Synoptic monthly gridded global and regional four-dimensional World Ocean Database and Global Temperature and Salinity Profile Programme (*T*, *S*, *u*, *v*) fields with the optimal spectral decomposition and P-vector methods

Peter C. Chu and Chenwu Fan

Naval Ocean Analysis and Prediction (NOAP) Laboratory, Department of Oceanography, Naval Postgraduate School, Monterey, CA, USA

Correspondence: P. C. Chu, Naval Postgraduate School, Monterey, CA 93943, USA, E-mail: pcchu@nps.edu

The research presented in this paper was funded by the National Oceanographic Data Center (now National Centers for Environmental Information) and the Naval Oceanographic Office.

Synoptic monthly gridded (SMG) world ocean temperature (*T*) and salinity (*S*) fields with (1° × 1°) horizontal resolution and 28 standard vertical levels from the ocean surface to 3000 m deep have been established from the observational (*T*, *S*) profiles from the NOAA National Centers for Environmental Information (NCEI) World Ocean Database (WOD) from January 1945 to December 2014 and the Global Temperature and Salinity Profile Programme (GTSPP) from January 1990 to December 2009 using the optimal spectral decomposition (OSD) method. The world ocean abstract geostrophic currents (*u*, *v*) are calculated from the 4D SMG-WOD and SMG-GTSPP (*T*, *S*) data using the P-vector inverse method. The SMG-WOD (*T*, *S*, *u*, *v*) and SMG-GTSPP (*T*, *S*, *u*, *v*) fields with a higher horizontal resolution (0.25° × 0.25°) have also been produced for the Mediterranean Sea, the Japan/East Sea, and the Gulf of Mexico.

Geosci. Data J. (2017), https://doi.org/10.1002/gdj3.48

Received: 12 March 2017, revised: 12 June 2017, accepted: 21 August 2017

Key words: WOD, Global Temperature and Salinity Profile Program, SMG-WOD, SMG-GTSPP, world ocean 4D (T, S, u, v) fields, OSD method, P-vector method, global heat content, global freshwater content

Dataset

(a) Identifier: NCEI Accession 0140938
Creator: NOAP Lab, Department of Oceanography, Naval Postgraduate School, Monterey, California, USA
Title: Synoptic monthly gridded three-dimensional World Ocean Database temperature and salinity from January 1945 to December 2014
Publisher: NOAA/NCEI, Silver Spring, Maryland, USA
Publication year: 2016
http://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0140938

(b) Identifier: NCEI Accession 0146195
Creator: NOAP Lab, Department of Oceanography, Naval Postgraduate School, Monterey, California, USA
Title: Synoptic monthly gridded WOD absolute geostrophic velocity (SMG-WOD-V) (January 1945–December 2014) with the P-vector method
Publisher: NOAA/NCEI, Silver Spring, Maryland, USA
Publication year: 2016
http://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0146195

(c) Identifier: NCEI Accession 0138647
Creator: NOAP Lab, Department of Oceanography, Naval Postgraduate School, Monterey, California, USA
Title: Synoptic monthly gridded GTSPP from January 1990 to December 2009
Publisher: NOAA/NCEI, Silver Spring, Maryland, USA
Publication year: 2016
http://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0138647

(d) Identifier: NCEI Accession 0157702
Creator: NOAP Lab, Department of Oceanography, Naval Postgraduate School, Monterey, California,

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

UNCORRECTED PROOF

G D J 3 Journal Code | 48 Manuscript No. | WILEY | Dispatch: 13.10.17 No. of pages: 22 | CE: Raja Rajeswari PE: Kaviarasi N.

USA
Title: Synoptic monthly gridded (0.25°) Mediterranean Sea (T, S, u, v) dataset (January 1960–December 2013) from the NOAA/NCEI WOD Profile Data
Publisher: NOAA/NCEI, Silver Spring, Maryland, USA
Publication year: 2016
https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0157702

(e) Identifier: NCEI Accession 0157703
Creator: NOAP Lab, Department of Oceanography, Naval Postgraduate School, Monterey, California, USA
Title: Synoptic monthly gridded (0.25°) Japan/East Sea (T, S, u, v) dataset (January 1960–December 2013) from the NOAA/NCEI WOD Profile Data
Publisher: NOAA/NCEI, Silver Spring, Maryland, USA
Publication year: 2016
https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0157703

(f) Identifier: NCEI Accession 0156423
Creator: NOAP Lab, Department of Oceanography, Naval Postgraduate School, Monterey, California, USA
Title: Synoptic monthly gridded (0.25°) Gulf of Mexico (T, S, u, v) dataset (January 1945–December 2014) from the NOAA/NCEI WOD Profile Data
Publisher: NOAA/NCEI, Silver Spring, Maryland, USA
Publication year: 2016
http://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0156423

(g) Identifier: NCEI Accession 0138646
Creator: NOAP Lab, Department of Oceanography, Naval Postgraduate School, Monterey, California, USA
Title: Volume transport stream function calculated from World Ocean Atlas 2013 (WOA13-VTSF) and climatological wind
Publisher: NOAA/NCEI, Silver Spring, Maryland, USA
Publication year: 2016
https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0138646

Introduction

Ocean observational temperature (T) and salinity (S) profiles after quality control are available in the NOAA/National Centers for Environmental Information (NCEI) for public use: World Ocean Database (WOD; Boyer *et al.*, 2013) and the Global Temperature and Salinity Profile Programme (GTSPP; Sun *et al.*, 2009). Both datasets are not evenly distributed in time and space. The NOAA/NCEI WOD was established more than five decades ago and has been serving the world oceanographic community ever since. There are 12 808 409 (T, S) stations in WOD13 dataset. Interested readers can find detailed information about WOD from the NOAA Atlas NESDIS 72 (Boyer *et al.*, 2013) or from the website: https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html. However, the GTSPP data are relatively new (Sun *et al.*, 2009). There are total 6 662 209 (4 273 991) GTSPP temperature (salinity) stations with yearly profile number <80 000 from 1990 to 1999, increasing exponentially to 1.7 million in 2009 from 1999 as the Argo floats into practice (Figure 1). The GTSPP stations are over the global oceans, as shown in Figure 2a for temperature and in Figure 2b for salinity.

The absolute geostrophic currents (u, v) can be determined from (T, S) profiles using the inverse methods (e.g. Stommel and Schott, 1977; Wunsch, 1978; Chu, 1995, 2006). To enhance their utility to the oceanographic and climate research, as well as operational environmental forecasting communities, establishment of objectively analysed synoptic monthly gridded (SMG) four-dimensional (T, S, u, v) datasets on regular spatial grid points and time step (1 month) is an urgent need in various aspects such as use as boundary and/or initial conditions in numerical ocean circulation models, verification of numerical simulations of the ocean, as a form of 'sea truth' for satellite measurements such as altimetry observations of sea surface height among others (Boyer *et al.*, 2013). The optimal spectral decomposition (OSD) method is used to establish 4D objectively analysed global SMG (T, S) datasets with monthly increment, $1^\circ \times 1^\circ$ ($0.25^\circ \times 0.25^\circ$ for the Mediterranean Sea, the Japan/East Sea, and the Gulf of Mexico) horizontal resolution, and 28 standard NCEI vertical levels from the ocean surface to 3000 m deep (Table 1). With such a vertical resolution, that is, 10 m in upper 50 m depth, 25 m between 50 m and 150 m, etc., it cannot well represent global ocean mixed layer and the thermocline variability. The authors are working on the derived WOD – isothermal (isohaline, isopycnal) layer temperature (salinity, density) and depth as well as thermocline (halocline, pycnocline) strength. The P-vector inverse method is used to establish 4D objectively analysed SMG absolute geostrophic current (u, v) datasets.

1. Data production methods

1.1. OSD method

The OSD method (Chu *et al.*, 2003a,b, 2004, 2015, 2016), used to produce SMG-(T, S) datasets, can be outlined as follows. Let $\mathbf{r} = (x, y)$ be the horizontal coordinates and z the vertical coordinate. The horizontal position vector ($\mathbf{r}$) is represented by $\mathbf{r}_n$ ($n = 1, 2, \ldots, N$) at grid points and by $\mathbf{r}^{(m)}$ ($m = 1, 2, \ldots, M$) at observational locations. Here, N is the total number of the grid points, and M is the total number of

UNCORRECTED PROOF

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

LOW RESOLUTION FIG

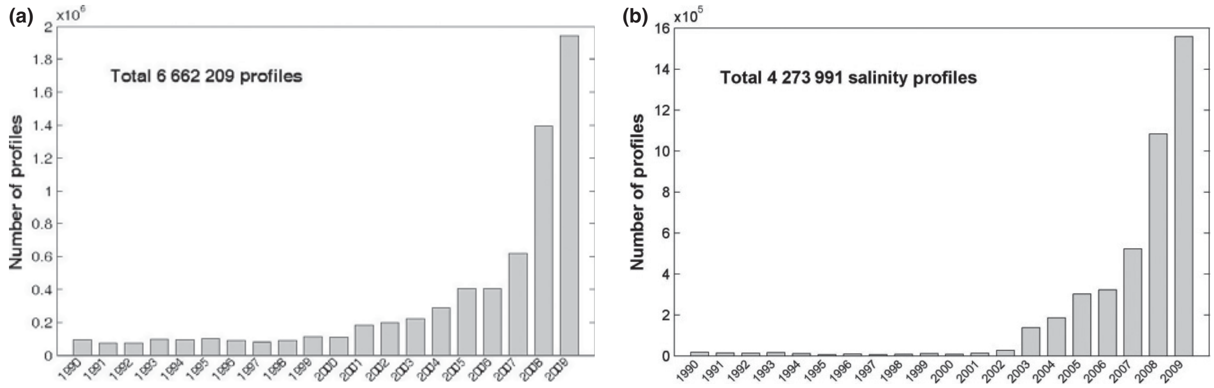


Figure 1. Temporal distribution of (a) Global Temperature and Salinity Profile Programme temperature and (b) salinity stations.

UNCORRECTED PROOF

Table 1. Vertical depths of the synoptic monthly gridded data

Layer	Depth (m)	Layer	Depth (m)	Layer	Depth (m)
1	0	11	250	21	1200
2	10	12	300	22	1300
3	20	13	400	23	1400
4	30	14	500	24	1500
5	50	15	600	25	1750
6	75	16	700	26	2000
7	100	17	800	27	2500
8	125	18	900	28	3000
9	150	19	1000		
10	200	20	1100		

observational points. Gridded temperature and salinity can be ordered by grid point and by variable, forming a single vector $\boldsymbol{c} = (T, S)$ of length NP with N the total number of grid points and P the number of variables. For example, the background field ($\mathbf{c}_b$) is on the grid points and represented as

$$\mathbf{c}_b^T = [c_b(\mathbf{r}_1), c_b(\mathbf{r}_2), \ldots, c_b(\mathbf{r}_N)] \tag{1}$$

where the superscript 'T' means transpose. The observation ($\mathbf{c}_o$) is on the observational points and represented by

$$\mathbf{c}_o^T = \left[c_o(\mathbf{r}^{(1)}), c_o(\mathbf{r}^{(2)}), \ldots, c_o(\mathbf{r}^{(M)})\right] \tag{2}$$

Usually, the data analysis and assimilation is to blend $\mathbf{c}_b$ (at the grid points $\mathbf{r}_n$) with observational data ($\mathbf{c}_o$) (at observational points $\mathbf{r}^{(m)}$) into the assimilated (or analysis) field ($\boldsymbol{c}_a$) at the grid points $\mathbf{r}_n$,

$$\mathbf{c}_a = \mathbf{c}_b + \mathbf{Wd}, \quad \mathbf{d} \equiv (\mathbf{c}_o - \mathbf{Hc}_b) \tag{3}$$

to represent the (unknown) 'truth' $\mathbf{c}_t$ with an analysis error (ε_a) and an observational error (ε_o)

$$\varepsilon_a = \mathbf{c}_a - \mathbf{c}_t, \quad \varepsilon_o \equiv \mathbf{H}^T\mathbf{c}_o - \mathbf{c}_t$$

Here, $\mathbf{H} = [h_{mn}]$ is the $M \times N$ linear observation operator matrix; $\mathbf{d}$ is the innovation (also called the observational increment) at the observational points $\mathbf{r}^{(m)}$; $\mathbf{W} = [w_{nm}]$, is the $N \times M$ weight matrix interpolating the innovation $\mathbf{d}$ into the grid points $\mathbf{r}_n$ (Figure 3). The background and observational error covariance matrices should be given as *a priori* in order to determine the weight matrix $\mathbf{W}$, such as in the optimal interpolation (OI), Kalman filter, and variational method (3DVar, 4DVar, ...).

On the other hand, the OSD has been developed without using the weight matrix (Chu *et al.*, 2003a,b). Existence of a lateral boundary (Γ) for an ocean domain (Ω) provides a great opportunity to use a spectral method in ocean data analysis and assimilation through decomposing the variable anomaly at the grid points $[c(\mathbf{r}_n) - c_b(\mathbf{r}_n)]$ into the spectral form (Chu *et al.*, 2015),

$$c_a(\mathbf{r}_n) - c_b(\mathbf{r}_n) = s_K(\mathbf{r}_n), \quad s_K(\mathbf{r}_n) \equiv \sum_{k=1}^{K} a_k\phi_k(\mathbf{r}_n) \tag{4}$$

where $\{\phi_k\}$ are basis functions; K is the mode truncation, which is determined using the steep-descending method (Chu *et al.*, 2015). The eigenvectors of the Laplace operator with the same lateral boundary condition of $(c - c_b)$ can be used as the basis functions $\{\phi_k\}$. The $K \times N$ basis function matrix $\boldsymbol{\Phi}$ is calculated by

$$\Phi = \{\phi_{kn}\} = \begin{bmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \ldots & \phi_K(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & \ldots & \phi_K(\mathbf{r}_2) \\ \ldots & \ldots & \ldots & \ldots \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & \ldots & \phi_K(\mathbf{r}_N) \end{bmatrix} \tag{5}$$

Figure 4 shows the first three basis functions for the Atlantic Ocean, Indian Ocean, and Pacific Oceans.

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.

LOW RESOLUTION FIG

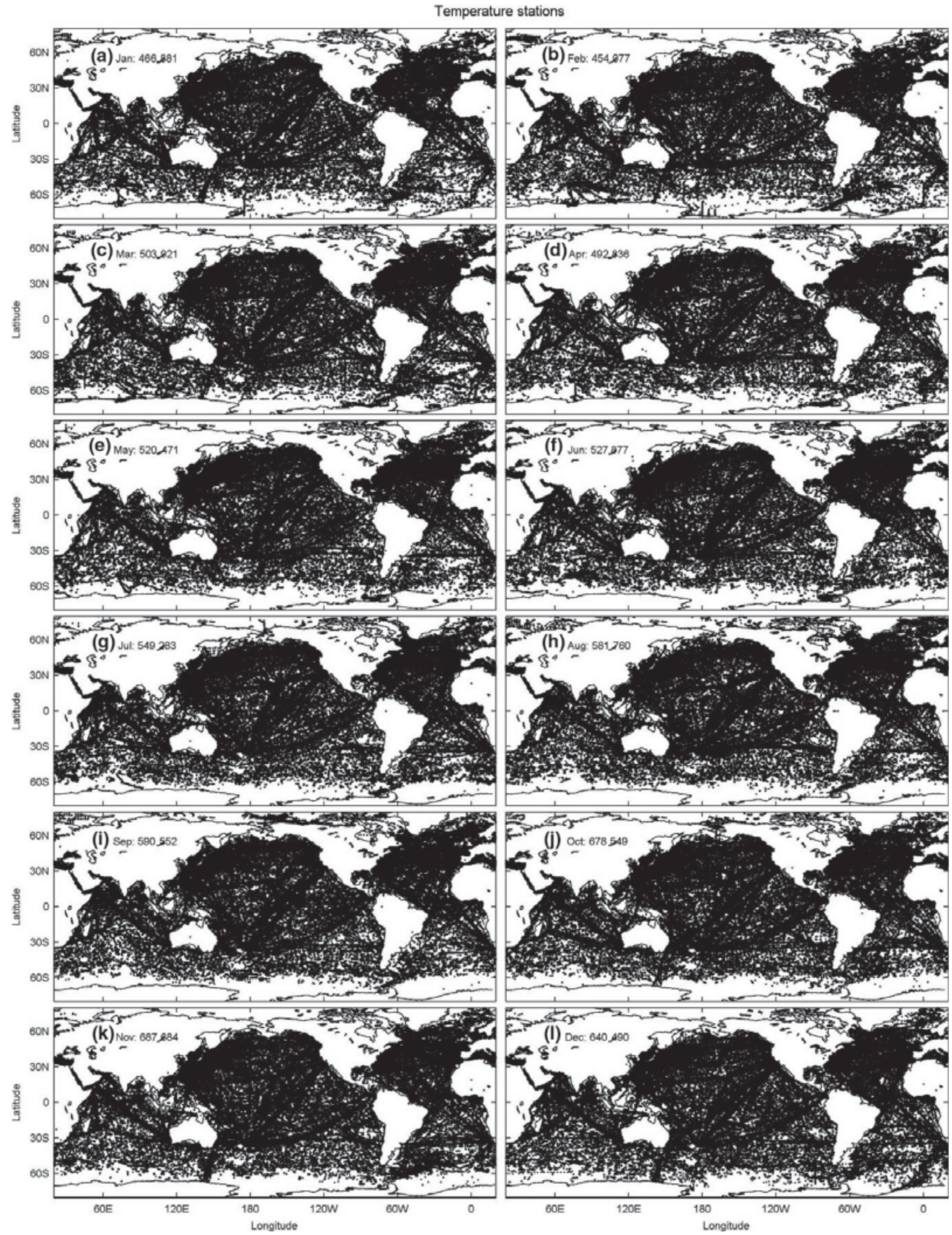


Figure 2. (a) Spatial distributions of monthly Global Temperature and Salinity Profile Programme (GTSPP) stations during 1990–2009. (b) Spatial distributions of monthly GTSPP stations during 2006–2009.

The first basis-function $\phi_1(\mathbf{x}_n)$ for all the three oceans shows a near-zonal structure. The second basis function $\phi_2(\mathbf{x}_n)$ shows a near-meridional structure for the Indian Ocean and Pacific Ocean and for the lower latitudes (30°N–30°S) of the Atlantic Ocean. The third basis-function $\phi_3(\mathbf{x}_n)$ shows the east–west slanted dipole pattern with opposite signs in northeastern and southwestern regions in the Indian Ocean and Pacific Ocean, and near-meridional structure in the North Atlantic and near-zonal structure in the South Atlantic. The higher order basis functions have more complicated variability structures. In producing the SMG-WOD and

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

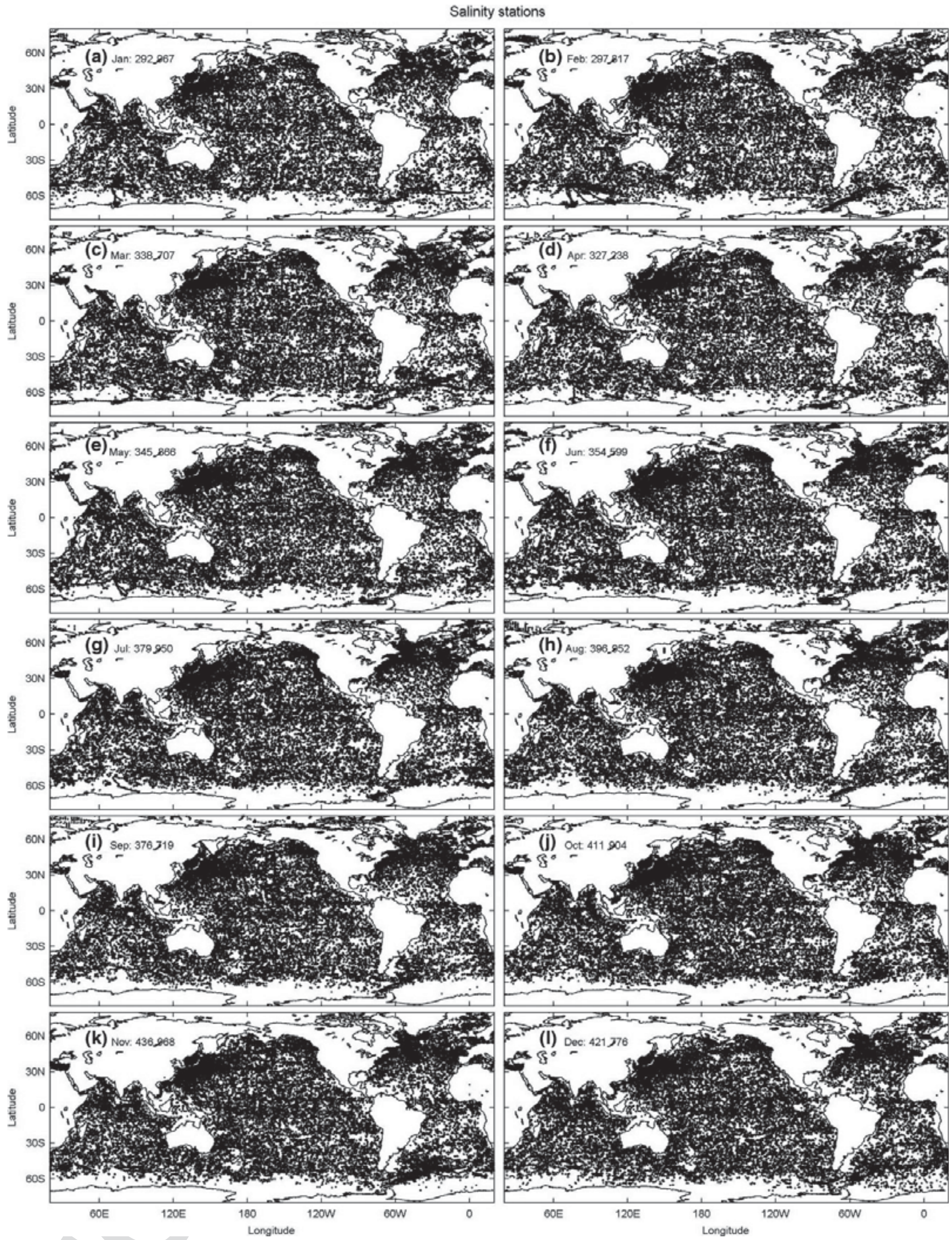


Figure 2. Continued

SMG-GTSPP (T, S) data, around 30 basis functions are used. The OSD data assimilation equation is given by (Chu *et al.*, 2016)

$$\mathbf{c}_a = \mathbf{c}_b + \mathbf{F}\Phi^T\left[\Phi\mathbf{F}\Phi^T\right]^{-1}\Phi\mathbf{H}^T\mathbf{d} \qquad (6)$$

where $\mathbf{F}$ is an $N \times N$ diagonal observational contribution matrix

$$\mathbf{F} = \begin{bmatrix} f_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & f_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ldots_i & 0 & 0 & 0 \\ 0 & 0 & 0 & f_n & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 0 & 0 & f_N \end{bmatrix}, \quad f_n \equiv \sum_{m=1}^{M} h_{nm} \qquad (7)$$

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.

LOW RESOLUTION FIG

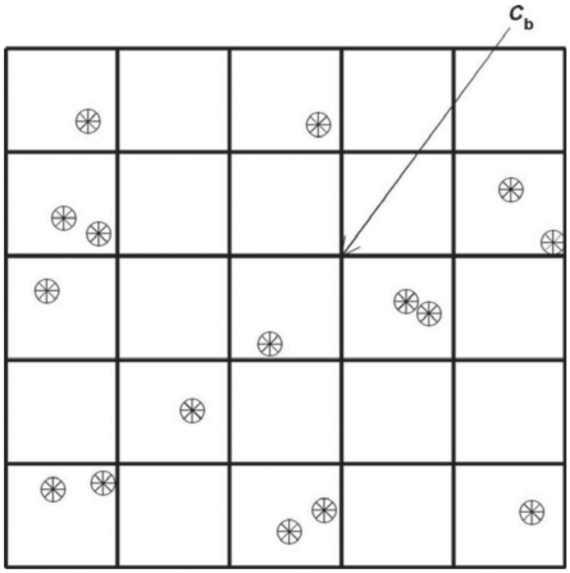


Figure 3. Illustration of ocean data assimilation with $\boldsymbol{c}_b$ located at the grid points, and $\boldsymbol{c}_o$ located at the points '*'. The ocean data assimilation is to convert the innovation, $\boldsymbol{d} = \boldsymbol{c}_o - \mathbf{H}\boldsymbol{c}_b$, from the observational points to the grid points.

The M-dimensional innovation vector [see (3)]

$$\mathbf{d}^T = \left[d(\mathbf{r}^{(1)}), d(\mathbf{r}^{(2)}), \ldots, d(\mathbf{r}^{(M)})\right]$$

at observational points can be transformed into the grid points

$$D_n \equiv D(\mathbf{r}_n) = c_o(\mathbf{r}_n) - c_b(\mathbf{r}_n) = \frac{\sum_{m=1}^{M} h_{nm} d^{(m)}}{f_n} \tag{8}$$

In computing SMG-WOD and SMG-GTSPP (T, S) data, the World Ocean Atlas 2013 (WOA13) monthly mean fields were used as the background $\mathbf{c}_b$. The analysis error (i.e. analysis c_a *vs* 'truth' c_t) at the grid points is given by

$$\begin{aligned}\varepsilon_a(\mathbf{r}_n) &\equiv c_a(\mathbf{r}_n) - c_t(\mathbf{r}_n)\\ &= [c_a(\mathbf{r}_n) - c_b(\mathbf{r}_n)] - [c_o(\mathbf{r}_n) - c_b(\mathbf{r}_n)] + [c_o(\mathbf{r}_n) - c_t(\mathbf{r}_n)]\\ &= s_K(\mathbf{r}_n) - D(\mathbf{r}_n) + \varepsilon_o(\mathbf{r}_n)\end{aligned} \tag{9}$$

Here, Equations (4) and (8) are used. The analysis error is decomposed into two parts

$$\varepsilon_a(\mathbf{r}_n) = \varepsilon_K(\mathbf{r}_n) + \varepsilon_o(\mathbf{r}_n) \tag{10}$$

with the truncation error given by

$$\varepsilon_K(\mathbf{r}_n) = s_K(\mathbf{r}_n) - D(\mathbf{r}_n) \tag{11a}$$

and the observational error given by

$$\varepsilon_o(\mathbf{r}_n) = c_o(\mathbf{r}_n) - c_t(\mathbf{r}_n) \tag{11b}$$

The analysis error variance over the whole domain is given by

$$\begin{aligned}E_a^2 &\equiv \left\langle\left[\varepsilon_a^T \mathbf{F} \varepsilon_a\right]\right\rangle \le \left\langle\left[\varepsilon_K^T \mathbf{F} \varepsilon_K\right]\right\rangle\\ &+ 2\sqrt{\left\langle\left[\varepsilon_K^T \mathbf{F} \varepsilon_K\right]\right\rangle}\sqrt{\left\langle\left[\varepsilon_o^T \mathbf{F} \varepsilon_o\right]\right\rangle} + \left\langle\left[\varepsilon_o^T \mathbf{F} \varepsilon_o\right]\right\rangle\\ &= E_K^2 + 2E_K\sqrt{M/N}e_o + (M/N)e_o^2\end{aligned} \tag{12}$$

where $E_K^2 = \left\langle\left[\varepsilon_K^T \mathbf{F} \varepsilon_K\right]\right\rangle$ is the truncation error, e_o^2 is the observational error variance, and the Cauchy–Schwarz inequality is used. The mode truncation is optimally determined from the relative analysis error reduction

$$\gamma_K = \ln\left[\frac{E_{K-1}^2 + 2E_{K-1}\sqrt{M/N}e_o + Me_o^2/N}{E_K^2 + 2E_K\sqrt{M/N}e_o + Me_o^2/N}\right] \quad K = 2, 3, \ldots \tag{13}$$

exceeding the threshold γ_{th},

$$K_{OPT} = \max_{\gamma_K \ge \gamma_{th}} (K) \tag{14}$$

After the mode truncation K_{OPT} is determined, the spectral coefficients (a_k, k = 1, 2, …, K_{OPT}) in Equation (4) can be calculated, and so as the truncation error variance $E_{K_{OPT}}^2$.

Advantages of the OSD *versus* other data assimilation methods such as the OI, Kalman filter, and variational method (3DVar, 4DVar, …) are no background error covariance matrix needed, capable of filtering out noises since the basis functions are the eigenvectors of the Laplace operator, and continuity of the temperature and salinity at the boundary connecting the regional and global products.

The third advantage is due to the same boundary condition used to obtain the basis functions from the regional sea side and the open ocean side, which guarantees the continuity of the basis functions and in turn the (T, S) fields at the open boundary. Total numbers of temperature and salinity observations increase dramatically after 2000 because of Argo project (see Figures 1 and 2), which indicated that the quality of SMG-WOD and SMG-GTSPP synoptic monthly gridded dataset in that time span is better than those before 2000s.

The OSD method has been proven an effective ocean data analysis method (Chu *et al.*, 2016). With it, several new ocean phenomena have been identified from observational data such as a bimodal structure of chlorophyll-*a* with winter/spring (February–March) and fall (September–October) blooms in the Black Sea (Chu *et al.*, 2005a), fall–winter recurrence of current reversal from westward to eastward on the Texas–Louisiana continental shelf from the current meter, near-surface drifting buoy (Chu *et al.*, 2005b),

UNCORRECTED PROOF

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

LOW RESOLUTION COLOR FIG

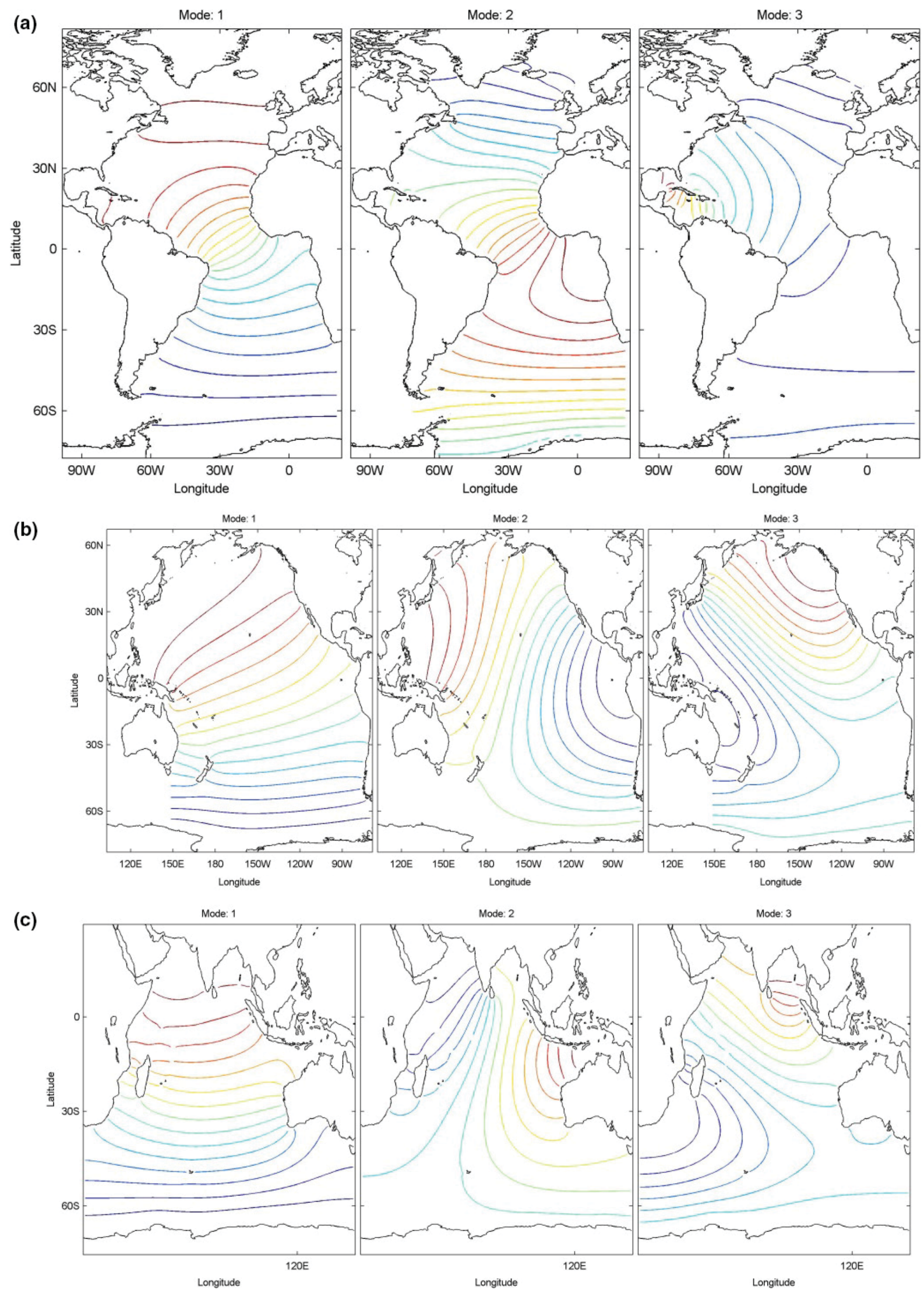


Figure 4. First three basis functions for the (a) Atlantic Ocean, (b) Pacific Ocean, and (c) Indian Ocean.

propagation of long Rossby waves at mid-depths (around 1000 m) in the tropical North Atlantic from the Argo float data (Chu *et al.*, 2007), and temporal and spatial variability of global upper ocean heat content (Chu, 2011).

1.2. P-vector method

The P-vector inverse method was first proposed by Chu (1995) and described in detail in a book by Chu (2006). It can also be found from the authors' previous paper (Chu and Fan, 2015) published in the *Geoscience Data Journal* (see the website: http://onlinelibrary.wiley.com/doi/10.1002/gdj3.31/full).

2. Quality control of WOD/GTSPP profile data

Quality control (QC) has been conducted on the WOD and GTSPP temperature and salinity profiles. The

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.

LOW RESOLUTION COLOR FIG

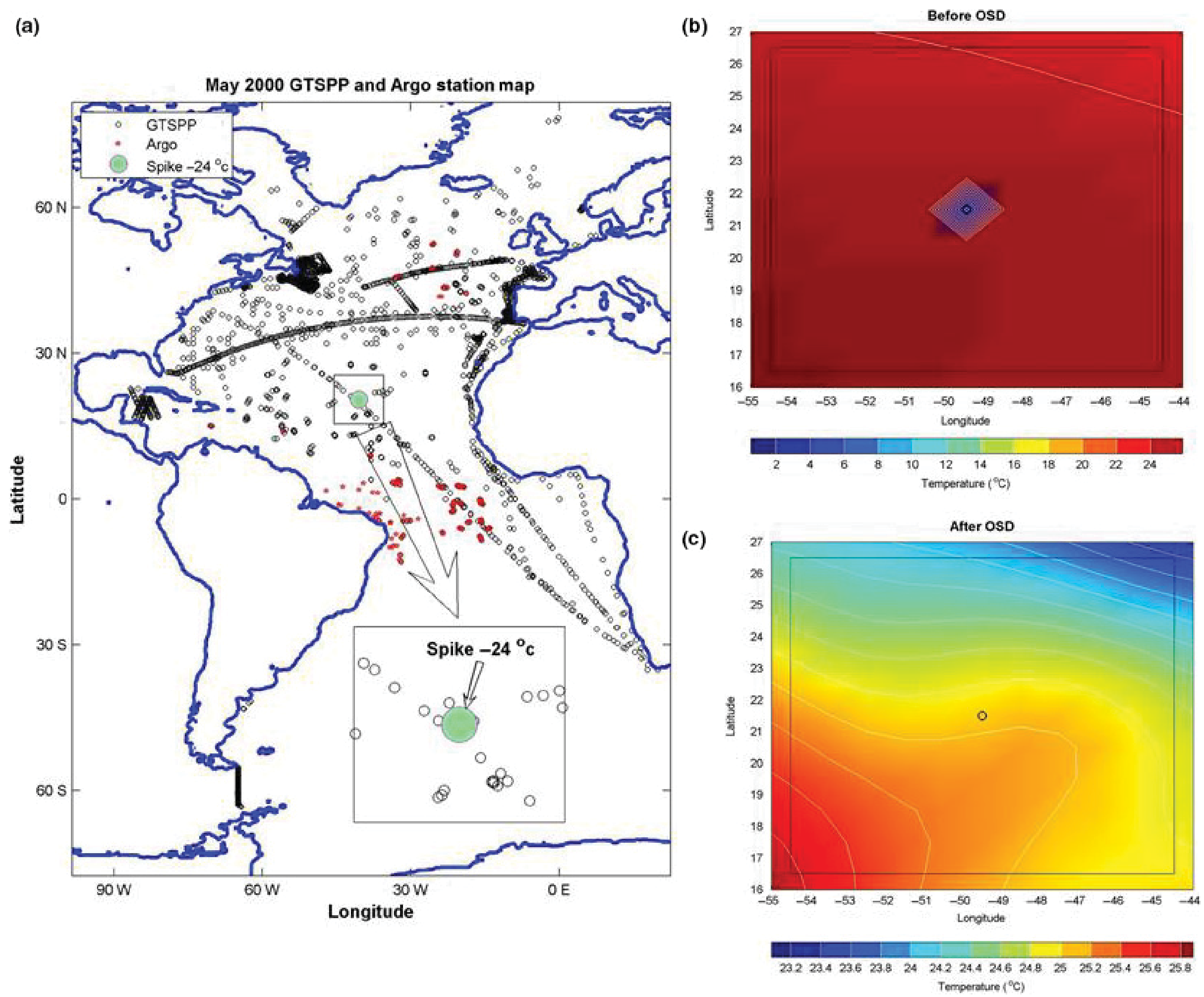


Figure 5. Capability of the optimal spectral decomposition (OSD) method to eliminate high noisy data (from Sun and Chu, 2013): (a) ocean temperature profile data from Global Temperature and Salinity Profile Programme (GTSPP)/Argo with green dot indicating the location of a fake profile by subtracting 24°C from its original one from top to bottom artificially, (b) sea surface temperature field before using the OSD method showing a closed up fake data point (green dot in a), and (c) the sea surface temperature filed after using the OSD method showing the elimination of the fake data point.

general procedures include position and time checks from platforms track looking, min/max range checks, spike checks (excessive gradients and inversions), stability checks, standard deviation outlier checks, and duplication checks either by having received the data more than once, or same profile with different resolutions. QC flags values carried with the observations represent particular QC tests failed, not a verdict on the overall quality of the data (Sun, 2013; Boyer, 2014). As mentioned by Boyer (2014), 'All QC (automatic and manual) for the WOD are performed with a specific purpose in mind – the calculation of the World Ocean Atlas (WOA) climatological mean fields and ocean heat/salt content calculations. The tests deliberately flag data which may be good observations, but do not represent a long-term mean or (for heat/salt calculations) a short-term large scale pattern'. However, the QC for GTSPP does not have such a constraint. The WOA climatology is taken as a measure for the statistical tests (Sun, 2013).

Table 2. Extracted data type and characteristics

Data type	Characteristics
char	8-bit characters intended for representing text
byte	8-bit signed or unsigned integers
short	16-bit signed integers
int	32-bit signed integers
Float/real	32-bit IEEE floating point
double	64-bit IEEE floating point

In addition to the general QC procedures, the WOD standard depth level data for the expendable bathythermograph (XBT) and mechanical

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

LOW RESOLUTION COLOR FIG

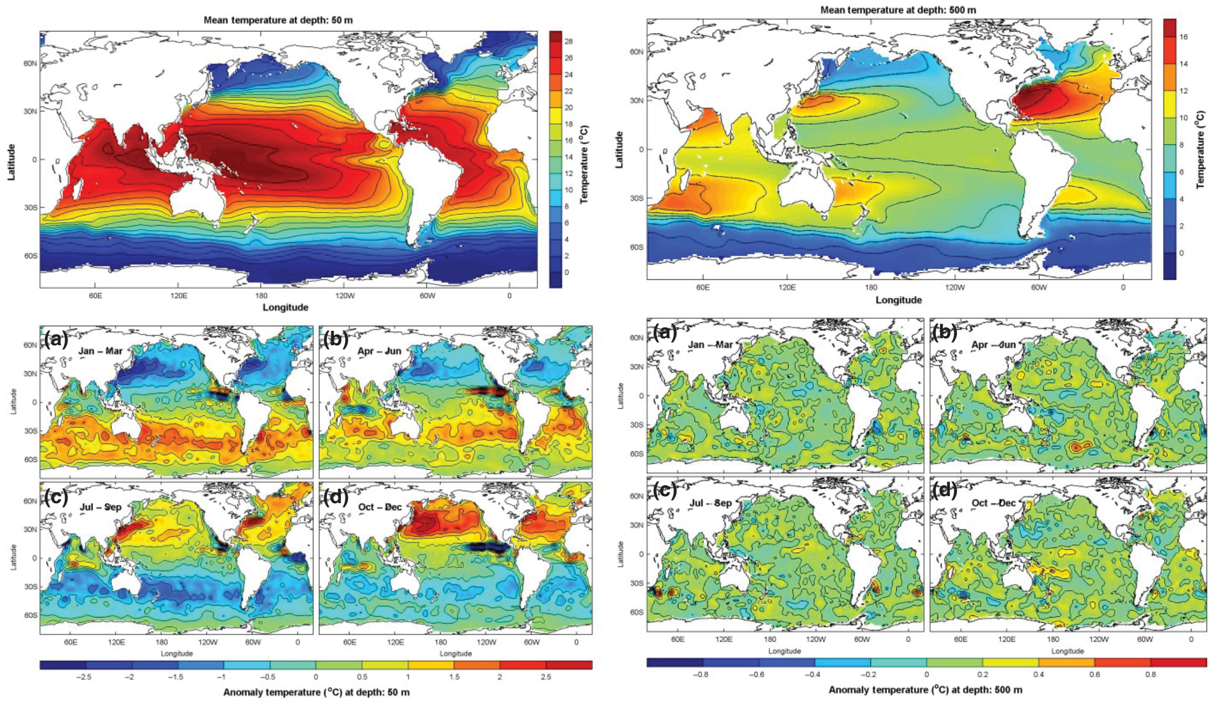


Figure 6. Total mean temperature ($\bar{\bar{T}}$) and seasonal anomalies ($\tilde{T}$) at (a) 50 m and (b) 500 m depths calculated from the SMG-WOD dataset.

bathythermograph (MBT) have bias correction through depth adjustment since 2009. Such an adjustment is to diminish the documented time-dependent temperature biases as described by Levitus *et al.* (2009), which is in addition to the correction of depths to the fall rate for all XBT data. The observed level data do not have any adjustment since only the standard level depth XBT and MBT data are contributed to construct the WOA climatological data. Interested readers are referred to the NCEI website https://www.nodc.noaa.gov/OC5/XBT_BIAS/xbt_bias.html for detailed information.

The GTSPP Argo salinity data have salinity drift. This is 'because 90% of the Argo floats have electrode-type conductivity cell (see Sea-Bird online at http://www.seabird.com/), the "prolonged and unattended" presence in the ocean make them susceptible to fouling–biofilm formation inside the cell, which alters the conductivity and thereby the salinity measurements (Thadathil *et al.*, 2012)'. As mentioned in the Argo data management (Wong *et al.*, 2013), statistical comparison methods that are used to determine conductivity sensor drift (e.g. Wong *et al.*, 2003; Owens and Wong, 2009) are suitable for correction of positive salinity drift; for example, pressure error of -20 dbar will cause a positive salinity error of approximately 0.01 PSS-78. There are no cycles that have salinity drift adjustment ($\Delta S > 0.001$ PSS-78 in bottom data to distinguish from thermal lag adjustment at shallower levels); no cycles following ones that have salinity drift adjustment ($\Delta S > 0.001$ PSS-78 in bottom data); and no cycles in the 6 months prior to salinity drift adjustment ($\Delta S > 0.001$ PSS-78 in bottom data).

Sun and Chu (2013) shows the capability of the OSD method in the QC of the GTSPP/Argo data. Take May 2000 GTSPP and Argo (*T*, *S*) profile data as an example (Figure 5a). In the track of CTD measurements, a fake spike of -24°C is added from the surface to the bottom of the temperature profile at the middle marked by a green spot (Figure 5a). The sea surface temperature filed is unrealistic with a strong closed fake cold data point (Figure 5b). After using the OSD method the fake data have been eliminated (Figure 5c).

3. SMG data

Both SMG-WOD and SMG-GTSPP datasets are in the Network Common Data Form (netCDF) (see the website: http://www.unidata.ucar.edu/software/netcdf/), which is an interface for array-oriented data access, a library for implementation of interface, and a machine-independent format for representing data. The netCDF software was developed at the Unidata (http://www.unidata.ucar.edu) Program Center in Boulder, Colorado. Each element is stored at a disc address which is a linear function of the array indices (subscripts) by which it is identified. Hence, these indices need not be stored separately (as in a relational database). This provides a fast and compact storage method. The external types supported by the netCDF interface are listed in Table 2. These types are chosen to provide a

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.

LOW RESOLUTION COLOR FIG

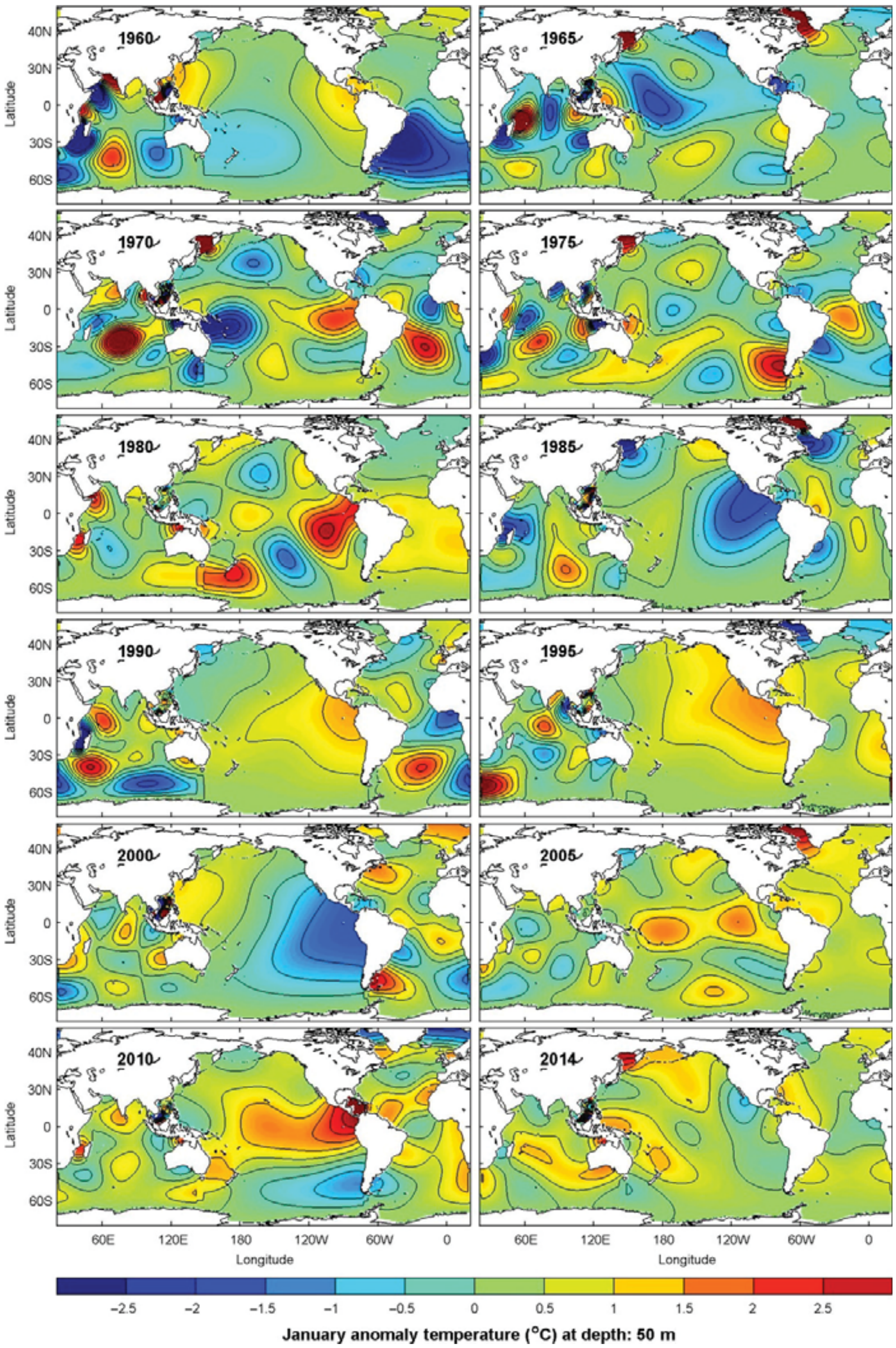


Figure 7. SMG-WOD January temperature anomaly (relative to climatological January mean) (°C) at 50 m depth in selected years: (a) 1960, (b) 1965, (c) 1970, (d) 1975, (e) 1980, (f) 1985, (g) 1990, (h) 1995, (i) 2000, (j) 2005, (k) 2010, (l) 2014.

reasonably wide range of trade-offs between data precision and number of bits required for each value. The external data types are independent from whatever internal data types are supported by a particular machine and language combination. These types of extracted data are called 'external', because they correspond to the portable external representation for netCDF data.

4. Data characteristics

4.1 Four-dimensional data

The SMG-WOD global (1° × 1°) and regional (Gulf of Mexico, Japan/East Sea, and Mediterranean Sea, 0.25° × 0.25°), and SMG-GTSPP (1° × 1°) datasets are four-dimensional with 28 vertical levels (Table 1)

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

LOW RESOLUTION COLOR FIG

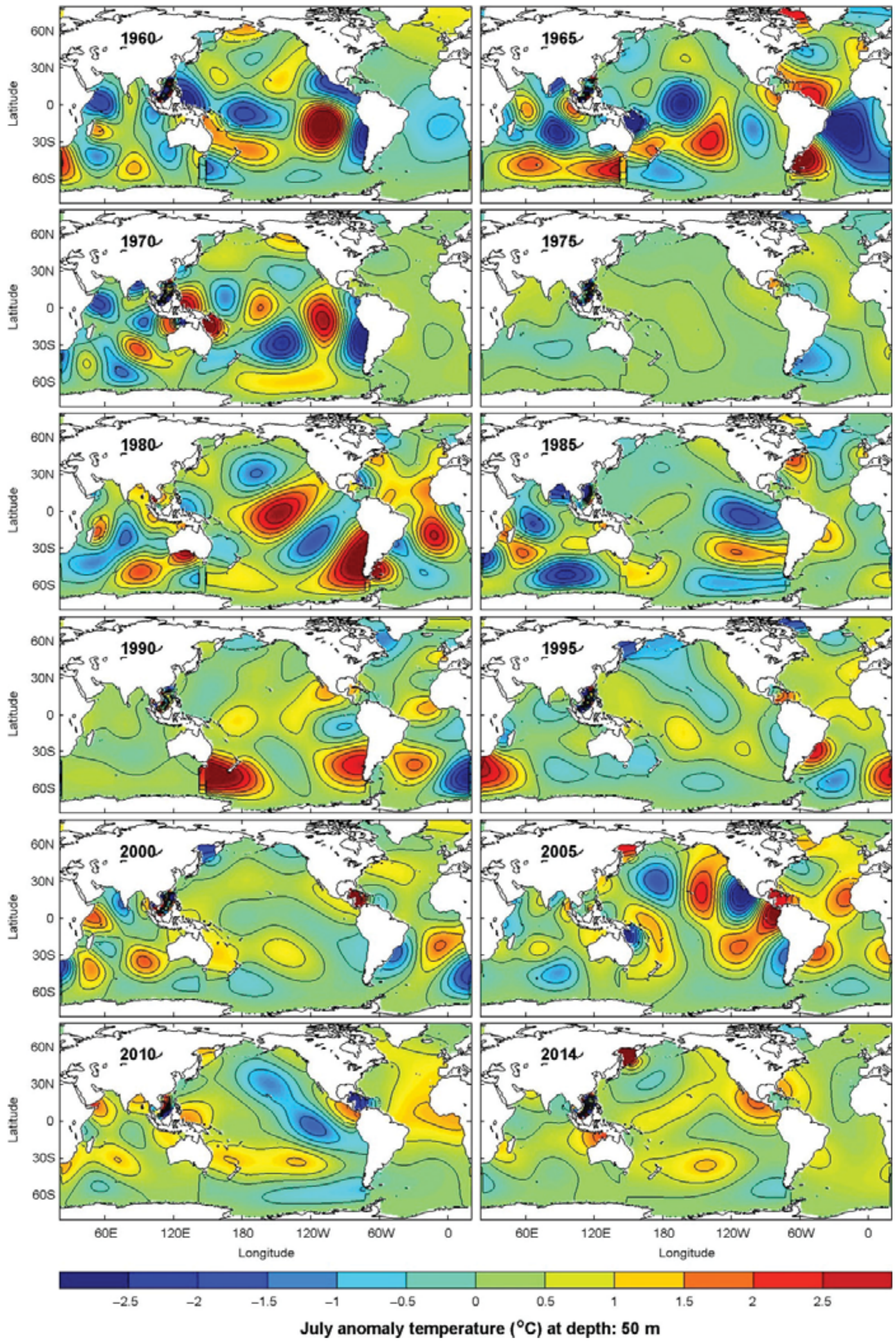


Figure 8. SMG-WOD July temperature anomaly (relative to climatological July mean) (°C) at 50 m depth in selected years: (a) 1960, (b) 1965, (c) 1970, (d) 1975, (e) 1980, (f) 1985, (g) 1990, (h) 1995, (i) 2000, (j) 2005, (k) 2010, (l) 2014.

and monthly time increment: 840 monthly instances (January 1945 to December 2014) for SMD-WOD and 240 monthly instances (January 1990 to December 2009) for SMG-GTSPP. Let $\tilde{t}_s$ be the monthly time, $s = 1, 2, \ldots, S$, with $S = 840$ for SMG-WOD and 240 SMG-GTSPP the total number of months and be represented by two parameters (τ_m, t_l), with τ_m ($m = 1, 2, \ldots, M$) = 1945, 1946, …, 2014 ($M = 70$) for SMG-WOD and 1990, 1991, …, 2009 ($M = 20$) for SMG-GTSPP, the time sequence in years, and $t_l = 1, 2, \ldots, 12$, the monthly sequence within a year. The four-dimensional data (T, S, u, v) can be represented by $\Psi(\mathbf{r}_n, z_j, \tilde{t}_s)$ with $\mathbf{r}_n$ the horizontal grid points, z_j ($j = 1, 2, \ldots, 28$) the vertical levels. Since mesoscale eddies have typical horizontal scales of <100 km and timescales on the order of a month, the SMG-WOD and SMG-GTSPP do not have any mesoscale signal such as the footprint of the mesoscale eddy. To

LOW RESOLUTION COLOR FIG

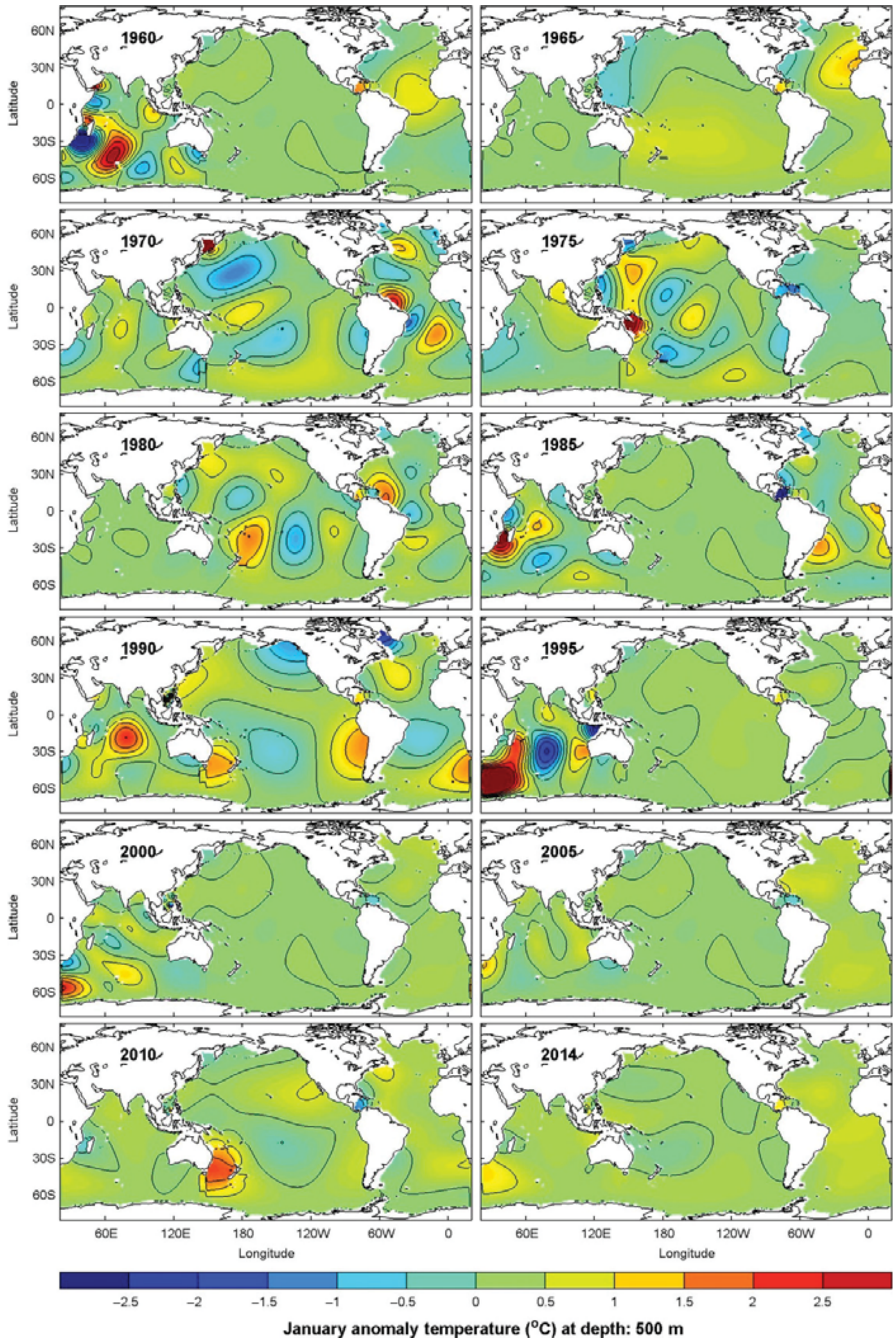


Figure 9. SMG-WOD January temperature anomaly (relative to climatological January mean) (°C) at 500 m depth in selected years: (a) 1960, (b) 1965, (c) 1970, (d) 1975, (e) 1980, (f) 1985, (g) 1990, (h) 1995, (i) 2000, (j) 2005, (k) 2010, (l) 2014.

resolve mesoscale eddy, finer resolution in space and time is needed.

Climatological monthly mean

$$\bar{\Psi}(\mathbf{r}_n, z_j, t_l) = \frac{1}{M}\sum_{m=1}^{M}\Psi(\mathbf{r}_n, z_j\tau_m, t_l) \qquad (15)$$

and total-time mean

$$\bar{\bar{\Psi}}(\mathbf{r}_n, z_j) = \frac{1}{12}\sum_{l=1}^{12}\bar{\Psi}(\mathbf{r}_n, z_j, t_l) \qquad (16)$$

are calculated. Subtraction of $\bar{\bar{\Psi}}(\mathbf{r}_n, z_j)$ from $\bar{\Psi}(\mathbf{r}_n, z_j, t_l)$,

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

LOW RESOLUTION COLOR FIG

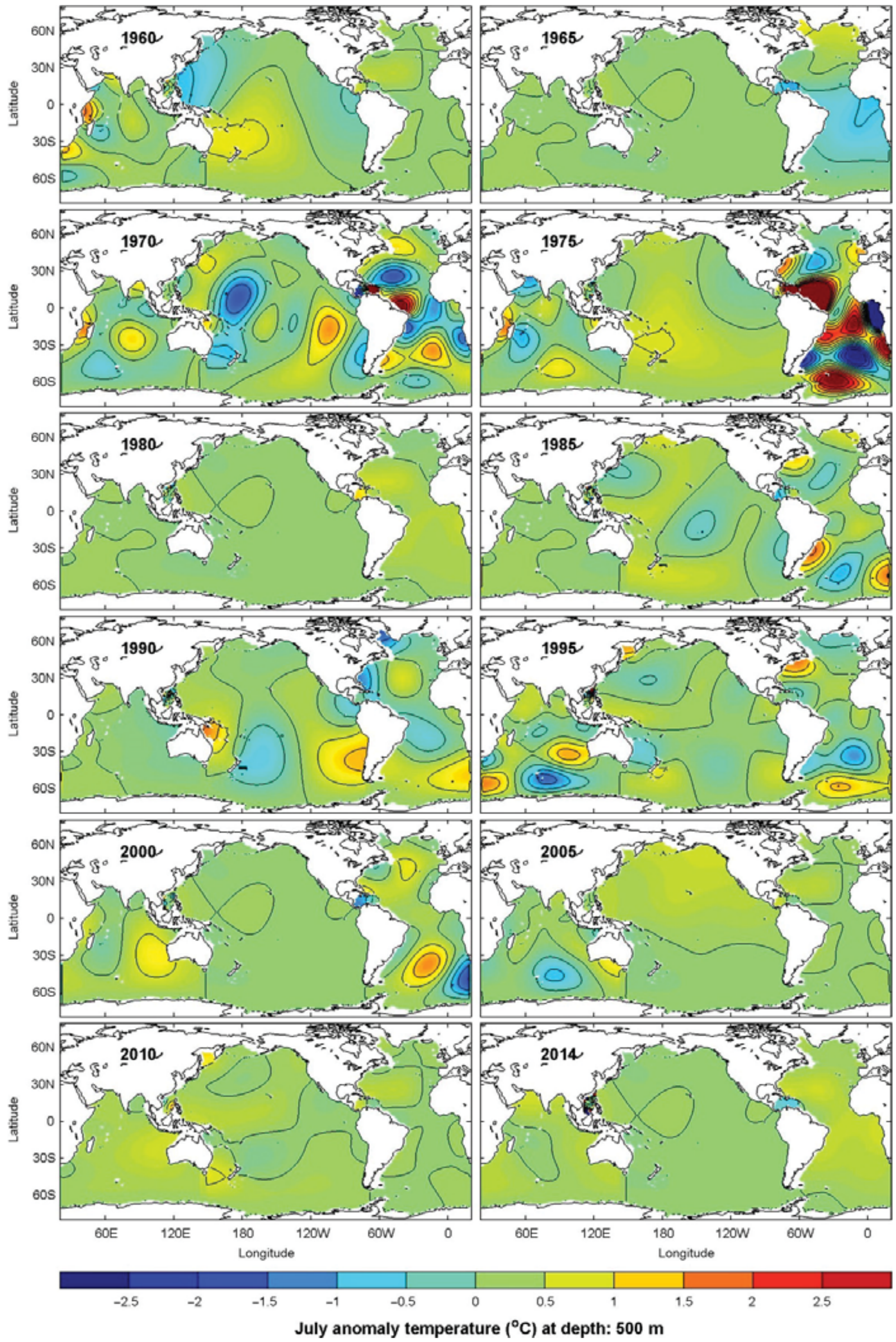


Figure 10. SMG-WOD July temperature anomaly (relative to climatological July mean) (°C) at 500 m depth in selected years: (a) 1960, (b) 1965, (c) 1970, (d) 1975, (e) 1980, (f) 1985, (g) 1990, (h) 1995, (i) 2000, (j) 2005, (k) 2010, (l) 2014.

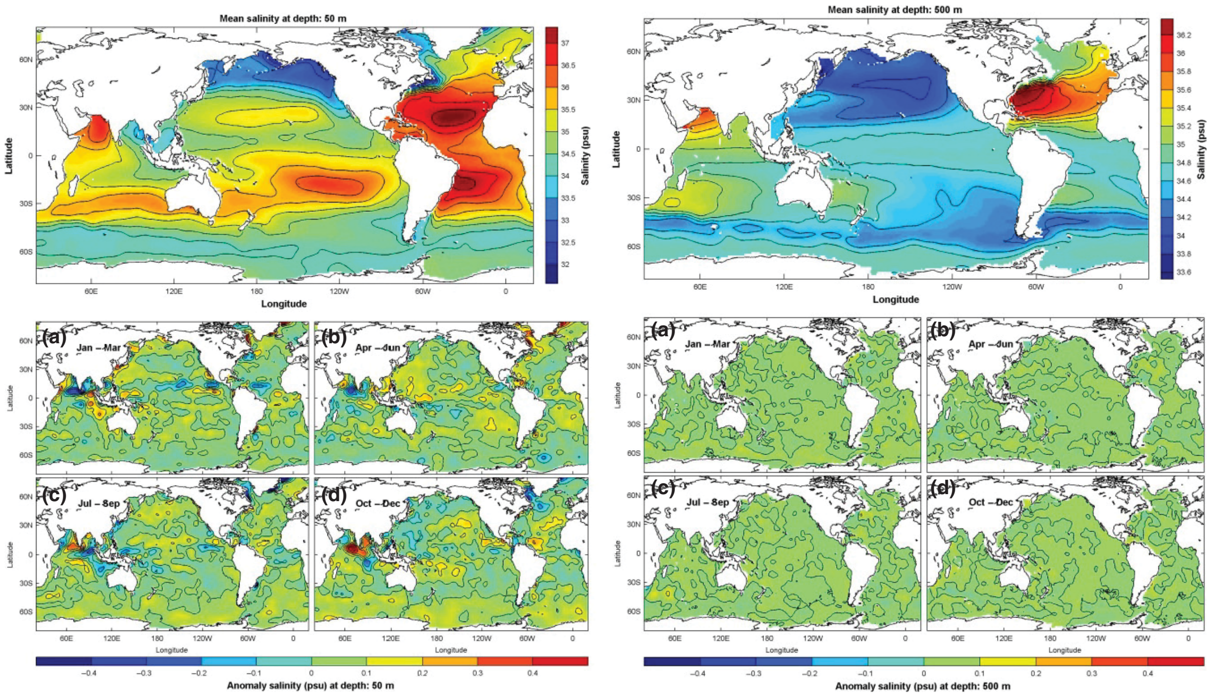
$$\delta\Psi(\mathbf{r}_n, z_j, t_l) = \bar{\Psi}(\mathbf{r}_n, z_j, t_l) - \bar{\bar{\Psi}}(\mathbf{r}_n, z_j) \qquad (17)$$

represents the mean seasonal variability. Subtraction of $\bar{\Psi}(\mathbf{r}_n, z_j, t_l)$ from $\Psi(\mathbf{r}_n, z_j\tau_m, t_l)$,

$$\Delta\Psi(\mathbf{r}_n, z_j\tau_m, t_l) = \Psi(\mathbf{r}_n, z_j\tau_m, t_l) - \bar{\Psi}(\mathbf{r}_n, z_j, t_l) \qquad (18)$$

is the monthly anomaly (relative to the climatological monthly mean) and represents the interannual to interdecadal variability. The characteristics of the SMG-WOD global datasets are presented for illustration. Since the WOA13 monthly (*T*, *S*) data were used as the background field ($\mathbf{c}_b$), the total mean and seasonal variability of the SMG-WOD (*T*, *S*) are the same as the WOA13 (*T*, *S*). In the following sub-sections, the temperature, salinity, and geostrophic currents are presented at 50 and 500 m depths. It is noted that in China adjacent seas such as the East China Sea, the

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.



23 **Figure 11.** Total mean temperature ($\bar{\bar{S}}$) and seasonal anomalies ($\tilde{S}$) at (a) 50 m, and (b) 500 m depths calculated from the SMG-WOD dataset.

water depth is shallower than 500 m or 50 m. The SMG-WOD (T, S, u, v) data are blank in these regional seas.

4.2. Temperature

4.2.1. Total mean and seasonal variability

At 50 m depth, the total-time mean (1945–2014) of temperature ($\bar{\bar{T}}$) shows the following features: (1) colder than 8°C north of 40°N and south of 40°S, (2) warm areas associated with the subtropical gyres such as a strong warm core (>26°C) from the South Atlantic with a triangular shape of the Brazilian-Argentine Coast (0° to 30°S) to the Gulf Stream region, extending eastward and reaching the northwest African Coast, a warm core located in the western Pacific including the Kuroshio and its extension regions, western equatorial/South Pacific, as well as the Indian Ocean north of 20°S. The mean seasonal variability of temperature (δT) shows the following features: (1) range of seasonal variability from −3°C to +3°C, (2) warm seasonal anomalies occurring in the Northern Hemisphere in summer (July to September) and fall (October to December) and in the Southern Hemisphere in winter (January to March) and Spring (April to June), (3) cold seasonal anomalies occur in the Northern Hemisphere in winter (January to March) and Spring (April to June) and in the Southern Hemisphere in summer (July to September) and fall (October to December) (Figure 6a).

At 500 m depth, the total-time mean (1945–2014) of temperature ($\bar{\bar{T}}$) shows the following features: (1) noticeable warm cores in the western part of middle latitudes (Northern and Southern Hemispheres) in the three oceans such as the Gulf Stream region in the North Atlantic, Kuroshio region in the North Pacific, and Arabian Sea in the Northern Indian Ocean, and (2) warm areas are associated with the subtropical gyres in the Atlantic and Pacific Oceans. The mean seasonal variability in temperature (δT) shows the following features: (1) the range of seasonal variability from −1°C to +1°C, (2) dominating eddy-like structure throughout the year, and (3) no evident differential Northern/Southern Hemispheric seasonal variability (Figure 6b). It is noted that 'eddy' is referred to large eddy from here on.

4.2.2. Interannual to interdecadal variability

Evident interannual to interdecadal variability is noted at 50 m depth from five-yearly January ΔT ($\mathbf{r}_n$, 50 m, τ_m, Jan) (Figure 7) and July ΔT ($\mathbf{r}_n$, 50 m, τ_m, Jul) (Figure 8), from 1960 to 2004. 12
The Pacific Ocean in January is characterized by alternating zonal-varying pattern and multi-eddy structure such as, a zonal-varying pattern in 1960 with two warm anomalies (near 1.5°C) occurring in the western North Pacific and eastern Pacific and a cold anomaly (around −1°C) appearing in the western South Pacific, a multi-eddy-like structure from 1965 to 1980, a zonal-varying pattern in 1985 with a weak anomaly sandwiched by two evident cold anomalies (−2.0°C) occurring in the western North Pacific and eastern Pacific, a zonal-varying pattern in 1990, 1995 with

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

LOW RESOLUTION COLOR FIG

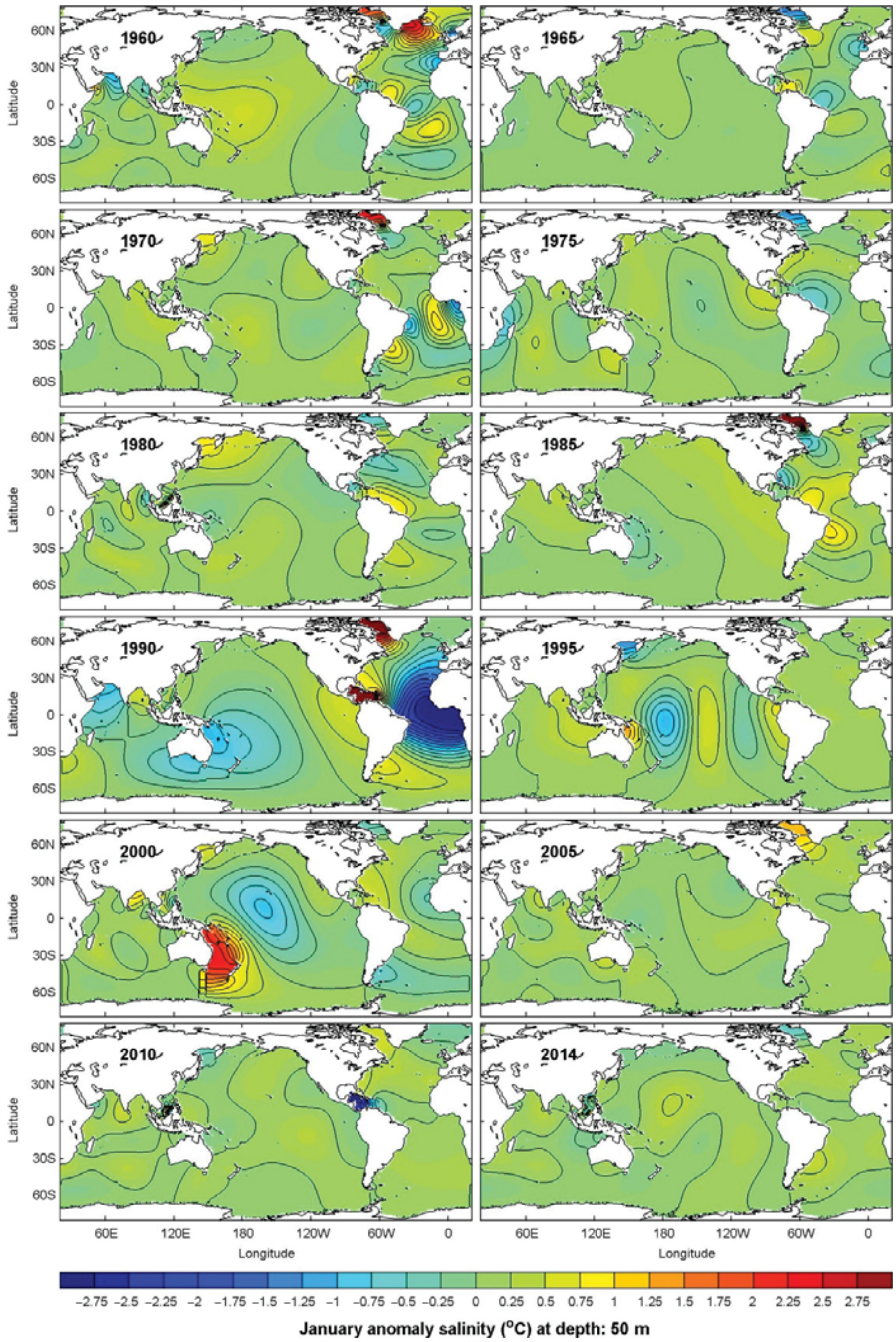


Figure 12. SMG-WOD January salinity anomaly (relative to climatological January mean) (psu) at 50 m depth in selected years: (a) 1960, (b) 1965, (c) 1970, (d) 1975, (e) 1980, (f) 1985, (g) 1990, (h) 1995, (i) 2000, (j) 2005, (k) 2010, (l) 2014.

weak cold anomaly (−0.5°C) in the western North Pacific and evident warm anomaly (1°C) in the eastern Pacific, a reverse zonal-varying pattern in 2000 with warm anomaly (1°C) in the western North Pacific and cold anomaly (<−1.5°C) in the eastern Pacific, evident warm anomaly in tropical Pacific in 2005, 2010, and a multi-eddy-like structure in 2014. Different from January, the Pacific Ocean in July (Figure 8) is characterized by a multi-eddy-like structure in most years except in 1975, when weak anomaly (−0.5°C to 0.5°C) dominates the whole Pacific Ocean. The Atlantic Ocean is characterized by alternating meridional-varying pattern and multi-eddy structure at 50 m depth in January (Figure 7) and July (Figure 8). In January, meridional-varying pattern in 1960 with strong cold anomaly (−2.5°C) occurring in the South

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.

LOW RESOLUTION COLOR FIG

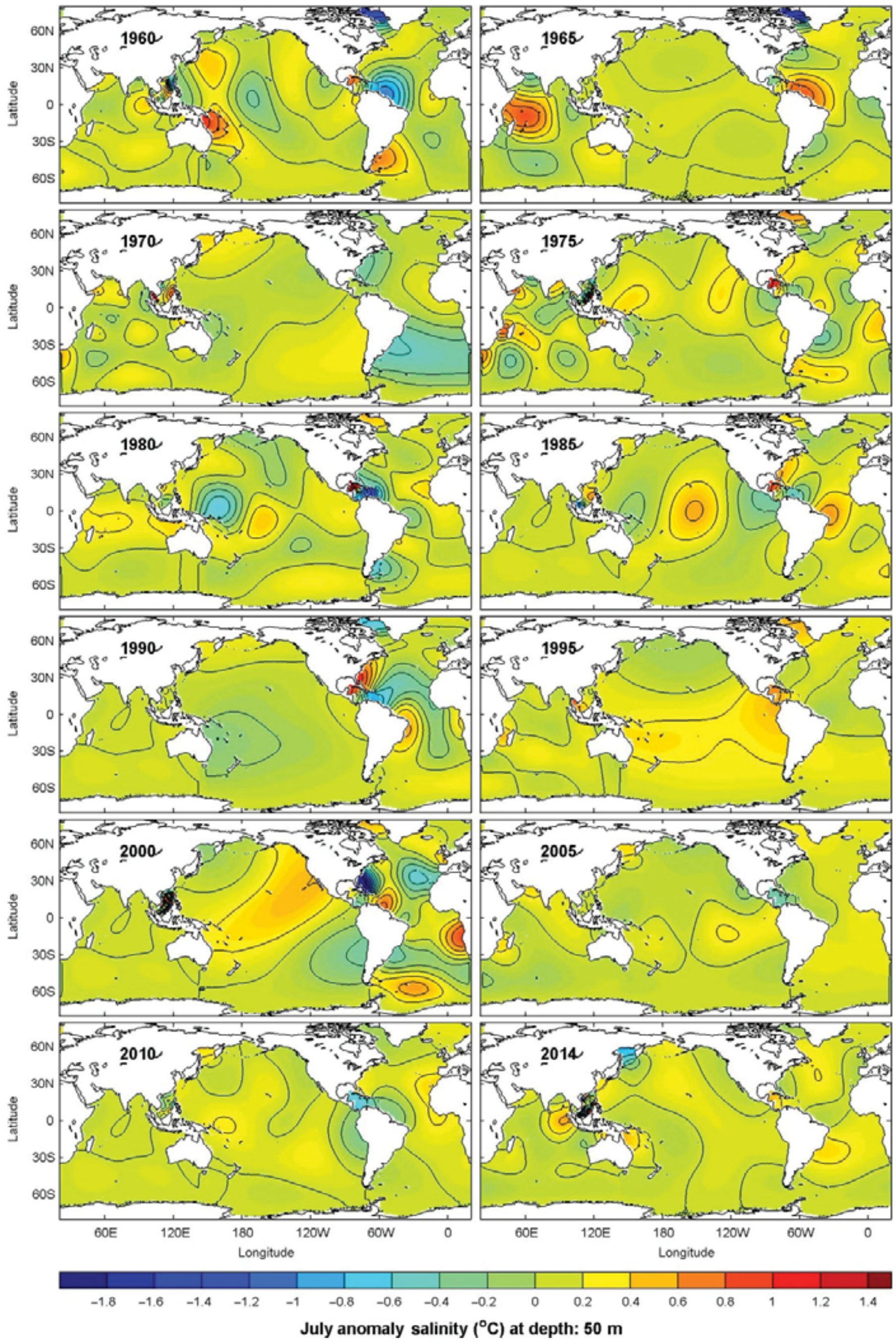


Figure 13. SMG-WOD July salinity anomaly (relative to climatological July mean) (psu) at 50 m depth in selected years: (a) 1960, (b) 1965, (c) 1970, (d) 1975, (e) 1980, (f) 1985, (g) 1990, (h) 1995, (i) 2000, (j) 2005, (k) 2010, (l) 2014.

Atlantic and weak warm anomaly (<0.5°C) appearing in the North Atlantic, a multi-eddy-like structure in 1965, 1970, 1975, a meridional-varying pattern in 1980 with warm anomaly (1.5°C) occurring in the South Atlantic and weak cold anomaly (−0.5°C) in the North Atlantic, a multi-eddy-like structure in 1985, 1990, 1995, 2000, 2010, and weak anomaly in 2005, 2014 (Figure 7). The Atlantic Ocean is characterized by a multi-eddy-like structure in most years (Figure 8). The Indian Ocean is characterized by a multi-eddy structure at 50 m depth in January (Figure 7) and July (Figure 8).

Reduced interannual to interdecadal variability is also noted at 500 m depth from five-yearly January $\Delta T(\mathbf{r}_n, 500\ m, \tau_m, \mathrm{Jan})$ (Figure 9) and July $\Delta T(\mathbf{r}_n, 500\ m, \tau_m, \mathrm{Jul})$ (Figure 10) from 1960 to 2004.

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

LOW RESOLUTION COLOR FIG

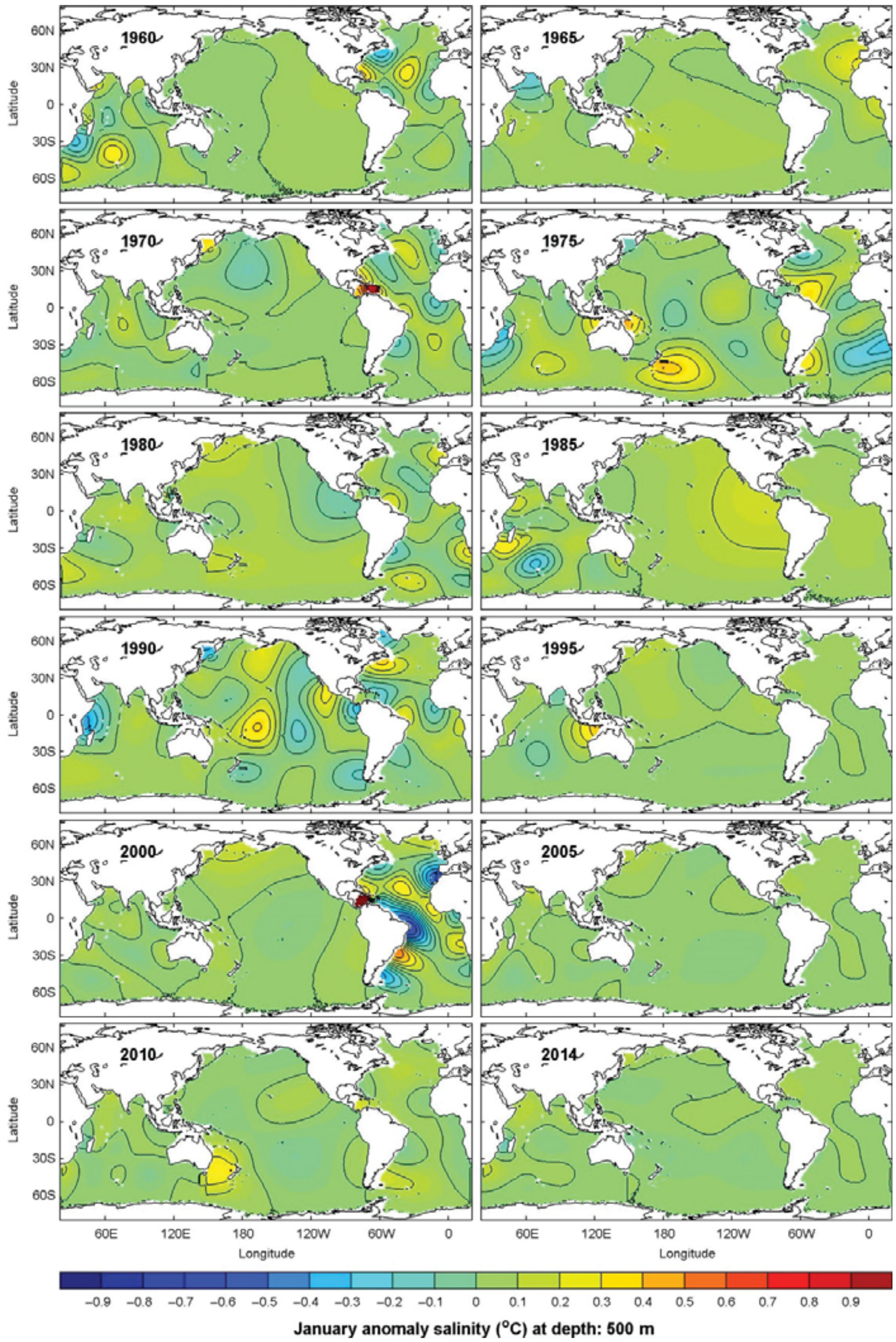


Figure 14. SMG-WOD January salinity anomaly (relative to climatological January mean) (psu) at 500 m depth in selected years: (a) 1960, (b) 1965, (c) 1970, (d) 1975, (e) 1980, (f) 1985, (g) 1990, (h) 1995, (i) 2000, (j) 2005, (k) 2010, (l) 2014.

The world oceans are characterized by alternating low variability (−1°C to 1°C) pattern and multi-eddy structure with variability larger than (−2°C to 2°C). The Pacific Ocean has a low variability pattern in 1960, 1965, 1985, 1995, 2000, 2005, 2014 in January (Figure 9) and in most years except in 1970 in July (Figure 10) and a multi-eddy-like structure in rest of years. The Atlantic Ocean has a multi-eddy structure in 1970, 1980, 1985, 1990 in January (Figure 9) and in 1970, 1975, 1985, 1990, 1995, 2000 in July (Figure 10) and a low variability pattern in rest of years. The Indian Ocean has a low variability pattern in 1965, 1975, 1980, 2010 in January (Figure 9) and in 1965, 1980, 1985, 1990 in July (Figure 10) and a multi-eddy structure in rest of years.

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.

LOW RESOLUTION COLOR FIG

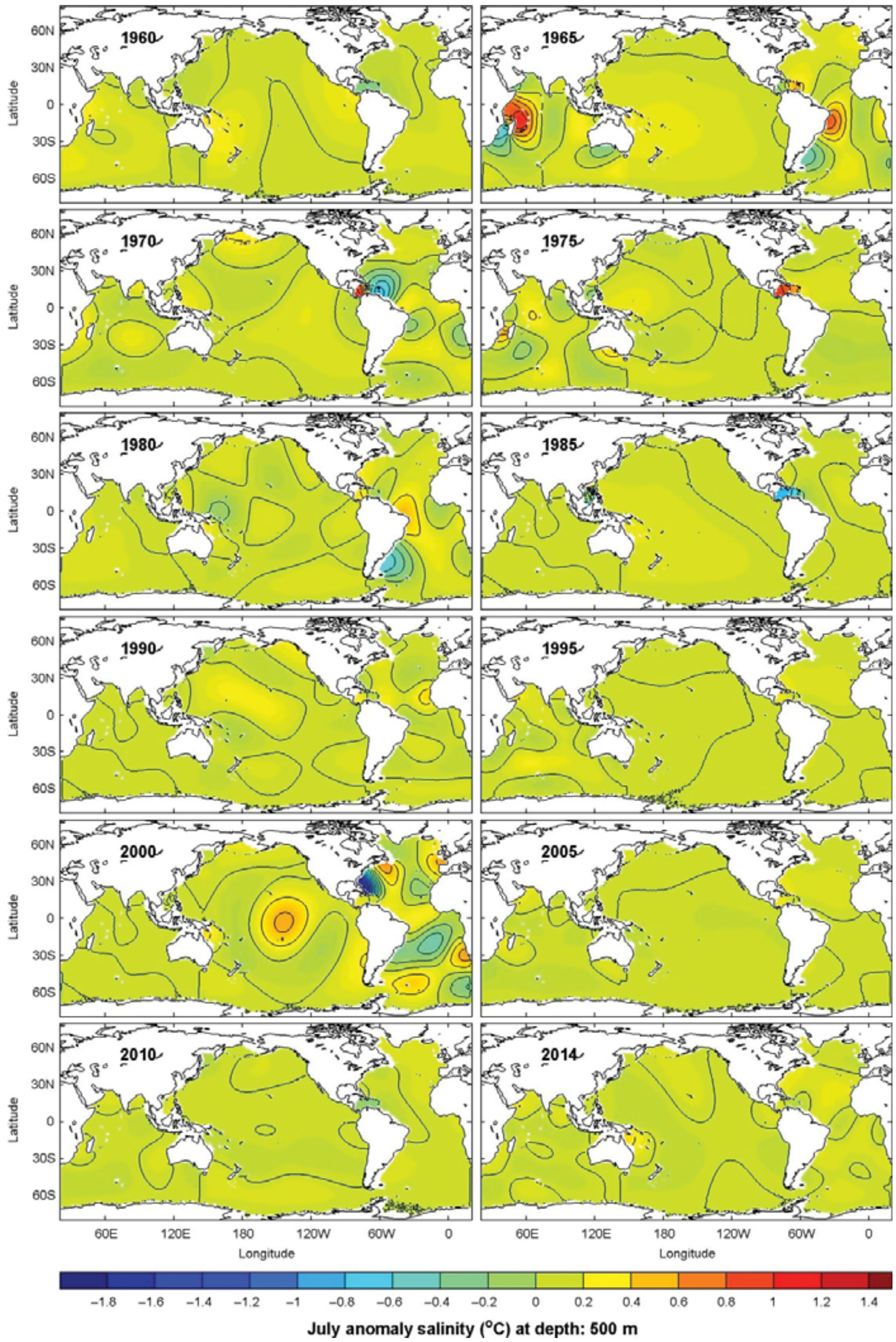


Figure 15. SMG-WOD July salinity anomaly (relative to climatological July mean) (psu) at 500 m depth in selected years: (a) 1960, (b) 1965, (c) 1970, (d) 1975, (e) 1980, (f) 1985, (g) 1990, (h) 1995, (i) 2000, (j) 2005, (k) 2010, (l) 2014.

4.3. Salinity

4.3.1. Total mean and seasonal variability

At 50 m depth, the total-time mean (1945–2014) salinity ($\bar{\bar{S}}$) shows following features: (1) strong salty ($\bar{\bar{S}} > 37.0$ psu) areas are associated with the North and South Atlantic subtropical gyres; (2) less salty ($37.0 > \bar{\bar{S}} > 35.0$ psu) areas are associated with the North and South pacific subtropical gyres, Arabian Sea, and Southern Indian subtropical gyre, and (3) least salty ($\bar{\bar{S}} < 33.0$ psu) region is associated with the North Pacific subarctic gyre. The mean seasonal variability of salinity (δS) at 50 m depth shows the following features: (1) the range of seasonal variability from -0.5 to 0.5 psu, (2) no hemispherical asymmetry in the Atlantic and Pacific Oceans, and (3) larger δS in the Indian Ocean than in the Atlantic and Pacific Oceans (Figure 11a). At 500 m depth, the total-time mean ($\bar{\bar{S}}$) shows following features:

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

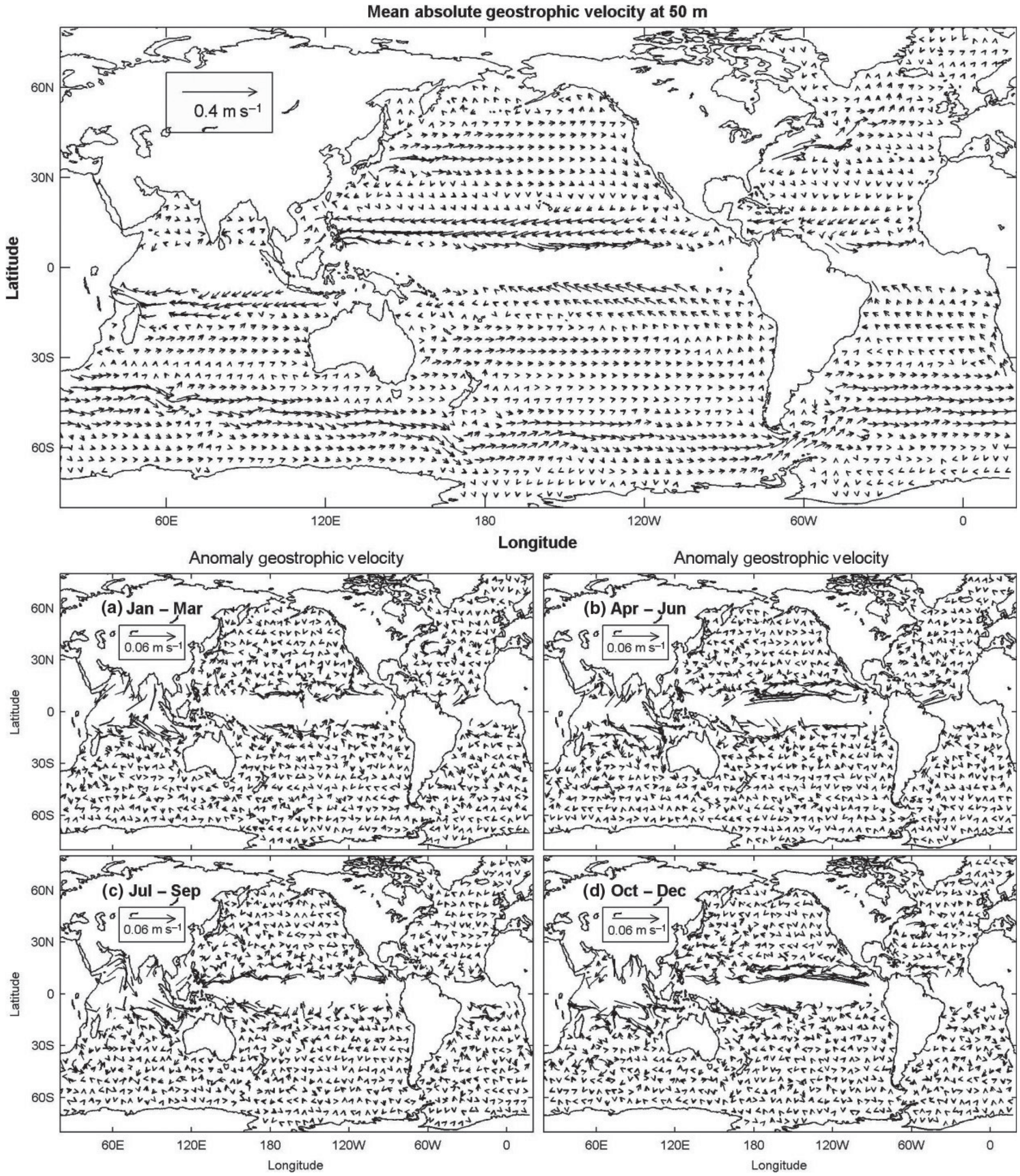


Figure 16. Total mean absolute geostrophic currents $(\bar{\bar{u}}, \bar{\bar{v}})$and seasonal anomalies $(\tilde{u}, \tilde{v})$ at (a) 50 m, and (b) 500 m depths calculated from the SMG-WOD dataset.

(1) strong salty ($\bar{\bar{S}} > 36.0$ psu) areas associated with the North Atlantic subtropical gyre, (2) least salty ($\bar{\bar{S}} < 34.0$ psu) region associated with the North Pacific subarctic gyre, and (3) less salty ($36.0 > \bar{\bar{S}} > 34.0$ psu) areas are rest of the world oceans. The mean seasonal variability of salinity (δS) shows very minor seasonal variability from -0.1 to 0.1 psu all over the world oceans (Figure 10b).

4.3.2. Interannual to interdecadal variability

Evident interannual to interdecadal variability from 1960 to 2004 is noted at 50 m depth from five-yearly January $\Delta S(\mathbf{r}_n, 50\ m, \tau_m, \text{Jan})$ (-3.0 to 3.0 psu) (Figure 12) and July $\Delta S(\mathbf{r}_n, 50\ m, \tau_m, \text{Jul})$ (-2.0 to 1.5 psu) (Figure 13). The Pacific Ocean is characterized by alternating zonal-varying pattern (1990, 1995, 2000) and low-variability pattern (-0.5 to 0.5 psu) in

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.

LOW RESOLUTION FIG

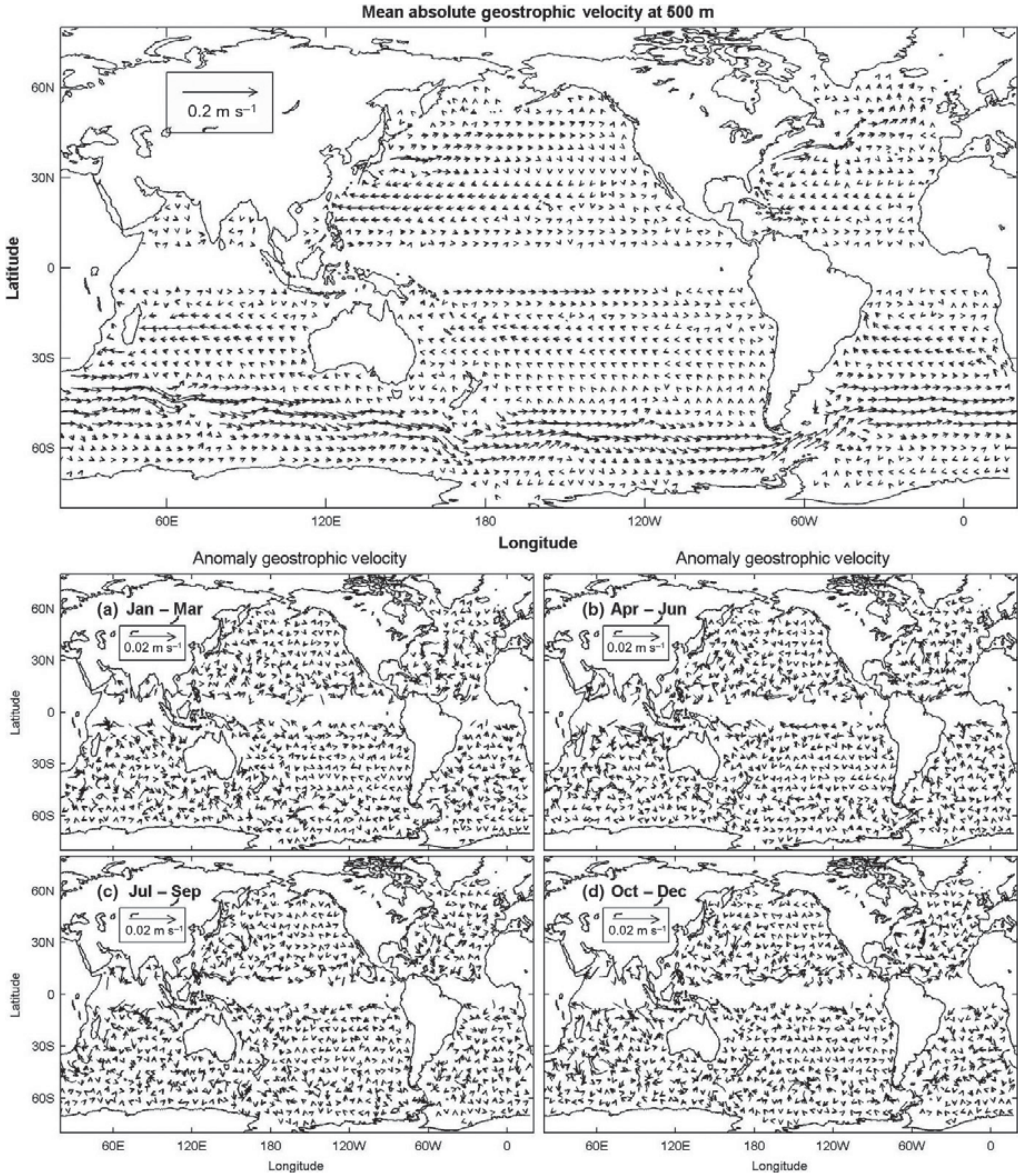


Figure 17. Total mean absolute geostrophic currents $(\bar{\bar{u}}, \bar{\bar{v}})$and seasonal anomalies $(\tilde{u}, \tilde{v})$ at (a) 500 m, and (b) 500 m depths calculated from the SMG-WOD dataset.

rest of years in January, and by a multi-eddy structure in 1960, 1975, 1980, 1985, and low-variability pattern (−0.5 to 0.5 psu) in rest of years in July. The Atlantic Ocean is characterized by alternating low-variability pattern (−0.5 to 0.5 psu) in 1995, 2005, 2014, and multi-eddy structure in rest of years in both January and July. The Indian has dominating low variability pattern in January and alternating a multi-eddy structure in 1965, 1970, 1975, 2014, and low-variability pattern in rest of years in July.

Reduced interannual to interdecadal variability from 1960 to 2004 is also noted at 500 m depth from five-yearly January $\Delta S(\mathbf{r}_n, 500\ m, \tau_m, \text{Jan})$ (−1.0 to 1.0 psu) (Figure 14), but not in July $\Delta S(\mathbf{r}_n, 500\ m, \tau_m, \text{Jul})$(−2.0 to 1.5 psu) (Figure 15). The world oceans are characterized by alternating low

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

variability (−0.3 to 0.3 psu) pattern and multi-eddy structure with larger variability. The Pacific Ocean is characterized by alternating multi-eddy structure in (1975, 1990) and low-variability pattern in rest of years in January, and by a multi-eddy structure in 2000, and low-variability pattern in rest of years in July. The Atlantic Ocean is characterized by a multi-eddy structure in 1965, 1970, 1980, 1990, 2000 and low-variability pattern in rest of years in July. The Indian has low variability pattern in 1965, 1970, 1980, 1995, 2000, 2005, and a multi-eddy structure in rest of years in January and by a multi-eddy structure in 1965, 1975, and low-variability pattern in rest of years in July.

4.4. Absolute geostrophic velocity

The equatorial region (−8°S–8°N) is excluded for the absolute geostrophic velocity data since the geostrophic balance is not valid there. Description of interannual to interdecadal variability of the absolute geostrophic velocity (u, v) from the datasets is beyond the scope of this paper. Only the total time mean and seasonal variability is presented. At 50 m depth, the total-time mean (1945–2014) absolute geostrophic velocity $(\bar{\bar{u}}, \bar{\bar{v}})$ with a maximum speed near 0.4 m/s shows the existence of subtropical gyres with major currents: the North Equatorial Current (NEC) flowing westward from the west coast of Africa to the Brazilian cost (North Atlantic), from the west coast of Baja California to west of Philippines and turning the direction to join the Kuroshio in the North Pacific and the Gulf Stream in the North Atlantic; the North Equatorial Counter Current (NECC) moving eastward south of the NEC (evident in the North Atlantic and North Pacific); the South Equatorial Current flowing westward from the eastern South Atlantic to the Brazilian coast, from the west coast of Africa to western South Pacific, and from northwest coast of Australia to east coast of Africa; eastward flowing South Atlantic Current, South Indian Current, South Pacific Current, and Antarctic Circumpolar Current. The mean seasonal variability (δu, δv) is almost an order of magnitude smaller than the total-time mean with a maximum of around 0.06 m/s (Figure 16).

At 500 m depth, the total-time mean (1945–2014) absolute geostrophic velocity $(\bar{\bar{u}}, \bar{\bar{v}})$ with a maximum speed near 0.2 m/s shows the existence of weaker subtropical gyres with evident Gulf Stream, Kuroshio, and Antarctic Circumpolar Current. The mean seasonal variability (δu, δv) is an order of magnitude smaller than the total-time mean with a maximum of around 0.02 m/s (Figure 17).

5. Data download

The data can be downloaded directly from the NCEI websites (see Datasets section). Please contact NCEI Customer Service if you need further assistance (http://www.nodc.noaa.gov/about/contact.html). To read the data, the free Netcdf package needs to be downloaded from the website: https://www.image.ucar.edu/GSP/Software/Netcdf/. The MATLAB (version 2008b and later) provides access to more than 30 functions in the netCDF interface. This interface provides an application program interface that you can use to enable reading data from and writing data to netCDF files (known as *datasets* in netCDF terminology). The MATLAB code is listed as follows to read the SMG-GTSPP data in netCDF as an illustration.

```
% open netcdf data
ncid=netcdf.open(GriddedMonthlyGTSPP_OSD.nc,
nowrite);
% get year
yr_id=netcdf.inqVarID(ncid, year);
yr=netcdf.getVar(ncid,yr_id);
% get month
mon_id=netcdf.inqVarID(ncid, Month);
mon=netcdf.getVar(ncid,mon_id);
% get longitude
lon_id=netcdf.inqVarID(ncid, Longitude);
lon=netcdf.getVar(ncid,lon_id);
% get latitude
lat_id=netcdf.inqVarID(ncid, Latitude);
lat=netcdf.getVar(ncid,lat_id);
% get vertical coordinate z
z_id=netcdf.inqVarID(ncid, Depth(m));
z=netcdf.getVar(ncid,z_id);
% get temperature data
t_id=netcdf.inqVarID(ncid, temperature);
% get the units
units=netcdf.getAtt(ncid,t_id, units);
% get salinity data
s_id=netcdf.inqVarID(ncid, salinity);
% get the units
units=netcdf.getAtt(ncid,s_id, units);
% get all the data
t=netcdf.getVar(ncid,t_id);
s=netcdf.getVar(ncid,s_id);
% get part of the data
% example: March 1991 at depth level k
t=netcdf.getVar(ncid,t_id,[0,0,k-1,3,2],
[179,360,1,1,1]);
```

Acknowledgements

Outstanding efforts of Dr. Charles Sun, Dr. Alexandra Grodsky, and Dr. Ricardo Locarnini to publish the datasets at the NCEI are highly appreciated. The research presented in this paper was funded by the National Oceanographic Data Center (now National Centers for Environmental Information) and the Naval Oceanographic Office. The Research Office of the Naval Postgraduate School supports the publication in the Geoscience Data Journal.

References

Boyer TP. 2014. Description of NODC tests. In *International Quality-Controlled Ocean Database 2nd*

UNCORRECTED PROOF

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.

Annual Workshop Report. NOAA: Silver Spring, MD, USA, June 4–6 2014, http://www.iquod.org/documents/IQuOD2-report_v2.0.pdf

Boyer TP, Antonov JI, Baranova OK, Coleman C, Garcia HE, Grodsky A, Johnson DR, Locarnini RA, Mishonov AV, O'Brien TD, Paver CR, Reagan JR, Seidov D, Smolyar IV, Zweng MM. 2013. *World Ocean Database 2013*. Levitus S, Ed.; Mishonov A, Technical Ed.; NOAA Atlas NESDIS 72; 209 pp

Chu PC. 1995. P-vector method for determining absolute velocity from hydrographic data. *Marine Technological Society Journal* **29**: 3–14.

Chu PC. 2006. *P-vector Inverse Method*. Springer: Berlin, Germany; 605 pp.

Chu PC. 2011. Global upper ocean heat content and climate variability. *Ocean Dynamics* **61**: 1189–1204.

Chu PC, Fan CW. 2015. Absolute geostrophic velocity inverted from World Ocean Atlas 2013 (WOAV13) with the P-vector method. *Geoscience Data Journal* **2**: 78–82.

Chu PC, Ivanov LM, Korzhova TP, Margolina TM, Melnichenko OM. 2003a. Analysis of sparse and noisy ocean current data using flow decomposition. Part 1: Theory. *Journal of Atmospheric and Oceanic Technology* **20**: 478–491.

Chu PC, Ivanov LM, Korzhova TP, Margolina TM, Melnichenko OM. 2003b. Analysis of sparse and noisy ocean current data using flow decomposition. Part 2: Application to Eulerian and Lagrangian data. *Journal of Atmospheric and Oceanic Technology* **20**: 492–512.

Chu PC, Ivanov LM, Margolina TM. 2004. Rotation method for reconstructing process and field from imperfect data. *International Journal Bifurcation and Chaos* **14**: 2991–2997.

Chu PC, Ivanov LM, Melnichenko OM. 2005a. Fall-winter current reversals on the Texas-Louisiana continental shelf. *Journal of Physical Oceanography* **35**: 902–910.

Chu PC, Ivanov LM, Margolina TM. 2005b. Seasonal variability of the Black Sea Chlorophyll-a concentration. *Journal of Marine Systems* **56**: 243–261.

Chu PC, Ivanov LM, Melnichenko OV, Wells NC. 2007. Long baroclinic Rossby waves in the tropical North Atlantic observed from profiling floats. *Journal of Geophysical Research* **112**: C05032.

Chu PC, Tokmakian RT, Fan CW, Sun CL. 2015. Optimal spectral decomposition (OSD) for ocean data assimilation. *Journal of Atmospheric and Oceanic Technology* **32**: 828–841.

Chu PC, Fan CW, Margolina TM. 2016. Ocean spectral data assimilation without background error covariance matrix. *Ocean Dynamics* **66**: 1143–1163.

Levitus S, Antonov JI, Boyer TP, Locarnini RA, Garcia HE, Mishonov AV. 2009. Global ocean heat content 1955–2008 in light of recently revealed instrumentation problems. *Geophysical Research Letters* **36**: L07608.

Owens WB, Wong APS. 2009. An improved calibration method for the drift of the conductivity sensor on autonomous CTD profiling floats by θ-S climatology. *Deep Sea Research* **1**: 450–457.

Stommel H, Schott F. 1977. The beta spiral and the determination of the absolute velocity field from hydrographic station data. *Deep Sea Research* **24**: 325–329.

Sun LC. 2013. Global Temperature and Salinity Profile Programme (GTSPP) data quality tests. Presented at the CLIVAR-GSOP Coordinate Quality-Control of Global Subsurface Ocean Climate Observations, Hobart, Tas., Australia, June 12–14, 2013.

Sun LC, Chu PC. 2013. Optimal spectral decomposition (OSD): an advanced approach for optimal estimation of ocean states and data QC tests. In *Presented at the CLIVAR-GSOP Coordinate Quality-Control of Global Subsurface Ocean Climate Observations*, Hobart, Tas., Australia, June 12–14, 2013, www.clivar.org/sites/default/files/documents/gsop/OSD_Sun_Chu-cut_downAT.ppt

Sun LC, Thresher A, Keeley R, Hall N, Hamilton M, Chinn P, Tran A, Goni G, Petit de la Villeon L, Carval T, Cowen L, Manzella G, Gopalakrishna V, Guerrero R, Reseghetti F, Kanno Y, Klein B, Rickards L, Baldoni A, Lin S, Ji F, Nagaya Y. 2009. The data management system for the Global Temperature and Salinity Profile Program (GTSPP). In Hall J, Harrison DE, Stammer D (eds). *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference* (Vol. 2). ESA Publication WPP-306: Venice, Italy, 21–25 September 2009.

Thadathil P, Bajish CC, Behera S, Gopalakrishna VV. 2012. Drift in salinity data from Argo profiling floats in the Sea of Japan. *Journal of Atmospheric and Oceanic Technology* **29**: 129–138.

Wong APS, Johnson GC, Owens WB. 2003. Delay-mode calibration of autonomous CTD profiling float salinity data by θ-S climatology. *Journal of Atmospheric and Oceanic Technology* **20**: 308–318.

Wong APS, Keeley R, Carval T, the Argo Data Management Team. 2013. *Argo Data Management*, http://www.argodatamgt.org/content/download/20685/142877/file/argo-quality-control-manual_version2.9.pdf

Wunsch C. 1978. The general circulation of the North Atlantic west of 50°W determined from inverse method. *Review of Geophysics* **16**: 583–620.

UNCORRECTED PROOF

Published 2017. This article is a U.S. Government work and is in the public domain in the USA.
Geoscience Data Journal published by Royal Meteorological Society and John Wiley & Sons Ltd.