



海洋观测数据质量控制技术研究现状及展望

谭哲韬^{1,6}, 张斌^{2,3}, 吴晓芬⁴, 董明媚⁵, 成里京^{1,2,6*}

1. 中国科学院大气物理研究所, 国际气候与环境科学中心, 北京 100029;
2. 中国科学院海洋大科学研究中心, 青岛 266071;
3. 中国科学院海洋研究所, 青岛 266071;
4. 自然资源部第二海洋研究所, 卫星海洋环境动力学国家重点实验室, 杭州 310012;
5. 国家海洋信息中心, 天津 300171;
6. 中国科学院大学, 北京 100049

* 通讯作者, E-mail: chenglij@mail.iap.ac.cn

收稿日期: 2021-04-11; 收修改稿日期: 2021-08-20; 接受日期: 2021-09-14; 网络版发表日期: 2021-12-02

中国科学院战略性先导科技专项(B类)项目(编号: XDB42040402)、自然资源部第二海洋研究所卫星海洋环境动力学国家重点实验室资助项目(编号: QNHX2133)、国家重点研发计划全球变化及应对专项项目(编号: 2017YFA0603202)、中国科学院海洋大科学中心自主部署项目(编号: COMS2019Q01)、国家自然科学基金项目(批准号: 42076202)和中国科学院青年创新促进会资助

摘要 海洋观测数据的质量控制是建立高质量海洋科学数据库的基础, 对于推动海洋及多学科交叉研究、预测预报、灾害预警等具有重要意义. 随着各种自动化观测平台(仪器)的出现, 海洋调查从近海逐渐拓展到开放大洋, 海洋科学已进入大数据时代. 目前, 国内外对如何获得高质量的现场观测数据越来越重视. 然而, 由于获取数据手段多样、数据质量千差万别、数据错情类型繁多等因素, 使得如何高效和精准地发现这些质量问题并对其进行质量控制是一个难点, 也是数据处理中的核心技术. 本文系统总结了海洋观测数据(温度、盐度、溶解氧等)质量控制的研究现状与最新进展, 对比了国内外不同机构质量控制的差异, 分析了现已取得的成果及存在的问题, 探讨了开展质量控制性能评估的可能性, 并就未来的发展方向提出了一些建议.

关键词 质量控制, 海洋观测, 温度, 盐度, 数据质量, 大数据

1 引言

现代海洋科学研究越来越依赖于系统的、高可信度的基本观测数据. 准确、可靠的海洋观测数据是建设高质量海洋科学数据库、进行多学科交叉研究的重要基础. 随着海洋观测、模拟手段的快速提升, 海洋数据呈爆炸式增长, 全方位、连续、多源的观测使得海洋数据目前存量已达到EB级别, 日增量也达到TB级

别, 海洋科学已进入大数据时代(刘帅等, 2020). 2017年, 联合国教科文组织发布的《全球海洋科学报告: 世界海洋科学现状》首次对全球海洋科学进行综合评估并提出八个高优先主题(UNESCO, 2017). 其中, “海洋观测与数据”是八个高优先主题之一, 是支撑海洋与气候、海洋健康等其他七个主题的重要部分. 在新时代建设海洋强国、发展海洋科学技术、建设高质量海洋科学数据库的背景之下, 当海洋观测成为中国获取近

中文引用格式: 谭哲韬, 张斌, 吴晓芬, 董明媚, 成里京. 2022. 海洋观测数据质量控制技术研究现状及展望. 中国科学: 地球科学, 52(3): 418-437, doi: 10.1360/SSTe-2021-0096

英文引用格式: Tan Z, Zhang B, Wu X, Dong M, Cheng L. 2022. Quality control for ocean observations: From present to future. Science China Earth Sciences, 65(2): 215-233, <https://doi.org/10.1007/s11430-021-9846-7>

海及大洋的物理、生化状况的第一手资料的时候, 准确、可靠的海洋现场观测数据(温度、盐度、溶解氧、叶绿素、pH等)已成为必然要求(李慧青等, 2015; 陈洪滨等, 2019).

但是, 目前正在流通的国际海洋数据集(库)以及中国一些近海观测的数据集中普遍存在着质量问题. 这些质量问题包括了由于仪器偏差、仪器故障、外界干扰、转码错误、通信错误或观测人员失误造成的过失误差等, 也包括了人为痕迹处理等问题. 各国获取的海洋的现场观测数据, 由于来自不同的数据提交机构、不同的航次、不同的汇交格式、不同的仪器、不同的采集(操作)标准和方式; 加之即使是同一个仪器, 也有可能是不同的型号及不同的采样分辨率, 这就导致了数据错情类型繁多. 例如, 对于物理海洋方面的温度、盐度、溶解氧、流速数据, 从有观测记录的年份(大约从20世纪初)开始, 仪器类型多达几十种, 例如机械式探温仪(mechanical bathythermographs, MBT)、投弃式温深计(expendable bathythermographs, XBT)、南森采水器(Nansen/Niskin bottles)、CTD温-盐-深仪(conductivity-temperature-depth sensors)、水下滑翔机(glidors)、动物载体传感器(animal mounted pinniped profiles, APB)、卫星遥感等(Guinehut等, 2009; Balmaseda等, 2013; Palmer等, 2017). 测量类型包括了测站定点观测(如浮标站、海岛站)、走航(大面廓线)观测、移动观测(如Argo、水下滑翔机)、遥感观测等. 观测方式包括了固定平台观测、机动式(移动式)平台观测等. 航线几乎覆盖了全球每一个角落, 成员国数量包括了美国、英国在内的近百个国家. 根据世界海洋数据集(WOD18)的最新统计(截至2021年4月), 目前海洋温度观测的数据量已达3000万个(条)、盐度观测的数据量已达2000万个(条)(Garcia等, 2018). 因此, 从多种观测手段、多种采样方式获取到的多种数据类型的数据中, 或多或少存在着不同类型的质量问题. 而对于海洋化学要素等以采样分析为主的数据而言(例如营养盐、pH等), 在数据录入、转换、标定(校正)、比对的时候, 也可能会由于人为操作不当或熟练度不高等原因而引入过失误差. 此外, 对于这些存在质量问题的数据, 如果元数据缺少相应的信息(例如观测记录、实验报告、操作记录缺失), 那么将会很难定位问题的来源.

考虑到上述的质量问题可能会影响数据应用研究

结果的准确性, 因此有必要开展数据的质量控制工作(quality control, QC). 如果不解决科研所需基础观测资料的质量问题, 海洋科学研究工作(如为预报模式所需的初始场资料等)就缺少了最基本的依据. 如图1所示, 以温度廓线观测为例, 质量控制可以筛选出很多异常的温度数据. 对已收集到的数据进行质量控制, 这一点在科学界已达成共识. 此外, 质量控制是数据驱动科学研究中的一个基础任务, 经由一个良好的质量控制系统产生的高质量观测数据集对于推动海洋科学多学科交叉研究、模式评估、业务预报和灾害预警等具有非常重要的意义(Graves, 2009; 张斌等, 2017; Brunton和Kutz, 2019). 例如, 在气候变化科学研究、物理海洋规律性研究、海洋模式开发与改进、海洋生态灾害监测与预警当中, 一般都需要用到经质量控制后的现场观测数据作为研究工作的数据基础(如Arthur, 1965; Talley, 1984; Cheng等, 2017; Zika等, 2018; 齐庆华, 2020; 于仁成等, 2020; Zou等, 2020). 因此, 对于物理海洋、海洋生态、海洋化学等(任杰, 2019)多维度的海洋观测数据(尤其是温度、盐度、溶解氧数据)的质量控制, 则显得尤为重要, 这也对发展更加完备的质量控制技术提出了更高的要求(Liu等, 2017a; 解静等, 2020; 刘帅等, 2020).

一般而言, 质量控制是指通过人工审查或计算机算法辅助的方式, 对海洋观测数据进行全方位的质量评估, 识别和剔除因各类错情产生的虚假观测数据, 其目的是提高数据的准确性和可用性. 在过去数十年间, 国内外众多知名的质量控制机构都在为解决这些数据的质量问题作出了很多的努力, 并且发布了许多著名的经质量控制后的海洋现场观测数据集(库). 例如在1994年, 美国国家海洋和大气管理局国家环境信息中心(NOAA/NCEI)率先发布了第一个包含了历史温度和盐度观测的数据集产品: 世界海洋数据库1994(World Ocean Database 1994). 随后的十余年间, 1998年(WOD98)、2001年(WOD01)、2005年(WOD05)、2009年(WOD09)、2013年(WOD13)的版本和最新的2018年版本(WOD18)相继发布, 变量从温度、盐度等基础物理海洋变量拓展到了溶解氧、流速、pH、磷酸盐、叶绿素等海洋生物化学的变量. 除此之外, 还有多个著名的全球和区域尺度的经不同标准的质量控制后的海洋现场观测数据集也应用于不同的科学研究场景. 例如全球海洋数据分析项目(GLODAP)数据集(OI-

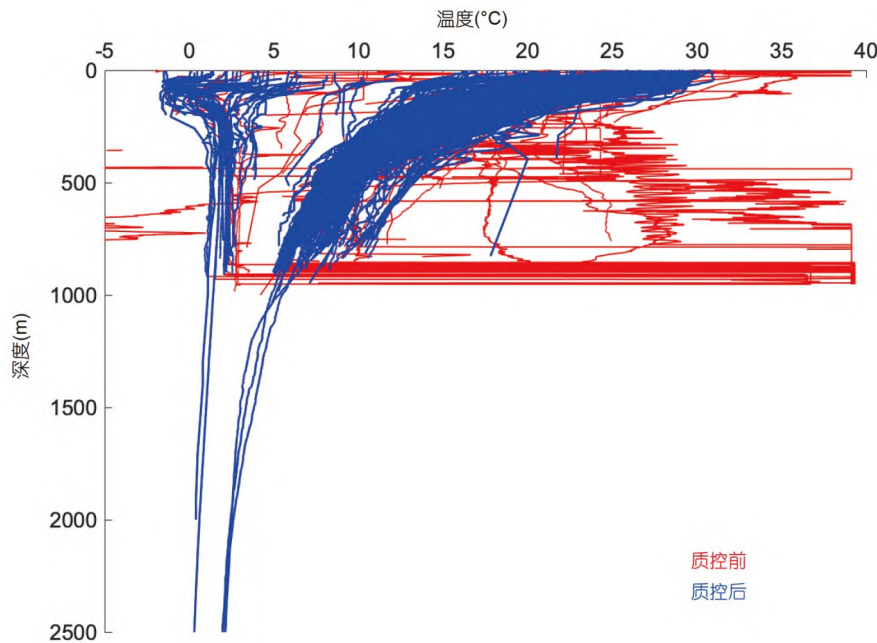


图 1 温度质量控制效果图

红色廓线为未经过质量控制的数据, 蓝色廓线为经过质量控制的正确数据(数据来源: QuOTA数据集印度洋部分数据)

sen等, 2016)、英国哈德莱气象中心EN数据集(Good等, 2013)、Copernicus海洋现场温盐(CORA)数据集(Szekely等, 2019)、海洋质量控制温度档案(QuOTA)数据集(Thresher等, 2008)、国际温盐剖面计划(GTSP)数据库(Sun等, 2010)、Argo计划数据库(Argo, 2000; Liu等, 2017b)、实时海洋质量控制(QAR-TOD)数据库(Bushnell, 2016)、世界大洋环流实验Argo全球水文气候数据集WOCE-WAGHC(Gouretski, 2018)等。几十年来, 这些数据集支撑了众多海洋和气候变化前沿科学研究(Domingues等, 2008; Palmer等, 2017; 邱国强等, 2018; Goni等, 2019; Fox-Kemper等, 2021)。

质量控制从时效性分类可以分为(近)实时质量控制(near real-time QC, NRQC)和延时质量控制(delayed mode QC, DMQC)。实时质量控制是指在浮标、船载仪器得到海洋观测数据之后立即在船基上或数据收发中心(如美国GTS中心)进行的第一次较为简单快速的质控(Goni等, 2019), 其延迟时间通常较短(1小时至1周)。而延时质量控制是采用更加标准化和更加严格的质控技术, 在实时质量控制的基础上进一步对海洋观测数据进行质量评估, 耗费的时间相较于实时质量控制更多, 通常有6~12个月的延迟(Roemmich等, 2019)。

目前, 通过Argo浮标观测得到的温度、盐度数据都经过NRQC和DMQC, 并已经广泛运用在海盆尺度范围内的观测和科学研究中(deYoung等, 2019; Smith等, 2019)。

质量控制从质控技术方式分类上还可分为自动化质控技术(automatic quality control, AutoQC)和(可视化)专家质控技术(visual/expert quality control, ExpertQC)。自动化质量控制一般是通过编写计算机程序, 对每一条剖面的观测物理参数进行自动检查, 随后对可疑的剖面进行自动质控标记(Bailey等, 1994; Thadathil等, 2001; Toma等, 2016; Gouretski, 2018)(图2)。而专家质控技术(又称“人工检查质控技术”)则是通过海洋学专家过往经验, 对经过自动化质控技术处理后的数据, 进行进一步的目视审查。在该项技术中, 由欧洲SeaDataNet中心开发的海洋数据视图软件(ocean data view)是目前最流行的用于海洋观测廓线可视化和专家质控的软件之一(<https://odv.awi.de>)。值得指出的是, 该项技术通常需要消耗大量的人力、物力、财力资源, 且需要的时间成本比较高, 因此专家质控技术的普适性比自动化质控技术要低。但是, 自动化质控技术和专家质控技术都是质控环节中不可缺少的一环, 两者是相辅相成、互相促进的。如图3所示,

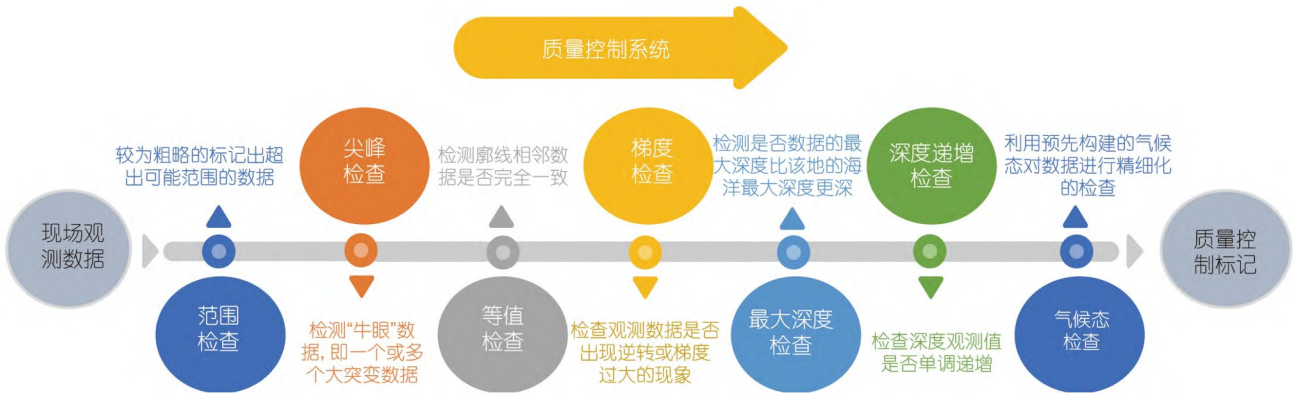


图 2 海洋温盐观测自动化质量控制(AutoQC)基本流程示意图

质量控制检查子模块包括但不限于图中所展示的检查模块

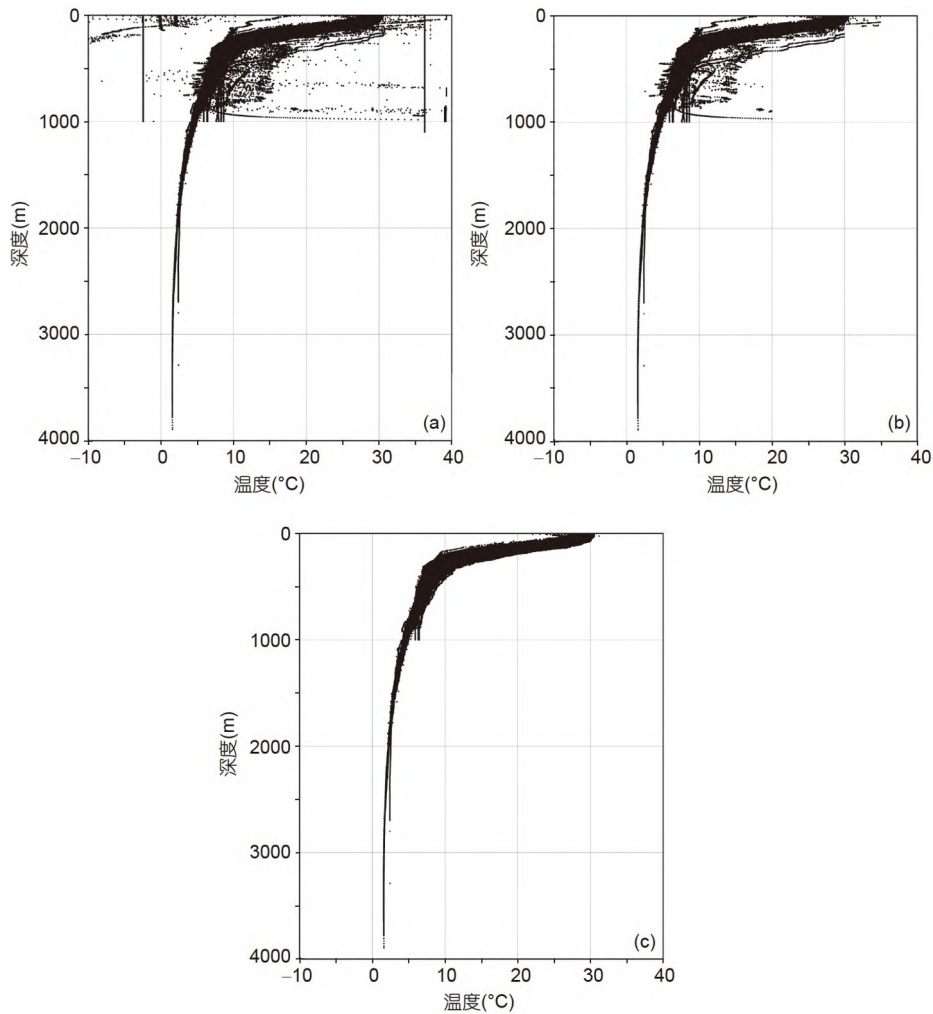


图 3 随机抽选的WOD18数据集中海洋温度观测数据(5°N~6°N, 126°E~127°E方区1940~2019年所有数据)

(a) 原始温度观测廓线; (b) WOD自动化质控(AutoQC)后标记为正确的廓线; (c) 专家质控(ExpertQC)后标记为正确的廓线

专家质控技术可以识别出很多在自动化质控技术中没有被识别出来的错误数据。

在20世纪90年代之前, 质量控制主要以专家(人工)质控为主(例如, Bailey等, 1994; Boyer和Levitus, 1994; Thadathil等, 2001)。进入21世纪, 随着高性能计算机的广泛使用和为了应对海洋实时数据量日益快速增加的情况, 耗时更少、普适性更强的自动化质控技术越来越受重视(例如, 纪风颖等, 2004; Foken等, 2005; Gronell和Wijffels, 2008; Cummings, 2011; 王辉赞等, 2012)。在2010年前后, 随着人工智能技术的飞速发展, 机器学习和深度学习的方法也被逐渐应用到质量控制当中(例如, Sugiura和Hosoda, 2020; Mieruch等, 2021)。

2 质量控制技术方法的介绍

由于国内外自动化质控技术(AutoQC)相比于专家质控技术(ExpertQC)应用得更加成熟、广泛且有客观的技术标准, 而专家质控技术更依赖于特定专家的经验, 透明度较差、主观性较强, 很难予以客观评判(Roemmich等, 2019)。因此本节将基于自动化质控技术, 进一步详细地介绍各个质量控制检查方法的原理(图3)。

2.1 范围检查

顾名思义, 范围检查即是根据人们对海洋最基本的认识, 检查所测量的参数及相应的元数据(年份、月份、日份、时间、时区、经纬度)是否在合理值区间内。如果某一观测值不在合理的区间内, 则该观测值将被标记为可疑值。范围检查一般是所有的质量控制程序中, 最先使用的一个子模块。其可以包括如下的检查:

(1) 极值检查。极值检查一般是根据对海洋最基本的认识进行的一项最简单的检查。例如海温一般取值在 $[-2^{\circ}\text{C}, 35^{\circ}\text{C}]$, 海表面水温的变化应在 $[-4^{\circ}\text{C}, 44^{\circ}\text{C}]$, 盐度变化范围一般取值在 $[0\text{psu}, 40\text{psu}]$, 纬度取值范围在 $[90^{\circ}\text{S}, 90^{\circ}\text{N}]$ 等。根据统计学理论和物理海洋学最基础的经验, 超出这个范围之外的数值, 都可以认为是几乎不可能发生的极小概率事件。

(2) 全球深度-极值检查(global profile envelop check)。UNESCO-IOC(2010)指出该检查是极值检查

的进一步延伸。对于温度(密度)观测而言, 由于下层平均海温(密度)一般而言比上层低(高), 因此该检查旨在确定观测值是否在随深度变化的合理范围内。如果某一观测值不在合理的区间内, 则标记为可疑值(图4a)。例如在WOD的质量控制手册当中, 300~400m的深度范围内, 温度取值范围只能为 $[-2.0^{\circ}\text{C}, 27.0^{\circ}\text{C}]$, 盐度取值范围只能为 $[3\text{psu}, 41\text{psu}]$, 该阈值的选取依赖于深度的选取(Garcia等, 2018)。

(3) 冰点检查。由于海水(除表层外)具有不结冰的特性, 因此该检查旨在确定是否在给定的深度和盐度下观测到的温度是否比计算得到的冰点温度低(Wong等, 2020)。冰点温度可以通过压力和对应的盐度计算得到, 如下所示(UNESCO-IOC, 2010):

$$T = -0.0575 \times S + 1.710523e^{-3} \times S^{\frac{3}{2}} - 2.154996e^{-4} \times S^2 - 7.53e^{-4} \times P, \quad (1)$$

式中, T 是计算出的冰点摄氏温度, S 是范围取值在27~35psu的盐度, P 是在给定的盐度下的压力值(dbar)。

2.2 相关性检验(一致性检验)

由于海水的流体连续性和海气相互作用等特性, 海洋不同观测变量之间存在着一定的联系。例如海表温度与气温之间的关系、海表盐度与降水关系(降水较多的海区海表盐度会相较于降水较少的海区要低)、海表温度与寒潮关系、风速与浪高的关系、次表层温度与季节的关系等。因此, 杨扬等(2017)指出质量控制可以根据在同一时刻、同一地点观测得到的数据间的相互关系(是否符合一定的物理联系)进行检验。如果某一要素的观测值与其他同时刻要素的观测值在一定的范围内不符合两者之间的关系, 则标记为可疑值。该检查一般用在能同时获得多个物理变量时间序列的浮标观测中(李学坤和李凤金, 1997)。

2.3 连续性检验

相对于船舶走航观测而言, 海洋站、浮标站、验潮站等观测信息具有定点、长时间观测序列的优势, 对于了解某一位置的海洋状况而言, 具有船舶走航观测无法企及的优势。对于定点、长时间序列的观测资料, 由于要素前后时刻的变化通常是连续的、有规律的(如海水特性的渐变过程、观测时间是递增的), 因此可以通过确定要素相邻两时刻间的差值, 对观测资

料进行校验, 超过一定的阈值通常为异常值, 这对于检验时间序列的突变值非常友好(例如, 史静涛和周智海, 2011; 刘首华等, 2016; 杨扬等, 2017).

2.4 等值检查

一些仪器由于仪器特性和故障, 要素观测的读数可能会出现停滞或暂停读数的现象, 导致了读数随深度不发生改变的错误. 因此, 等值检查作为众多质控机构都采纳的检查之一, 其旨在检查廓线的观测值在一定的深度范围内是否完全相同(Daneshzadeh等, 1994; UNESCO-IOC, 2010; Gouretski, 2018; Wong等, 2020). 如果在一定的深度范围内出现了多个相同的观测值, 在排除等温层(对于温度观测而言)的因素之后, 则将这些观测值标记为可疑值(图4b).

2.5 统计特性检验

海洋站定点、长时间观测的资料在理论上往往服从于一定的概率统计分布. 根据数据的分布特征, 一方面可以建立分布拟合函数, 将拟合值与实际观测值进行比对(一般为卡方拟合优度检验); 另一方面针对(近似)服从正态分布的大样本数据, 可以采用莱茵达准则判断其合理性. 该方法常用在潮汐资料、浮标资料的质量控制(温玉波, 2014). 例如, 莱茵达准则(又称“ 3σ 准则”)经常用在海表温度(SST)的质控当中(史静涛和周智海, 2011; 杨扬等, 2017; 任焕萍等, 2021).

2.6 逆温检验

一般而言, 海水温度随深度的增加逐渐降低, 根据这一特性, 若出现逆温, 且逆温超过一定的阈值, 则可以认为是虚假逆温(UNESCO-IOC, 2010). 但是需要注意的是, Marshall和Schott(1999)指出在深对流区域或一些海盆(例如地中海), 存在一定范围的真实逆温, 这些逆温反映的是真实的海洋状况, 因此需要谨慎进行逆温检验的标记.

2.7 尖峰检验(牛眼检查)

在海洋现场观测中, 由于仪器性能或操作不当等原因, 可能在某一个深度位置存在明显不合理的较大突变, 即“牛眼”(图4d; UNESCO-IOC, 2010; Wong等, 2020). 例如, XBT探头在投放的时候, 电流可能由于导线绝缘不良而产生泄漏, 造成了在某一个深度上温度

观测值突然出现较大的突变(Barton和Gonzalez, 2016). 因此该检查旨在确定温度或盐度廓线是否存在较大突变(尖峰). 例如, Gouretski(2018)对于定义温度和盐度廓线的“牛眼”采用了如下的方法:

$$\begin{aligned} V_1 &= |D_{k+1} - (D_k + D_{k+2}) \times 0.5|, \\ V_2 &= |(D_{k+2} - D_k) \times 0.5|, \\ V &= V_1 - V_2, \end{aligned} \quad (2)$$

式中, D_k 、 D_{k+1} 、 D_{k+2} 分别代表相邻的三个深度层(k 、 $k+1$ 、 $k+2$)的温度或盐度观测(摄氏度或psu), V 表示由相邻三个观测计算出来的数据差异大小. 如果 V 超出了一定的阈值(该阈值随深度变化而变化, 由统计结果得出), 则 D_{k+1} 被判定为“牛眼”, 标记为异常值.

2.8 深度递增检查

现场观测的温度、盐度等剖面, 其深度均是从海表面0m单调往下递增. 但是由于仪器记录系统、操作人员记录不当等原因, 可能会出现某一点的深度值, 比它上、下相邻的观测值要大(小), 这就造成了深度不是单调增加或随观测时间保持不变. 因此该检查旨在测试测量的深度是否是单调增加(Argo, 2000; Ingleby和Huddleston, 2007; Garcia等, 2018). 若不满足单调增加, 则标记其为可疑值, 必要时对深度的顺序进行订正.

2.9 密度递增检查

海洋不同于大气, Pawlowicz(2013)指出海洋具有层结稳定的特性, 在绝大多数海区(除深对流区域外)下层海水的密度总是大于上层海水的密度. 若随着深度的增加, 海洋的密度出现反转的现象, 则出现反转的位置的温度或盐度则可以判定为可疑值(Argo, 2000; UNESCO-IOC, 2010; Good等, 2013; Garcia等, 2018). 其中, 密度可以由海水状态方程计算得出.

2.10 陆地位置检查

海洋观测资料的经纬度总是定位在水上(海洋或者内陆湖), 但是可能由于人工记录不当等原因, 记录的经纬度可能是经过人为修改或是错误的. 因此该检查旨在确定剖面的观测位置是否是在陆地上: 通过将经纬度与已知的陆地经纬度文件进行比较, 若位置出现在陆地上, 则标记整条廓线为可疑值. 几乎所有的

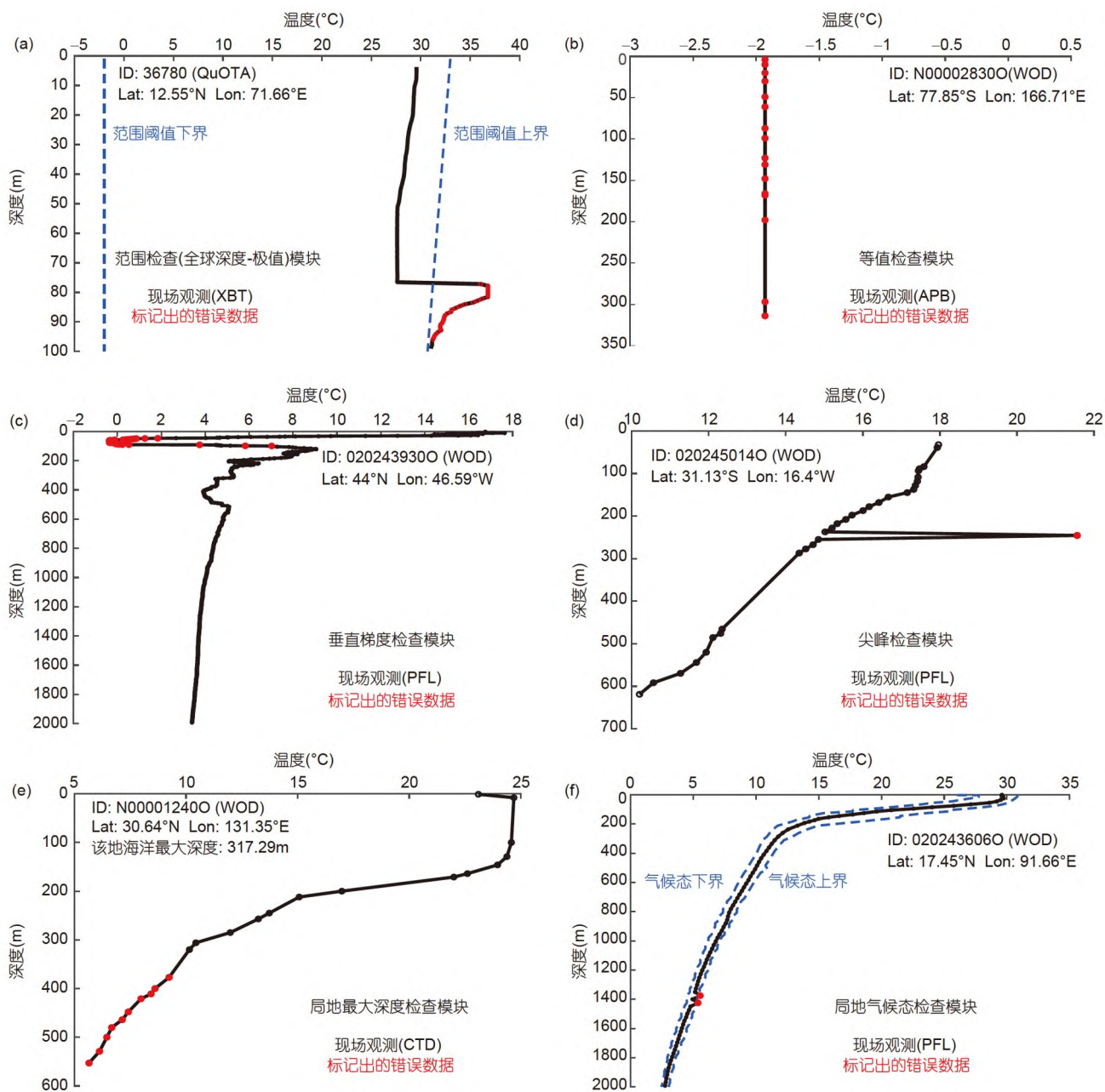


图4 不同模块的质量控制示意图

黑色线为温度现场观测廓线, 红色点为被质量控制识别出的错误(可疑)观测值

质控机构中都包含了这个检查模块(表1).

2.11 垂直梯度检查

由于一些廓线的深度分辨率较低, 相邻两个观测点的深度差非常大, 这就造成了要素观测可能在两点间存在急剧升高或降低. 根据梯度的定义, 该检查旨

在测试相邻的两个深度值对应的观测值的垂直梯度是否超出了一定范围(阈值依赖于深度). 该梯度范围阈值可以通过统计规律得出. 当垂直梯度落在该阈值之外时, 相邻的两个观测值都将被标记为可疑值(图4c). 例如, WOD质控方法中采用的盐度最大梯度值在400m以浅为 9.000psu m^{-1} , 400m以深为 0.050psu m^{-1}

表1 不同机构或数据集提供的质量控制方法的整理归纳^{a)}

子模块类型	WOD	Argo	GTSP	EN4	WAGHC	杭州Argo野外站	国家海洋信息中心			
							大面	定点	移动	
格式转换	0		0			0	0	0	0	
预检查 (基础检查)	测站识别码检查		1	1		1	2	2	2	
	日期时间检查	1	1	1		1	3	3	3	
	经纬度检查	1	1	1		1	4	4	4	
	陆地位置检查	1	1	1	5	4	1	5	5	
	流速/速度检查	2	1	1	3				6	
	声速检查			1						
	深度(压力)递增检查	3	3	4	6	2	3	8	8	9
范围检查	全球范围检查		2	2	1	3		6	6	7
	区域参数范围检查	5	2	3			2	7	7	8
	全球深度-温/盐范围检查	5		5		3	7			
	冰点检查		11	7				14	14	15
剖面检查	尖峰(牛眼)检查		4	8	2	8	8	9	9	10
	垂直梯度检查	6		9		9	4			
	距离中值(毛刺)检查		5					10	10	11
	密度翻转检查	7	8	10	4			13	13	14
	数字翻转检查		6					11	11	12
	等值检查		7	6		6	5	12	12	13
	极值检查					7	6			
	相邻廓线检查			14	7					
仪器类型检查	逆温检查			12		10				
	Argo灰名单检查		9			9			16	
	仪器最大深度检查				5				1	
	盐度漂移检查(Argo)		10			13			17	
统计特征检查	莱茵达准则检查	8						15		
	拟合优度检查									
	局地气候态检查			13		9	12			
	温度-盐度图检验						11	15	18	
	相邻观测数据检查				9				16	19
局地最大深度(压力)检查		0	11	5	4		1	1		
航线检查			15	3			16		20	
背景场检查				8						
高精度廓线配对检查	4									

a) 数字代表对应的检查模块在相应机构的质控程序中的先后顺序, 无数字代表对应的机构中没有该检查模块. 缩写如下: WOD, 美国国家海洋和大气管理局世界海洋数据集NOAA-WOD18; EN4, 英国哈德来环流中心EN4数据集; GTSP, 全球温盐剖面计划GTSP; Argo, 全球海洋观测阵列Argo计划; WAGHC, 世界大洋环流实验全球水气候数据集WOCE/WAGHC-ICDC; 杭州Argo野外站, 杭州全球海洋Argo系统野外科学观测研究站(中国Argo实时资料中心); 国家海洋信息中心(NMDIS)

(Garcia等, 2018).

2.12 局地最大深度检查

有一些观测仪器在触碰到海底的时候, 可能没有立刻暂停工作(如XBT、MBT), 这就导致了观测要素的读数有可能在不断地增加. 同时, 由于人为处理等原因, 汇交到国际数据库的一些廓线的深度观测值可能会经过数倍放大/缩小. 因此, 该检查旨在确定深度观测值是否比该地的海底深度(通过电子测深仪获得)更深(例如: UNESCO-IOC(2010); Good等(2013); Gour-etski(2018)). 若大于该地的最大允许深度, 则将大于该深度的所有测量值标记为可疑值, 不再用于进一步的分析(图4e).

2.13 局地气候态检查

局地海洋观测获取到的数值通常在一个气候平均态确定的范围内波动, 例如太平洋暖池的SST大约在28℃上下范围内波动, 并且从来不会低于15℃和高于35℃(Schmidtko等, 2013; Takeshita等, 2013; Udaya Bhaskar等, 2017; 吴晓芬等, 2019). 因此, 局地气候态检查旨在将观测到的温度、盐度、溶解氧等剖面与预先构建的该地大量历史观测值(气候态)及其波动范围进行比对(如使用由Locarnini等(2019)研发的2018年世界海洋图集(WOA18)气候态作为基准等). 根据历史观测数据的统计分布特征或概率分布(先验猜测)来界定一个阈值范围, 若某一观测值超过设定的阈值, 则将其标记为可疑值(图4f). 对于正态分布的数据, 可以采用莱茵达准则(极值检验)确定阈值(气候态上下界); 对于偏态分布的数据, 可以采用分位数或者由Vanderviere和Huber(2004)提出的“调整箱型图”等方法确定阈值. 气候态检查本质上属于2.1节中的范围检查, 但相比之下考虑了区域的差异, 使得检查更具精细化、准确化. 根据IQuOD(2019)和Gouretski(2018)的研究, 局地气候态检查是质量控制系统的核心, 相比于其他检查模块, 该模块可以识别至少50%的错情数据.

2.14 温度-盐度图检验(盐-密模态定量分析)

Emery和Meinke(1986)指出由于水团在不同地区和季节往往有较为确定的模式, 可以用温度-盐度气候态曲线($T-S$ 图)来描述水团的局地特征. 因此可以通过将观测的温-盐曲线与事先构建的气候态温-盐曲线进

行比较, 通过从辨识水团特征的角度去检验资料中含有的异常信息来标记可疑数据. 其本质上属于气候态范围检查的范畴(汪德元, 1988; 纪风颖等, 2004).

当然, 质量控制的技术方法远远不止文中所提及的这14种, 有关其他未提及或者未被应用的技术方法, 可以进一步参考Cummings(2011).

3 国内外研究现状

本节将着重对国内外现有的主流质控系统进行归纳整理并做简要介绍和对比, 并讨论不同系统之间的差异. 由于延时质量控制的子模块检查一般包含了实时质量控制子模块, 因此本节将主要讨论的是延时质控方法(除特别说明外), 不再刻意区分两者之间的差异.

目前, 国际上系统开展海洋观测数据质量控制的研究机构有美国国家海洋和大气管理局国家环境信息中心(NOAA/NCEI)、英国气象局哈德莱中心(Met Office Hadley Centre)、全球温盐剖面计划(GTSPP)、Argo全球海洋观测阵列计划、世界大洋环流实验室(WOCE)等. 在中国, 主要的研究机构有杭州全球海洋Argo系统野外科学观测研究站(中国Argo实时资料中心)(刘增宏等, 2016; Liu等, 2017b)、国家海洋信息中心(杨扬等, 2017)、中国科学院海洋研究所(Zhang和Wang, 2019)等. 表1对上述机构所涉及的质量控制子模块进行了系统的整理和归纳. 从表1中不难发现:

(1) 各个机构子模块检查顺序不尽相同. 但通常为了提高效率, 几乎所有的研发机构中都把预检查置于质量控制系统中的第一部分.

(2) 范围检查作为预检查的进一步延伸, 在大多数研发机构当中, 都作为质控检查的第二个环节. 而剖面检查、仪器类型检查、局地最大深度检查作为进一步延伸, 通常将其作为第三个环节, 置于范围检查之后, 并且这三大检查之间也无明显的先后顺序. 而统计特性检查, 尤其是局地气候态检查和温度-盐度图($\theta-S$ 检查), 由于其可以通过概率分布的方法有效识别异常剖面, 通常作为质量控制中的重中之重置于质量控制中的最后一个环节.

(3) 日期时间检查、经纬度检查、陆地位置检查、深度递增检查、尖峰检查、全球范围检查在所有的机构当中出现的频次最多; 局地最大深度检查、流

速检查、垂直梯度检查、等值检查、局地气候态检查、冰点检查、极值检查、盐度漂移检查的出现频次依次减少。

(4) 有一些检查只适用于特定的变量类型。如 Waterfall 检查(即相邻廓线比对)只适用于盐度变量(UNESCO-IOC, 2010)。

(5) 有一些检查只适用于特定的仪器类型。如盐度/深度漂移检查只适用于Argo浮标获取到的盐度观测数据, 热滞后校正只适用于CTD获取到的盐度观测数据。

(6) 有一些检查一般只应用于特定的站面类型。如流速/速度检查、拟合优度检查、莱茵达准则、测站识别码检查只适用于浮标、海岛站等具有定点、长时间连续观测特征的数据, 不适用于走航观测获取到的数据。

(7) 考虑到不同的机构开展质量控制研发的目的不一致, 有一些检查模块只在某一研发机构当中单独出现, 例如, 盐度/深度漂移检查只出现在了Argo的质控系统中, 这主要是因为Argo浮标由于仪器的老化, 在长时间使用之后, 会出现盐度、深度漂移的现象, 而这种现象, 在其他仪器和机构当中, 是相对较少的。

当然, 表1中也可以发现许多差异和区别。例如, 细网格、高分辨率的局地气候态检查作为一种较严格的质控方法, 一直被认为是可以识别异常值的最有效、最重要的方法(Gouretski, 2018)。但是该方法没有出现在WOD数据集和Argo实时数据集的质控中(WOD使

用的是 $5^{\circ} \times 5^{\circ}$ 粗网格莱茵达准则), 这主要是因为这些数据集的使用目的和受众不仅面向不同的学科和科研人员, 还面向多学科交叉以及普通大众, 所以这些数据集相对于其他数据集而言, 需要牺牲一些准确率以尽可能多地保留真实有效的数据。因此有些科学家有充足的理由认为WOD的方法是一种较为宽松的质控方法(如图3b所示, 经WOD质控之后的温度数据仍然有不少的错误数据没有被剔除)。相反, WAGHC-ICDC的质量控制使用了月平均的 $0.5^{\circ} \times 0.5^{\circ}$ 的高分辨率气候态数据进行检查, 因此该质量控制相较于WOD和Argo的质量控制而言更加严格: 通过宁愿牺牲一些优良数据的思路也要以尽可能多地标记错误数据。

此外, 在质量控制标识符(flag)方面, 不同的质控研发机构之间存在着不同思路的标记方式(表2): 有0和1的二分类方式(如WAGHC-ICDC), 也有多分类方式(如GTSP、Argo), 还有以质控的名称进行分类标记的方式(如WOD18)。因此, 有必要努力寻找一致的、协调的、客观的质量控制标识符(Domingues和Palmer, 2015), 以期形成统一的质量控制标识的行业标准。

总而言之, 质量控制因其使用目的不同, 而出现了不同的标识符、不同模块之间的排列组合, 除此之外, 即使是相同名称的子模块, 内部的参数设置也可能未必是一样的, 这就导致了不同的质控方法可能在具有不同特征(时间尺度、空间尺度)的数据集中会表现得不一样。每个质控方法都有它的优势, 也有它的劣势,

表2 不同质量控制机构或数据集提供的对廓线逐个温度、盐度观测值的质量控制标识符(flag)的标记含义^{a)}

标记符	WOD18	Argo	GTSP	WAGHC-ICDC
0	正确值	没有经过质控	没有经过质控	正确值
1	未通过范围检查	正确值	正确值	异常值
2	未通过逆温检查	可能是正确值	可能是正确值	-
3	未通过梯度检查	可能是异常值	可能是异常值	-
4	未通过尖峰检查	异常值	异常值	-
5	未通过逆温检查、梯度检查	被修改过	被修改过	-
6	未通过范围检查、逆温检查	-	-	-
7	未通过范围检查、梯度检查	-	-	-
8	未通过范围检查、尖峰检查	估算值	-	-
9	未通过范围检查、梯度检查、逆温检查	缺测值	缺测值	-

a) WOD18, 美国国家海洋和大气管理局世界海洋数据集NOAA-WOD18; Argo, Argo全球海洋观测阵列; GTSP, 全球温盐剖面计划GTSP; WAGHC-ICDC, 世界大洋环流实验全球水文气候数据集WOCE/WAGHC-ICDC

因此需要将更多资源投入到质量控制程序的比对当中(下一节中详细阐述).

4 质量控制的评估

如第3节所述, 不同机构的质控方法各有差异, 因此有必要开展质量控制系统性能的评估, 以验证质控方法的准确性、可靠性、普适性.

4.1 质量控制评估的重要性

正如每个质控方法都有它的优势, 也有它的劣势, 有的在一些特定的数据集会表现良好, 而有的只适用于某些特定的基础科学研究. 开展质量控制的评估工作, 寻找最佳且最优的质量控制方法, 将有助于建设高质量、高标准的海洋信息数据库, 从而为海洋科学研究提供更加可靠的数据集来源(Wedd等, 2015). 目前, 国内外在AutoQC上已有很好基础, 但是尚未有一种国内外公认的最佳方案. 这主要是因为: (1) 不同机构研发的质量控制的使用目的、侧重点(科学问题、应用问题)不同; (2) 目前暂无可靠稳健的统一度量标准; (3) 缺少经过人工检验的高质量基准数据集作为开展评估工作的参考数据.

因此, 有必要寻找一个度量标准, 来衡量一组特定的质量控制检查的性能情况. 即如何评估一个质量控制方法的性能? 其基准指标是什么? 针对这一问题, 目前国际上已有一些组织尝试开展了质量控制的评估. 例如, 国际质量控制海洋数据库计划(IQuOD)通过制定国际可接受、一致的质控标准来协调每个温度数据集的基础数据质量(Domingues和Palmer, 2015; Cowley等, 2021), 其主要的任务包括进一步发展自动化质控技术(AutoQC)、专家质控技术(ExpertQC)、排重技术(duplicate checking)、颁布统一的质量控制符(quality flag), 以及统一的数据格式等, 最终发布一致性和完整性的无偏海洋观测数据(Goni等, 2019). 而如果落脚到具体的科学问题上, 例如对于海洋热含量监测, Argo全球海洋观测组网系统发展的目标也是旨在寻找国际认可的适用于监测海洋热量含量的数据质量控制体系(Argo Science Team, 1999). 因此, 不管是从发展质量控制技术本身而言, 还是从解决实际科学问题的角度而言, 开展质量控制性能的评估工作, 具有必然性、紧迫性和先进性.

4.2 质量控制评估的方法

质量控制评估方法依据侧重点不同, 可以分为基于独立质控系统的综合评估(4.2.1节)和基于独立数据的对比评估(4.2.2节). 前者为定量评估, 后者为定性评估.

4.2.1 基于独立质控系统的综合评估

综合评估是指针对包含多个质量控制子模块的大型质控系统开展的综合评估.

由IQuOD(2019)和Mieruch等(2021)使用的去伪概率(true positive rate, TPR)和弃真概率(false positive rate FPR)为开展综合评估提供了可行的思路. 运用该指标, 可以通过使用高质量的经人工审查的质量控制基准数据集作为真值(例如由Thresher等(2008)研发的QuOTA数据集)去评价质量控制系统的性能, 这是一种直接而又定量的方法. 其中去伪概率(TPR)的定义如下(Yerushalmy, 1947):

$$TPR = 100\% \times \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (3)$$

式中, N_{TP} 为坏的数据被正确识别出来的总个数(true positive), N_{FN} 为坏的数据没有被正确识别出来的总个数(false negative). TPR的物理意义可以解释为“有多大比例的坏数据被正确剔除”. 显而易见, 质量控制程序应尽可能多地拒绝坏数据, 因此TPR应越大越好. 若TPR越小, 则代表去伪率偏低(即系统过于宽松、漏判数量过多).

而弃真概率(FPR)的定义如下(Yerushalmy, 1947):

$$FPR = 100\% \times \frac{N_{FP}}{N_{TN} + N_{FP}}, \quad (4)$$

式中, N_{TN} 为好的数据被真实保留的总个数(true negative), N_{FP} 为好的数据被判别为错误数据的总个数(false positive). FPR的物理意义可以解释为“有多大比例的好数据被错误剔除了”. 显而易见, 质量控制程序应尽可能多地保留好数据, 因此FPR应越小越好. 若FPR越大, 则代表弃真率偏大(即系统过于严格、错判数量过多).

基于上述定义, 一个最优的质量控制系统, 应该是去伪概率尽可能地大, 同时弃真概率尽可能地小. 如图5所示, 基于该计算方法, 通过对多个国际上主要的质量控制机构(NOAA大西洋海洋和气象实验室

AOML/Argo实时质量控制系统Argo-RTQC/澳大利亚联邦科学与工业研究组织CSIRO/WAGHC-ICDC)的温度自动化质量控制系统进行综合评估,发现AOML和ICDC的去伪概率(TPR)最高.但同时,不难发现AOML和ICDC虽然去伪概率最高,但是弃真概率也比较高,表明AOML和ICDC(在针对某些数据集中)是较为严格的质量控制程序,即在去除众多错误的数据的同时也把少部分的优良数据给去除.实际上,目前即使是最优的质量控制系统,也很难保证去伪概率接近于100%的同时弃真概率接近于0%.

该方法的优点是普适性比较强,只要有足够多的样本(基准数据集),就能较好地反应质量控制程序的性能差异.但是正是因为需要经人工判断(专家质控)的数据集作为样本,所以基准数据集的不完美也会对结果造成一定的影响.因此,需要更多客观且高精度的基准数据集作为参考.

4.2.2 基于独立数据的对比评估

不同的质控系统之间的比较,还可以通过和其他独立的数据进行对比,通过比较质控前和质控后数据集的差异,或对同一个数据集使用不同的质控方法来理解质控系统的性能,这是一种间接而又定性的方法.

例如,由于海洋变暖导致的比容海平面上升是区域海平面变动的主要原因,因此可以通过用质控后的观测数据计算海洋热含量(OHC),进而推算区域海平面高度的变化.通过其与卫星高度计直接观测的海平面高度对比,可揭示OHC的准确性,进而评估所使用特定质量控制方案(方案)的有效性.已有研究表明,对于用温度估算过去海洋热含量变化趋势,质量控制的性能是影响海洋热含量估算的一个重要因子(Boyer等, 2016).那么,在保证其他因子不变的情况下,以海平面高度变化的观测作为真值,用不同机构的质控系统计算出来的热含量去推算海平面高度变化,来探讨不同质控系统下的结果差异是存在可能的.如图6所示,当采用不同的质量控制方案的数据集在估算全球海洋热含量异常(OHCA)时,虽然在大尺度(海盆尺度)上差异不是很明显,但是在分辨海洋中小尺度信号上具有不同的表现:位于热带东太平洋地区的海洋热含量冷异常在采用WOD质量控制的数据集里会比在采用EN3v2a质量控制数据集里要大,最大可相差 $5 \times 10^{19} \text{J}$;且使用EN3v2a质量控制可以分辨出比WOD质量控制

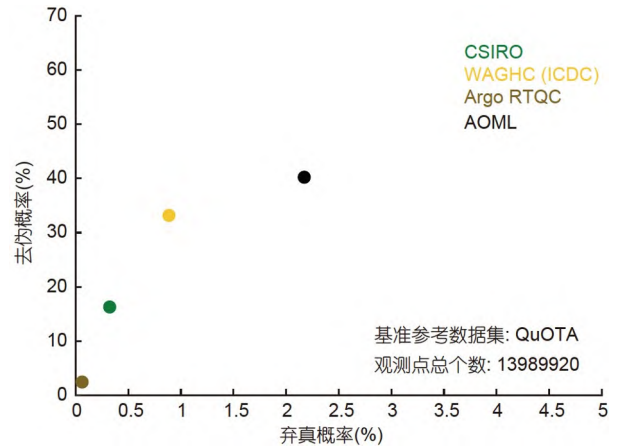


图5 采用QuOTA温度观测基准数据集对澳大利亚联邦科学与工业研究组织(CSIRO)、世界大洋环流实验全球水气候数据集(WAGHC-ICDC)、Argo实时质量控制(Argo RTQC)、NOAA大西洋海洋和气象实验室(AOML)的温度质量控制性能对比评估

横坐标为弃真概率(FRR),纵坐标为去伪概率(TPR)

更多的小尺度特征(Boyer等, 2016).这给独立数据进行对比评估提供了一种可行的思路.

通过面向特定的基础科学问题,用独立的数据作对比是一种可行的思路.该方法的优点是可以透过具体的科学问题来评价不同的质量控制方法(基于数据的结果是否能更准确地反映已知的物理现象).但是缺点是:它是一种间接的方法,无法做到直接的、定量的评估,并且该评价的结果可能只适用于特定的科学问题,普适性不强.

5 质量控制技术的不足与展望

如何快速、准确识别数据错情,是现代海洋科学研究中数据处理面临的难点.随着未来海洋观测数据的持续性增长,维护和推进海洋数据管理,发展和改进海洋观测的质量控制,是跟上不断增长的数据流、构建具有实时联动的海洋数据库的新要求.本节将首先简述质量控制技术存在的不足之处,在此基础上提出一些改进海洋观测数据质量控制的可能建议,以推进质量控制技术进一步向前发展.

5.1 质量控制中存在的不足

5.1.1 统计特性检查的不足

统计特性检查(例如莱茵达准则、局地气候态检

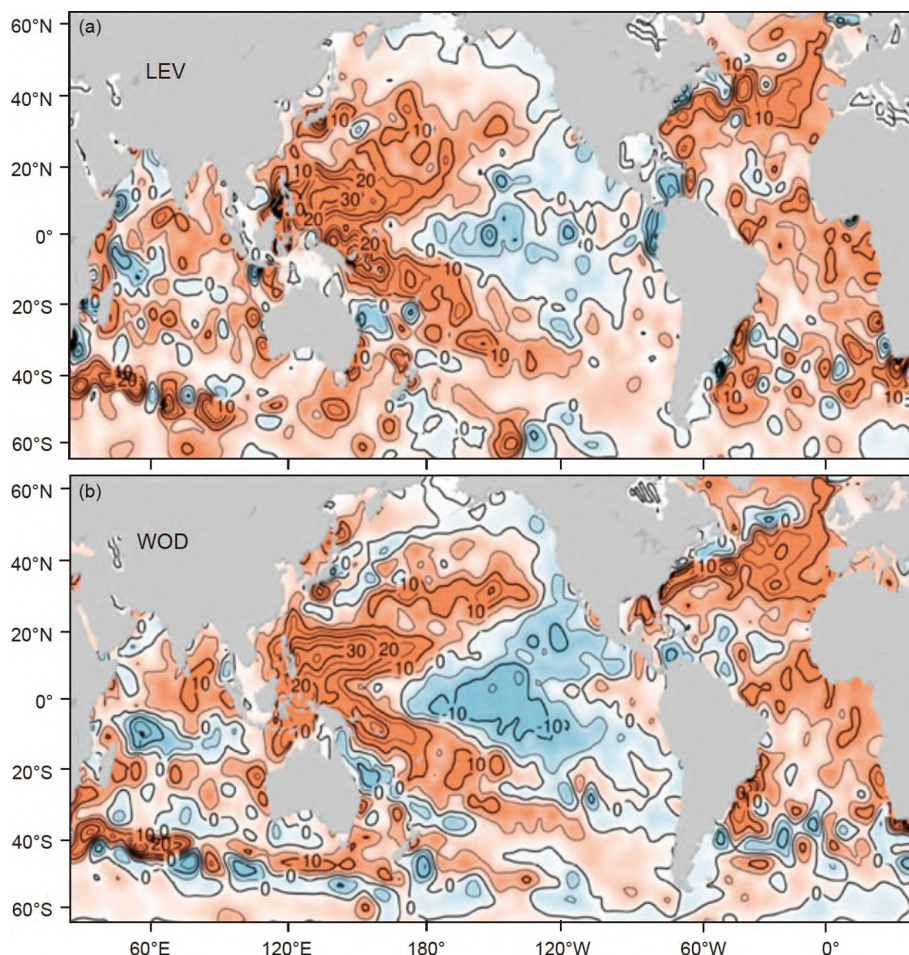


图 6 全球海洋热含量异常(OHCA)分布

两张图均使用相同年份(2008年)相同数据来源(WOD09)的数据和相同的格点化(LEV mapping)方法,但是(a)使用的是EN3v2a的质量控制方法,而(b)使用的是WOD质量控制方法.等高线间隔为 $5 \times 10^{18} \text{J}$,红色表示正异常,蓝色表示负异常.深灰色阴影部分表示未计算OHCA的区域.图引自Boyer等(2016)

查、温度-盐度图检查)通过概率分布的方法,来判断异常剖面的存在,其可以抽象为阈值问题.在21世纪初Argo浮标被广泛使用之前,海洋现场观测的数据量较少,没有办法做到大规模的统计(如概率密度分布函数),因此无法真实并且尽可能准确地还原一个区域的数据分布特征.而随着步入大数据时代,Argo等海量数据实时进入数据库,因此可以通过大数据的方法进行大规模的统计,实现(区域)水团分布的特征的重构.以往,较常规的方法是设定一个概率阈值(比如正态分布的 $\pm 3\sigma$ 准则,若观测值落在 $\pm 3\sigma$ 之外,则可以判断为异常值),但是这种方法也有它的局限性,例如一些区域的观测不满足近似正态分布的要求,因此也有方法使用了统计分位数的方法确定阈值(Gouretski, 2018).

而局地气候态检查一直被认为是识别异常观测剖面最有效的方法,该方法一般是通过将待检查的剖面 and 预先构建的历史气候态数据(背景场)做比较,若超过一定的阈值,则标记为可疑值(参阅第2节).因此,一个准确无偏的背景场对于局地气候态检查而言非常重要.在前人关于构建网格化的历史气候态数据的研究当中,一般受限于观测剖面的数量:在一些较少观测甚至是无观测的区域,无法准确构建该地区的历史气候态(Cheng和Zhu, 2015).目前,国际上构建数量少(无)观测区域的历史气候态数据(例如WOA18)的方法一般是采用反距离加权平均的方法来进行历史气候态的重构(Lyman和Johnson, 2014; Garcia等, 2018).但是该方法存在如下可能的缺陷:(1)使用该方法所假设的

前提是海洋环流(水团)性质为各向同性,但是实际上海洋环流更多的是各向异性,也就是重构的区域可能反映的是两个甚至多个不同特性的水团。(2)该方法没有考虑环流在特殊地形下的空间不连续,存在跨岛屿选取了空间上不连续的水团的现象,这显然不符合海洋环流的特征,这一点在一些岛屿林立、地形复杂、环流多样的边缘海区域显得尤其重要。(3)重构的数据在空间上无全覆盖:例如,WOA18的气候态的统计平均场、统计标准差场无法做到空间上的全覆盖,尤其在2000m以下的深海区域(UNESCO-IOC, 2010; Gouretski, 2018; Locarnini等, 2019)。因此,如果无法准确地还原一个区域的数据分布特征,那么将会影响到构建局地气候态检查阈值的准确性。

5.1.2 梯度检查的不足

过去的垂直梯度检查一般是通过统计学的方法粗略设定梯度的阈值(例如WOD-QC、ICDC-QC),也有通过基于廓线的滑动中位数和滑动标准差来设定阈值(即MEDSTD方法)。但是这些方法的稳健性通常不高,往往存在较高概率的漏判和错判。例如,这些方法一般只适用于检查单个观测点的梯度突增或突减的情况(如尖峰),而对于多个点的连续突增或连续突减的情况(如廓线“抖动”、“阶梯状”廓线)则表现不佳。此外,一些边缘海地区(例如马尔马拉海、阿曼湾)存在深海强逆温的现象(Barton和Gonzalez, 2016),通常在这些海区的廓线往往会在梯度检查中发生错判,其本质上可以归结为不完备的梯度检查。因此,Argo数据管理团队提出了使用距离中值(median distance, MMED)的方法替代梯度检查(Cabanes等, 2021)。其相对于传统的梯度检查而言,虽然在运算速度上比较耗时,但是能够结合密度观测大幅提高准确性和降低错判概率。目前,MMED方法已在Argo的实时质控系统(RTQC)当中取代原本的梯度检查(Wong等, 2020)。即便如此,在实际运用当中,仍会发现该方法在识别其他仪器(如XBT、MBT)的异常廓线中则表现略逊一筹,这主要是因为MMED阈值的选取依赖于同时刻的海水密度的数据,而大多数XBT、MBT仅有温度的观测。

而在海洋大面观测当中,廓线数据的很多问题,若反映在检查模块的分类当中,本质上都可归结为廓线的梯度(形状)问题。例如尖峰检查,本质上是梯度出现了大幅突增(突减)的情况(图4c);逆温检查,本质上是

梯度出现了不合理的负值(图4d);而等值检查,本质上是梯度出现连续为0的情况(图4b)。从某种意义上,改进尖峰检查、逆温检查、等值检查、距离中值检查(MMED),本质上可以归结为改进梯度检查。因此,未来可以考虑构建一个稳健的局地(温度、盐度)梯度气候态检查,以期提升现有的梯度检查模块的性能。

5.1.3 仪器特性检查的不足

如表1所示,目前的仪器特性检查,只有部分研发机构中包括了仪器最大深度的检查和Argo的盐度、深度漂移特性检查。但实际上,不同的观测仪器可能有自己独有的仪器内部设计的问题,这就会导致有一些异常剖面,只会出现在特定的仪器当中。例如,除了普遍知道的Argo浮标有盐度、深度漂移等现象之外(吴晓芬等, 2019),XBT也可能存在导线漏电、电线拉伸而出现异常剖面等现象(Kroner和Blumenthal, 1978; Cheng等, 2016; Tan等, 2021)。此外,CTD可能会有压力数据偏差、盐度数据热滞后的现象,当然,MBT、APB可能也有由其仪器导致的数据偏差问题(Gouretski和Cheng, 2020)。在过去,识别这些异常的剖面,通常是通过人工目视审查(ExpertQC)的方式,逐一对剖面进行质控,需要消耗大量人力、物力、财力。目前,除了Argo仪器以外,国内外几乎没有一个机构有关于特定仪器类型的自动化检查(表1)。因此,在日益增长的数据时代,有必要进一步发展针对仪器特性的检查。

除了上述几个可能存在的不足之外,海洋观测数据的质量控制子模块还有很多需要改进的地方,尤其是对于不同的质量控制,没有一个国际公认的评判标准去评估其优劣(第4节所述)。因此需要集中更多的人力、物力、财力,在一定的客观评判标准之下,进一步解决质量控制中的不足(5.2节)。

5.2 质量控制技术的改进与展望

5.2.1 改进质量控制子模块检查顺序

不同子模块检查的先后顺序可能会造成不同的结果(UNESCO-IOC, 2010)。同时,目前几乎所有的子模块都是互相独立,没有前后依赖关系(表1)。因此,有必要测试子模块出现的先后顺序,以评估什么样的排列组合可以使质控结果最优、运算最快。另外,是否有可能将前后的子模块合并在一起,也是可能进一步优化质控系统的思路。

5.2.2 推进海洋生物、海洋化学要素的质量控制研发工作

目前, 海洋观测质量控制主要集中在温度和盐度这两个物理要素, 流速和溶解氧的质量控制次之。而对于海洋生物、海洋化学方面等要素观测, 例如pH、二氧化碳分压、叶绿素、营养盐、初级生产力等要素的质量控制则少之又少。目前, 仅Argo数据管理团队、美国国家海洋与大气管理局综合海洋观测系统(NOAA-IOOS)在叶绿素观测、光学观测、海浪观测、营养盐观测、浮游植物观测中有少量研究(Catherine等, 2014; U.S. Integrated Ocean Observing System, 2015a, 2015b, 2017a, 2017b)。但是, 随着近几年生物地球化学Argo(BGC-Argo)等大型浮标在开放大洋的广泛布放, 随着中国近海生态要素监测逐步完善, 生化方面的观测数据量将会逐渐增多, 传统依靠专家质控的方法已不适用。因此, 研发适用于中国国情的海洋生化要素观测的质量控制系统尤为重要和紧迫。

5.2.3 提升统计特征检查的性能

如5.1.1节所述, 提升统计特征检查, 尤其是局地气候态检查, 可以从以下四个方面入手构建精确的气候态背景场(阈值), 以辅助基于气候平均态检查的质量控制: (1) 考虑海洋环流的各项异性的特征, 提高气候态数据的准确性。(2) 考虑岛屿和海底地形之间的差异, 避免气候态数据空间上存在不连续。(3) 考虑数据的概率分布特点: 统计样本的频率分布可能不是严格的正态分布, 而是偏态分布。

5.2.4 运用机器学习方法辅助质量控制

随着海洋观测组网和每天有越来越多的海量(TB级)观测数据进入到数据收发中心并发布于全球, 这就使得专家质量控制的人力资源负担越来越重。例如, Argo数据的延时质量控制系统中包括了经人工审查的过程(即延时质量控制), 而随着每天有大约400GB的剖面数据增长(Roemmich等, 2019), 这其中的人力、物力、财力、时间资源的消耗性越来越大, 因此国内外主要质控机构也在考虑是否可以通过使用机器学习等方法来解决这一负担, 从而提升质控的精度和普适性。此外, 传统的自动化质量控制系统(AutoQC)也不是十全十美, 仍然可能存在很多没有识别到的异常数据(漏

判), 或者有些正确的数据被错误地标记为异常数据(错判)。

在这样的背景之下, 通过机器学习(machine learning, ML)的方法融合到现有的质量控制技术成为了可能。机器学习的本质是对数据进行分类、识别和预测(Schapiro, 2003; Lebedev和Nicolelis, 2006; Pereira等, 2009), 而质量控制的本质是对海洋观测数据进行识别和分类(“好”、“坏”数据)。鉴于机器学习与质量控制这两者间的共性, 因此可以将机器学习的思维运用到海洋观测的质量控制当中。

在技术方面, 机器学习任务所涉及的分方法, 例如神经网络(neural network)、聚类(clustering)、特征识别(feature selection)、决策树(decision tree)、贝叶斯学习(Bayesian learning)、深度神经网络(DNN)等技术均可尝试用于对海洋观测数据的质量进行分类。在训练样本方面, 目前已经经过人工审查(专家质控)的高质量数据(例如SeaDataNet、CoTeDe数据、Argo DMQC数据)都可以作为样本的训练集和测试集, 进行监督学习, 这将有助于了解异常剖面的特点并在训练中设计适当的参数。目前, 国际上已逐步开展了基于机器学习的海洋数据质量控制方面的分类标记尝试(Ono等, 2015; Castelao, 2020)。例如, Smith等(2012)开展了基于动态贝叶斯网络(DBN)对海洋传感器数据质量的在线评估工作, 该方法的优势在于不需要考虑传感器的硬件条件、部署情况、现场测量情况等较难以量化的因素。在这之后, 自组织映射(SOM)、条件随机场(CRF)、支持向量机(SVM)、决策树、方法签名(signature method)等多种机器学习技术也不断被应用于探索Argo浮标数据的质量控制(Udaya Bhaskar等, 2017; Sugiura和Hosoda, 2020)。与基于传统统计学方法不同, 基于机器学习的方法一般只依赖于不同地区的廓线形状, 例如方法签名(signature method)可以把廓线形状看成是不同的非线性函数进行相应的监督学习, 其优势是不需要考虑传统统计学方法的各种限制性约束条件(如样本数量、经纬度、深度等)对结果的影响(Sugiura和Hosoda, 2020)。此外, 融合传统的自动化质量控制技术和人工智能技术的海洋数据质量控制的深度学习的思路也逐渐被提及(Leahy等, 2018; Castelao, 2020)。例如, Mieruch等(2021)将人工神经网络技术运用到传统的质量控制算法当中, 在地中海大型数据集的测试中, 能够正确检测出超过90%的错误温度

观测数据. 这些尝试性工作可以作为未来运用机器学习来改进质量控制的基础和范本.

但是, 目前的机器学习辅助质量控制的研究仍处于起步阶段, 暂未大规模运用到海洋数据收发中心等实时业务机构, 而且研究的质控变量较少, 主要集中在温度观测中. 此外, 目前还没有公认的可用于机器学习的训练集和测试集. 另外, 在有监督学习的基础上, 可以考虑使用无监督学习的方法提高质控系统的性能.

5.2.5 基于海洋物理过程改进质量控制

Foken等(2005)指出在考虑海洋物理过程的基础上, 质量控制可以通过使用其他独立的数据(例如对温度进行质量控制, 则可以使用非温度的观测数据进行校准), 对质控算法进行改进. 使用独立数据的前提是数据之间的内在联系必须是建立在物理过程基础之上的.

例如, 如果需要对温度进行质量控制, 考虑到海洋浅层运动与海面大气运动存在着物理机制上的联系, 那么海表面风速、风向的观测、海表净热通量的观测、台风多要素的观测、海表面湿度的观测可以用来进行辅助判断(Hall等, 1991). 基于上述的大气与海洋之间的联系, 一些看似错误的的海况数据, 其实可能是海况的真实反映, 这些数据应予以保留, 例如:

(1) 台风: 台风过境时的观测数据, 如海表面风速、SST数值、上层10m近表层温盐数据和海流数据都可能会发生急剧变化. 例如, 有学者指出大多数台风过境的时候, 海表和混合层的盐度、温度将会显著降低, 甚至能达到 5°C 和1psu的变化幅度(Liu等, 2007). 这些数据在常规的检查中会有较大概率落入到阈值范围之外, 被错误地标记为异常数据, 但实际上, 这些数据大部分都是异常海况的真实反映. 因此, 有必要结合历史上独立的台风(飓风)的观测资料(如台风路径附近海域的海表面10m气温、风场), 在考虑与海洋浅层温度、盐度的物理联系基础之上, 仔细鉴别看似错误但实际上是正确的温、盐数据.

(2) 强寒潮: 在季风海区(如中国南海、中国东海、中国黄海、中国渤海、墨西哥湾等), 由于强寒潮南下过境, 可能会导致海洋浅层(混合层为主)的急剧(短时间)降温, 甚至在低纬度海区出现逆温现象. 应配合气象机构提供的已有寒潮数据图集, 仔细鉴别正确

的异常温、盐数据.

5.2.6 发展平台端数据质量控制技术

目前的质量控制技术多集中在数据收发中心或用户端(终端)进行的, 而在获取第一手资料的平台端进行的质量控制或数据处理则比较少. 但是随着科技的进步, 现在的海洋观测手段不仅仅满足于固定平台的观测, 而是快速向水面/水下漂流等机动式平台发展. 多类型的平台端的发展给质量控制技术提出了快速、实时、易操作的更高要求. 现有的平台端数据质量控制技术仅仅维持在一些较为简单的处理, 例如进行简单的抽样或平均, 或对一些明显错误的数据进行人为快速标记或剔除(例如: 刘淑雅等, 2013; 王丽芳等, 2021). 因此, 如果能在平台端发展更高质量、更快速的质量控制技术并开展数据处理, 同时借助现场观测人员的经验、对现场和元数据情况的全面了解和快速判断, 将极大提高数据的准确性和可利用价值.

6 总结

本文简要介绍了与质量控制相关的技术方法、阐述了国内外的研究现状, 随后讨论了质量控制技术性能的评估, 最后阐述了质控方法的不足、改进和展望. 主要认识总结如下:

(1) 自海洋科学进入大数据时代以来, 国内外在质量控制研究中, 尤其是对温度、盐度的质量控制有了较为深入的认识和发展, 研究框架也逐渐成熟, 研发了大量适用于不同场景的质量控制系统.

(2) 延时质量控制技术(DMQC)是在实时质量控制技术(RTQC)的基础上发展的一个更加精准的环节, 而自动化质量控制技术(AutoQC)与专家质量控制技术(ExpertQC)则是相辅相成. 不同的质控分类方法本质上都是为了提高数据质量, 都应予以重视. 而质控环节中涉及的不同子模块, 均有其优点和缺点. 开展质量控制性能的综合评估和基于独立数据的对比评估工作, 在对比不同机构的质控方法中通过调试、改进、发展不同子模块的性能将有助于提升质控的稳健性和普适性.

展望未来, 海洋观测数据的质量提升还有很大的发展空间, 例如: (1) 提升传统质量控制检查模块的性能, 开展更广泛的质量控制系统性能的评估工作; (2)

推进溶解氧、二氧化碳分压、pH、叶绿素、营养盐等要素的质量控制研发工作; (3) 开展基于人工智能(机器学习)的质量控制研发工作; (4) 推进质量控制深度融合到实时、联动的海洋数据库业务化工作中, 为中国海洋实时监测预警提供坚实支撑。

自19世纪50年代有观测记录以来, 海洋科学历史数据已积累了大约170年。随着观测组网的越来越丰富, 观测数据量也随之越来越大, 建立中国高质量的海洋科学数据库(尤其是物理海洋、海洋生物化学方面的参数)十分迫切和重要。而若要建立高质量的海洋科学数据库, 则必须首先体现在数据的高质量上。质量控制的目的就是要严把进入数据库的资料质量, 确保资料的权威性。因此, 一个科学、全面、可持续和高质量的数据管理工作是进行数据质量控制的基础。只有提供高质量的数据, 才能更好地推动海洋科学及其相关交叉学科的发展。

最后, 我们需要加强国际间的合作, 在一定框架下分享开源的质量控制程序, 提供良好的人才资源和经费支持, 服务于应用性能更优的海洋(实时)数据库管理, 这是“数据驱动科学”最好的体现之一。

致谢 感谢国家海洋信息中心、中国科学院海洋科学大数据中心、国际Argo数据管理团队、美国国家海洋与大气管理局(NOAA)、英国气象局哈德莱中心等机构提供的数据支撑和帮助。感谢为本文内容提供建设性建议的匿名审稿人。更要感谢所有奋斗在海洋观测、海洋数据共享服务、海洋监测预警一线的科研工作者。

参考文献

陈洪滨, 李军, 马舒庆, 胡树贞. 2019. 海洋气象观测技术研发进展. 科技导报, 37: 91-97

纪凤颖, 苗春葆, 张书东. 2004. 基于气候性温盐关系模型对Argo数据进行质量控制的研究. 海洋通报, 6: 8-15

李慧青, 李燕, 刘玉, 王宁, 王强. 2015. 海洋科学观测及其技术发展浅析与思考. 海洋开发与管理, 32: 63-67

李学坤, 李凤金. 1997. 海洋水文气象实时数据质量控制. 海洋预报, 3: 72-80

刘首华, 陈满春, 董明媚, 高志刚, 张建立, 武双全, 林峰竹. 2016. 一种实用海洋浮标数据异常值质控方法. 海洋通报, 35: 264-270

刘淑雅, 刘春颖, 延鹏, 黄政然, 彭彤, 周玉娟, 丁海兵. 2013. 胶州湾及青岛近海海水pH值的分光光度法研究. 海洋湖沼通报, 3: 108-

114

刘帅, 陈戈, 刘颖洁, 田丰林. 2020. 海洋大数据应用技术分析与趋势研究. 中国海洋大学学报(自然科学版), 50: 154-164

刘增宏, 吴晓芬, 许建平, 李宏, 卢少磊, 孙朝辉, 曹敏杰. 2016. 中国Argo海洋观测十五年. 地球科学进展, 31: 445-460

齐庆华. 2020. 海洋中尺度过程与区域海洋环境和气候安全. 海洋开发与管理, 10: 12-20

邱国强, 王海黎, 邢小罡. 2018. BGC-Argo浮标观测在海洋生物地球化学中的应用. 厦门大学学报(自然科学版), 57: 827-840

任焕萍, 张斌, 谭哲韬, 李富超. 2021. 一种精细化的海洋浮标数据质量控制方法. 海洋科学, 45: 93-103

任杰. 2019. 海洋观测技术(高等院校海洋科学专业规划教材). 广州: 中山大学出版社. 251

史静涛, 周智海. 2011. 海洋站数据质量控制技术探讨. 海洋技术, 30: 114-117

汪德元. 1988. 盐密模与浅海水文资料质量控制. 海洋通报, 3: 67-71

王辉赞, 张韧, 王桂华, 安玉柱, 金宝刚. 2012. Argo浮标温盐剖面观测资料的质量控制技术. 地球物理学报, 55: 577-588

王丽芳, 黄韬, 杜川军, 郭香会. 2021. 不同海水营养盐现场连续观测系统的比较研究. 热带海洋学报, 40: 103-113

温玉波. 2014. 海洋环境观测数据的质量控制研究. 农业网络信息, 2: 35-38

吴晓芬, 周慧, 曹敏杰, 刘增宏, 孙朝辉, 卢少磊. 2019. 船载CTD仪与自动剖面浮标观测资料质量初探. 海洋与湖沼, 50: 278-290

解静, 常江, 孙家文, 祁祥礼, 范东华, 郭松林. 2020. 海洋水文观测实时共享技术与应用研究. 海洋环境科学, 39: 302-308

杨扬, 苗庆生, 韦广昊, 董明媚, 东成. 2017. 海洋站观测资料的质量控制方法及其应用. 海洋开发与管理, 34: 109-113

于仁成, 吕颂辉, 齐雨藻, 周名江. 2020. 中国近海有害藻华研究现状与展望. 海洋与湖沼, 51: 768-788

张斌, 王凡, 成里京, 王鑫, 张冬生, 赵苑, 李学刚, 陈昇, 王秋璐. 2017. 海洋环境变化关键参数观测数据处理方法和产品研制. 中国基础科学, 19: 6-11

Argo. 2000. Argo Float Data and Metadata from Global Data Assembly Centre (Argo GDAC). SEANOE

Argo Science Team. 1999. Report of The Argo Science Team Meeting (Argo-1). Melbourne: GODAE International Project Office

Arthur R S. 1965. On the calculation of vertical motion in eastern boundary currents from determinations of horizontal motion. *J Geophys Res*, 70: 2799-2803

Bailey R, Gronell A M, Phillips H, Tanner E, Meyers G. 1994. Quality Control Cookbook for XBT Data. CSIRO Marine Laboratories Reports 221. Hobart: CSIRO Marine Laboratories

Balmaseda M A, Trenberth K E, Källén E. 2013. Distinctive climate signals in reanalysis of global ocean heat content. *Geophys Res Lett*, 40: 1754-1759

- Barton Z, Gonzalez I. 2016. AOML High Density XBT System Setup Instructions and Troubleshooting Manual. Washington: National Oceanic and Atmospheric Administration
- Boyer T, Domingues C M, Good S A, Johnson G C, Lyman J M, Ishii M, Gouretski V, Willis J K, Antonov J, Wijffels S, Church J A, Cowley R, Bindoff N L. 2016. Sensitivity of global upper-ocean heat content estimates to mapping methods, XBT bias corrections, and baseline climatologies. *J Clim*, 29: 4817–4842
- Boyer T P, Levitus S. 1994. Quality Control and Processing of Historical Oceanographic Temperature, Salinity, and Oxygen Data. NOAA Technical Report NESDIS 81. Washington: National Oceanic and Atmospheric Administration
- Brunton S L, Kutz J N. 2019. Data-driven Science and Engineering: Machine Learning, Dynamical Systems, and Control. Cambridge: Cambridge University Press. 320
- Bushnell M. 2016. Quality assurance/quality control of real-time oceanographic data. In: Proceedings of the OCEANS 2016 MTS/IEEE Monterey, Monterey
- Cabanes C, Angel-Benavides I, Buck J, Coatanoan C, Dobler D, Herbert G, Klein B, Maze G, Notarstefano G, Owens B. 2021. DMQC cookbook for core Argo parameters. <https://doi.org/10.13155/78994>
- Castelao G P. 2020. A framework to quality control oceanographic data. *J Open Source Softw*, 5: 2063
- Catherine S, Herve C, Antoine P, Fabrizio D O. 2014. Bio-Argo Quality Control Manual for The Chlorophyll—A concentration. IFREMER
- Cheng L J, Zhu J. 2015. Influences of the choice of climatology on ocean heat content estimation. *J Atmos Ocean Tech*, 32: 388–394
- Cheng L J, Trenberth K E, Fasullo J, Boyer T, Abraham J, Zhu J. 2017. Improved estimates of ocean heat content from 1960 to 2015. *Sci Adv*, 3: e1601545
- Cheng L J, Abraham J, Goni G, Boyer T, Wijffels S, Cowley R, Gouretski V, Reseghetti F, Kizu S, Dong S, Bringas F, Goes M, Houpert L, Sprintall J, Zhu J. 2016. XBT science: Assessment of instrumental biases and errors. *Bull Am Meteorol Soc*, 97: 924–933
- Cowley R, Killick R E, Boyer T, Gouretski V, Reseghetti F, Kizu S, Palmer M D, Cheng L J, Storto A, Le Menn M, Simoncelli S, Macdonald A M, Domingues C M. 2021. International quality-controlled ocean database (IQuOD) v0.1: The temperature uncertainty specification. *Front Mar Sci*, 8: 689–695
- Cummings J A. 2011. Ocean data quality control. In: Schiller A, Brassington G B, eds. Operational Oceanography in The 21st Century. Berlin: Springer. 91–121
- Daneshzadeh Y H C, Festa J F, Minton S M. 1994. Procedures Used at AOML to Quality Control Real Time XBT Data Collected in The Atlantic Ocean. Miami: Atlantic Oceanographic Meteorological Laboratory
- de Young B, Visbeck M, de Araujo Filho M C, Baringer M O, Black C, Buch E, Canonico G, Coelho P, Duha J T, Edwards M, Fischer A, Fritz J S, Ketelhake S, Muelbert J H, Monteiro P, Nolan G, O'Rourke E, Ott M, Le Traon P Y, Pouliquen S, Sousa-Pinto I, Tanhua T, Velho F V, Willis Z. 2019. An integrated all-Atlantic ocean observing system in 2030. *Front Mar Sci*, 6: 428
- Domingues C, Palmer M. 2015. The IQuOD initiative: Towards an international quality controlled ocean database. *CLIVAR Exch*, 67: 38–40
- Domingues C M, Church J A, White N J, Gleckler P J, Wijffels S E, Barker P M, Dunn J R. 2008. Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature*, 453: 1090–1093
- Emery W J, Meincke J. 1986. Global water masses—Summary and review. *Oceanol Acta*, 9: 383–391
- Foken T, Göckede M, Mauder M, Mahrt L, Amiro B, Munger W. 2005. Post-field data quality control. In: Lee X, Massman W, Law B, eds. Handbook of Micrometeorology. Dordrecht: Springer Netherlands. 181–208
- Fox-Kemper B, Hewitt H T, Xiao C, Aðalgeirsdóttir G, Drijfhout S S, Edwards T L, Golledge N R, Hemer M, Kopp R E, Krinner G, Mix A, Notz D, Nowicki S, Nurhati L S, Ruiz L, Sallée J-B, Slangen A B A, Yu Y. 2021. Ocean, cryosphere and sea level change. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the 6th Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press
- Garcia H E, Boyer T P, Locarnini R A, Baranova O K, Zweng M M. 2018. World Ocean Database 2018: User's Manual. Silver Spring: National Oceanic and Atmospheric Administration
- Goni G J, Sprintall J, Bringas F, Cheng L J, Cirano M, Dong S F, Domingues R, Goes M, Lopez H, Morrow R, Rivero U, Rossby T, Todd R E, Trinanes J, Zilberman N, Baringer M, Boyer T, Cowley R, Domingues C M, Hutchinson K, Kramp M, Mata M M, Reseghetti F, Sun C, Bhaskar T U, Volkov D. 2019. More than 50 years of successful continuous temperature section measurements by the global expendable bathythermograph network, its integrability, societal benefits, and future. *Front Mar Sci*, 6: 452
- Good S A, Martin M J, Rayner N A. 2013. EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J Geophys Res Oceans*, 118: 6704–6716
- Gouretski V. 2018. World ocean circulation experiment—Argo global hydrographic climatology. *Ocean Sci*, 14: 1127–1146
- Gouretski V, Cheng L J. 2020. Correction for systematic errors in the global dataset of temperature profiles from mechanical bathythermographs. *J Atmos Ocean Tech*, 37: 841–855

- Graves S. 2009. CODATA, and international activities. In: Proceedings of the European Geosciences Union General Assembly 2009, Vienna
- Gronell A, Wijffels S E. 2008. A semiautomated approach for quality controlling large historical ocean temperature archives. *J Atmos Ocean Tech*, 25: 990–1003
- Guinehut S, Coataoan C, Dhompas A L, Le Traon P Y, Larnicol G. 2009. On the use of satellite altimeter data in Argo quality control. *J Atmos Ocean Tech*, 26: 395–402
- Hall C D, Ashcroft J, Wright J D. 1991. The use of output from a numerical model to monitor the quality of marine surface observation. *Meteorol Mag*, 120: 137–149
- Ingleby B, Huddleston M. 2007. Quality control of ocean temperature and salinity profiles—Historical and real-time data. *J Mar Syst*, 65: 158–175
- IQuOD. 2019. IQuOD (International Quality-controlled Ocean Database) 6th Annual Workshop Report. http://www.iquod.org/documents/IQuOD2019_Report_v2.0.pdf
- Kroner S M, Blumenthal B P. 1978. Guide to Common Shipboard Expendable Bathythermograph (SXB T) Recording Malfunctions. Naval Oceanographic Office Reference Publication
- Leahy T P, Llopis F P, Palmer M D, Robinson N H. 2018. Using neural networks to correct historical climate observations. *J Atmos Ocean Tech*, 35: 2053–2059
- Lebedev M A, Nicoletis M A L. 2006. Brain-machine interfaces: Past, present and future. *Trends Neurosci*, 29: 536–546
- Liu Y J, Qiu M, Liu C, Guo Z W. 2017a. Big data challenges in ocean observation: A survey. *Pers Ubiquit Comput*, 21: 55–65
- Liu Z H, Xu J P, Zhu B K, Sun C H, Zhang L F. 2007. The upper ocean response to tropical cyclones in the northwestern Pacific analyzed with Argo data. *Chin J Ocean Limnol*, 25: 123–131
- Liu Z H, Wu X F, Xu J P, Li H, Lu S L, Sun C H, Cao M J. 2017b. China Argo project: Progress in China Argo ocean observations and data applications. *Acta Oceanol Sin*, 36: 1–11
- Locarnini R A, Mishonov A V, Baranova O K, Boyer T P, Zweng M M, Garcia H E, Reagan J R, Seidov D, Weathers K W, Paver C R, Smolyar I V. 2019. World Ocean Atlas 2018, Volume 1: Temperature. NOAA Atlas NESDIS 81. Silver Spring: National Oceanic and Atmospheric Administration
- Lyman J M, Johnson G C. 2014. Estimating global ocean heat content changes in the upper 1800 m since 1950 and the influence of climatology choice. *J Clim*, 27: 1945–1957
- Marshall J, Schott F. 1999. Open-ocean convection: Observations, theory, and models. *Rev Geophys*, 37: 1–64
- Mieruch S, Demirel S, Simoncelli S, Schlitzer R, Seitz S. 2021. SalaciaML: A deep learning approach for supporting ocean data quality control. *Front Mar Sci*, 8: 611–742
- Olsen A, Key R M, van Heuven S, Lauvset S K, Velo A, Lin X H, Schirnack C, Kozyr A, Tanhua T, Hoppema M, Jutterström S, Steinfeldt R, Jeansson E, Ishii M, Pérez F F, Suzuki T. 2016. The global ocean data analysis project version 2 (GLODAPv2)—An internally consistent data product for the world ocean. *Earth Syst Sci Data*, 8: 297–323
- Ono S, Matsuyama H, Fukui K, Hosoda S. 2015. A preliminary study on quality control of oceanic observation data by machine learning methods. In: Proceedings of the 18th Asia Pacific Symposium on Intelligent and Evolutionary Systems. Berlin: Springer. 679–693
- Palmer M D, Roberts C D, Balmaseda M, Chang Y S, Chepurin G, Ferry N, Fujii Y, Good S A, Guinehut S, Haines K, Hernandez F, Köhl A, Lee T, Martin M J, Masina S, Masuda S, Peterson K A, Storto A, Toyoda T, Valdivieso M, Vernieres G, Wang O, Xue Y. 2017. Ocean heat content variability and change in an ensemble of ocean reanalyses. *Clim Dyn*, 49: 909–930
- Pawlowicz R. 2013. Key physical variables in the ocean: Temperature, salinity and density. *Nat Educ Knowl*, 4: 13
- Pereira F, Mitchell T, Botvinick M. 2009. Machine learning classifiers and fMRI: A tutorial overview. *NeuroImage*, 45: S199–S209
- Roemmich D, Alford M H, Claustre H, Johnson K, King B, Moum J, Oke P, Owens W B, Pouliquen S, Purkey S, Scanderbeg M, Suga T, Wijffels S, Zilberman N, Bakker D, Baringer M, Belbeoch M, Bittig H C, Boss E, Calil P, Carse F, Carval T, Chai F, Conchubhair D Ó, d’Ortenzio F, Dall’Olmo G, Desbruyeres D, Fennel K, Fer I, Ferrari R, Forget G, Freeland H, Fujiki T, Gehlen M, Greenan B, Hallberg R, Hibiya T, Hosoda S, Jayne S, Jochum M, Johnson G C, Kang K R, Kolodziejczyk N, Körtzinger A, Traon P Y L, Lenn Y D, Maze G, Mork K A, Morris T, Nagai T, Nash J, Garabato A N, Olsen A, Pattabi R R, Prakash S, Riser S, Schmechtig C, Schmid C, Shroyer E, Sterl A, Sutton P, Talley L, Tanhua T, Thierry V, Thomalla S, Toole J, Troisi A, Trull T W, Turton J, Velez-Belchi P J, Walczowski W, Wang H L, Wanninkhof R, Waterhouse A F, Waterman S, Watson A, Wilson C, Wong A P S, Xu J P, Yasuda I. 2019. On the future of Argo: A global, full-depth, multi-disciplinary array. *Front Mar Sci*, 6: 439
- Schaphire R E. 2003. The boosting approach to machine learning: An overview. In: Denison D D, Hansen M H, Holmes C C, Mallick B, Yu B, eds. Nonlinear Estimation and Classification. Lecture Notes in Statistics. Berlin: Springer
- Schmidtko S, Johnson G C, Lyman J M. 2013. MIMOC: A global monthly isopycnal upper-ocean climatology with mixed layers. *J Geophys Res Oceans*, 118: 1658–1672
- Smith D, Timms G, de Souza P, D’Este C. 2012. A Bayesian framework for the automated online assessment of sensor data

- quality. *Sensors*, 12: 9476–9501
- Smith N, Kessler W S, Cravatte S, Sprintall J, Wijffels S, Cronin M F, Sutton A, Serra Y L, Dewitte B, Strutton P G, Hill K, Gupta A S, Lin X P, Takahashi K, Chen D K, Brunner S. 2019. Tropical Pacific observing system. *Front Mar Sci*, 6: 31
- Sugiura N, Hosoda S. 2020. Machine learning technique using the signature method for automated quality control of Argo profiles. *Earth Space Sci*, 7: e2019EA001019
- Sun C, Thresher A, Keeley R, Hall N, Hamilton M, Chinn P, Tran A, Goni G, la Villeon L, Carval T. 2010. The data management system for the global temperature and salinity profile programme. In: *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society*, Venice. 9: 86
- Szekely T, Gourrion J, Pouliquen S, Reverdin G. 2019. The CORA 5.2 dataset for global *in situ* temperature and salinity measurements: Data description and validation. *Ocean Sci*, 15: 1601–1614
- Takeshita Y, Martz T R, Johnson K S, Plant J N, Gilbert D, Riser S C, Neill C, Tilbrook B. 2013. A climatology-based quality control procedure for profiling float oxygen data. *J Geophys Res Oceans*, 118: 5640–5650
- Talley L D. 1984. Meridional heat transport in the Pacific Ocean. *J Phys Oceanogr*, 14: 231–241
- Tan Z T, Reseghetti F, Abraham J, Cowley R, Chen K Y, Zhu J, Zhang B, Cheng L J. 2021. Examining the influence of recording system on the pure temperature error in XBT data. *J Atmos Ocean Tech*, 38: 759–776
- Thadathil P, Ghosh A K, Sarupria J S, Gopalakrishna V V. 2001. An interactive graphical system for XBT data quality control and visualization. *Comput Geosci*, 27: 867–876
- Thresher A, Cowley R, Wijffels S. 2008. QuOTA dataset (Quality-controlled Ocean Temperature Archive). v3. CSIRO. Data Collection., <https://doi.org/10.25919/5ec357563bd3e>
- Toma D M, Benadí A G, Manuel-Gonzalez B J, Del-Rio-Fernandez J. 2016. Systematic quality control for long term ocean observations and applications. *ACTA IMEKO*, 5: 64–68
- U.S. Integrated Ocean Observing System. 2015a. Manual for real-time quality control of dissolved nutrients data: A guide to quality control and quality assurance of coastal and dissolved nutrients observations., <http://doi.org/10.7289/V5TT4P7R>
- U.S. Integrated Ocean Observing System. 2015b. Manual for real-time quality control of *in-situ* surface wave data: A guide to quality control and quality assurance of *in-situ* surface wave observations. <http://doi.org/10.7289/V5K91T>
- U.S. Integrated Ocean Observing System. 2017a. Manual for real-time quality control of phytoplankton data: A guide to quality control and quality assurance of phytoplankton observations., <http://doi.org/10.7289/V56D5R6S>
- U.S. Integrated Ocean Observing System. 2017b. Manual for real-time quality control of ocean optics data version 1.1: A guide to quality control and quality assurance of coastal and oceanic optics observations., <https://doi.org/10.25923/v9p8-ft24>
- Udaya Bhaskar T V S, Venkat Shesu R, Boyer T P, Pattabhi Rama Rao E. 2017. Quality control of oceanographic *in situ* data from Argo floats using climatological convex hulls. *MethodsX*, 4: 469–479
- UNESCO. 2017. *Global Ocean Science Report: The Current Status of Ocean Science Around The World*. Paris: UNESCO
- UNESCO-IOC. 2010. *GTSP Real-time Quality Control Manual*. Revised Edition 2010. Paris: UNESCO
- Vanderviere E, Huber M. 2004. An adjusted boxplot for skewed distributions. In: *Proceedings of Computational Statistics*. Berlin: Springer. 1933–1940
- Wedd R, Stringer M, Haines K. 2015. Argo real-time quality control intercomparison. *J Oper Oceanogr*, 8: 108–122
- Wong A, Keeley R, Carval T. 2020. Argo Quality Control Manual for CTD and Trajectory Data. <https://archimer.ifremer.fr/doc/00228/33951/32470.pdf>
- Yerushalmy J. 1947. Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. *Public Health Rep*, 62: 1432–1439
- Zhang B, Wang F. 2019. Advances in quality control system of ocean thermohaline data. In: *Proceedings of the 21st EGU General Assembly*, Vienna
- Zika J D, Skliris N, Blaker A T, Marsh R, Nurser A J G, Josey S A. 2018. Improved estimates of water cycle change from ocean salinity: The key role of ocean warming. *Environ Res Lett*, 13: 074036
- Zou L W, Zhou T J, Tang J P, Liu H L. 2020. Introduction to the regional coupled model WRF4-LICOM: Performance and model intercomparison over the Western North Pacific. *Adv Atmos Sci*, 37: 800–816

(责任编辑: 吴立新)