

海洋环境大数据系统建设研究进展

张彭辉¹, 杜建平¹, 吴帅², 程文华¹, 侯健¹, 孔祥峰³, 张丽³

(1. 中国船舶集团有限公司第七二五研究所海洋腐蚀与防护重点实验室 青岛 266101; 2. 西南技术工程研究所 重庆 400039;
3. 齐鲁工业大学(山东省科学院) 山东省科学院海洋仪器仪表研究所 青岛 266061)

摘要:海洋环境大数据是诸多学科研究领域的基础,也是实施海洋强国战略和建设海洋强国奋斗目标的先决条件。文章从全球海洋环境观测计划及系统出发,探讨现阶段全球海洋环境大数据观测、国家数据库建设、数据质控方法以及数据应用现状,阐明了海洋环境大数据在国家发展中的重要性,客观地分析了目前海洋环境大数据发展在基础设施、数据处理及数据质控能力建设等方面面临的关键问题与挑战,以期将来海洋环境大数据在多尺度、多学科、多行业的深入发展与深度融合提供可靠的资料与建议。

关键词:海洋环境监测;大数据;数据质控;数据应用

中图分类号:P76;TP31

文献标志码:A

文章编号:1005-9857(2021)11-0104-09

Research Progress on the Construction of Marine Environment Big Data System

ZHANG Penghui¹, DU Jianping¹, WU Shuai², CHENG Wenhua¹, HOU Jian¹,
KONG Xiangfeng³, ZHANG Li³

(1. State key laboratory for marine corrosion and protection, Luoyang ship material research institute, Qingdao 266101, China; 2. Southwest Technology and Engineering Research Institute, Chongqing 400039, China; 3. Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences), Qingdao 266061, China)

Abstract: Big data on the marine environment is the foundation of many disciplines and research fields, and it is also a prerequisite for implementing the strategy of a maritime power and the goal of building a maritime power. Starting from the global marine environment observation plan and system, this paper discussed the current global marine environment big data observation, national database construction, data quality control methods and data application status, clarified the importance of marine environment big data in national development, and objectively analyzed the key issues and challenges on capacity building of infrastructure, data processing and quality control faced by the current development of marine environmental big data, to provide reliable in-

收稿日期:2020-11-26;修订日期:2021-09-27

基金项目:西南技术工程研究所合作基金项目(WZDW592019020201);山东省自然科学基金(ZR2019BD044);山东省科学院国际合作项目(2019GHPY15)。

作者简介:张彭辉,工程师,硕士研究生,研究方向为环境试验

通信作者:孔祥峰,高级工程师,博士,研究方向为海洋环境监测技术

formation and suggestions for the in-depth development and in-depth integration of marine environmental big data in multi-scale, multi-disciplinary, and multi-industry in the future.

Keywords: Marine environment monitoring, Big data, Data quality control, Data application

0 引言

“大数据”这一概念起源于 20 世纪 80 年代,最早被提出用来描述容量非常大的数据集,美国科学家 John R. Mashey 在“Big Data and the Next Wave of InfraStress”的演讲中即提到这一概念^[1]。随着摩尔定律的应用与技术的发展,大数据不仅是指简单的数据集,而是通过一系列特征准确定位的专业术语,即“4V”特征:大数量(Volume)、多样性(Variety)、高速度(Velocity)和高价值(Value)^[2]。海洋环境观测在海洋探测中起着至关重要的作用。随着信息技术的指数增长及海洋观测手段的日益发展,海洋科学已正式进入大数据时代。海洋观测数据的获取通常建立在多时空尺度、多平台的采集与分析上,具有数据量大、维度广、类型多样、测量持续、利用潜力大的特点,是一种典型的大数据,即海洋环境大数据。

作为海洋大国,我国大陆海岸线长达 1.8 万 km 余,岛屿岸线长达 1.4 万 km 余,海岸线总长度超过 3.2 万 km,有 300 万 km² 余的管辖海域,海洋空间范围广阔,海洋环境复杂。随着我国海洋监测立体网络的建立与发展,“908 近海海洋综合调查与评价”等专项的顺利实施,全球海洋 Argo 系统的布放与观测,多种监测平台与高端传感器的开发与应用,海洋环境数据呈现了暴发式增长。海洋环境大数据是生物科学、地球科学、海洋和大气科学等诸多研究领域的基础,也是维护国家海洋权益、保护海洋生态环境、开发海洋资源、发展海洋经济、实施海洋强国战略、建设海洋强国奋斗目标的先决条件。与常规大数据不同,海洋环境大数据因其独特的地理特性,呈现了与环境属性、时空差异等因素的高度耦合性^[3],在大数据时代的背景下,如何有效利用海洋环境大数据的特点挖掘更深层的科学信息,解决大数据存储、算法突破、组织分析方法、支持决策应用系统等关键技术问题,为未来海洋领域发展、海洋管理决策、全球海洋预报、灾害预警、海洋生产开发等奠定坚实基础,仍是海洋环境大数据发展过程

中需要攻克的重点以及难点。

本研究从全球海洋环境观测计划及系统出发,探讨了现阶段全球海洋环境大数据观测、国家数据库建设、数据质控方法以及数据应用现状,阐明了海洋环境大数据在国家发展中的重要性,客观分析了目前海洋环境大数据发展面临的关键问题与挑战,以期为将来海洋环境大数据在多尺度、多学科、多行业的深入发展与深度融合提供可靠的资料与建议。

1 海洋环境大数据的获取

海洋环境大数据包括海洋气象、水文、化学以及海底地形地貌等多个方面,其获取需要长期计划和持续发展的观测技术的稳定支持。在全球一体化的大背景下,成立于 1991 年的全球海洋观测系统(Global Ocean Observing System, GOOS)通过政府间合作,开展地转海洋学实时观测阵(Array for Real-time Geostrophic Oceanography, Argo)计划、数据浮标合作组(Data Buoy Cooperation Panel, DBCP)项目、系泊与船基时间序列(OceanSITES)项目、全球海洋船基水动力学调查项目(the Global Ocean Ship-based Hydrographic Investigation Program, GO-SHIP)、全球海平面观测系统(the Global Sea Level Observing System,GLOSS)以及船观测队(Ship Observation Team, SOT)等项目,对全球海洋水文、气象、化学和生物等多类关键海洋变量(Essential Ocean Variables, EOVs)进行观测^[4]。南极科学委员会(Scientific Committee on Antarctic Research, SCAR)组织实施的南大洋观测系统计划(Southern Ocean Observing System, SOOS)、北极理事会(Arctic Committee, AC)和国际北极科学委员会(International Arctic Scientific Committee, IASC)联合开展北极持续观测网计划(Sustaining Arctic Observation Net, SAON)开始对南北两极海洋环境要素进行长期观测^[5]。美国和加拿大于 1998 年和 1999 年启动海王星(NEPTUNE)计划,即东太平洋时间系列海底网络实验(North-East

Pacific Time-series Undersea Networked Experiments),旨在建立区域性、长期实时的交互式海洋观测平台^[6]。此外,多种深海观测计划,如热带海洋与全球大气实验计划(TOGA)、综合大洋钻探计划(IODP)以及国际大洋中脊计划(InterRidge)也正在开展当中^[7]。上述计划的实施,将为海洋环境大数据的获取提供便利条件。

海洋环境大数据的获取技术包括调查船调查、浮标潜标台站等海洋实地观测技术,以遥感为主的卫星观测技术以及多学科融合为基础的海洋观测网络。调查船调查是最早使用的获取原位海洋环境数据的方式之一,包括全球海洋观测系统的 GO-SHIP 等在内的多个项目,对全球海洋近岸与大洋区域水体环境开展数百次海洋调查。目前,相关航次公开发布 CTD 数据包括北极相关航次 94 个,调查时间从 1972 年 7 月至 2015 年 9 月;大西洋相关航次 116 个,调查时间从 1972 年 7 月至 2020 年 3 月;太平洋相关航次与时间观测连续观测结果 657 个,调查时间从 1973 年 8 月至 2020 年 3 月;印度洋相关航次 96 个,调查时间从 1977 年 12 月至 2019 年 5 月;南大洋相关航次 193 个,调查时间从 1972 年 7 月至 2019 年 5 月(<https://cchdo.ucsd.edu/>)。1958—1960 年以及 2004—2009 年,我国先后两次开展了全国海洋综合调查,编绘海洋物理、海洋化学、海洋生物以及海洋地质地貌图集和图志等,摸清我国近海海洋资源储备情况,为我国近岸海域海洋大数据提供宝贵资料^[8-9]。

浮标、潜标和观测站观测也是获取原位海洋环境数据的重要手段。作为全球海洋观测系统的一部分的 Argo 计划,利用剖面浮标为全球海洋剖面水温和盐度数据的实时获得提供了重要的支撑^[10]。截至 2020 年 11 月 5 日,布放在全球海洋中仍处于工作状态的 Argo 剖面浮标已达 3 717 个。同时,随着海洋技术的发展以及深海观测需求的增大,Argo 剖面浮标的探测深度已经由之前的 0~2 000 m 拓展到 0~6 000 m^[11]。通过集成生物化学传感器,在获得剖面温度和盐度数据的同时,浮标还可实时观测悬浮颗粒、光强、pH、硝酸盐浓度、叶绿素 a 浓度以及溶解氧浓度等多种环境数据,为多参数环境大数据的获得提供了极为便利的条件^[4,11]。此外,同样属于全球海洋

观测系统的 OceanSITES、DBCP 和 GLOSS 等项目提供的由台站、系泊浮标、潜标和卫星追踪漂流浮标观测获得的气象、物理海洋和生物化学等多种参数数据,也是海洋环境大数据的重要来源^[4]。

自美国 1997 年发射第一颗专用海洋水色卫星 Seastar 以来,遥感技术在海洋研究中的应用蓬勃发展。通过卫星遥感,可以估算全球初级生产力、观测海水化学参数(如盐度等)和上层水体动力学过程(如海平面高度变化等),为海洋环境大数据提供了丰富资料^[12]。

海底观测网络是通过综合运用水下机器人、海床光纤电缆等技术手段,实现全天候、原位、长期、连续、实时、高效的新型海洋观测平台。2016 年 6 月 6 日,美国大型海洋观测计划(Ocean Observatories Initiative, OOI)正式投入使用,83 个实验平台将投放大西洋和太平洋的 7 个观测阵列中,用来观测海表至海底的物理、化学、地质和生物过程^[13]。此外,加拿大海底观测网(Ocean Networks Canada, ONC)、欧洲多学科海底及水体观测系统(European Multidisciplinary Seafloor and Water-Column Observatory, EM-SO)、日本地震和海啸海底观测密集网络(DONET, DONET2)和日本海沟海底地震海啸观测网(S-net)以及我国的南海海底观测网试验系统,均已投入建设或开始使用^[7]。美国 OOI 系统的数据向公众免费开放,有效推动了海洋大数据研究的进步。

2 海洋数据类型

目前,对于浩瀚庞大的海洋来说,已经进入大数据时代,并且形成了范围较广的海洋监测体系,收集了大量的海洋自然科学数据,在传感器技术、海洋观测技术、通信技术的快速发展下,海洋环境观测数据的采集频率不断增加,产生了相应的应用和产品。根据海洋数据的监测方法不同和表达方式的不同,数据分为现场观测监测数据、海洋遥感数据、海洋再分析产品数据、海洋模拟数据等^[14]。

海洋实测调查包括船基观测、定点观测和移动观测等观测方式,不同观测方式采集的数据不同,例如船基观测的数据采集主要包括海洋气象、物理海洋、海洋物理、海洋化学、海洋生物等;定点观测根据平台的不同,又分为岸基雷达站、河口水文站、海洋气象站、验潮站、浮标、潜标等^[15]。海洋移动观测能够覆

盖更大的区域,包括水上、水下移动观测平台,能够提供水下几千米的温度、盐度、深度等^[16]。

海洋遥感数据包括卫星遥感数据和航空遥感数据,其中航空遥感又可分为有人机航空遥感和无人机航空遥感^[17]。航空遥感由于监测范围广、监测频次高、监测数据连续,适用于重点区域的高精度监测,如近海海洋调查、海岸带制图、资源勘测、海洋动态监管、海洋突发情况应急响应、海洋资源环境监测等。此外,海洋遥感数据定量化应用的不断深入和个性化服务的不断完善,也是未来海洋数据发展的重要方向^[18]。

海洋数值模拟是以现实海洋为基本物理背景,以计算机为载体,按照物理规律,建立数学模型,从而对海洋状态(包括海温、盐度、海流、海浪、潮汐等要素)进行模拟,参数化、定量化地描述海洋的具体状况。国内外比较常用的数值模拟产品网站有:中国科学院海洋研究所海洋大数据中心、国家海洋信息中心、国家极地科学数据中心、普林斯顿海洋模型(Princeton Ocean Model, POM)、非结构有限元近海海洋模型(AnUnstructured Grid, Finite-Volume Coastal Ocean Model, FVCOM)、汉堡陆架海洋模型(Hamburg Shelf Ocean Model, HAM-SOM)等^[19]。数据分析是大数据价值链中最后也是最重要的一个阶段,其目的是提取潜在的有用价值,提供建议或决策,有效地处理和利用现有观测资料,重建高质量、长时间序列和高分辨率的海洋资料集,最终形成海洋再分析产品。目前,国内外科研机构先后公开发布了数十个时间序列长、要素较为完整、统一协调的全球海洋再分析数据产品。包括美国的全球海洋再分析产品,欧洲中期天气预报中心的全球海洋再分析产品,法国全球海洋再分析产品,日本的全球海洋再分析产品,中国的全球海洋再分析产品^[20]。这些海洋再分析产品时间跨度不同,时间分辨率也不同,变量一般包括盐度、温度、经纬向海流速度和海表面高度等,其中日本的MOVE-G2 海洋再分析产品还提供了一些生物变量等。海洋再分析产品已经在诸多研究领域得到广泛应用,例如海洋对气候变化的响应、大洋环流与温盐结构、海洋动力学过程、监测方案的设计、数值预报和模式评估等。

3 海洋数据质控方法

随着海洋环境监测数据种类和量级的不断增加,数据质量控制的重要性日益突出^[21]。海洋环境监测数据的质量,将直接影响对海洋环境的综合评估、生态保护、灾害预警、管理决策和资源开发等。海洋数据的质量控制,决定了海洋数据的准确性和可靠性,具有重要的科学意义和社会价值^[22-24]。目前采用的质量控制方法主要分为两类:一类是根据数据的基本信息进行合理性辨别,如合理性检验和一致性检验等;另一类是基于数据处理算法进行质量管理,包括相关性检验、异常值检验和统计学习检验等方法(图 1)。

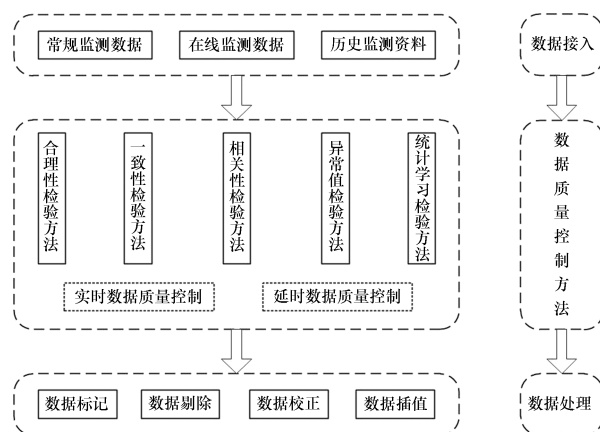


图 1 海洋监测数据质量控制

3.1 合理性检验方法

合理性校验主要是对数据基本信息进行初步检验,判断其数值是否合理,如数据的日期时间、位置坐标、记录格式、观测要素等^[25]。在海洋数据的质量管理中,数据日期不能超前于当前日期,年、月、日、时、分、秒均应在其规定进制内;数据位置的经度在 $-180^{\circ}\sim 180^{\circ}$,纬度在 $-90^{\circ}\sim 90^{\circ}$;数据记录的起始位置、长度、类型等要素,均满足规定要求;观测要素应为已知观测范围内的要素等^[26]。

3.2 一致性检验方法

海洋监测数据的一致性检验主要包括值域一致性检验和逻辑一致性检验。值域一致性检验是根据每个数据参数在该环境中的经验值域范围对数据进行检验,判断其是否在经验值域范围内^[27]。经验值域范围可以根据被监测数据参数的理化性质、最低检

出限、最大污染等级、时间空间分布规律、历史统计数据和相关文献资料等进行划定。逻辑一致性检验是根据某些检测参数间存在的逻辑关系进行质量控制,包括同一参数在不同形态下的关系(如总氮和无机氮、有机氮之间的关系)、同一物质的不同参数之间的关系(如有机物在 COD 的锰法检测中氧化程度低于在铬法检测中的氧化程度)、不同参数之间的内在关系(如 pH 值和总碱度的一致性关系)等^[28]。

3.3 相关性检验方法

相关性检验是根据海洋观测数据各元素之间的自相关或互相关关系来判断数据的质量。如针对某一参数,一天内数据的最大值应高于其平均值;两个相邻时间或位置间的观测值,其差值应在一定范围内;海水的温度、盐度数据与其密度数据之间的关系等^[29]。如:罗冬莲^[30]研究得出深沪湾溶解氧与浮游植物数量呈二次抛物线关系,与悬浮物呈线性负相关关系;戴文娟等^[31]利用 Aprior 算法对滨海观测数据进行了相关性检验,检测出东海大桥海洋站的一批温度、盐度和波浪数据不符合相关性检验,系机器故障导致,应作缺测处理。

3.4 异常值检验方法

在海洋环境数据的监测过程中,由于人为、仪器、通信等方面问题造成的错误、误差或噪声,会体现异于正常数据分布规律的显著性差别,该类数据视为异常值,应通过数据质量控制予以标记或处理。海洋数据质控过程中用到的异常值检验方法有 Grubbs 检验、Dixon 检验、莱特准则检验、局地检验、K-means 检验等。在实际应用中,可以根据监测数据的特征选择合适的检验方法,或进行方法的结合、改进,以对数据进行更好的质量控制^[32]。如:郑琳等^[33]分析数值分布较大的营养盐要素适合选用 Grubbs 检验,而针对数值差异较小的温度、pH 和 COD 等参数的数据,选用 Dixon 检验能得到更好的效果;刘首华等^[34]通过结合 Grubbs 检验和局地异常值检验,建立了一种针对海洋浮标波高监测数据异常值的质控方法;蒋华等^[35-36]对 K-means 检验算法进行了改进,以提高海洋数据异常值的检测率。

3.5 统计学习检验方法

在传统海洋数据质量控制中,有些数据需利用

专家经验进行人工检验。随着智能学习算法的发展,学者们开始采用统计学习或智能学习等相对复杂的算法来进行数据质量控制^[37]。如:李飞等^[38]探索引入统计学习算法(Support Vector Domain Description, SVDD),利用历史观测数据进行训练,建立可用于海洋站多要素数据质量控制的模型;王振华等^[39]引入梯度模糊数抽样检验模型,解决具有不确定质量参数的海洋数据质量检验问题;陈括等^[40]将粗集理论引入海洋数据质量检验中,用于对不同质量检验方案进行优化选择。

在实际应用中,海洋监测数据的质量控制往往根据参数的自身特点和应用需求,选用多个质量控制方法相结合,并根据质量控制方法的复杂性和时效性,分批进行质量检验^[41-42]。以国际 Argo 计划的监测数据为例,其质量控制通常分为实时质量控制和延时质量控制,实时质量控制在浮标观测后的 24 h 内完成,进行基本的合理性、一致性等质量检验,突出实时性;延时质量控制在 6 个月内完成,可根据历史资料统计分析等方法,对数据进行检验和校正,突出其准确性^[43-45]。

4 海洋大数据应用

4.1 海洋生态环境保护

海洋是人类社会发展重要的战略资源基地,蕴含着丰富的海洋资源。良好的海洋生态环境是海洋资源可持续开发利用的基本条件。近年来,海洋环境监测技术不断发展,已经从单一特征性参数监测,发展为多任务、多区域、多站位、多参数的综合调查监测^[46],为海洋资源的开发利用、海洋生态环境保护、海洋经济可持续发展提供了重要支撑。

海洋环境监测内容丰富,包括海洋气象观测、水质分析、沉积物分析、生物监测和海洋放射性监测等^[47]。随着海洋监测数据的不断获取和更新,已积累形成海量的数据资源,传统的数据分析与信息挖掘等技术已经难以满足海量数据的分析需求,而基于大数据的海洋信息技术,能快速有效地对海洋生态监控区的指标参数进行整合与优化^[48]。例如,利用遥感与声学数据,对海洋底栖生物的生物群落和物种分布进行监测,为海洋生态保护提供科学参考^[49]。利用海洋水色遥感探测如叶绿素、悬浮物、

黄色物质、污染物等信息,对海洋水质监测和近岸水体污染预警具有重要意义^[50]。基于海洋遥感大数据和生物多样性监测技术,实现对典型海洋生态系统、海洋生态功能区、敏感区和脆弱区的连续监测,及对重要海洋经济物种专项监测^[51-52]。通过海洋大数据分析与应用,进一步提升海洋生态调查精度,推动海洋水质环境改善,保护海洋生物多样性,提高海洋生态环境综合治理水平。

4.2 海洋防灾减灾

中国是世界上受海洋灾害影响最为严重的国家之一,灾害种类多、分布广、频率高,影响正常的海洋生产活动并危害国民生命财产安全。2001—2017年,中国海洋灾害(风暴潮、海浪、海冰、赤潮及其他海洋灾害)造成的直接经济损失累计超过 2 075 亿元^[53]。改革开放以来,人口、城市在中国沿海地区高度聚集,沿海地区城镇化与工业化快速发展,海洋灾害对中国沿海地区的经济社会发展带来严峻挑战。

近年来,我国海洋环境监测技术研究与应用已经取得巨大的进步,逐步建立起由海洋监测平台、多参数浮标、调查船、卫星遥感等组成的海洋环境立体监测网络^[47]。随着海洋自动监测技术的发展,基于海洋大数据的数据处理系统对海洋监测数据进行分析、处理和挖掘,为海洋灾害预报预警提供支持。2000 年基本建成的“海洋环境立体监测和信息服务系统”上海示范区是我国第一套自行研发的海洋环境立体监测系统,为上海地区海洋经济发展、海洋环境保护、减轻海洋灾害损失提供海洋环境信息服务^[47]。2002 年我国发射了第一颗海洋卫星——“海洋一号 A”卫星,主要开展海洋水色要素(包括叶绿素、悬浮泥沙、可溶性有机物)、水温、污染物及浅海水深和水下地形的探测工作,并对全球海洋环境进行遥感观测,对我国近海的海洋灾害如赤潮、绿潮等进行监测和预警^[54]。“九五”以来,我国加大了海洋环境监测高新技术的研发投入,攻克了一系列关键技术。集成自主研发仪器设备构建的区域性海洋立体监测网于 2008 年投入业务化运行,产生的近海监测数据服务地方经济发展,提高海洋防灾减灾水平^[47]。近年来,我国陆续建成多个海洋环境立体监测系统与近海海洋观察研究网络,

实现对我国东海、黄海和南海北部海域的长期综合监测,提高了所在区域对赤潮、绿潮、溢油、核辐射、环境突发事件的监控能力和预警水平,并为我国近海变化规律预测、生态环境保护、海洋经济发展提供了大量的海洋综合性基础资料^[55]。

4.3 全球气候变化研究

全球气候变化影响人类活动与全球生态系统,并产生诸多严重后果,如冰川消融、海平面上升、厄尔尼诺现象、旱涝灾害增加、严寒热浪等极端天气增多。海洋水文气象监测为全球气候变化研究提供基础数据,也是海洋环境监测的重要内容。随着海洋观测和信息技术的发展,海洋水文气象监测已经形成了以遥感、遥测自动化为重点的海洋水文气象自动监测系统^[56]。基于海洋遥感大数据,海洋水文气象监测涵盖诸多海洋环境要素,具有广泛的应用领域。例如,根据美国 NASA 卫星观测结果,北极海冰自 20 世纪 80 年代以来每 10 年以 13.1% 的速度下降,南极洲和格陵兰岛海冰分别以每 10 年 148 亿 t 和 279 亿 t 的速度减少,全球海平面每年平均升高 3.3 mm^[57]。利用 Argo 数据对不同深度处的海水温度年周期数据进行分析,获取全球变暖理想观测点,为海面浮标布局提供科研参考^[58]。通过卫星遥感监测到的海洋风场、海洋高度数据,可以应用于海平面变化观测、海洋极端降雨预报;海表盐度、海表温度数据可服务于大洋环流、台风预报、全球气候变化与预测等^[59]。

未来,全球化的海洋观测系统是人类认识海洋、掌握海洋环境变化规律、开展全球气候变化研究的发展方向。基于此,全球海洋观测系统(GOOS)在海洋环境监测和资源共享方面发挥重要作用,促进国际间开展海洋数据搜集、海洋数据应用与合作,为全球气候变化背景下的海洋多要素变化与关联分析提供平台^[47]。而从海洋大数据中挖掘的大气—海洋互作信息、生态系统—海洋理化参数耦合变异过程、气候变化对极端气候事件发生频次、强度及空间分布特点的影响,将成为未来利用海洋环境监测数据开展全球气候变化分析的研究热点^[16]。

4.4 海洋经济(航运、养殖、捕捞等)

我国不断加大海洋环境监测力度,基于海洋环境

监测的大数据分析服务国家海洋经济的快速发展。随着多功能海洋监测平台的建设,基于海洋环境监测系统的大数据分析在环境科学研究、航运、渔业、能源探测等诸多海洋活动中得到运用并提高效率和生产力。受益于卫星技术(如通信、遥感、导航等)的发展,海洋监测技术平台为船舶航运业提供了全面的风速、水流、浪潮、水深、水温变化等数据资料,为船型开发设计、航线航速优化提供支持^[60]。如2014年1月通过对卫星遥感数据与海洋气象数据的快速综合分析,中国“雪龙”号极地考察船在南极冰海成功救援被困的俄罗斯“绍卡利斯基院士”号考察船,成为海洋大数据指导极地航行船只脱困的典型案例^[61]。

此外,传统渔业和海水养殖业通过依靠海洋环境监测技术、生物技术、先进材料技术、传感器技术的发展实现产业的转型升级。如利用海洋生物监测技术,对养殖区域重金属、有机污染物、赤潮毒素、生化耗氧量、硫化物、氨氮等指标进行检测与实时监测^[46],提高鱼虾贝藻等养殖品种的生长品质与安全品质,提高养殖产量,并改善养殖生态环境,实现冷水名贵鱼种的增殖增产^[62-63]。新型海洋牧场在传统海洋牧场原有人工鱼礁、网箱养殖、增殖放流的基础上引入物联网、大数据、通信传感等信息技术手段,对渔场海况信息(如叶绿素浓度、海表温度、海洋表面盐度、海洋表面高度、海洋水色信息等)进行在线监测,对海洋初级生产力、海洋生物存量分布及其变化进行分析,提高对渔期、渔场、鱼群数量和质量等渔情预报的准确性^[64]。近年来,海洋环境监测系统在我国新型海洋牧场建设中发挥重要作用,如在山东海洋牧场区域海域布设海底观测站,实现对海洋牧场水文生态及生物资源状况的实时在线监测,建立山东省海洋牧场观测预警数据中心,提升区域海洋安全管理和风险预报预警能力。

4.5 国防建设

近年来,各沿海国家正积极地应用和发展海洋监测高新技术,从空间、水面、水下对海洋进行立体监测,加速构建海洋环境立体监测系统,维护国家安全和海洋权益^[65]。目前,海洋环境立体监测系统的核心是依托遥感卫星组成的天基海洋环境监测平台,同时,海洋巡航飞机、无人航空遥感飞机组成空基平台,

固定监测站和高频地波雷达站组成岸基平台,浮标、潜标、船舶组成海基监测平台,水下固定监测站、水下水声探测阵组成海底监测平台,上述多平台共同构成海洋环境的立体监测系统,能够对特定海域实施海天一体化的立体监测^[65-66]。例如,美军航母编队在作战巡航海域以“岸基保障为主,舰基平台为辅”为原则,通过卫星、雷达遥感、遥测浮标、航母编队舰船,完成海洋环境特征参数的获取、分析、预测和评估^[67]。此外,为满足海洋环境监测和海上国防建设需求,一系列军事技术指标被应用于海洋环境监测装备的研制中。如,高频地波雷达能够超视距探测海平面视线以下的舰船、飞机、导弹和冰山等移动目标,因其独特的性能优势被各海洋国家竞相研制与部署,以抵御战争威胁并应用于海洋开发与研究^[68]。水下航行器因具有适应复杂环境、监测范围广、作业时间长、隐蔽性强等特点,成为世界各国高度重视的海洋科技^[69]。目前水下航行器正向远航程、深海型、多传感器、混合式、智能化方向发展,特别利用图像识别、人工智能、大数据分析实现信息处理、精密导航定位等技术要求,在民用领域开展海底地形勘察、海洋资源调查、水文参数测量和生物监测,在军用领域用于侦察、扫雷、援潜救生等^[70]。

5 结论与展望

人类社会文明、科学技术的发展与海洋一直息息相关,国际政治、经济、军事等各个方面也与海洋有着密切联系。未来海洋事业将会更加深入地贯穿到人类科技、文化、经济、政治的各个层面,从多产业、多环境、全方位影响国家核心力量建设。海洋环境大数据平台的建设、海洋环境大数据的发展也将为建设海洋强国提供更强有力的支撑。因此,仍处在发展阶段的海洋环境大数据建设,尚存在一系列问题亟待解决。

5.1 基础软硬件设施的能力建设

在全球各地观测站点不断更新观测数据的当下,海洋环境数据量已达到前所未有的规模,传统以单一数据输入与查询为中心的数据管理系统已不能满足科学研究的需要。日益增长的数据量迫切需求更高效、稳定的数据存储与管理系统的支持,因此对基础的软硬件设施提出了更高的建设要求。

5.2 数据处理能力建设

海洋环境大数据发展的核心在于数据的有效

处理、提取、分析及反馈。海量数据资源的产生既给科研人员带来了极大的挑战,也给科学研究带来了更大的机遇。如何深度挖掘数据、优化算法,决定了从复杂数据中获取信息的全面性与有效性,也是实现海洋环境大数据价值的重要手段。

5.3 数据质量控制能力建设

海洋大数据的数据量大,数据计算速度快,数据种类繁多,由于数据读取的不准确性和不完整、数据传输或计算的意外错误,可能引起数据处理结果的严重误差,极大地影响数据的有效利用,浪费传输、存储和计算资源。如何在第一时间使数据系统及时自检错误,并准确修复,也是海洋环境大数据建设研究中一个挑战性难题。

参考文献

- [1] MASHEY J R. Big Data and the Next Wave of InfraStress [C]// Computer Science Division Seminar. Berkeley: University of California, Berkeley, 1997.
- [2] BARWICK H. The“four Vs”of Big Data: Implementing Information Infrastructure Symposium[EB/OL].[2012-10-02].http://www.computerworld.com.au/article/396198/iis_four_vs_big_data/
- [3] 钱程, 陈戈. 海洋大数据科学发展现状与展望[J]. 学科与领域, 2018, 33(8):884-891.
- [4] SLOYAN B M, Roughan M, HILL K. The global ocean observing system[C] // New Frontiers in Operational Oceanography, E. Chassignet, A. Pascual, J. Tintoré, and J. Verron, Eds., GODAE OceanView, 2018, 75-90.
- [5] 何剑锋, 张芳, 林凌. 我国极地海底观测系统的发展与展望[J]. 地球科学进展, 2013, 28(5): 566-571.
- [6] 罗续业, 李彦. 海王星海底长期观测系统的技术分析[J]. 海洋技术, 2006, 25(3): 15-18.
- [7] 李风华, 路艳国, 王海斌, 等. 海底观测网的研究进展与发展趋势[J]. 中国科学院院刊, 2019, 34(3): 321-330.
- [8] 徐渡. 1958-1960年: 全国海洋综合调查[J]. 海洋科学, 2010, 34(4):109-110.
- [9] 崔爱菊, 曲媛媛, 韩京云. 浅谈“908专项”档案整理工作[J]. 海洋开发与管理, 2013, 30(12): 46-48.
- [10] 苏纪兰. 如何正确认识 Argo 计划[J]. 海洋技术, 2001, 20(3):1-2.
- [11] JAYNE S R, ROEMMICH D, ZILBERMAN N, et al. The Argo Program: Present and Future [J]. Oceanography, 2017, 30(2):18-28.
- [12] 陈戈, 杨杰, 张本涛, 等. 新一代海洋科学卫星的思考与展望[J]. 中国海洋大学学报(自然科学版), 2019, 49(10): 110-117.
- [13] 吴自军. 美国大型海底观测计划(OOD)正式启动运行[J]. 海洋地质与第四纪地质, 2016, 36(6): 134.
- [14] CHEN C, ZHANG Y, SHI X Y. Research on Detection and Analysis of Marine Ecological Environment Based on Big Data [J]. IOP Conference Series Earth and Environmental Science, 2020, 514:032053.
- [15] WANG Z A, MOUSTAHFID H, MUELLER A V, et al. Advancing Observation of Ocean Biogeochemistry, Biology, and Ecosystems With Cost-Effective in situ Sensing Technologies[J]. Frontiers in Marine Science, 2019, 6:519.
- [16] 侯雪燕, 洪阳, 张建民, 等. 海洋大数据: 内涵、应用及平台建设[J]. 海洋通报, 2017, 36(4): 361-369.
- [17] 罗续业. 海洋技术进展 2014 [M]. 北京: 海洋出版社, 2015: 295.
- [18] 徐京萍, 赵建华. 遥感技术在海域使用动态监测中的应用[J]. 卫星应用, 2016, 6: 35-39.
- [19] 郑沛楠, 宋军, 张芳苒, 等. 常用海洋数值模式简介[J]. 海洋预报, 2008, 25(4): 108-120.
- [20] 王世红, 赵一丁, 尹训强, 等. 全球海洋再分析产品的研究现状[J]. 地球科学进展, 2018, 33(8): 794-807.
- [21] QIAN C, LIU A, HUANG R, et al. Quality control of marine big data-a case study of real-time observation station data in Qingdao[J]. Journal of Oceanology and Limnology, 2019, v.37(06):195-205.
- [22] GUAN C, YANG Y. Research on computer control method of marine environmental monitoring data quality [J]. IOP Conference Series Earth and Environmental Science, 2019, 252:042067.
- [23] 姜欢欢, 张威, 马芳, 等. 浅谈海洋环境监测质量控制及质量保证技术[J]. 海洋开发与管理, 2014, 31(4):58-61.
- [24] 张铁艳, 王化仁, 杨鲲, 等. 海洋调查观测资料的质量控制[J]. 水道港口, 2006(1):48-50.
- [25] 路文海, 向先全, 杨翼, 等. 海洋环境监测数据处理技术流程与方法研究[J]. 海洋开发与管理, 2015, 32(2):58-62.
- [26] 杨扬, 苗庆生, 韦广昊, 等. 海洋站观测资料的质量控制方法及其应用[J]. 海洋开发与管理, 2017, 34(10):109-113.
- [27] 张立辉. 海洋大数据质量检测中的嵌套循环算法[J]. 舰船科学技术, 2019, 41(6):152-154.
- [28] 向先全, 路文海, 杨翼, 等. 海洋环境监测数据集质量控制方法研究[J]. 海洋开发与管理, 2015, 32(1):88-91.
- [29] 温玉波. 海洋环境观测数据的质量控制研究[J]. 农业网络信息, 2014(2):35-38.
- [30] 罗冬莲. 深沪湾溶解氧的分布及其与浮游植物悬浮物的相关性研究[J]. 海洋通报, 2002(1):31-36.
- [31] 戴文娟, 雒伟民, 陈靛瑜, 等. Aprior 算法在海滨观测数据相关性检验中的应用[J]. 海洋信息, 2014(4):35-38.
- [32] 周旭聪, 胡剑. 海洋站自动观测系统数据异常的检查方法研究[J]. 海洋信息, 2019, 24(4):30-34.

- [33] 郑琳,刘艳,崔文林,等. 海洋监测数据质量评估研究[J]. 海洋通报, 2014, 33(2): 228—234.
- [34] 刘首华,陈满春,董明媚,等. 一种实用海洋浮标数据异常值质控方法[J]. 海洋通报, 2016(3): 264—270.
- [35] 蒋华,武尧,王鑫,等. 改进 K 均值聚类的海洋数据异常检测算法研究[J]. 计算机科学, 2019, 46(7): 211—216.
- [36] 蒋华,季丰,王慧娇,等. 改进 Kmeans 算法的海洋数据异常检测[J]. 计算机工程与设计, 2018, 39(10): 3132—3136.
- [37] MORELLO E B, GALIBERT G, SMITH D, et al. Quality Control (QC) procedures for Australia's National Reference Station's sensor data-Comparing semi-autonomous systems to an expert oceanographer[J]. *Methods in Oceanography*, 2014, 9: 17—33.
- [38] 李飞,卢勇夺,林波,等. 初探支持向量机算法在海洋站观测数据质量控制中的应用[J]. 海洋预报, 2016, 33(2): 66—73.
- [39] 王振华,周雪楠,黄冬梅. 不确定海洋数据的质量抽样检验模型研究[J]. 计算机科学, 2015, 42(2): 182—184.
- [40] 陈括,匡翠萍,王蕾,等. 基于粗集理论的海洋数据质量检验方案多属性决策法[J]. 海洋环境科学, 2019, 38(2): 233—237.
- [41] 于婷,刘玉龙,杨锦坤,等. 实时和延时海洋观测数据质量评估方法研究[J]. 海洋通报, 2013, 32(6): 610—614.
- [42] INGLEBY B, HUDDLESTON M. Quality control of ocean temperature and salinity profiles-Historical and real-time data [J]. *Journal of Marine Systems*, 2007, 65(1—4): 158—175.
- [43] 卢少磊,许建平. Argo 剖面浮标观测资料质量控制方法研究概述[C]// 第八届全国海洋资料同化研讨会暨第二届 Argo 科学研讨会, 舟山: 国家海洋局卫星海洋环境动力学国家重点实验室, 2013: 64—75.
- [44] 王辉赞,张韧,王桂华,等. Argo 浮标温盐剖面观测资料的质量控制技术[J]. 地球物理学报, 2012, 55(2): 577—588.
- [45] GAILLARD F, AUTRET E, THIERRY V, et al. Quality Control of Large Argo Datasets[J]. *Journal of Atmospheric & Oceanic Technology*, 2007, 26(2): 337—351.
- [46] 曾容,许艳,杨翼,等. 海洋环境监测数据统计研究 [J]. 海洋开发与管理, 2017, 34(4): 32—7.
- [47] 吴夏青,王巧宁. 海洋环境监测概况 [C]// 陈令新,王巧宁,孙西艳. 海洋环境分析监测技术, 北京: 科学出版社, 2018: 3—25.
- [48] DANOVARO R, FANELLI E, CANALS M, et al. Towards a marine strategy for the deep Mediterranean Sea: Analysis of current ecological status [J]. *Marine Policy*, 2020, 112.
- [49] BROWN C J, SMITH S J, LAWTON P, et al. Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques [J]. *Estuarine, Coastal and Shelf Science*, 2011, 92(3): 502—20.
- [50] 潘刚,段舜山,徐宁. 海洋赤潮水色遥感技术研究进展[J]. 生态科学, 2007, 26(5): 460—5.
- [51] 林金兰,陈彬,黄浩,等. 海洋生物多样性保护优先区域的确定[J]. 生物多样性, 2013, 21(1): 38—46.
- [52] 马克平. 生物多样性监测依赖于地面人工观测与先进技术手段的有机结合[J]. 生物多样性, 2016, 24(11): 1201—2.
- [53] 叶攀. 本世纪以来中国海洋灾害致直接经济损失超 2000 亿元[N]. 中新社, 2018—05—11(1).
- [54] 杨保华. 构建中国海洋卫星体系 提升海洋环境与灾害监测能力[J]. 中国空间科学技术, 2011, 31(5): 1.
- [55] 李健,陈荣裕,王盛安,等. 国际海洋观测技术发展趋势与中国深海台站建设实践[J]. 热带海洋学报, 2012, 31(2): 123—133.
- [56] 李帅,郭俊如,姜晓轶,等. 海洋水文气象多时空尺度资料来源分析[J]. 海洋通报, 2020, 39(1): 24—39.
- [57] NASA. GLOBAL CLIMATE CHANGE [EB/OL]. [2020—10—13] https://www.nasa.gov/mission_pages/noaa-n/climate/index.html.
- [58] CHEN G, ZHANG H, WANG X. Annual amphidromic columns of sea temperature in global oceans from Argo data[J]. *Geophysical Research Letters*, 2014, 41(6): 2056—62.
- [59] 洪阳,侯雪燕. 海洋大数据平台建设及应用[J]. 卫星应用, 2016 (6): 26—30.
- [60] 徐文,鄢社锋,季飞,等. 海洋信息获取, 传输, 处理及融合前沿研究评述[J]. 中国科学: 信息科学, 2016, 46(8): 1053—85.
- [61] 张林,李春花,柴先明,等. 2014 年初雪龙船在南极被海冰围困期间海洋气象环境分析[J]. 极地研究, 2014, 26(4): 487—95.
- [62] 张学雷,朱明远,李瑞香,等. 贝类养殖环境的多参数同步连续监测[J]. 海洋科学进展, 2004, 22(3): 340—5.
- [63] 尤锋,张培军,相建海,等. 海水养殖鱼类遗传多样性的保护[J]. 海洋科学, 2003, 27(12): 10—3.
- [64] 朱文东. 智慧渔业背景下智慧型海洋牧场发展研究[D]. 舟山: 浙江海洋大学, 2019.
- [65] 孙西艳,付文龙,温国义,等. 海洋环境立体监测系统[C]// 陈令新,王巧宁,孙西艳. 海洋环境分析监测技术, 北京: 科学出版社, 2018: 315—65.
- [66] 漆随平,厉运周. 海洋环境监测技术及仪器装备的发展现状与趋势[J]. 山东科学, 2019, 32(5): 21—30.
- [67] 于宇,黄孝鹏,崔威威,等. 国外海洋环境观测系统和技术发展趋势[J]. 舰船科学技术, 2017, 39(12): 179—83.
- [68] 陈钊龙. 地波雷达海洋环境监测中的技术及应用[J]. 中国新通信, 2014, 16(19): 121—2.
- [69] KINSEY J C, EUSTICE R M, WHITCOMB L L. A survey of underwater vehicle navigation: Recent advances and new challenges[C]// *Proceedings of IFAC Conference on Maneuvering and Control of Marine Craft*, Lisbon, Portugal, 2006, 9.
- [70] 李硕,刘健,徐会希,等. 我国深海自主水下载机器人的研究现状[J]. 中国科学: 信息科学, 2018, 48(9): 1152—64.