还会再有156颗卫星发射,届时总数将达到271颗。

(5) **海洋观测网络**。美国的Seaweb是世界上最早

① Spatiotemporal database. [2018-07-20]. http://en.wikipedia.org/wiki/Spatiotemporal_database.

② AUV. [2018-07-20]. https://en.wikipedia.org/wiki/Autonomous_underwater_vehicle.

③ 国家深海基地管理中心. [2018-07-20]. <http://www.ndsc.org.cn/>.

部署和应用的海洋观测网络；Argo 计划则是由美国海洋科学家于 1998 年倡导发起的一个监测全球海洋的大型网络，全世界几十个国家参与其中。截至 2018 年 7 月，布放在全球海洋中仍处于工作状态的 Argo 剖面浮标已达 3 762 个^④。未来，Argo 剖面浮标将增加到近 4 000 个；在维持现有 Argo 观测内容的基础上，新的 Argo 浮标观测范围将扩大到海面 2 000 m 以下甚至海底，同时携带安装生物、地球、化学等新型传感器。

在海洋大数据的获取方面，虽然已实现了基于空基-地基-海基的多元立体实时化发展，但是亟待突破深海、极端环境和高分辨率的大数据获取技术及平台的发展脉络，同时如何基于空间数据的时空耦合与地理关联特性，面向空间研究对象合理布设、高效利用观测手段成为数据获取阶段的挑战。

2 海洋大数据分析技术现状

海洋大数据从上游获取后，在中游主要涉及存储管理、挖掘及表达可视化等分析技术。

(1) **海洋大数据存储管理**。世界各主要海洋国家均有负责数据处理和管理的海洋数据中心。美国国家航空航天局（NASA）的地球观测中心建立了地球观测系统数据和信息系统，存储和管理全部数据，采用的是分布式开放的系统架构^[15]；欧洲航天局（ESA）也建立了采取基于任务的分布式存储的数据中心^[16]。中国目前海洋卫星遥感数据的存储采用由千兆交换机连接构成的 NAS（网络接入存储）三级存储体系，主要采用磁盘阵列加光盘存储的方式^[17]。国外的海洋大数据存储采用了逻辑上集中，物理上分散的分布式服务器集群存储架构；而国内的海洋大数据还属于地域上的集中式服务器存储——随着数据量的增长，在线存储资源有限，难以实现在线存储资源的动态扩展和灵活配置，离线数据获

取耗时，无法在线直接访问任意数据。

(2) **海洋大数据挖掘分析**。目前已有 MapReduce、Storm、StreamBase、Pregel 等先进的并行计算框架^[18-20]^⑤，且在各领域中得到广泛应用。海洋大数据在信息挖掘过程中也从传统的经验模态正交法（EOF）发展到了具有时空解耦特性的四维谐波提取法（4D-HEM）^[21-24]。但是由于海洋大数据的时空耦合及地理关联特性，导致传统的数据挖掘算法无法有效地进行时空解耦与地理分解，使得挖掘算法成为海洋大数据科学全链条运转环节中亟待改进与调整的重要屏障^[25]。

(3) **海洋大数据表达可视化**。利用科学可视化技术展示海洋数据以及更进一步地利用可视化分析技术挖掘时空数据规律，是建立从感知到认知的关键技术桥梁。海洋矢量场可视化算法主要有图表法、几何法^[26,27]、纹理法^[28]、拓扑法^[29]等。标量场可视化算法在大规模体绘制^[30-32]、实时光照^[33,34]、多变量提特征提取^[32,35]、二维时空可视化等方面都取得了重要成果^[36-39]。但是随着海洋数据体量的继续增大，对可视化表达方式、处理效能等方面都提出了非常高的要求，需要一方面尽可能真实地反映数据的特性，另一方面充分提供系统的承载能力和处理能力，提高数据的更新和绘制能力。

纵观国内外海洋大数据分析技术研究，中国在数据存储管理及挖掘方面仍处于跟跑阶段，但在可视化分析方面已实现并跑。

3 海洋大数据应用现状

海洋大数据的应用主要为社会经济发展及气候预测等提供决策支撑。目前，世界各国都在积极投入“数字海洋”的建设，并为进一步建设“智慧海洋”平台奠定基础，如美国和加拿大制定的“海王星”计划、日本的“ARANA”计划、非洲沿海 25 国的“非洲近海资源数

④ Argo. [2018-07-20]. <http://www.argo.ucsd.edu/>.

⑤ Twitter Storm. [2018-07-20]. <http://www.oschina.net/p/twitter-storm>.

⑥ Digital Ocean. [2018-07-20]. <https://cn.wikipedia.org/wiki/DigitalOcean>.

据和网络信息平台”以及中国的“iOcean”平台等^⑥。海洋大数据在气候预警决策支撑方面，主要是建立在高性能集群基础上的完备数值预报体系^[40]。例如：美国大气海洋局（NOAA）计划在2023年推出WoF（Warn-on-Forecast）系统。该系统可为美国及其临近海域提供精细化天气预报和灾害预警——美国本土计算网格大小精细至3—10 km，全球区域内网格精细至15 km；该系统的计算需求高达1万亿亿次。中国系列海洋卫星产品在赤潮/绿潮监测、海冰监测、渔业生产和水质调查等方面也得到了全面的业务化应用。其中，HY-2产品应用于中国与欧盟的数值模式预报及多源融合产品中，而高分辨率海面温度产品在马航失联客机海上搜救保障、极地大洋航线中提供了重要的支撑保障^[17]。

目前，海洋相关行业都在积极推动海洋大数据在行业中的应用，关注海洋科学领域的新发现和新发明，并推动产业化落地。然而，行业应用在大数据转型中毕竟是个新形态、新过程，可以借鉴和参考的经验不多，海洋大数据应用的行业落地过程还有不少问题需要解决。随着大数据相关技术的研究不断取得突破，传统行业如何重新审视自己的发展战略，积极拓展行业内部（上、下游）之间、行业之间的数据更好地融合与利用是海洋大数据应用中的一项挑战。

4 未来发展方向

围绕国家海洋发展战略，明确未来5—10年海洋大数据科学发展的方向，确定海洋科学领域应用大数据的关键技术瓶颈，提出推进海洋数据科学发展的关键步骤和重点支持领域，实现海洋数据从“数据大”困境到“大数据”时代的战略性转变。

（1）研究海洋科学与数据科学融合发展的主要方向与理论。分析海洋大数据的特点，结合海洋科学各领域的发展现状和趋势，探索海洋科学与数据科学融合发展的核心问题，明确未来5—10年海洋数据科学的发展方向；以海洋科研需求推动数据科研体系发展，建立

有效推动海洋科研的数据驱动方法。

（2）探索支撑海洋大数据发展的重点观测和探测计划。根据海洋科学发展的历史和现状、结合海洋科学发展趋势与国家海洋战略发展需求，分析研究支撑未来5—10年海洋大数据发展的重点观测和探测区域，从海、陆、空、天、时5个维度深入探索形成海洋大数据的关键基础和能力。

（3）研究适应大数据特点的海洋科学和信息科学发展趋势。针对海洋科学的综合与交叉学科特性，分析海洋大数据在物理、化学、生物、地质等主要海洋学科发展过程中的作用和影响，探索海洋大数据与各学科交叉融合过程中的关键技术瓶颈，以及云存储、物联网、人工智能、泛在计算、交互可视、混合现实等前沿信息技术在海洋中的应用前景，为构建面向现代海洋科学的大数据分析理论与大数据海洋学知识发现体系提供指导。

（4）研究海洋大数据的共享机制和协同创新平台建设需求。根据“海洋强国”“一带一路”“海陆统筹”和“军民融合”等国家战略，以及经济社会发展对海洋科学各领域的具体需求，分析海洋大数据的共享机制，研究建设海洋大数据协同创新平台的关键技术和建设规划，形成产、学、研、用的有机融合。

（5）探索应用海洋大数据的新兴产业发展趋势与科技需求。以大数据感知、计算、信息产品三大类数据服务为基础，围绕海洋科学、海洋健康与生物多样性、全球气候变化、海洋水产品食品安全、海洋污染与人类健康、海洋灾难与海事安全、蓝色经济等各个领域的需求，分析应用海洋大数据的新兴产业发展趋势，推进海洋科技与蓝色经济的深度融合。

5 结语

海洋大数据时代的到来，机遇与挑战并存。国际社会已经认识到海洋大数据科学对于人类社会发展的必要性，世界各国也意识到海洋大数据科学对国家核心竞

争力的提升具有重要标志性意义。海洋大数据科学应从上、中、下游逐步攻坚克难,真正将其发展成为保障“海洋强国”“一带一路”“海陆统筹”和“军民融合”等国家战略实施的重要支撑,以及维护中国国家与人民利益的重要保障。

参考文献

- 1 Sue N. Big data: The harvard computers. *Nature*, 2008, 455: 36-37.
- 2 Cukier K. Data, data everywhere: A special report on managing information. *Economist*, 2010, 2: 10-11.
- 3 Overpeck T J, Meehl A G, Bony S, et al. Climate data challenges in the 21st Century. *Science*, 2011, 331(11): 700-702.
- 4 Jonathan O T, Meehl G A, Bony S, et al. Climate data challenges in the 21st century. *Science*, 2011, 331: 700-702.
- 5 Manyika J, Chui M, Brown B, et al. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, 2011:1-137.
- 6 Agrawal D, Bernstein P, Bertino E, et al. Challenges and opportunities with big data. A Community White Paper Developed by Leading Researchers Across the United States, 2012. [2018-07-20]. <http://cra.org/ccc/docs/init/bigdatawhitepaper.pdf>.
- 7 Tankard C. Big data security. *Network Security*, 2012, (7): 5-8.
- 8 郭华东, 王力哲, 陈方, 等. 科学大数据与数字地球. *科学通报*, 2014, 59(12): 1047-1054.
- 10 吴朝晖, 陈华钧, 杨建华. 空间大数据信息基础设施. 杭州: 浙江大学出版社, 2013: 9-10.
- 11 于志刚. 海洋技术. 北京: 海洋出版社, 2009.
- 12 戴洪磊, 牟乃夏, 王春玉, 等. 我国海洋浮标发展现状及趋势. *气象水文海洋仪器*, 2014, 31(2): 118-121.
- 13 林明森, 张有广, 袁欣哲. 海洋遥感卫星发展历程和趋势展望. *海洋学报*, 2015, 37(1): 1-10.
- 14 冯士筌, 李凤歧, 李少菁. 海洋科学导论. 北京: 高等教育出版社, 1999.
- 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): 74-76.
- 17 Alverson K, Baker D J. Taking the pulse of the oceans. *Science*, 2006, 314: 1657.
- 18 蒋兴伟, 林明森, 张有广. HY-2卫星地面应用系统综述. *中国工程科学*, 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 Pan E. Stream base announces hadoop integration to provide real-time analytics for big data. [2018-07-20]. <https://searchbusinessanalytics.techtarget.com/feature/Streaming-data-systems-take-big-data-analytics-into-real-time-realm>.
- 21 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.
- 22 Lorenz K Z. The comparative method in studying innate behavior pattern. In: *Physiological Mechanisms in Animal Behavior*. Oxford: Academic Press, 1950: 221-268.
- 23 Chen G. A novel scheme for identifying principal modes in geophysical variability with application to global precipitation. *Journal of Geophysical Research*, 2006, 111: D11103.
- 24 Chen G, Wang X. Vertical structure of upper-ocean seasonality: Annual and semiannual cycles with oceanographic implications. *Journal of Climate*, 2016, 29: 37-59.
- 25 Chen G, Wang X, Qian C. Vertical structure of upper-ocean seasonality: Extratropical spiral versus tropical phase lock. *Journal of Climate*, 2016, 29: 4021-4030.
- 26 李德仁, 王树良, 李德毅. 空间数据挖掘理论与应用. 北京: 科学出版社, 2013: 10-13.

- 27 McLoughlin T, Laramée R S, Peikert R, et al. Over two decades of integration-based, geometric flow visualization. *Computer Graphics Forum*, 2010, 29: 1807-1829.
- 28 Weiskopf D. GPU-Based Interactive Visualization Techniques. [2018-07-20]. <http://www.scopus.com/inward/record.url?eid=2-s2.0-84892054283&partnerID=tZOTx3yl>.
- 29 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.
- 30 Hansen C D, Johnson C R. *Visualization Handbook*. Netherlands: Academic Press, 2011.
- 31 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.
- 32 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.
- 33 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.
- 34 Zheng L, Chaudhari A J, Badawi R D, et al. Using global illumination in volume visualization of rheumatoid arthritis CT data. *IEEE Computer Graphics and Applications* 2014, 34: 16-23.
- 35 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.
- 36 Xie H, Maz K L. Visualizing large 3D geodesic grid data with massively distributed GPUs. Paris: 2014 IEEE 4th Symposium on Large Data Analysis and Visualization (LDAV), 2014: 3-10. DOI: 10.1109/LDAV.2014.7013198.
- 37 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..
- 38 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.
- 39 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.
- 40 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.
- 41 Wang B. A typical type of high-performance computation: earth system modeling. *Physics*, 2009, 38: 569-574.

Big Data Science for Ocean: Present and Future

QIAN Chengcheng^{1,3} CHEN Ge^{2,3*}

(1 North China Sea Marine Forecasting Center of State Oceanic Administration, Qingdao 266061, China;

2 College of Information Science and Engineering, Ocean University of China, Qingdao 266000, China;

3 Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266000, China)

Abstract With major improvements in ocean observations and modeling, as well as in data science development, current oceanography has gone through three critical transformations: theory-driven, technology-driven, and data-dominant. Oceanography big data are indispensable in humankind's journey to better learn the seas. The importance of big data in society development and national core competence is increasingly recognized by all countries. This paper addresses the overall development of big data acquiring, analysis, and application. Big data now can be obtained from platforms based in the space, the air, the land, and the ocean. However, breakthroughs are needed in such conditions as deep seas, extreme environments, and where high resolution is required. Compared to other nations, big data storage, management, and mining in China are relatively preliminary. On the other hand, we are leading in data visualization. Many areas are seeking for the application of ocean big data and paying attention to the latest discoveries. Still, connections between the applications and the upstream data acquiring and analyses are sometimes lost. Ocean big data have such features as being spatial-temporal coupled and geographically related when it comes to the fields of marine food safety, ocean pollution and mankind health, marine accidents, marine biodiversity, blue economy, etc. It is a general trend to apply ocean big data in the newly established fields. First, being spatial-temporal coupled is one of the most important features for air-land-sea data sets, which are also known as multi-dimensional data. As the observational techniques developing increasingly fast, it is getting more critical to get as high resolution data as possible in every dimension. Therefore, it is necessary to start analyzing the data sets from both temporal and spatial aspects, while this could pose a serious challenge to the ocean big data deep analyses since multiple factors should be considered in time and space. Another character of air-land-sea data is that they are geographically related. Different from being randomly distributed like the big data in other areas, ocean big data are all influenced by adjacent pixels. Since pixels sitting close to each other typically has linear or non-linear relationships with each other, the models on different temporal-spatial scales are presented in different ways. As a result, it is still very challenging in terms of ocean big data obtaining, analysis and application. Therefore, it becomes a critical issue which needs to be addressed by closely combining oceanography and data field in order to deal with the scientific or technical or engineering or even humanity challenges. This study proposes the possible research topics and key techniques in ocean big data field in 5 to 10 years: (1) investigating the potential areas and methods of the fusion of ocean science and big data field; (2) exploring the possible observation plans in support of ocean big data development; (3) determining the oceanography and information development trends which can fit the most features of big data; (4) studying the techniques and platforms of ocean big data sharing; and (5) exploring the development and technical requirements of possible new areas of ocean big data applications. In the future, step by step, we hope to keep investigating the methods of obtaining ocean big data with high resolution in extreme conditions such as deep sea trenches, and to find new analyses theories and techniques to make possible breakthroughs in terms of the multi-dimensional analyses of ocean big data. We will also try to realize the establishment of a new platform where ocean big data can be freely shared. Another goal is to satisfy the different needs in areas like marine food safety, marine pollution and society health, marine accidents, biodiversity, etc. and to ensure the new ocean big data based industries being smoothly established. This paper will be a significant contribution in understanding the oceans and improving our marine forecast abilities.

Keywords ocean big data, research actuality, future direction, key technologies

*Corresponding author



钱程程 国家海洋局北海预报中心工程师，中国海洋大学地图学与地理信息系统专业博士，海洋科学博士后，曾赴美国加州大学欧文分校联合培养1年。主要研究方向为海洋大数据应用与海洋遥感应用。曾主持或参与过青岛市博士后研究基金项目、中国海洋大学博士创新研究基金项目、国家重大研发计划、国家自然科学基金重点专项等10余项项目；在国内外重要学术刊物上发表发表论文14篇，其中SCI论文12篇。E-mail: qianchengcheng@bhfj.gov.cn

QIAN Chengcheng Attended the Ocean University of China (OUC) where she obtained her Ph.D. degree in Cartography and Geography Information System in 2015. After graduation she was a Postdoctoral Fellow at OUC (2015–2017). Since 2017, she has been an Engineer of Ocean Big Data Application at the North China Sea Marine Forecasting Center of State Oceanic Administration. As the leader or participant she has participated in more than 10 projects, such as China Postdoctoral Science Foundation, Doctoral Candidate Creative Foundation of Ocean University of China, and National Natural Science Foundation of China. She has published 14 papers in internationally recognized journals, which 12 of them are SCI indexed. E-mail: qianchengcheng@bhfj.gov.cn



陈戈 中国海洋大学高等海洋研究院副院长，教授，博士生导师；教育部“长江学者奖励计划”特聘教授，“国家杰出青年科学基金”获得者。先后任国务院学位委员会“海洋学科评议组”、国家“863”计划“地球观测与导航领域专家组”“海洋环境监测技术主题专家组”成员，教育部“海洋科学类专业教学指导委员会”秘书长，教育部“科学技术委员会信息学部”委员，“观澜号”海洋科学卫星首席科学家。长期从事卫星海洋遥感与海洋信息技术的科研和教学工作，近年来的研究方向拓展到大数据海洋学等新兴交叉领域，取得了一系列具有自主知识产权的原创性成果。在遥感、海洋、大气和信息领域学术刊物上发表论文220余篇，出版专著3部。E-mail: gechen@ouc.edu.cn

CHEN Ge Attended the Ocean University of China (OUC) where he obtained his B.S. degree in Marine Physics in 1988, M.S. degree in Satellite Oceanography in 1990, and Ph.D. degree in Physical Oceanography in 1993. After graduation he was a Postdoctoral Fellow at IFREMER (The French Research Institute for the Exploitation of the Sea), France (1994–1996). Since 1997, he has been a Professor of Satellite Oceanography and Meteorology at OUC. He served as the Executive Secretary of the International Pan Ocean Remote Sensing Conference (PORSEC) Association during 1998 and 2002. In 2001 he received the National Natural Science Funding for Outstanding Young Scientists by the National Natural Science Foundation of China, and became the Chair Professor of Cheung Kong Scholars Program nominated by the Chinese Ministry of Education. He is the Deputy Dean of the Institute for Advanced Marine Sciences at OUC, and was the Dean of the College of Information Science and Engineering during 2004 and 2017. Prof. Chen is the author of more than 100 peer-reviewed scientific papers published in internationally recognized journals. Prof. Chen is a member of the Expert Committee on Ocean Technology for the National High Technology Program (the “863” Program) of China nominated by the Chinese Ministry of Science and Technology. His current research interests include satellite remote sensing of the ocean and big data oceanography. E-mail: gechen@ouc.edu.cn

■ 责任编辑：岳凌生