

基于梯度依赖OI的全球多参数Argo数据集的构建与验证

王丹阳¹, 张春玲^{1,2}, 卢少磊^{3,4*}, 李兆钦^{3,4}, 刘增宏^{3,4}

(1.上海海洋大学海洋科学学院,上海201306;2.自然资源部海洋生态监测与修复技术重点实验室,上海201306;3.自然资源部第二海洋研究所,浙江杭州310012;4.卫星海洋环境动力学国家重点实验室,浙江杭州310012)

摘 要: 利用Argo资料,基于梯度依赖最优插值客观分析系统,重构2004—2020年空间分辨率为 $1^\circ \times 1^\circ$ 的全球多参数Argo网格化数据集,并通过置信区间估计、实测数据检验、与其他数据集对比等方式,对该数据集进行一系列的验证。结果表明:重构的Argo数据集在95%的统计概率下,90%以上的温度、盐度重构结果可信,且与实测数据的温度、盐度最大偏差不超过 $\pm 1.0^\circ\text{C}$ 、 ± 0.02 。该数据集所反映的大尺度信号与现有数据集一致,并且可以保留较多中小尺度信号,分析结果与实际观测更接近。

关键词: 梯度依赖最优插值;Argo;多参数;网格化数据集;客观分析

中图分类号: P715.2 **文献标识码:** A **文章编号:** 1003-0239(2023)02-0077-12

0 引言

地转海洋学实时观测阵(Array for Real-time Geostrophic Oceanography, Argo)俗称“全球海洋观测网”。类似于陆地上多个气象站组建成的气象网,Argo是一个由多个浮标在海洋上组成的实时观测网,旨在获取准实时、大范围、高分辨率的全球海洋次表层观测资料^[1-2]。自2007年成立以来,截止到2020年12月,Argo获取的温度和盐度剖面数量已经超过200万条,并且仍以每年超过14万条的速度递增。Argo剖面数量已经大大超过了过去一百多年由海洋调查船获取的全部剖面数量。

由于Argo浮标获取温盐信息是一个悬停漂移的过程,因此观测剖面在空间上具有随机性,这使得Argo剖面资料在应用于海洋预报和气象预报等研究时受到一定的限制。随着Argo观测资料的不断积累,许多Argo成员国也陆续开发了相应的数据分析系统,研制了各自的Argo网格数据产品。例如:英国气象局(Met Office)采用业务化海洋大气预

报模式FOAM(Forecasting Ocean Assimilation Model)提供的初始场和最优插值(Optimal Interpolation, OI)同化方法,构建了全球月平均、空间分辨率为 $1^\circ \times 1^\circ$ 的Argo网格数据集^[3];日本海洋科技中心(Japan Marine Science and Technology Center, JAMSTEC)基于OI客观分析系统,研制了全球海洋空间分辨率为 $1^\circ \times 1^\circ$ 的月平均温度、盐度数据集^[4];美国斯克利普斯海洋研究所(Scripps Institution of Oceanography, SIO)^[5]完全基于Argo观测资料本身,通过OI构建了海洋中上层(2.5~1 975 m,垂向58层)月平均的气候态温度、盐度分布场,空间分辨率为 $1^\circ \times 1^\circ$,数据定期更新;国际太平洋研究中心(International Pacific Research Center, IPRC)则利用三维变分同化技术,生成了月平均的全球温度、盐度及动力高度网格资料集。我国作为Argo成员国之一,在Argo网格数据的研制方面也做了大量工作。自然资源部杭州全球Argo系统野外科学观测研究站(网址:<http://www.argo.org.cn/>,以下简称“Argo野外站”)利用Barnes逐步

收稿日期:2022-04-25;修回日期:2022-06-12。

基金项目:国家自然科学基金项目(42106090、U1811464)。

作者简介:王丹阳(1996-),女,硕士在读,主要从事Argo数据处理研究工作。E-mail: wdy1996wdy@163.com

*通信作者:卢少磊(1988-),男,工程师,硕士,主要从事Argo数据处理研究工作。E-mail: lsl324004@163.com

订正法,构建的水平分辨率为 $1^\circ \times 1^\circ$ 、垂向 58 层的月平均数据集已在国际 Argo 管理中心正式发布,数据可实时获取并定期更新^[6];该观测研究站还陆续开发了一系列衍生要素网格数据集^[7-8]。国家海洋环境预报中心和中国科学院大气物理研究所也分别基于集合 OI 同化方案,研发了相应的 Argo 资料同化再分析数据集^[9-11]。

上述数据集均能较准确地刻画全球中上层海洋的热盐结构特征,通过将时空规则化,有效扩大了 Argo 数据的应用范围。但是,目前大部分 Argo 数据集仅包括温度、盐度等主要要素,全球温跃层参数、声速场等实用性较强的衍生要素还有待于精确确定。另外,OI 虽然在构建网格数据集方面效果明显,但如何确定有效的空间相关尺度是一个关键问题。大量研究人员利用观测资料和数值模式结果来估计背景误差协方差矩阵的相关尺度^[12-14]。这种方法主要利用观测信息相对于背景信息的偏差,通过函数拟合来获得背景误差方差及相关尺度,其依赖于足够密集且可以提供多种尺度信息的观测网,所估计的相关尺度通常为常数。但是,由于海洋要素场一般具有不同的相关尺度,利用目前的海洋观测资料还很难准确地估计相关尺度的变化。

因此,采用一种可以充分提取观测信息的客观分析/数据同化方法,构建更精确、多参数的 Argo 网格化产品,可为我国 Argo 网格数据集的研发提供有效补充。ZHANG 等^[15]根据海洋要素水平梯度的变化,提出了一种基于各向异性的梯度依赖相关尺度法 (Gradient - Dependent Correlation Scale Method, GDCSM),其本质上是对 OI 客观分析的一种改进。该方法经过几次应用改进^[16-17],形成了可以有效提取 Argo 观测信息的梯度依赖 OI 方法。本文立足于 Argo 资料本身(未同化其他观测资料),基于梯度依赖 OI 方法和热盐统计模型,构建了全球海域上层海洋多参数(包括温度、盐度、深度、声速、跃层等)的 Argo 网格数据集,并系统验证了该数据集的可靠性和先进性。

1 Argo 数据集的研制

1.1 原始剖面数据

Argo 原始散点剖面数据由 Argo 野外站提供,

考虑到剖面质量和数量,我们选用了 2004 年 1 月—2020 年 12 月全球海域的温度、盐度剖面资料。这些观测剖面虽然经过 Argo 野外站的实时和部分延时质量控制,但里面还存在小部分有质量问题的剖面。为了确保数据集的可靠性,统一对这些 Argo 资料进行温盐范围检验、冰点检验、尖峰检验、温盐梯度检测及水深递增检验等质量再控制分析^[18-19],将检测到的可疑资料剔除,接着利用 AKIMA^[20]的方法将数据垂向插值到 57 个标准层,最终获得全球海域 204 个月的 Argo 剖面 1 703 152 条(见图 1)。全球大部分海域 $1^\circ \times 1^\circ$ 网格内的 Argo 剖面数均超过 20 条,南(北)纬 60° 以南(北)高纬度海域的 Argo 观测结果相对稀疏。

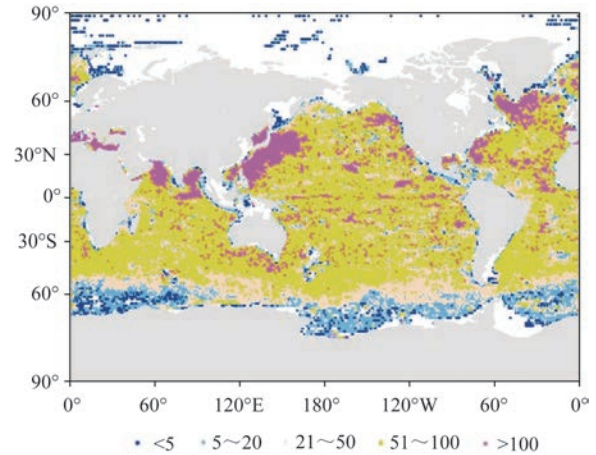


图1 2004年1月—2020年12月全球 $1^\circ \times 1^\circ$ 网格内的 Argo 观测剖面数量(白色区域无数据)

Fig.1 Number of Argo profiles within $1^\circ \times 1^\circ$ boxes from 1st January 2004 to 31st December 2020 in the global ocean

1.2 客观分析方法

1.2.1 梯度依赖 OI

梯度依赖 OI 是指根据海洋要素水平梯度变化规律自动调整分析点的相关系数,基于各向异性相关尺度,给出的一种能够有效提取多种尺度海洋信息的背景误差协方差公式。方程如下^[17,21]:

$$V_i^a = V_i^b + \sum_j^M w_{ij} \delta y_j^0 \quad (1)$$

$$\sum_{j=1}^M w_{ij} \mu_{jk} + \eta_k w_{ik} = \mu_{ik} \quad k = 1, 2, \dots, M \quad (2)$$

$$\mu_{ik} \sim \exp \left[-\frac{(x_i - x_k)^2}{(L_\phi / G_x)^2} - \frac{(y_i - y_k)^2}{(L_\phi / G_y)^2} \right] \quad (3)$$

$$G_x = 1 + \frac{|\partial v / \partial x|}{E(|\partial v / \partial x|)}, G_y = 1 + \frac{|\partial v / \partial y|}{E(|\partial v / \partial y|)} \quad (4)$$

式中: V_i^a 为第 i 个格点的分析场, V_i^b 为对应的背景场; w_{ij} 为第 i 个格点与第 j 个观测点之间的最优权重, 可通过求解线性方程 (2) 来确定; $\delta y_j^0 = y_j^0 - H(v_j^b)$ 为观测增量, y_j^0 为第 j 个观测点的观测值, v_j^b 为第 j 个观测点的初猜值, H 是从分析格点到观测点的双线性插值算子; μ_{jk}, μ_{ik} 分别为两个观测站点 j 和 k 、观测站点 k 和格点 i 间的背景误差相关系数; $\eta_k = \sigma_{ok}^2 / \sigma_{bi}^2$ 是观测误差与背景误差的平方比, 通常设置为可调常数 η , 实验表明, 在全球海域, 当 η 取值为 0.5 时, 温度、盐度分析结果的均方根误差均最小; 背景误差相关系数与距离成反比, 且满足式 (3) 的指数分布^[16]; x 和 y 是经度和纬度; L_ϕ / G_x 与 L_ϕ / G_y 分别是经向和纬向相关尺度, L_ϕ 由相关尺度常数 (通常为 $5^\circ \sim 10^\circ$ ^[17]) 与分析格点所处纬度的乘积来确定; $|\partial v / \partial x|$ 和 $|\partial v / \partial y|$ 分别为经向和纬向的分析变量绝对梯度, $E(|\partial v / \partial x|)$ 和 $E(|\partial v / \partial y|)$ 是对应的水平梯度均值。考虑到 Argo 数据观测密度以及计算量问题, 在计算年际平均场和季节平均场时, 相关尺度常数取为 5° , 计算月平均场时, 相关尺度常数则设为 10° 。

1.2.2 表层反演模型

本文用于反演表层温度和盐度场的统计模型是基于一个六参数的海洋内部剖面结构构建的^[21-23]。该模型在计算密度跃层深度的基础上, 通过确定混合层深度 (同温跃层上界深度)、温跃层下界深度和温跃层梯度来推算海表温度 (Sea Surface Temperature, SST), 并在混合层深度范围内反演海表盐度 (Sea Surface Salinity, SSS)。这里分别选取 10 m、20 m、30 m、50 m、75 m、100 m、125 m、150 m 和 200 m 这 9 个起始反演水层, 进而将每个格点上的 9 个温度值进行加权平均。公式如下:

$$T_{ss} = \frac{P_0}{P_z} T_z, \quad S_{ss} = S_z \quad (0 \leq z \leq D_{ML}) \quad (5)$$

$$T_{ss} = T_z - G_{TT}(z - D_{ML}) \quad (D_{ML} \leq z \leq D_{TB}) \quad (6)$$

$$T_{ss,i} = \frac{\sum_{j=1}^9 b_{ij} T_{ss,i,j}}{\sum_{j=1}^9 b_{ij}}, \quad b_{ij} = \exp(-(r_{ij} - \bar{r}_i)^2 / L^2) \quad (7)$$

式中: T_{ss} 为海表温度; S_{ss} 为海表盐度; G_{TT} 为温跃层梯度; D_{ML} 和 D_{TB} 分别为混合层深度和温跃层下界深

度; T_z 和 S_z 是 Argo 剖面上 z 深度处对应的温度和盐度; P_0 和 P_z 分别为海表面和次表层 (z 深度处) 的压力; $T_{ss,i}$ 是第 i 个 Argo 剖面上经过加权平均后得到的 SST; $T_{ss,i,j}$ 为第 i 个 Argo 剖面根据第 j ($j = 1, 2, \dots, 9$) 个参考水层推算的 SST; b_{ij} 为对应的权重系数; r_{ij} 是 $T_{ss,i,j}$ 相对于全球温盐剖面计划 (Global Temperature and Salinity Profile Project, GTSP) 数据的均方根误差; L 为温度相关尺度常数, 全球海域通常设为 2° ^[21]。

1.3 研制流程

全球 Argo 网格数据集的研制包括以下 4 步: ① 进行剖面资料的质量检验和筛选, 并通过最大角度法计算每个剖面对应的跃层参数 (包括温跃层上界深度、温跃层下界深度、温跃层梯度) 和声速。② 通过计算 $1^\circ \times 1^\circ$ 网格内的算术平均值, 融合 2004 年以来所有的观测剖面数据, 形成多年气候态初猜场, 进而利用传统 OI 分析得到多年气候态背景场, 并以此计算梯度依赖相关尺度参数; 再以气候态背景场为初猜场, 通过梯度依赖 OI 构建多年季节平均的背景场, 以此类推, 构建多年月平均的背景场。③ 利用梯度依赖 OI 分析得到 2004 年 1 月—2020 年 12 月逐年、逐月次表层的温度、盐度和声速数据以及温跃层参数的网格数据。④ 基于表层参数统计模型反演推算表层温度和盐度^[21], 进而计算表层声速场。

基于梯度依赖 OI 客观分析系统构建的全球 Argo 数据集 (以下简称 GDCSM-Argo) 主要包括温度、盐度、声速、混合层深度、温跃层下界深度、温跃层梯度、经度、纬度和深度等要素; 空间覆盖范围为 $179.5^\circ \text{W} \sim 179.5^\circ \text{E}$, $89.5^\circ \text{S} \sim 89.5^\circ \text{N}$; 覆盖时段为 2004 年 1 月—2020 年 12 月; 空间分辨率为 $1^\circ \times 1^\circ$; 时间分辨率为月平均、多年月平均、多年季节平均和多年气候态。

2 Argo 数据集验证

2.1 置信区间估计

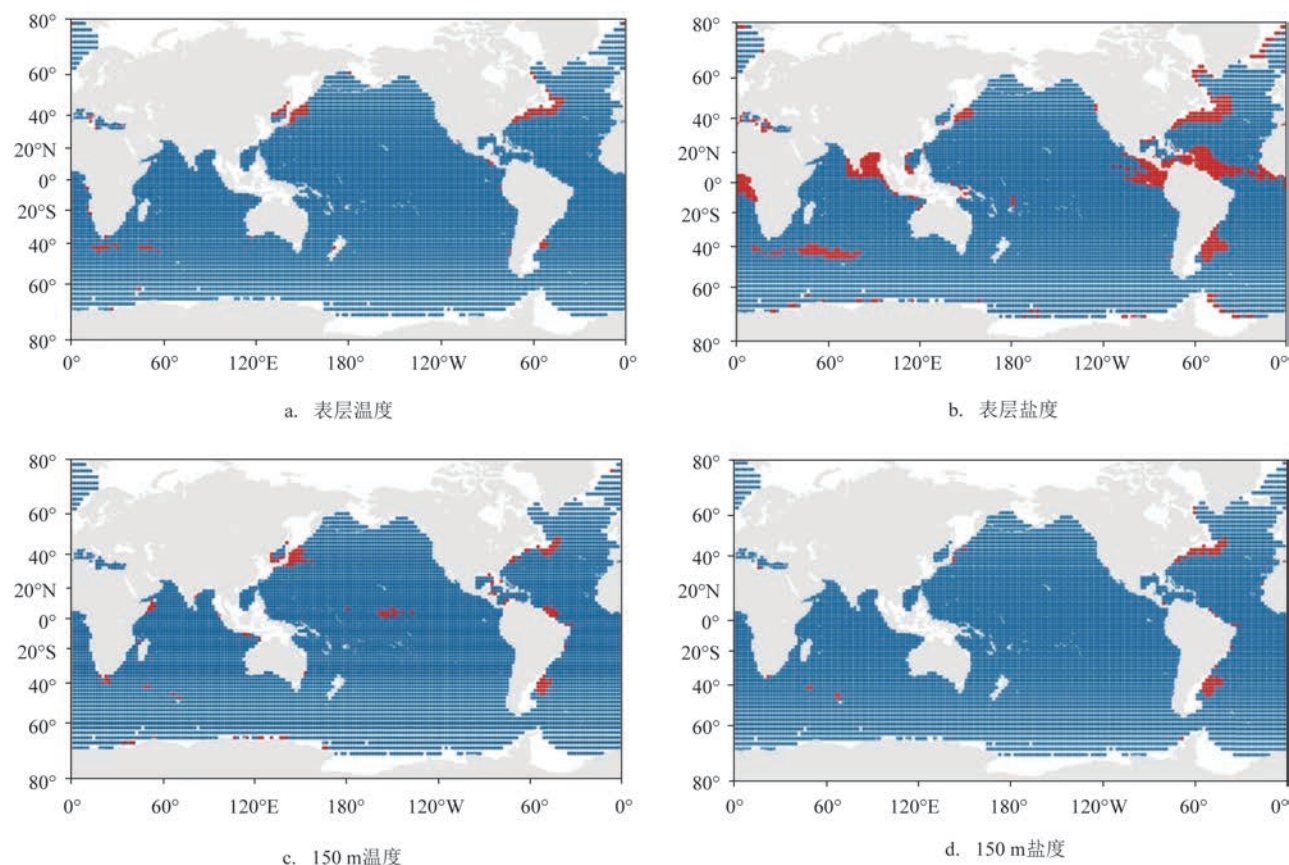
我们常用置信区间来判定分析结果在一定的统计概率 (90% 或 95%) 下是否可信。这里我们以每个分析格点搜索半径 5° 范围内的观测值空间序列, 作为该格点上分析值的一次取样, 通过 W 检验 (Shapiro-Wilk 检验)^[24-25] 确定服从正态分布的观测

序列 X , 进而由 $\frac{-\bar{X}}{S/\sqrt{n}}$ 服从自由度为 $(n-1)$ 的 t 分布, 可得格点分析值置信区间为 $\bar{X} \pm t_{\alpha}(n-1) \cdot S/\sqrt{n}$ 。式中: α 为显著性水平, 对应的置信概率为 $P = 1 - \alpha$; 在一定的自由度下, S 为均方根误差; t 值随着置信概率的大小而变化, 即 P 越大, t 值越大。本文统一取 $\alpha=0.05$, 即置信概率为 95%。由图 2a 可见, 表层温度 95% 以上的分析值是可信的, 只有少数分析值落在置信区间外, 且这些格点主要分布在温度水平梯度较大的 40°S 附近海域、大西洋湾流区、西北太平洋黑潮区以及观测资料较稀疏的亚极地海域和海-陆边界处。150 m 层 (见图 2b) 温度分析值的非可信格点比表层稍有增加, 这些非可信点仍然集中分布在观测资料少且温度梯度大的急流区, 但大部分格点的分析值在 95% 的统计概率下仍都是可信的。从盐度结果可以看出 (见图 2c—d), 150 m 的非可信点明显比表层多, 但尽管如此, 表层

盐度中 85% 以上的分析值都是可信的。表层盐度的非可信点多集中在边界流区域、赤道海域以及孟加拉湾附近海域, 这可能与当地的热盐输送及淡水通量有关。

2.2 实测数据检验

考虑到时间和空间上的连续性, 我们采用全球热带大洋锚系浮标观测阵 (Global Tropical Moored Buoy Array, GTMBA, 网址: <https://www.pmel.noaa.gov/gtmba/>) 中位于太平洋的热带大气海洋观测序列 (Tropical Atmosphere Ocean/Triangle Trans-Ocean Buoy Network, TAO/TRITON) 进行实测数据检验。由于浮标存在起伏不定、故障等问题, 导致观测深度不一、某些年份观测值缺失等缺点。这里仅选取观测深度较深且时间覆盖 2004 年 1 月—2020 年 12 月的代表站点 ($0^{\circ}, 147^{\circ}\text{E}$) 的温度和盐度观测资料加以对比。由图 3 可见, 2004—2020 年站点的温度偏



注: 蓝色点表示分析值落在置信区间内, 红色点表示分析值落在置信区间外

图2 95% 的置信概率下气候态分析结果的区间估计结果

Fig.2 Interval estimation results of climate analysis results under 95% confidence probability

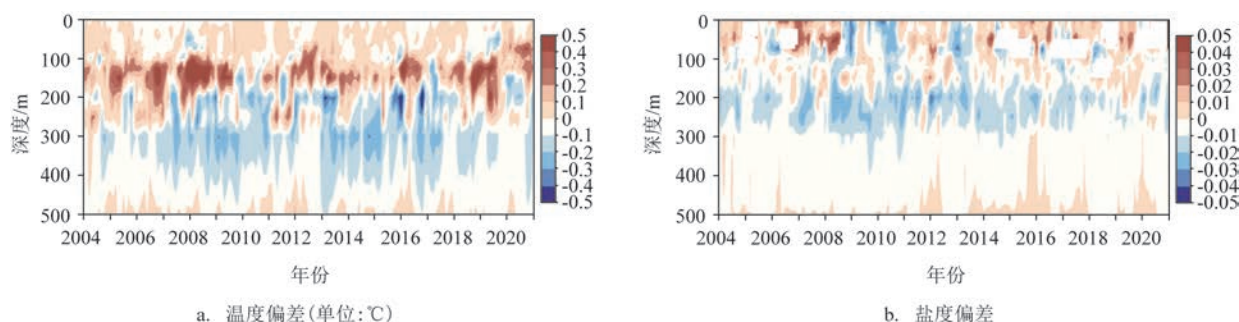


图3 2004年1月—2020年12月太平洋代表站点(0°, 147°E)GDCSM-Argo与TAO浮标观测的温度和盐度偏差

Fig.3 Temperature and salinity differences between GDCSM-Argo and TAO at (0°, 147°E) from January 2004 to December 2020

差均随深度的增加先增大后减小。100 m以浅跃层深度上每个月温度偏差均小于 ± 0.2 °C; 100~200 m的温度偏差增大至 $\pm 0.2 \sim \pm 1.0$ °C; 200 m以下的温度偏差又逐渐减小, 最大偏差不超过 ± 0.4 °C。盐度偏差从表层的约 ± 0.02 到500 m处逐渐减小, 300 m以深处的盐度偏差基本为0。

2.3 与其他数据对比验证

首先以150 m水层为例, 将本文构建的GDCSM-Argo数据集与WOA18(World Ocean Atlas 2018)气候态数据集进行对比, 对大洋温跃层深度处的温度和盐度构建结果进行气候态检验, 进而将GDCSM-Argo与英国气象局(Met Office)的EN4-Argo、国际太平洋中心(International Pacific Research Center)的IPRC-Argo、美国SIO的RG-Argo等数据集进行偏差对比和厄尔尼诺-南方涛动(El Niño-Southern Oscillation, ENSO)信号检验。

2.3.1 水平分布气候态检验

图4显示了150 m层GDCSM-Argo多年气候态

数据与WOA18数据的温度和盐度偏差。从图中可以看出, 两种数据在大部分海域的温度偏差在 ± 0.5 °C以内, 边界流区域(如北大西洋湾流区)的温度偏差相对较大, 达到1 °C, 且GDCSM-Argo数据的温度普遍高于WOA18数据(见图4a); 两种数据的盐度偏差不超过 ± 0.2 , 除了在北大西洋湾流区盐度偏差较大外, 大部分海域的盐度偏差为 $\pm 0.05 \sim \pm 0.1$ (见图4b)。

2.3.2 断面偏差对比

图5a和5b分别展示了GDCSM-Argo数据在纵贯太平洋海域的180°E和横跨三大洋的0.5°N断面上, 500 m以浅的温度气候态分布特征。在180°E断面、20°S~5°N之间, 存在一个暖水层(水温 > 28 °C), 厚度在120 m左右。暖水层之下的水温持续下降, 有较强的温度梯度, 垂直温度梯度在100~200 m之间达到最大, 反映了主温跃层的存在。在沿0.5°N的经度-深度断面上, 也存在一个西薄东厚的暖水层(水温 > 25 °C)。暖水层之下的水温先快速下降, 最后趋于稳定, 300 m左右水深处等温线倾斜

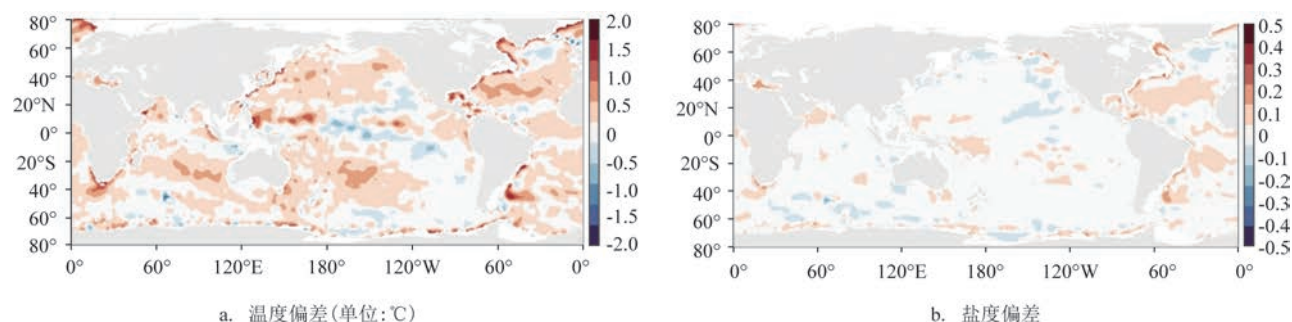


图4 150 m层GDCSM-Argo数据与WOA18数据的温度和盐度偏差对比

Fig.4 Temperature and salinity difference between GDCSM-Argo and WOA18 data at the depth of 150 m

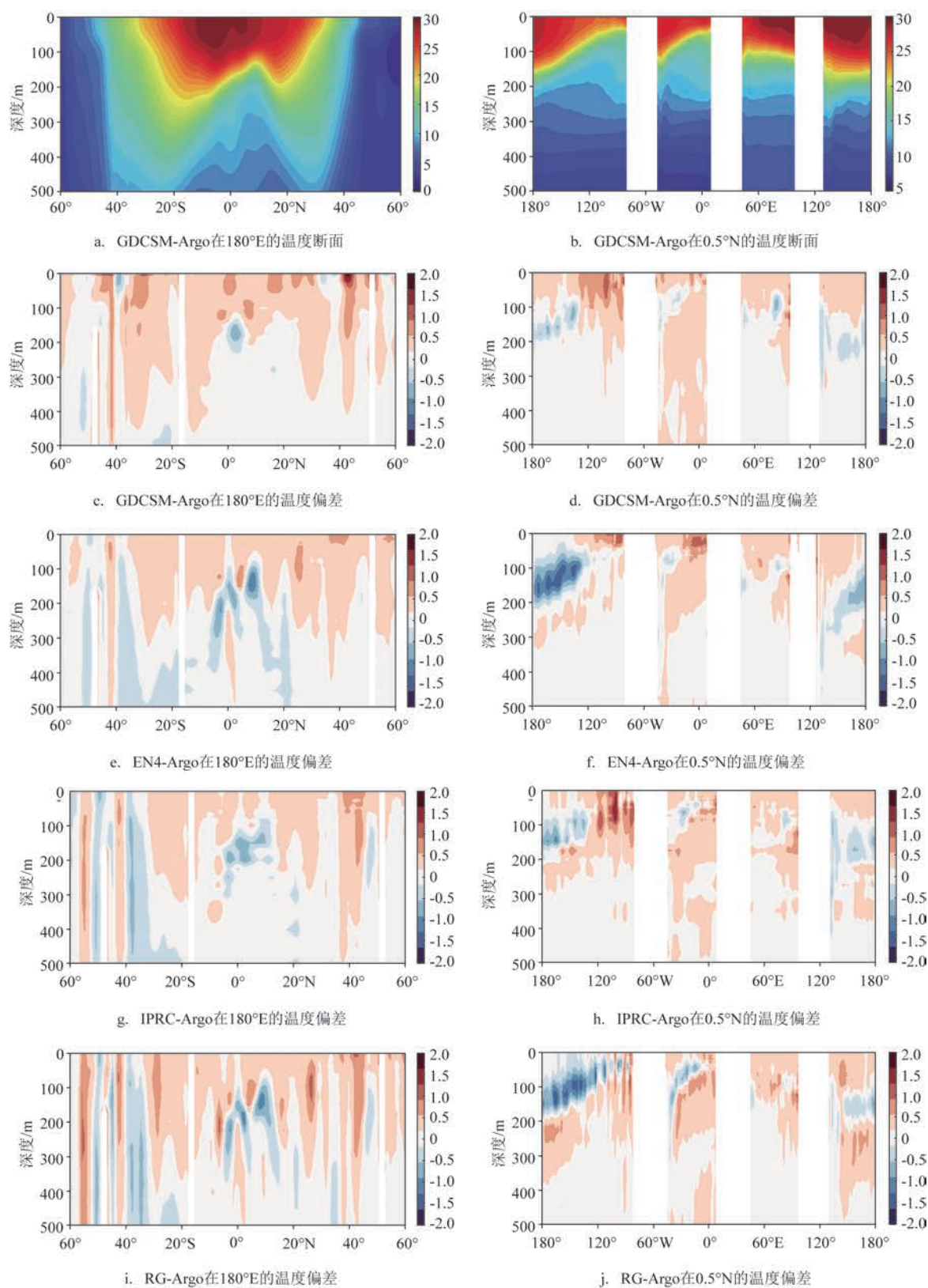


图5 沿180°E和0.5°N的GDCSM-Argo温度断面分布和不同数据集相对于WOA18数据的温度偏差分布(单位:℃)

Fig.5 GDCSM-Argo temperature section distribution and the distribution of temperature difference from different datasets relative to WOA18 data along the sections of 180°E and 0.5°N (unit: °C)

方向在各个大洋都与上层有不同程度的反向。这种结构与热带地区信风和洋流的分布密切相关。

图5c—j分别给出了两个断面上不同数据集相对于WOA18数据的温度偏差。由图可见,在180°E断面、5°S~5°N之间赤道附近海域100~250 m深度处,GDCSM-Argo数据的温度值较WOA18数据偏低0.2~0.5℃,同样的低温特征在其他3个数据集的偏差对比中也同样存在。几个数据集的对比结果表明,GDCSM-Argo数据的低温范围最小、影响深度最浅。在南半球40°S附近海域500 m以浅,4个数据的温度偏差均约为±0.5℃;相较之下,GDCSM-Argo数据的偏差高值水平范围较小,在300~500 m深度范围内,大部分温度偏差均接近0℃。在40°N附近海域,GDCSM-Argo数据的近表面温度偏差约为0.5~1℃,其他3个数据集在该海域也存在-0.5~0.5℃的温度偏差。除此之外,GDCSM-Argo数据在该断面上的温度偏差均在0~0.2℃之间,其偏差低值范围明显大于其他数据集。而在沿0.5°N纬度断面上(见图5d、5f、5h、5j),300 m以深处几个数据集的温度偏差均小于±0.2℃,在太平洋120°E~120°W之间的海域,GDCSM-Argo的温度数据较WOA18偏低0.2~0.5℃,IPRC-Argo与GDCSM-Argo数据的结果类似,而EN4-Argo和RG-Argo数据的温度偏差则约为0.5~1.0℃。GDCSM-Argo的温度数据在大西洋和印度洋赤道海域大多偏高0.2~0.5℃,其他3个数据则均偏高0.5~1.0℃。整体而言,在两个代表断面上,GDCSM-Argo数据集大部分结果的温度偏差是4个数据集中最小的。

纵贯太平洋的GDCSM-Argo盐度断面图显示(见图6a),整个断面的盐度从左到右大致以低-高-低-高-低的形式分布,在南北半球分别存在一个高盐(盐度>35.0)区。南半球的高盐区位于0°~40°S之间,北半球的高盐区位于15°~30°N之间,且北半球的高盐区范围明显小于南半球。两个高盐区之间的热带海域、0°~40°S之间和0°~40°N之间的亚极地区是3个低盐(盐度<34.5)区。在38°N附近和42°S附近,等盐线分布密集,且有较强的盐度梯度,这可能与高、低浓度的盐水在此汇集有关。从沿0.5°N的断面上可以看出(见图6b),三大洋在200 m以浅均存在一个高盐中心(盐度>35),其中大西洋的最高盐度可达36.5以上,明显高于太平洋和印度

洋,印度洋的高盐区范围可深达500 m,太平洋的盐度相对最低,高盐中心的影响深度也相对较小。由图6c可见,在沿180°E的断面上,GDCSM-Argo数据集的盐度除了在0°~20°S及20°N附近海域的近表层(深度<100 m)偏低0.05~0.1,在其他绝大部分海洋区域两者的差异小于±0.05。其他3个数据集的盐度偏差分布也具有相似的特征,但偏差大于±0.05的海域和水层明显比GDCSM-Argo数据集的范围大。在沿0.5°N的断面上(见图6d、6f、6h、6j),4个数据集相对于WOA18数据集的盐度偏差高值(盐度偏差>0.05)多集中在200 m以浅的海域。GDCSM-Argo数据集的盐度偏差仅在赤道太平洋边界流和大西洋西边界区域较大(最大约为0.1),200 m以深海域的盐度偏差大部分接近0。IPRC-Argo与GDCSM-Argo数据集的结果类似,而EN4-Argo和RG-Argo数据集的盐度偏差高值区明显较大。

2.3.3 ENSO信号检验

由GDCSM-Argo、RG-Argo、IPRC-Argo、EN4-Argo这4个数据集计算的Nino3.4区指数与美国气候预报中心(Climatic Prediction Center, CPC)提供的ENSO指数的相关系数分别为0.973 3、0.969 7、0.958 8、0.953 9。前两种数据集采用的观测数据、构建的背景场以及使用的客观分析方法都较为相似,而后两种数据集的背景场均为数值模拟结果,在构建过程加入了Argo以外的其他观测资料。不同数据集对应的热带太平洋海域(120°E~80°W, 20°S~20°N) SST异常的经验正交函数 EOF的分解结果见图7和图8。

由图7可见,4个数据集第一模态的贡献率均超过80%,远远超过其他模态。第一模态空间分布均能明显反应出热带太平洋SST异常的“跷跷板”现象(见图7a),即东太平洋为正,西太平洋为负。相较而言,GDCSM和RG数据集能保留更多的中小尺度信息,GDCSM数据集的贡献率最大,约为83.2%,虽然IPRC数据集的贡献率也达82%以上,但其空间分布相对较平滑。第一模态对应的时间序列具有明显的季节周期性(见图7b),且研究期间的年际震荡趋于稳定。4个结果的时间周期也十分相似,存在明显的正负相位变化,反映出厄尔尼诺(EI Niño)与拉尼娜(La Niña)事件交替出现。GDCSM数据集的时间系数最小,其次是EN4和RG数据集,

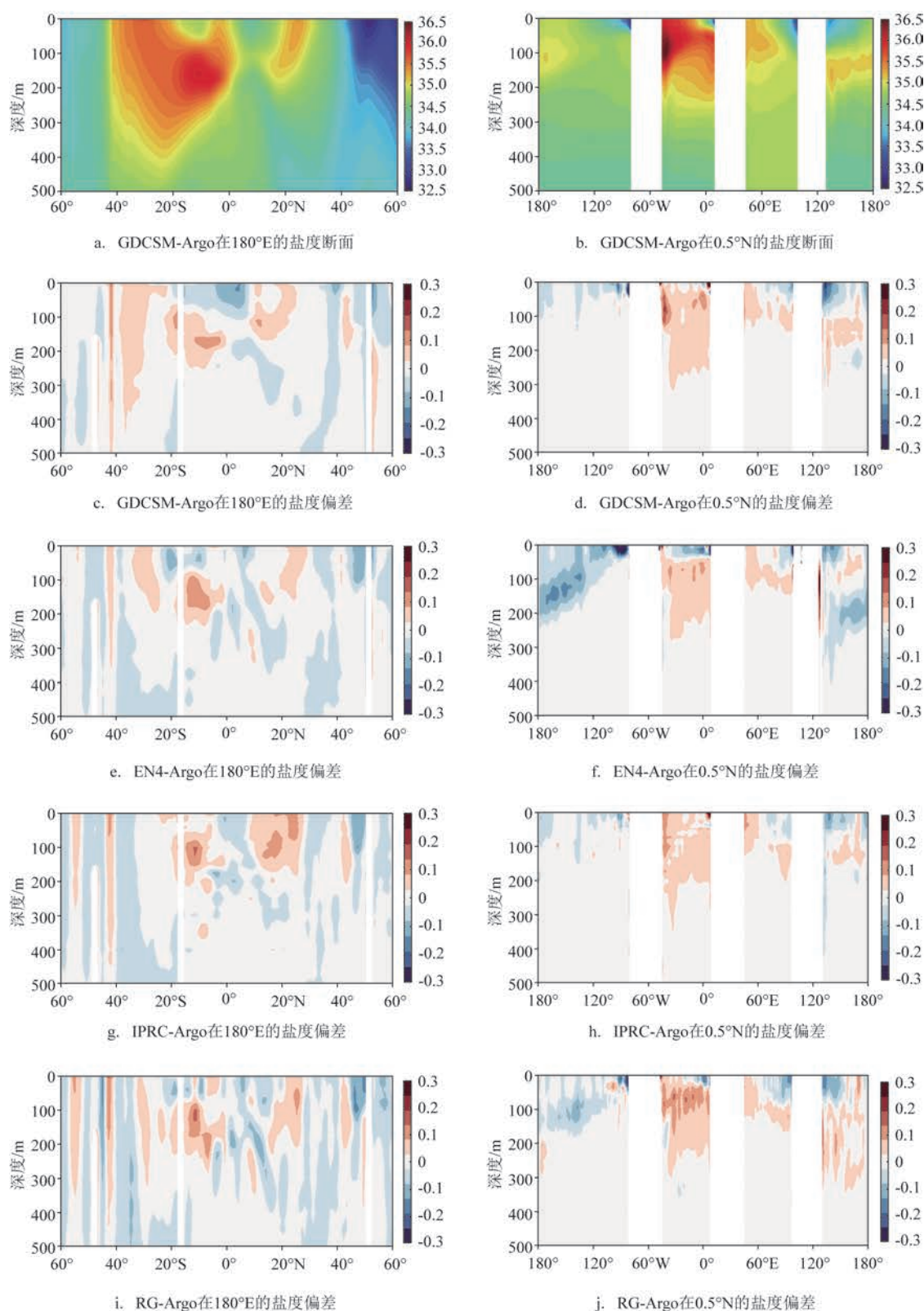


图6 沿180°E和0.5°N的GDCSM-Argo盐度断面分布和不同数据集相对于WOA18数据的盐度偏差分布

Fig.6 GDCSM-Argo salinity section distribution and the distribution of salinity difference from different datasets relative to WOA18 data along the sections of 180°E and 0.5°N

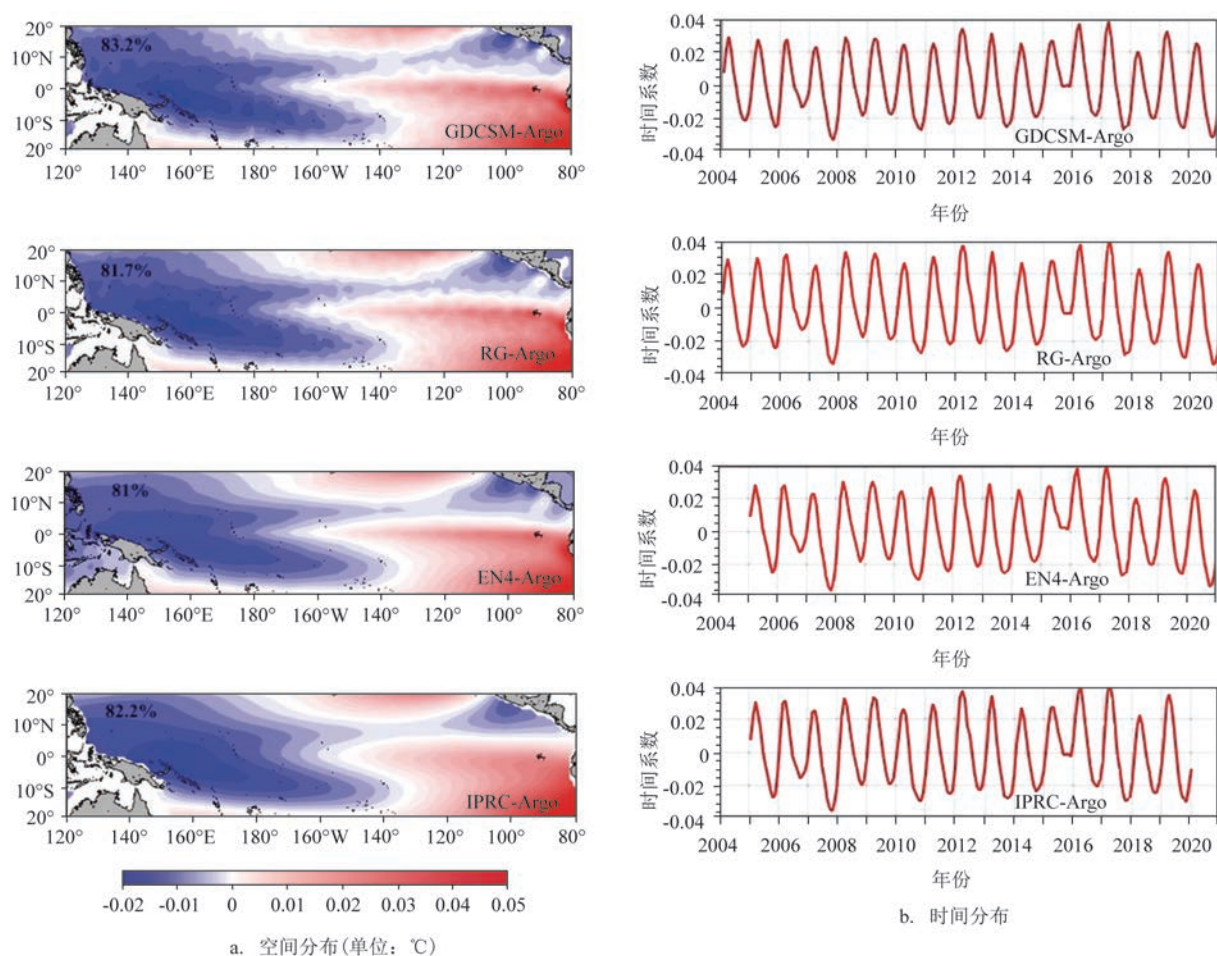


图7 不同数据集对应的热带太平洋海域海表温度距平的第一模态时空分布

Fig.7 The spatio-temporal distribution of the first mode of sea surface temperature anomalies in the tropical Pacific Ocean corresponding to different datasets

IPRC数据集的时间系数最大。从图8a可以看出,4个数据集第二模态的贡献率也均超过10%,其空间分布主要体现了暖水体积模态特征,热带太平洋南、北两侧呈现反相变化。与第一模态类似,GDCSM-Argo和RG-Argo数据集能够展现出更多的中小尺度信息,时间系数也存在明显的半年周期(见图8b)。由此可见,在反映年际变化信号时,GDCSM-Argo数据集与现有数据集具有较高的一致性,并且可以有效地提取更多中小尺度信号。

3 结论

全球海洋观测网(Argo)自建立以来,每年都以惊人的速度提供来自深海大洋中的温度、盐度信

息,但由于Argo浮标具有随波逐流的特性,观测得到的剖面数据具有时空不均匀性。本文采用一种改进的最优插值客观分析(梯度依赖OI)方法,在未加入其他观测资料的前提下,重构了2004年1月—2020年12月全球海洋Argo三维网格数据集(GDCSM-Argo)。该数据集包括温度、盐度、声速、跃层等要素,空间分辨率为 $1^{\circ} \times 1^{\circ}$,垂向(0~2 000 m)上分为不等间隔的57层。

置信区间估计结果表明,在95%的统计概率下,GDCSM-Argo数据集在全球海域绝大多数格点的分析结果是可信的,即使在水平梯度变化较大的海域,数据集也有约90%以上的温度、盐度分析结果可信。相对TAO浮标连续观测,GDCSM-Argo的数据在500 m以浅的温度、盐度最大偏差分别约

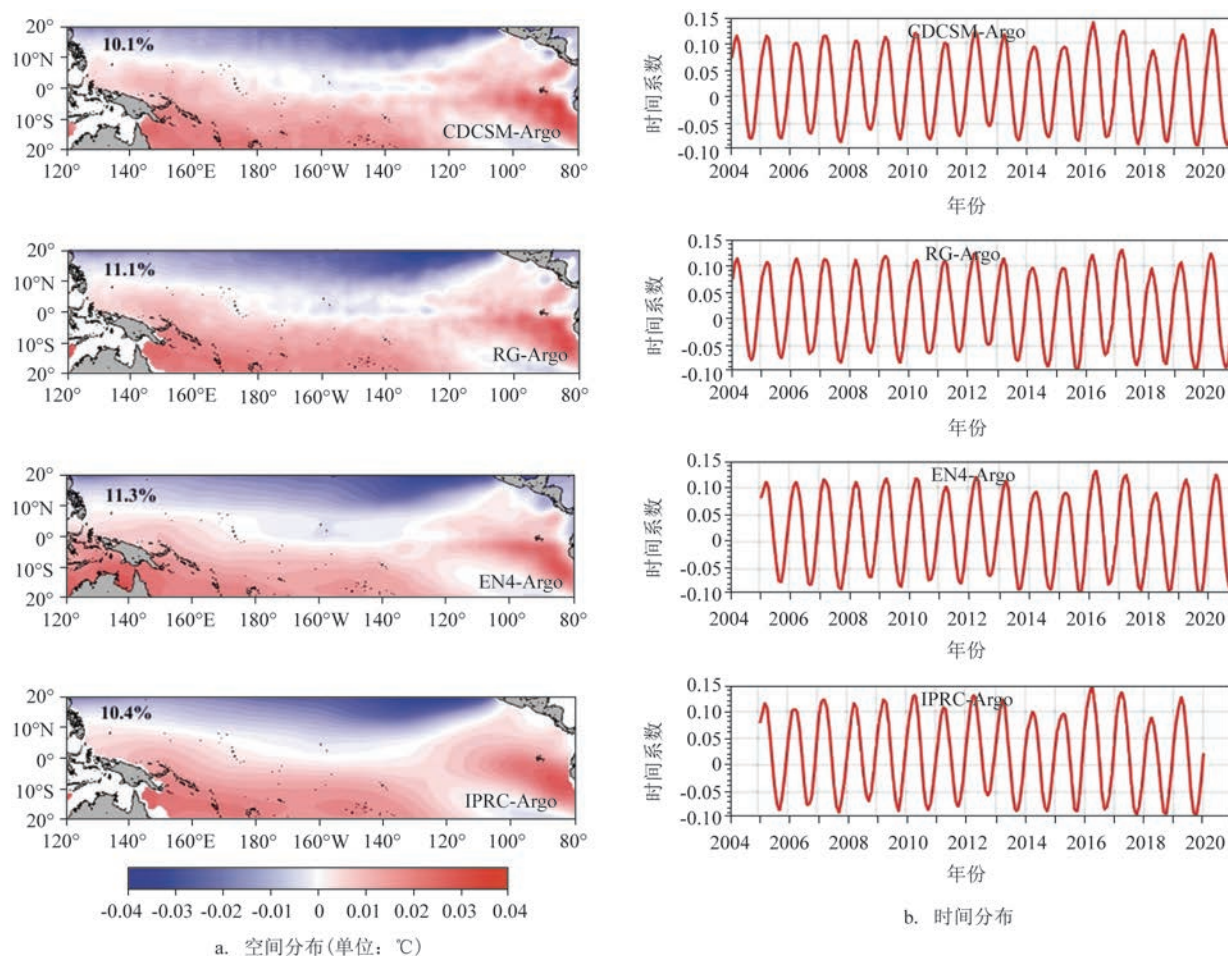


图8 不同数据集对应的热带太平洋海域海表温度距平的第二模态时空分布

Fig.8 The spatio-temporal distribution of the second mode of sea surface temperature anomalies in the tropical Pacific Ocean corresponding to different datasets

为 $\pm 1.0\text{ }^{\circ}\text{C}$ (100~200 m)和 ± 0.02 (表层);偏差随深度增加逐渐减小,300 m以下的温度偏差大部分小于 $\pm 0.2\text{ }^{\circ}\text{C}$,盐度偏差基本为0。通过不同网格数据集间的比较发现,150 m层GDCSM-Argo数据集相对于WOA18数据集的温度偏差在 $\pm 0.5\text{ }^{\circ}\text{C}$ 内,盐度偏差绝大部分小于 ± 0.1 ,两者的温度、盐度偏差均小于RG-Argo、IPRC-Argo、EN4-Argo等数据集与WOA18数据集的结果偏差。此外,GDCSM-Argo数据集所反映的大尺度海洋信息与其他数据集一致,而各向异性的背景误差相关计算方案使其同时保留了更多的中小尺度海洋特征。

由此可见,理论估计和对比验证均表明GDCSM-Argo数据集的温度、盐度重构结果是可靠、可信的,由于跃层和声速等衍生要素是基于温

度、盐度结果得到的,限于篇幅,其验证结果不再逐一展示。该数据集将每年进行实时更新,并可从Argo野外站免费下载使用。

参考文献:

- [1] ROEMMICH D, JOHNSON G C, RISER S, et al. The Argo program: observing the global ocean with profiling floats[J]. Oceanography, 2009, 22(2): 34-43.
- [2] Argo Science Team. On the design and implementation of Argo: an initial plan for a global array of profiling floats. International CLIVAR Project Office Report Number 2. GODAE report No. 5 [R]. Melbourne: GODAE International Project Office, 1998: 32.
- [3] GOOD S A, MARTIN M J, RAYNER N A. EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates[J]. Journal of Geophysical

- Research, 2013, 118(12): 6704-6716.
- [4] HOSODA S, OHIRA T, NAKAMURA T. A monthly mean dataset of global oceanic temperature and salinity derived from Argo float observations[J]. JAMSTEC Report of Research and Development, 2008, 8: 47-59.
- [5] ROEMMICH D, GILSON J. The 2004-2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo program[J]. Progress in Oceanography, 2009, 82(2): 81-100.
- [6] 卢少磊, 刘增宏, 李宏, 等. 全球海洋 Argo 网格资料集(BOA-Argo)用户手册[Z]. 2021: 23.
- LU S L, LIU Z H, LI H, et al. User manual of global ocean Argo gridded dataset (BOA_Argo)[Z]. 2021: 23.
- [7] 吴晓芬, 许建平, 李宏, 等. 西太平洋海域 Argo 衍生数据产品(集)用户手册[Z]. 2017: 18.
- WU X F, XU J P, LI H, et al. User manual of derived products from Argo dataset of the western Pacific Ocean[Z]. 2017: 18.
- [8] 杨小欣, 许建平, 吴晓芬, 等. 热带太平洋海域 Argo 衍生数据产品(热盐含量)用户手册[Z]. 2017: 21.
- YANG X X, XU J P, WU X F, et al. User manual of Argo derived dataset (heat and salt content) of the tropical Pacific Ocean[Z]. 2017: 21.
- [9] 梅山. 西太平洋海域 Argo 资料同化再分析数据集用户手册[Z]. 2017.
- MEI S. User manual of Argo data assimilation reanalysis dataset of the western Pacific Ocean[Z]. 2017.
- [10] CHENG L J, TRENBERTH K E, GRUBER N, et al. Improved estimates of changes in upper ocean salinity and the hydrological cycle[J]. Journal of Climate, 2020, 33(23): 10357-10381.
- [11] Cheng L J, TRENBERTH K E, FASULLO J, et al. Improved estimates of ocean heat content from 1960 to 2015[J]. Science Advances, 2017, 3(3): e1601545.
- [12] HOLLINGSWORTH A, LÖNNBERG P. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: the wind field[J]. Tellus A, 1986, 38A(2): 111-136.
- [13] MEYERS G, PHILLIPS H, SMITH N, et al. Space and time scales for optimal interpolation of temperature — Tropical Pacific Ocean[J]. Progress in Oceanography, 1991, 28(3): 189-218.
- [14] BONEKAMP H, VAN OLDENBORGH G J, BURGERS G. Variational Assimilation of Tropical Atmosphere-Ocean and expendable bathythermograph data in the Hamburg Ocean Primitive Equation ocean general circulation model, adjusting the surface fluxes in the tropical ocean[J]. Journal of Geophysical Research, 2001, 106(C8): 16693-16709.
- [15] ZHANG C L, XU J P, BAO X W, et al. An effective method for improving the accuracy of Argo objective analysis[J]. Acta Oceanologica Sinica, 2013, 32(7): 66-77.
- [16] 张春玲, 许建平, 鲍献文. 基于 Argo 资料的梯度依赖相关尺度方法[J]. 解放军理工大学学报(自然科学版), 2015, 16(5): 476-483.
- ZHANG C L, XU J P, BAO X W. Gradient-dependent correlation scale method based on Argo[J]. Journal of PLA University of Science and Technology (Natural Science Edition), 2015, 16(5): 476-483.
- [17] ZHANG C L, WANG Z F, LIU Y. An Argo-based experiment providing near-real-time subsurface oceanic environmental information for fishery data[J]. Fisheries Oceanography, 2021, 30(1): 85-98.
- [18] LI Z Q, LIU Z H, LU S L. Global Argo data fast receiving and post-quality-control system. IOP Conference Series: Earth and Environmental Science, 2020, 502: 012012.
- [19] LIU Z H, LI Z Q, LU S L, et al. Scattered dataset of global ocean temperature and salinity profiles from the international Argo program[J]. Journal of Global Change Data & Discovery, 2021, 5(4): 312-321.
- [20] AKIMA H. A new method of interpolation and smooth curve fitting based on local procedures[J]. Journal of the ACM, 1970, 17(4): 589-602.
- [21] ZHANG C L, ZHANG M L, WANG Z F, et al. Thermocline model for estimating Argo sea surface temperature[J]. Sustainable Marine Structures, 2022, 4(1): 1-12.
- [22] DERBER J, ROSATI A. A global oceanic data assimilation system [J]. Journal of Physical Oceanography, 1989, 19(9): 1333-1347.
- [23] CHU P C, TSENG H C, CHANG C P, et al. South China Sea warm pool detected in spring from the Navy's Master Oceanographic Observational Data Set (MOODS) [J]. Journal of Geophysical Research, 1997, 102(C7): 15761-15771.
- [24] 张福昌. 产品验收中的数理统计方法[M]. 北京: 中国对外经济贸易出版社, 1987.
- ZHANG F C. Mathematical statistical methods in product acceptance[M]. Beijing: China Foreign Economic Relations and Trade Press, 1987.
- [25] 梁冯珍. 应用概率统计[M]. 天津: 天津大学出版社, 2004.
- LIANG F Z. Applied probability and statistics[M]. Tianjin: Tianjin University Press, 2004.

Construction and validation of global multi-parameter Argo dataset based on gradient-dependent OI

WANG Danyang¹, ZHANG Chunling^{1,2}, LU Shaolei^{3,4*}, LI Zhaoqin^{3,4}, LIU Zenghong^{3,4}

(1. College of marine sciences, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Marine Ecological Monitoring and Restoration Technologies, Ministry of Natural Resources, Shanghai 201306, China; 3. Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China; 4. State Key Laboratory of Satellite Ocean Environment Dynamics, Ministry of Natural Resources, Hangzhou 310012, China)

Abstract: Based on the gradient-dependent optimal interpolation objective analysis system, only, a global multi-parameter Argo gridded dataset with a spatial resolution of $1^\circ \times 1^\circ$ from 2004 to 2020 is constructed using the Argo observation in this paper. A series of validations are made for this dataset including confidence interval estimation, observation inspection and comparison with other datasets. The results show that more than 90% of the reconstructed temperature and salinity are reliable under the statistical probability of 95% with the maximum bias from observations are less than $\pm 1.0^\circ\text{C}$ and ± 0.02 , respectively. The large-scale signals reflected in this dataset are consistent with the existing datasets, and more small and medium-scale signals can be retained. The analysis results are closer to the observations.

Key words: gradient-dependent optimal interpolation; Argo; multi-parameter; gridded dataset; objective analysis