Acta Oceanol. Sin., 2022, Vol. 41, No. 2, P. 65–79
https://doi.org/10.1007/s13131-021/1858-3
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

Evaluating the contribution of satellite measurements to the reconstruction of three-dimensional ocean temperature fields in combination with Argo profiles

Jian Chen[1, 2*], Hengqian Yan[3], Senliang Bao[3], Xindong Cui[1], Chengzu Bai[1], Huizan Wang[3]

[1] Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

[2] Beijing Institute of Applied Meteorology, Beijing 100029, China

[3] College of Meteorology and Oceanography, National University of Defense Technology, Nanjing 211101, China

Received 1 March 2021; accepted 20 May 2021

Abstract

Assimilation systems absorb both satellite measurements and Argo observations. This assimilation is essential to diagnose and evaluate the contribution from each type of data to the reconstructed analysis, allowing for better configuration of assimilation parameters. To achieve this, two comparative reconstruction schemes were designed under the optimal interpolation framework. Using a static scheme, an *in situ*-only field of ocean temperature was derived by correcting climatology with only Argo profiles. Through a dynamic scheme, a synthetic field was first derived from only satellite sea surface height and sea surface temperature measurements through vertical projection, and then a combined field was reconstructed by correcting the synthetic field with *in situ* profiles. For both schemes, a diagnostic iterative method was performed to optimize the background and observation error covariance statics. The root mean square difference (RMSD) of the *in situ*-only field, synthetic field and combined field were analyzed toward assimilated observations and independent observations, respectively. The rationale behind the distribution of RMSD was discussed using the following diagnostics: (1) The synthetic field has a smaller RMSD within the global mixed layer and extratropical deep waters, as in the Northwest Pacific Ocean; this is controlled by the explained variance of the vertical surface-underwater regression that reflects the ocean upper mixing and interior baroclinicity. (2) The *in situ*-only field has a smaller RMSD in the tropical upper layer and at midlatitudes; this is determined by the actual noise-to-signal ratio of ocean temperature. (3) The satellite observations make a more significant contribution to the analysis toward independent observations in the extratropics; this is determined by both the geographical feature of the synthetic field RMSD (smaller at depth in the extratropics) and that of the covariance correlation scales (smaller in the extratropics).

Key words: satellite measurements, Argo, three-dimensional reconstruction, ocean temperature

Citation: Chen Jian, Yan Hengqian, Bao Senliang, Cui Xindong, Bai Chengzu, Wang Huizan. 2022. Evaluating the contribution of satellite measurements to the reconstruction of three-dimensional ocean temperature fields in combination with Argo profiles. Acta Oceanologica Sinica, 41(2): 65–79, doi: 10.1007/s13131-021/1858-3

1 Introduction

Satellite remote sensing measurements, such as sea surface temperature (SST), sea surface height (SSH), and sea surface salinity (SSS), provide global high-resolution observations, but they cannot directly detect the ocean's interior. On the other hand, *in situ* observation systems such as the Argo network can reveal ocean three-dimensional (3-D) internal structures, but they undersample ocean variability with a coverage of only 3° by 3°. Therefore, many studies are devoted to combining Argo with satellite measurements to better reconstruct small-scale ocean interior variabilities.

There are generally two steps to assimilate satellite and Argo data at the same time, with either a dynamical method (Cooper and Haines, 1996; Alves et al., 2001) or a statistical method (Guinehut et al., 2004; Larnicol et al., 2006; Guinehut et al., 2012). First, subsurface state variables are estimated from satellite observations. This is achieved by exploiting the coupling (or regression) relationship between sea surface signals and inner structures. These vertical projection methods include, to name a few, a linear regression (Fox et al., 2002; Guinehut et al., 2004, 2012; Nardelli et al., 2012), a single empirical orthogonal function (EOF) reconstruction (sEOF-R) (Carnes et al., 1994), a bivariate EOF reconstruction (Maes and Behringer, 2000), a multivariate EOF reconstruction (mEOF-R) (Nardelli and Santoleri, 2005), a coupled pattern reconstruction (CPR) (Nardelli and Santoleri, 2004), a gravest empirical mode (GEM) method (Meinen and Watts, 2000; Mitchell et al., 2004), and intelligent methods such as genetic algorithms (GA) (Agarwal et al., 2007) and neural networks (NN) (Ballabrera-Poy et al., 2009; Wu et al., 2012). Second, the real-time Argo temperature (T) and salinity (S) profiles (as observations) are used to correct the synthetic T and S fields obtained in the first step (as a first guess or background). This is achieved through, e.g., optimal interpolation (OI) or 3-D variational data assimilation (3DVAR). This two-step approach has been proven to provide an improved estimate of the ocean compared to those from only satellite or *in situ* observations. This ap-

Foundation item: The National Natural Science Foundation of China under contract Nos 41706021 and 41976188.
*Corresponding author, E-mail: chenj03@126.com

proach is also complementary to results provided by ocean reanalysis combining model forecasts and observations through an assimilation method (Lellouche et al. 2018).

It is known that OI or 3DVAR algorithms rely on the theory of least-variance in linear statistical estimation. In this theory, each set of information is given a weight proportional to the inverse of its predefined error covariance. These pieces of information are typically given by observations and background estimates. The former refers to the scattered Argo observations, and the latter refers to the synthetic fields on model grids from satellite observations. Analysis systems are then dependent on appropriate statistics, or a priori specifications, for observation and background errors. Those statistics, however, are not perfectly known, and their determination remains a major challenge in assimilation systems**.** Under the two-step scheme, Guinehut et al. (2012) applied a spatially correlated error of 20% to the synthetic fields to correct large-scale errors introduced in the first step, in addition to the conventional use of white noise measurements representing 10% of the signal variance. These values, however, are estimated subjectively and possibly fail to accurately reflect the actual characteristics of satellite measurements or the impacts of the retrieval methods. One reason is proposed by Nardelli et al. (2012), who noted that the use of satellite surface measurements potentially changes the spatial scales that the 3-D synthetic field is effectively able to resolve. The differences in effective resolution tend to result in different statistical errors of the synthetic field, i.e., the background error.

Fortunately, one source of information on the observation and background errors is contained in the conditions of the innovation vector. They are defined as the differences between observations and their background or analysis counterparts, including observation minus-background (O–B), observation-minus-analysis (O–A) and background-minus analysis (B–A). Much work has been conducted in which the accumulated statistics of the innovation vector have been used together with external hypotheses to determine the statistical quantities required by least-variance linear estimation. Hollingsworth and Lönnberg (1986) have shown, assuming that background errors can be spatially correlated while observation errors are not (and that those two sets of errors are unbiased and mutually uncorrelated), that all remaining required quantities can be determined from the spatial covariance of the innovation. Dee and da Silva (1999) used maximum likelihood estimation to determine the required statistics by fitting an a priori model of covariances to the innovation statistics. Desroziers and Ivanov (2001) proposed an approach based on a consistency criterion of the analysis relying on statistics of observation-minus-analysis differences to adapt observation-error statistics. The consistency criterion used in this method was defined by Talagrand (1999). Chapnik et al. (2004) investigated the properties of the algorithm and showed that it was equivalent to a maximum likelihood method, although less expensive to implement. The algorithm introduced by Desroziers et al. (2005) is commonly used to adapt the observation error from the background and analysis residuals calculated in the bias correction. Chapnik et al. (2006), Lellouche et al. (2013), and Lellouche et al. (2018) also applied the same algorithm in an operational framework to tune observation-error variances. One advantage of these diagnostics techniques is that they can be applied to any analysis scheme.

Most of the work presented in this paper is devoted to finding the true background and observational errors in the two-step scheme. This is achieved by tuning the incorrectly predefined error covariance (or variance) matrix using a posteriori diagnostics. These diagnostics can potentially provide information on imperfectly known observations and background error statistics. They can thus help us to evaluate the contribution of satellite measurements, along with Argo profiles, to the reconstruction of ocean 3-D temperature fields. The method by Desroziers et al. (2005) was first briefly presented and then applied in the framework of the *in situ* Analysis System (ISAS) (Gaillard, 2012) from the French Laboratory of Ocean Physics. The ISAS, however, is modified by considering the synthetic fields as a background.

It must be emphasized that these consistency diagnostics can only be fulfilled in observational space (Desroziers et al., 2005). In other words, these diagnostics only reflect the information on the hindcast (i.e., assimilated) observations, not on the independent observations. Therefore, another aim of this paper is to make independent comparisons with a subset of *in situ T* profiles. This work will also further discuss the distribution of the root mean square difference (RMSD) of the reconstructed *T* fields in various suboceans and depths, as well as the underlying rationale. This allows us to more deliberately evaluate the contribution of the satellite and *in situ* measurements to the analysis. The diagnostics are complementary to the work of Guinehut et al. (2012), which has confirmed the advantage of merging two types of measurements.

Both unresolved issues mentioned above are of importance to better configure the reconstruction parameters and need to be well studied. Thus, the primary goals of this paper are as follows.

(1) To provide an additional consistency check on the two-step reconstruction scheme by using a set of diagnostics based on combinations of innovation differences, and to present some first implementation results of the variances tuned by the diagnostic method in the framework of the ISAS, which is adapted to a synthetic background configuration.

(2) To make comparisons between the reconstructed *T* fields (derived from only *in situ* observations, only satellite measurements, or both) and independent validation to show a detailed error pattern and then to explain the respective roles of *in situ* and satellite observations in shaping such a pattern.

The rest of the paper is organized as follows: the model's basic configuration and adjustment are described in Section 2. In Section 3, the analysis toward assimilated observations based on consistency diagnostics and tuning are evaluated. The analysis toward independent observations is evaluated in Section 4. Conclusions and discussion are presented in Section 5.

2 Model configuration and adjustment

The ISAS (Gaillard, 2012) was adjusted to better combine satellite and *in situ* observations. The ISAS is based on an OI method, in which the analyzed field, $\boldsymbol{x}^{\mathrm{a}}$, is a linear least square estimator, obtained as the linear combination of the innovation that minimizes the statistical error:

$$\begin{aligned} \boldsymbol{x}^{\mathrm{a}} &= \boldsymbol{x}^{\mathrm{b}} + \boldsymbol{K}\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}}, \\ \boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}} &= \boldsymbol{y}^{\mathrm{o}} - \boldsymbol{y}^{\mathrm{b}} = \boldsymbol{y}^{\mathrm{o}} - \boldsymbol{H}\boldsymbol{x}^{\mathrm{b}}, \\ \boldsymbol{K} &= \boldsymbol{P}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^{\mathrm{T}} + \boldsymbol{R})^{-1}, \end{aligned} \tag{1}$$

where $\boldsymbol{x}^{\mathrm{a}}$ is the analysis vector on model grids; $\boldsymbol{x}^{\mathrm{b}}$ is the background vector on model grids; $\boldsymbol{d}$ is the innovation vector; $\boldsymbol{K}$ is the gain matrix; $\boldsymbol{y}^{\mathrm{o}}$ is the observation vector on observational points; $\boldsymbol{y}^{\mathrm{b}}$ is the background vector on observational points; $\boldsymbol{H}$ is the observation operator matrix (or mapping matrix); $\boldsymbol{P}$ is the background error covariance matrix between model grids; $\boldsymbol{P}\boldsymbol{H}^{\mathrm{T}}$ and

$\boldsymbol{HPH}^{\mathrm{T}}$ are the background error covariance matrix between model grids and observational points and that between observational points and observational points, abbreviated as $\boldsymbol{C}_{\mathrm{ao}}$ and $\boldsymbol{C}_{\mathrm{oo}}$, respectively; $\boldsymbol{R}$ is the observation error covariance matrix.

In situ profiles: The *in situ* profiles are taken from the CORA3 dataset of the Coriolis data center (http://www.coriolis.eu.org). Only Argo profiles that contain both temperature and salinity (used to calculate a dynamic height anomaly, i.e., DHA, as a portion of a sea level anomaly, i.e., SLA) and that are valid to depths of 1 000 m (to obtain enough Argo measurements available for regression calculation) are selected. These observations were preprocessed according to the Argo recommendations for data quality control and interpolated to standard levels. They are then used as training data to compute the statistics that relate the surface and subsurface fields (2000–2010) (see "background field") and as assimilated observations in OI (90% of 2011–2012, see Section 3) or independent validation (10% of 2011–2012, see Section 4).

Background field: In the original ISAS, the background field is a monthly mean climatology computed by averaging over 2004–2010 the gridded fields from a previous analysis, and the resulting analysis field from Argo profiles are called an *in situ*-only field. In the two-step combination scheme, the climatology background is replaced by a synthetic field estimated from surface satellite observations, and the resulting analysis field is a combination of satellite and *in situ* observations and is thus called a combined field. This is similar to using a previous model output as the background field. The "model output" here however, does not involve thermodynamic integration but only statistical regression and thus greatly improves computational efficiency. A retrieval method similar to Guinehut et al. (2012) is adopted here. The synthetic temperature T_{synt} at a given subsurface position can be estimated using a multiple linear regression:

$$\delta T(z) = \alpha(z) \cdot \delta\mathrm{SLA} + \beta(z) \cdot \delta\mathrm{SST}, \tag{2}$$

where δ denotes anomalies with respect to monthly climatology, i.e., $\delta T = T_{\mathrm{synt}} - T^{\mathrm{b}}$, $\delta\mathrm{SLA} = \mathrm{SLA}^{\mathrm{o}} - \mathrm{SLA}^{\mathrm{b}}$, and $\delta\mathrm{SST} = \mathrm{SST}^{\mathrm{o}} - \mathrm{SST}^{\mathrm{b}}$, superscripts o and b denote the observational value and background value, respectively; the parameters α and β are regression coefficients of δSLA onto δT and δSST onto δT, respectively, which vary with depth, location, and time (only depth, z, dependency is shown for clarity) (Fig. 1). These season-dependent regression coefficients are computed on each 1° grid using all historical (2000–2010) *in situ* observations in a radius of influence of 5° latitude and 10°–25° longitude to reach a minimum of 500 profiles. The altimeter SLA product, with daily 0.25° resolutions, is from the AVISO project (http://www.aviso.oceanobs.com/duacs/). The differences between DHA calculated using a reference level at 1 000 m and altimeter SLA, assuming negligible measurement errors, represent the circulation at the reference level, or in other words, the deep baroclinic (below 1 000 m) and barotropic components of the circulation. At 1 000 m, the circulation is nonzero and must be removed from altimeter height before applying the vertical regression. This is obtained here using regression coefficients calculated for the DHA computed from Argo *T/S* profiles with a 1 000-m reference level, and the SLA altimeter collocated measurements. The method is very similar to that described by Guinehut et al. (2006), except that the reference level is taken at 1 000 m. The satellite SST product is the Reynolds L4 daily analysis on a 0.25° grid (Reynolds et al., 2007). The daily SLA and SST maps during the period of 2011–2012 are averaged into monthly composites to ensure adequate *in situ* profiles for one analysis. For the SSS measurements, it has been demonstrated that RMSDs are the smallest when only two variables (SLA and SST) are considered in temperature reconstructions (Guinehut et al., 2012; Nardelli et al., 2012) and therefore, SSS measurements are not used here.

Thus, the reconstruction schemes carried out in this paper are summarized in Fig. 1. By the static scheme, the *in situ*-only T field (derived from only *in situ* observations) was directly reconstructed by correcting the climate T field with *in situ* T profiles in OI. By the dynamic scheme, (1) the synthetic T field (derived from only satellite measurements) from the surface down to 1 000 m depth was first derived from altimeter SLA and SST measurements through a multiple linear regression method and regression covariances calculated from historical Argo thermohaline profiles; (2) the combined T field (derived from both *in situ* observations and satellite measurements) was then reconstructed by combining the synthetic T field and *in situ* T profiles in OI.

Model Resolution: In the original ISAS, the horizontal grid is a 0.5° Mercator, and the vertical resolution increases from 20 m at 2 000 m to 5 m in the upper layer. To match the remote sensing data, the horizontal resolution is increased to 0.25°; to achieve sufficient profiles for training regression relationships, the depth is truncated to 1 000 m above.

Background error covariance ($\boldsymbol{P}$, $\boldsymbol{C}_{\mathrm{ao}}$, and $\boldsymbol{C}_{\mathrm{oo}}$): The background error covariance between two points can be expressed as:

$$\begin{aligned} E[(x_i^{\mathrm{b}} - x_i^{\mathrm{t}})(x_j^{\mathrm{b}} - x_j^{\mathrm{t}})] &= \boldsymbol{C}_{ij}(\mathrm{d}x, \mathrm{d}y, \mathrm{d}t) \\ &= \sigma_{Li}\sigma_{Lj}\exp\left[-\left(\frac{\mathrm{d}x^2}{2L_x^2} + \frac{\mathrm{d}y^2}{2L_y^2} + \frac{\mathrm{d}t^2}{2L_t^2}\right)\right] + \\ &\quad \sigma_{Mi}\sigma_{Mj}\exp\left[-\left(\frac{\mathrm{d}x^2}{2M_x^2} + \frac{\mathrm{d}y^2}{2M_y^2} + \frac{\mathrm{d}t^2}{2M_t^2}\right)\right], \end{aligned} \tag{3}$$

where i, j is an observational point or a model grid; x^{b} and x^{t} are the background value and its corresponding true value; dx, dy, dt are the latitudinal, longitudinal, time separations between the points of i, j; L_x, L_y, L_t are the corresponding large-scale correlation scales; M_x, M_y, M_t are the corresponding mesoscale correlation scales; and σ_L, σ_M are the large-scale and mesoscale standard deviation (STD).

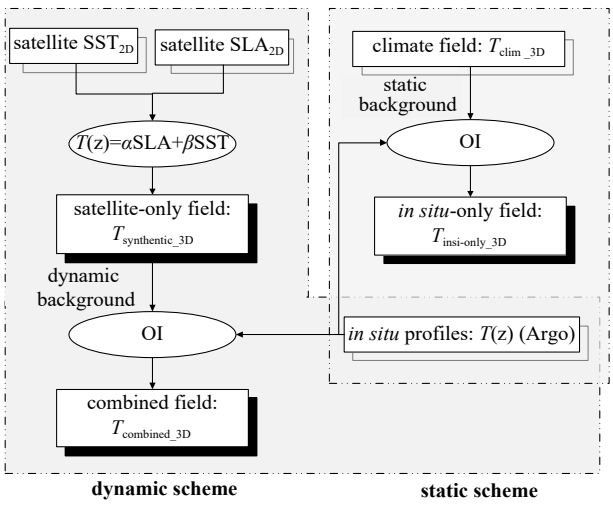

Fig. 1. The two schemes for reconstructing the 3-D ocean temperature fields.

The observation error covariance/variance ($\boldsymbol{R}$) between two points can be expressed as

$$E[(x_i^{\mathrm{o}}-x_i^{\mathrm{t}})(x_j^{\mathrm{o}}-x_j^{\mathrm{t}})]=\boldsymbol{R}_{ij}=\begin{cases}\sigma_{\mathrm{UR}i}^2+\sigma_{\mathrm{ME}i}^2, & i=j\\ 0, & i\neq j\end{cases}, \quad (4)$$

where i, j are observational points; x^{o} and x^{t} are the observed value and its corresponding true value; σ_{UR}^2 is the representative error variance due to scales the analysis cannot resolve, also known as the forward interpolation error; and σ_{ME}^2 is the instrument error variance due to inaccuracy in measurements. The error matrix combining the measurement error and the representation error is assumed to be diagonal, which means that the errors seen by the observation system are uncorrelated (Gaillard et al., 2016). This is only a crude approximation since both errors are likely to be correlated for measurements obtained with the same instrument or within the same area and time period.

Covariance scale: L_x and L_y are set to 300 km, i.e., the feature resolution of Argo floats, and L_t to 30 days, i.e., the feature duration time of large-scale oceanic processes; M_x and M_y are proportional to (four times) the Rossby radius calculated from the annual climatology of full depth, which decreases with increasing latitudes, with an upper limit of 300 km and a lower limit of twice the model grid, and M_t is set to 10 days, i.e., the feature duration time of mesoscale oceanic processes.

The total variance can be decomposed into four parts:

$$\sigma^2=\sigma_L^2+\sigma_M^2+\sigma_{\mathrm{UR}}^2+\sigma_{\mathrm{ME}}^2, \quad (5)$$

where σ^2 is the total variance; σ_{ME}^2 is the instrument error variance; $\sigma_L^2+\sigma_M^2+\sigma_{\mathrm{UR}}^2$ is ocean variance ($\sigma_{\mathrm{ocean}}^2$), which includes large-scale, mesoscale and representative error variances, namely, $\sigma_L^2=W_1\sigma_{\mathrm{ocean}}^2$, $\sigma_M^2=W_2\sigma_{\mathrm{ocean}}^2$, $\sigma_{\mathrm{UR}}^2=W_3\sigma_{\mathrm{ocean}}^2$, $W_1+W_2+W_3=1$; $\sigma_L^2+\sigma_M^2$ is the field variance, which appears in the matrix of $\boldsymbol{P}$, $\boldsymbol{C}_{\mathrm{ao}}$, and $\boldsymbol{C}_{\mathrm{oo}}$ and determines the proportions of large-scale and mesoscale signals accounting for the ocean variance; and $\sigma_{\mathrm{UR}}^2+\sigma_{\mathrm{ME}}^2$ is the error variance, which appears in the matrix of $\boldsymbol{R}$ and is considered noise. In the original ISAS, to account for under sampling in some areas (in particular the Southern Ocean), an expansion coefficient (f) is generally imposed on the ocean variance where $f\in[1, 2]$. In the dynamic scheme, this coefficient is possibly below 1 in some cases, as illustrated in Section 3.1.3.

3 Evaluation of the analysis toward assimilated observations

3.1 *Consistency diagnostics and tuning experiments using hindcasts*

3.1.1 *Diagnosis of observation, background, and analysis-error statistics*

Under the linear assumption and with correctly predefined $\boldsymbol{C}_{\mathrm{oo}}$ and $\boldsymbol{R}$, the following relationship holds (Desroziers et al., 2005) as follwos.

Background error covariance is

$$E[\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}}(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}]=\boldsymbol{C}_{\mathrm{oo}}. \quad (6)$$

Observation error covariance is

$$E[\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}}(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}]=\boldsymbol{R}. \quad (7)$$

Analysis error covariance is

$$E[\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}}(\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}})^{\mathrm{T}}]=\boldsymbol{C}_{\mathrm{oo}}+\boldsymbol{R}, \quad (8)$$

where $\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}}=\boldsymbol{y}^{\mathrm{a}}-\boldsymbol{y}^{\mathrm{b}}=\boldsymbol{H}\boldsymbol{x}^{\mathrm{a}}-\boldsymbol{H}\boldsymbol{x}^{\mathrm{b}}$, $\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}}=\boldsymbol{y}^{\mathrm{o}}-\boldsymbol{y}^{\mathrm{a}}=\boldsymbol{y}^{\mathrm{o}}-\boldsymbol{H}\boldsymbol{x}^{\mathrm{a}}$, and $\boldsymbol{y}^{\mathrm{o}}$, $\boldsymbol{y}^{\mathrm{b}}$, $\boldsymbol{y}^{\mathrm{a}}$ are the observation vector and its background and analysis counterparts in the observation space. The left side of Eqs (6)–(8) denote the diagnostic (or a posteriori) values for the background, observation, and analysis-error covariance (in observation space), and the right parts are the predefined (or a priori) counterparts.

The diagnostics shown above potentially provide information on the full covariances of observation, background, and analysis errors in observation space. One application of these diagnostics is to diagnose the diagonal parts of the covariance matrices. For any subset of observations i with N_i observations, it is possible simply to compute the quantities:

$$(\sigma_{\mathrm{b}}^{\mathrm{d}})_i^2=(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}})_i^{\mathrm{T}}(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}})_i/N_i=\sum_{j=1}^{N_i}(y_j^{\mathrm{a}}-y_j^{\mathrm{b}})(y_j^{\mathrm{o}}-y_j^{\mathrm{b}})/N_i, \quad (9)$$

$$(\sigma_{\mathrm{o}}^{\mathrm{d}})_i^2=(\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}})_i^{\mathrm{T}}(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}})_i/N_i=\sum_{j=1}^{N_i}(y_j^{\mathrm{o}}-y_j^{\mathrm{a}})(y_j^{\mathrm{o}}-y_j^{\mathrm{b}})/N_i, \quad (10)$$

$$(\sigma_{\mathrm{a}}^{\mathrm{d}})_i^2=(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}})_i^{\mathrm{T}}(\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}})_i/N_i=\sum_{j=1}^{N_i}(y_j^{\mathrm{a}}-y_j^{\mathrm{b}})(y_j^{\mathrm{o}}-y_j^{\mathrm{a}})/N_i, \quad (11)$$

where $y_j^{\mathrm{o}}, y_j^{\mathrm{b}}, y_j^{\mathrm{a}}$ are the observation, background, and analysis values at point j. The diagnosed values of background ($(\sigma_{\mathrm{b}}^{\mathrm{d}})_i^2$), observation ($(\sigma_{\mathrm{o}}^{\mathrm{d}})_i^2$), and analysis ($(\sigma_{\mathrm{a}}^{\mathrm{d}})_i^2$) errors can be different from the predefined values (denoted as $(\sigma_{\mathrm{b}}^{\mathrm{p}})_i^2$, $(\sigma_{\mathrm{o}}^{\mathrm{p}})_i^2$, and $(\sigma_{\mathrm{a}}^{\mathrm{p}})_i^2$). They correspond to the use of relations (6)–(8), respectively, but only focus on the mean diagonal elements of those matrices. These computations are shown to be nearly cost-free since they only combine quantities available after the analysis.

3.1.2 *Tuning of observation and background-error variances*

In the original ISAS configuration, the background and observation error covariances are exclusively determined by the weight coefficients (W_i) and the expansion coefficient (f), both of which are constant (W_1:W_2:W_3=1:2:3, f=1 or larger). For the dynamic scheme, the synthetic field derived from real-time satellite observations is introduced. Constant background and observation errors are expected to be incorrectly specified in OI. Thus, a method must be envisaged to tune them. The rationale of such a tuning procedure is to determine the values of $(\sigma_{\mathrm{b}})_i^2$ and $(\sigma_{\mathrm{o}})_i^2$, such that those values fulfil the relations $(\sigma_{\mathrm{b}})_i^2=(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}})_i^{\mathrm{T}}(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}})_i/N_i$ and $(\sigma_{\mathrm{o}})_i^2=(\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}})_i^{\mathrm{T}}(\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}})_i/N_i$. This is a nonlinear problem since $\boldsymbol{d}_{\mathrm{b}}^{\mathrm{a}}$, $\boldsymbol{d}_{\mathrm{a}}^{\mathrm{o}}$, and $\boldsymbol{d}_{\mathrm{b}}^{\mathrm{o}}$ depend themselves on the $(\sigma_{\mathrm{b}})_i^2$ and $(\sigma_{\mathrm{o}})_i^2$ values. The form of those nonlinear equations suggests the use of an iterative fixed-point method to solve this tuning problem (Desroziers et al., 2005).

To compare the predefined errors with the diagnosed errors, two scalar parameters were introduced:

$$S_{\mathrm{b}}=(\sigma_{\mathrm{b}}^{\mathrm{d}})^2/(\sigma_{\mathrm{b}}^{\mathrm{p}})^2, \quad S_{\mathrm{o}}=(\sigma_{\mathrm{o}}^{\mathrm{d}})^2/(\sigma_{\mathrm{o}}^{\mathrm{p}})^2. \quad (12)$$

One should redo the analysis iteratively. That is, the factors S_{b} and S_{o} are applied to error covariances ($\boldsymbol{C}_{\mathrm{oo}}$ and $\boldsymbol{R}$) by adjusting W_i and f to obtain the new estimates of S_{b} and S_{o}. However, it has

been shown by the cited authors that the iteration converges quickly, such that the first estimate is not too far from the final converged value. Moreover, the speed of convergence increases with the number of observations. Instead of region- or depth-dependent tuning, constant (global) tuning strategy was chosen. That is, these errors were calculated from all observations worldwide (60°S to 60°N) and averaged for all horizontal layers. Therefore, they represent the mean of error estimates from different regions and layers.

3.1.3 *Results of tuning experiments*

Tuning is performed for both the dynamic scheme and the static scheme. A superscript * is added to the dynamic diagnostics, namely, $\sigma_b^{p^*}$, $\sigma_o^{p^*}$, $\sigma_b^{d^*}$, and $\sigma_o^{d^*}$, to be distinguished from the static diagnostics, namely, σ_b^{p}, σ_o^{p}, σ_b^{d}, and σ_o^{d}. Convergence was achieved after two iterations for both schemes. The initial and final diagnostic profiles are shown in Figs 2 and 3. Initially, for both schemes, W_1:W_2:W_3 equals to 1:2:3 and f equals to1.00. Finally, for the static scheme, W_1:W_2:W_3 equals to 1:2:2.23 and f equals to1.81; for the dynamic scheme, W_1:W_2:W_3 equals to 1:2:2.99 and f equals to 0.73.

Figures 2 and 3 show that the final diagnostic statistics agree better with the predefined statistics after iteration, although some discrepancies still exist. For the static scheme (Fig. 2), after iterat-

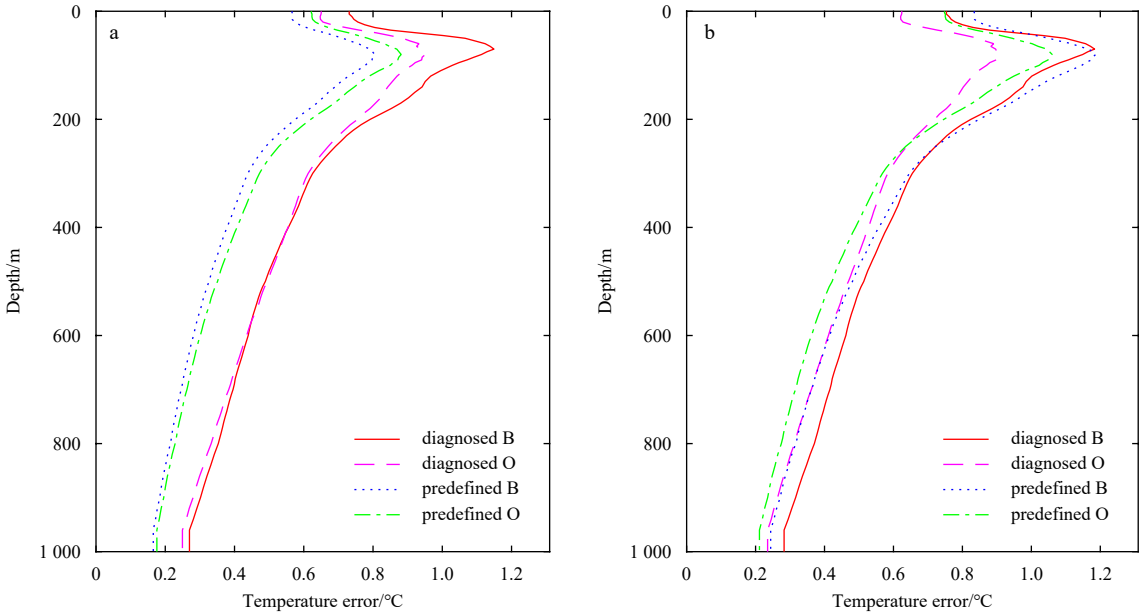

Fig. 2. The profiles of diagnosed B (σ_b^d), diagnosed O (σ_o^d), predefined B (σ_b^p), and predefined O (σ_o^p): initially (a) and after (b) two iterations in the static scheme. B denotes background error and O denotes observation error.

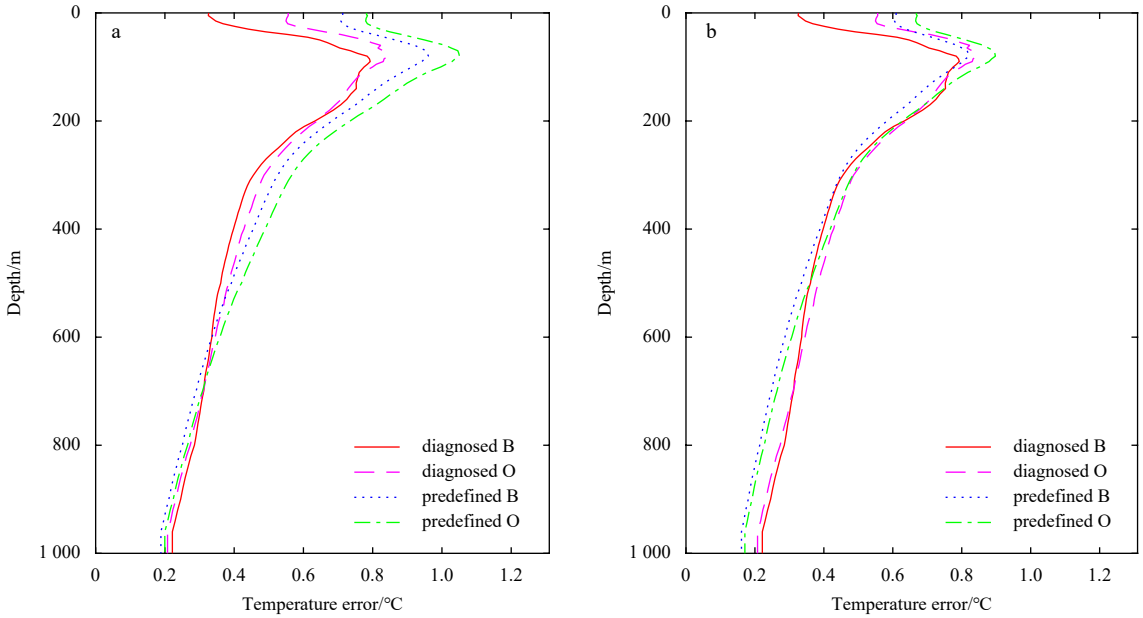

Fig. 3. The profiles of $\sigma_b^{d^*}$, $\sigma_o^{d^*}$, $\sigma_b^{p^*}$, and $\sigma_o^{p^*}$: initially (a) and after (b) two iterations in the dynamic scheme. B denotes background error and O denotes observation error.

ive tuning, the maximum difference between diagnostic and predefined values (for both background and observation error) on all layers was reduced to less than 0.2°C (Fig. 2b) from greater than 0.4°C (Fig. 2a). In such a way, the RMSD of the *in situ*-only field was reduced by ~ 0.1°C on average (not shown). Additional iterations led to a negligible reduction in error (<0.03°C), which implies that the iterative process was over. For the dynamic scheme (Fig. 3), after iterative tuning, the maximum difference between diagnostic and predefined values was reduced to be <0.3°C (Fig. 3b) from >0.4°C (Fig. 3a) such that the average RMSD of the combined field was reduced by ~ 0.1°C (not shown).

It can be seen from the values of f that both background and observation errors seemed to be underestimated in the static scheme and overestimated in the dynamic scheme before tuning. For the static scheme, the predefined (with an initial f of 1.00) background errors $(\sigma_b^p)^2$ and observation errors $(\sigma_o^p)^2$ were underestimated by almost 50% (with a tuned f of 1.81). This is mainly because of the lack of profiles used to calculate marine variance σ^2_{ocean}, which has been found by several studies (Gaillard et al., 2009; Gaillard, 2012; Von Schuckmann et al., 2009). For the dynamic scheme, the predefined (with initial f of 1.00) $(\sigma_b^{p*})^2$ and $(\sigma_o^{p*})^2$ have been overestimated slightly (with tuned f of 0.73). This is probably due to the slight discrepancies in the temporal spectra of the two data sources. Specifically, the synthetic field is real-time and monthly, but the historical *in situ* profiles around it are over a long period from 2004 to 2010. The latter contains ocean temperature viabilities on broader timescales than the former. Therefore, the marine variance calculated from their difference would be greater than the actual variance. Note that all diagnostics mentioned afterward refer to the finely tuned diagnostics.

On the other hand, the changes in values of W_3: (W_1+W_2) indicate the changes in the ratio between observation and background errors, known as the noise-to-signal ratio (NSR). The NSR seemed to be overestimated in the static scheme and roughly reasonable in the dynamic scheme before tuning.

3.2 *Contributions of satellite and Argo measurements to analysis improvement*

3.2.1 *Comparison between RMSD and diagnostic errors*

(1) Patterns of synthetic, *in situ*-only, and combined field RMSD

The RMSDs of the climatology field, synthetic field, *in situ*-only field, and combined field toward assimilated observations are denoted as σ_{clim}, σ_{synt}, σ_{insi}, and σ_{comb}, respectively. Among them, $\sigma_{clim}=((\boldsymbol{T}_{clim}-\boldsymbol{T}_{obs})\cdot(\boldsymbol{T}_{clim}-\boldsymbol{T}_{obs})^T/N_{obs})^{1/2}$; likewise, for the other three quantities) represents the temperature climate variability "signal"; the latter three quantities represent the "noise" of the corresponding analysis fields. Therefore, the latter three can be denoted as percentages of the former, namely, $p_{synt}=\sigma_{synt}/\sigma_{clim}$, $p_{insi}=\sigma_{insi}/\sigma_{clim}$, and $p_{comb}=\sigma_{comb}/\sigma_{clim}$. They represent the "analysis improvement" contributed from the synthetic field, *in situ*-only field, and combined field, respectively, on positions of assimilated observations.

First, the profiles of RMSD averaged over different sea areas (i.e., globally, in the extratropics, in the tropics) are shown in Fig. 4. Globally (Figs 4a and b), the RMSD of the synthetic field ranges from 0.6°C at the surface, with a maximum of 1.2°C in the thermocline layer depth decreasing to 0.3°C down to 1 000 m. For the *in situ*-only field and the combined field, their RMSD ranges from ~0.5°C at the surface, with a maximum of 0.8°C in the thermocline layer depth decreasing to 0.2°C down to 1 000 m. As a result, the synthetic field reduces RMSD by 20%–40% relative to the climate field, but the combined field reduces RMSD by <5% relative to the *in situ*-only field. That is, the use of satellite data has little impact on the final analysis fields of assimilated observations. In extratropical oceans (Figs 4c and d), the order of the RMSD amplitudes of the four fields are the same as that of the global case. However, the RMSD fluctuates more smoothly in the thermocline with reduced magnitude. In tropical oceans (Figs 4e and f), the order of the RMSD amplitudes of the four fields are the same as that of the global case. However, the temperature fluctuates

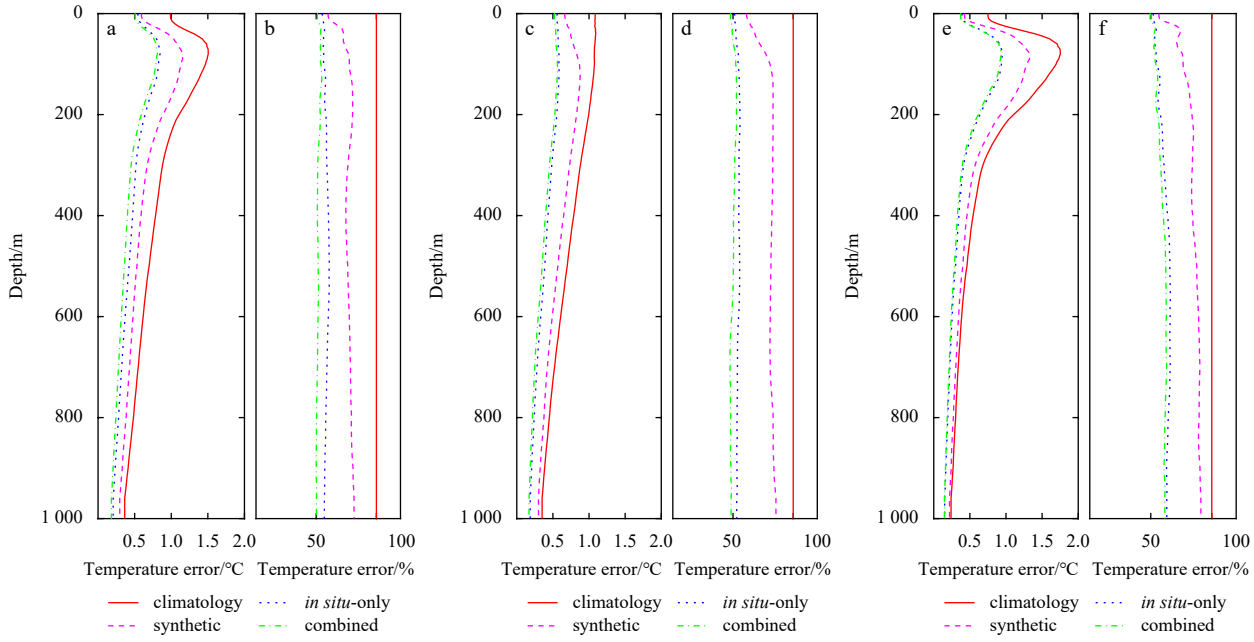

Fig. 4. Profiles of temperature error (RMSD) toward assimilated observations over the global ocean (60°S–60°N): σ_{clim}, σ_{synt}, σ_{insi}, and σ_{comb} (a); p_{synt} (=$\sigma_{synt}/\sigma_{clim}$), p_{insi} (=$\sigma_{insi}/\sigma_{clim}$), and p_{comb} (=$\sigma_{comb}/\sigma_{clim}$) (b) . c–d. The same as a and b but over the extratropical ocean (30°–60°S, 30°–60°N). e–f. The same as a and b but over the tropical ocean (30°S–30°N).

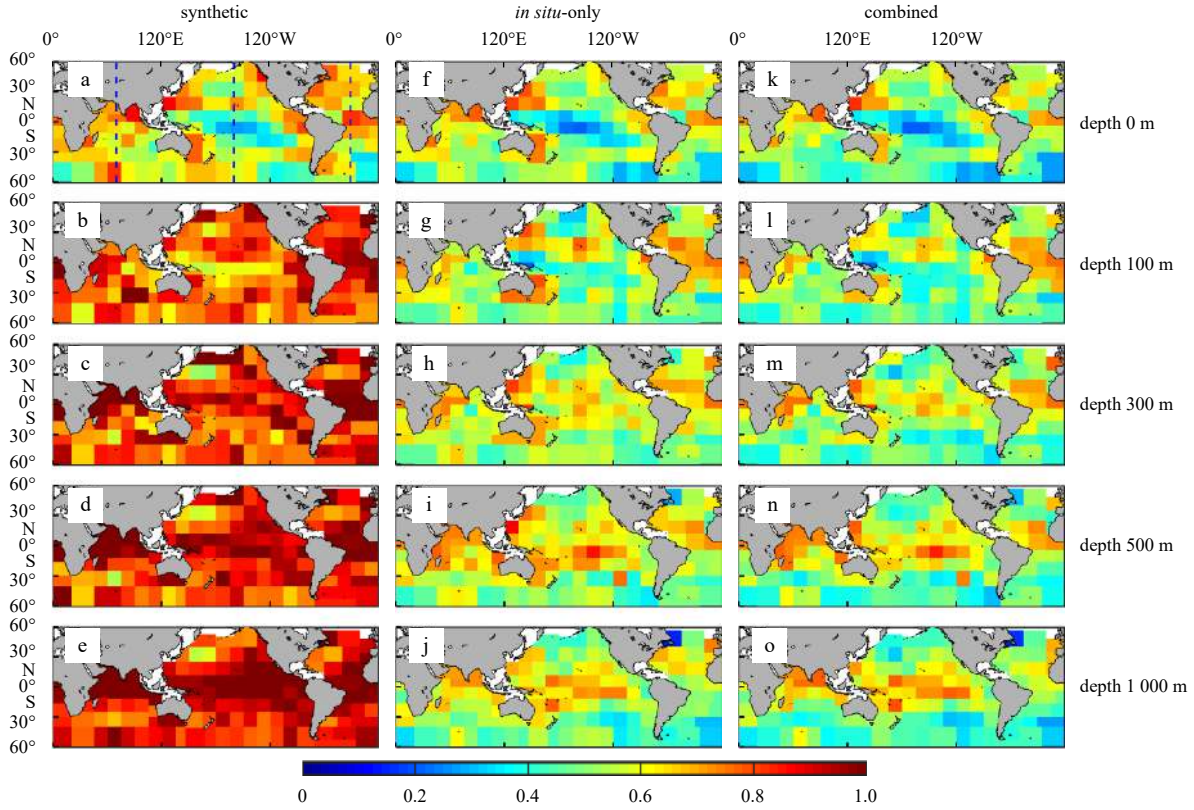

Fig. 5. Global maps of p_{synt} (=$\sigma_{synt}/\sigma_{clim}$) (a–e), p_{insi} (=$\sigma_{insi}/\sigma_{clim}$) (f–j), and p_{comb} (=$\sigma_{comb}/\sigma_{clim}$) (k–o)toward assimilated observations at each horizontal layer. The dash straight lines in a show the locations of the three subregions selected in Fig. 8.

more sharply in the thermocline with amplified magnitude.

Second, the horizontal maps of RMSD at various horizontal layers are shown in Fig. 5. For the synthetic field (left column of Fig. 5) at the surface, the RMSD is smaller (<0.5) in tropical oceans and larger in extratropical oceans (>0.5); at 100 m, low-RMSD areas almost disappear along equatorial oceans; at 300 m, low-RMSD areas move to extratropical oceans; at 500 m and below, low-RMSD areas are concentrated around the northwest and southwest Pacific, northwest and southwest Atlantic, South Indian, and other oceans. For the *in situ*-only field (central column of Fig. 5), at the surface, the RMSD is smaller in the tropics and larger in the extratropics; at 100 m, low-RMSD areas in the tropics significantly shrink, and the RMSD is significantly smaller than that of the synthetic field as a whole, suggesting that the *in situ* profiles begin to become more influential on the analysis downward from this layer; at 300 m, low-RMSD areas move to the extratropics; at 500 m and below, low-RMSD areas are gradually concentrated around the middle of the North Pacific, Southeast Pacific, South Atlantic, South Indian, and other oceans. For the combined field (right column of Fig. 5), the RMSD is roughly the same as the *in situ*-only field case and thus is not rephrased.

(2) Patterns of diagnostic errors in the observational space

The finely tuned diagnostics in the observational space introduced in Section 3.1, i.e., the diagnostic observation and background errors (σ_b^d, $\sigma_o^d$and $\sigma_b^{d^*}$, as well as $\sigma_o^{d^*}$), are used to account for the patterns of RMSD above and to evaluate the contribution of various data sources. The diagnosed background error of the static scheme (σ_b^d) is used as a benchmark representing the temperature signal STD. In this way, the other three diagnosed quantities are expressed as a ratio of σ_b^d, namely, p_b=$\sigma_b^{d^*}/\sigma_b^d$, p_o=σ_o^d/σ_b^d, and p_o^*=$\sigma_o^{d^*}/\sigma_b^d$.

Figure 6 shows the patterns of three ratios calculated from the diagnostic quantities, i.e., p_b, p_o, and p_o^*. Comparing Figs 5 and 6, it can be seen that three pairs of consistent patterns between p_b and p_{synt}, p_o and p_{insi}, p_o^*, and p_{comb}. This consistence is even clearer in Fig. 7. All scatter plots in Fig. 7 show a significant correlation between those three groups of quantities (p_b and p_{synt}; p_o and p_{insi}; p_o^* and p_{comb}) and uniformly provide a rationale behind the RMSD features in Fig. 5. That is, p_b, p_o, and p_o^* seem to be able to represent the relative RMSD of the synthetic *T* field (p_{synt}), *in situ*-only *T* field (p_{insi}) and combined *T* field (p_{comb}) toward assimilated observations, respectively, as shown below.

3.2.2 *Explaining patterns of RMSD toward assimilated observations*

(1) Rationale behind the synthetic-field RMSD feature

For the synthetic field, no direct diagnostics in Section 3.1 can be used since it is synthesized on the model grids, not in the observational space. Alternatively, another diagnostic, i.e., the explained variance (EV) inferred from the regression formulation (Eq. (2)), can be used to account for the RMSD feature of the *synthetic* field. The EV is defined as the ratio of the regression variance to its forecast variance, and its unbiased estimate on the level *z* is

$$\begin{aligned} \mathrm{EV}(z) &= \hat{R}^2(z) = \hat{\sigma}_r^2(z)/\hat{\sigma}_f^2(z) \\ &= 1 - [(n-1)/(n-r-1)][1-\sigma_r^2(z)/\sigma_f^2(z)], \end{aligned} \tag{13}$$

where *r* is the number of regression variables and is equal to 2 here; the forecast variance σ_f^2 (also called signal variance) is

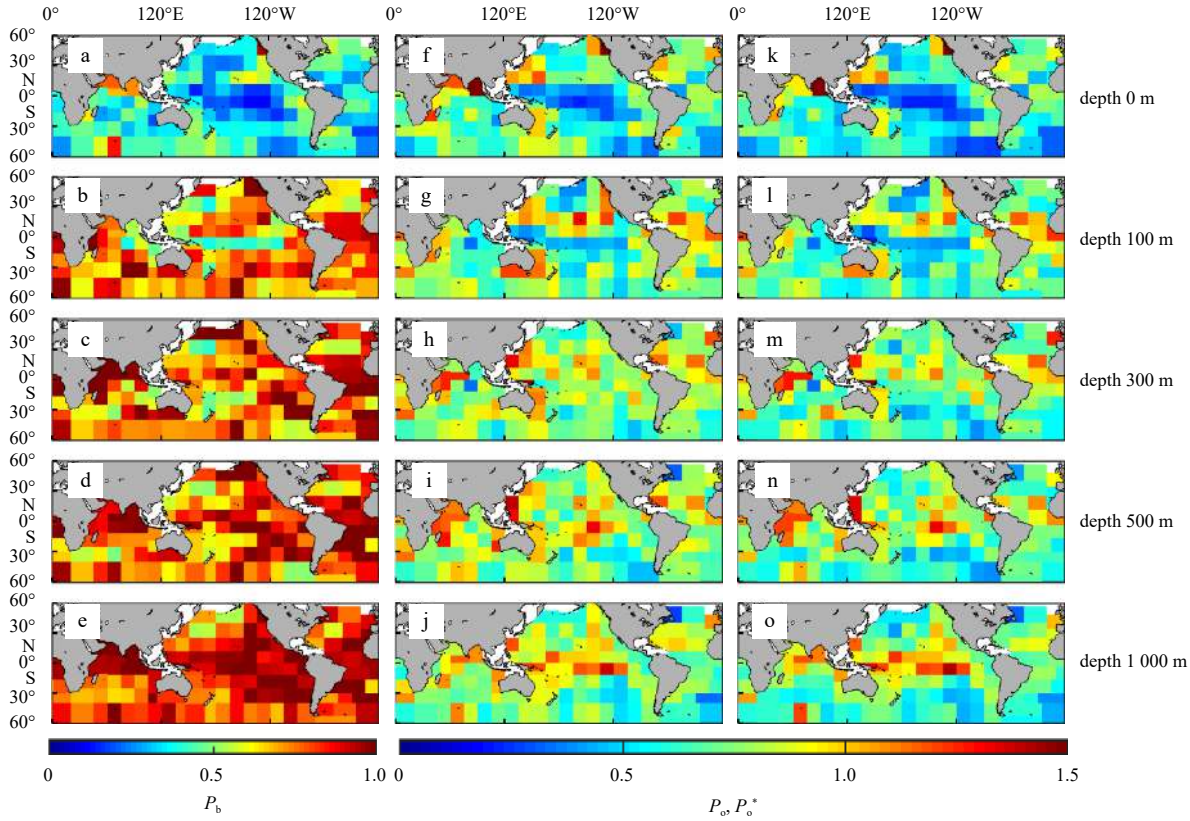

Fig. 6. Global maps of p_b (=$\sigma_b^{d^*}/\sigma_b^d$) (a–e), p_o (=σ_o^d/σ_b^d) (f–j), and p_o^* (=$\sigma_o^{d^*}/\sigma_b^d$) (k–o) at each horizontal layer.

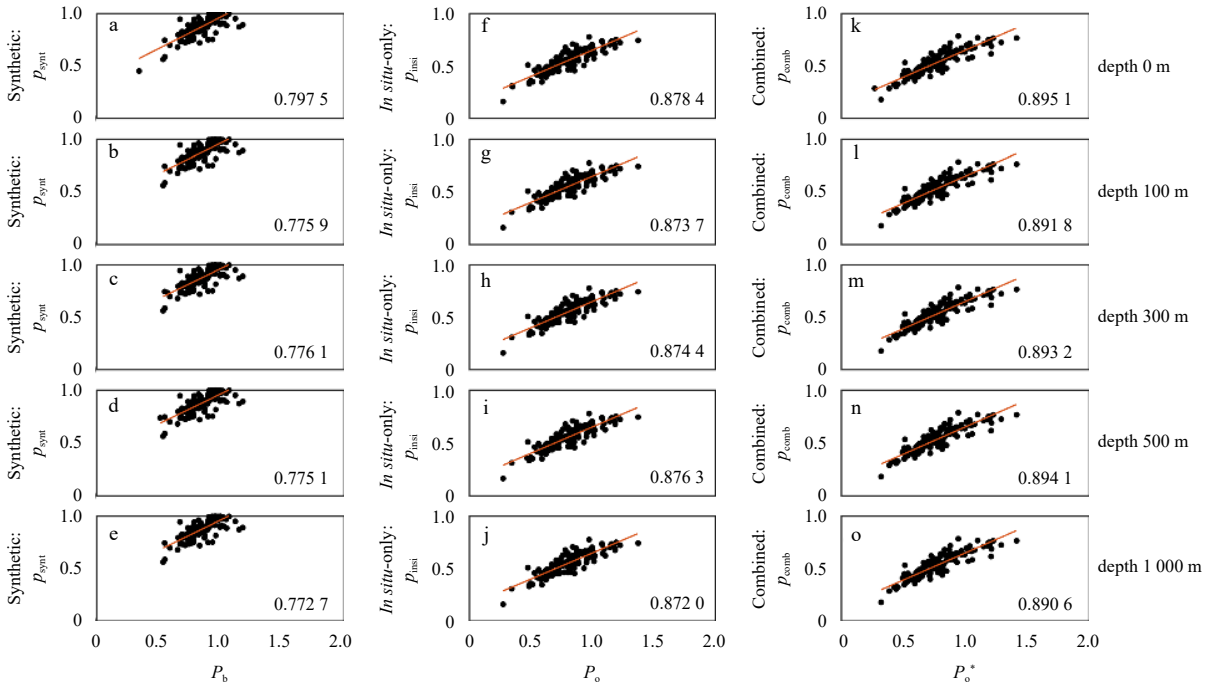

Fig. 7. Scatter plots of p_{synt} (as displayed in Figs 5a–e) against p_b (as displayed in Figs 6a–e) (a–e), p_{insi} (as displayed in Figs 5f–i) against p_o (as displayed in Figs 6f–i) (f–j), and p_{comb} (as displayed in Figs 5k–o) against p_o^* (as displayed in Figs 6k–o) (k–o) considering geographic dependence at each horizontal layer. The correlation coefficient corresponding to each scatter plot is also given on the bottom right of each panel, all of which are significant at the 1% significance level.

roughly equal to σ^2_{clim}; and the regression variance is expressed as

$$\sigma_r^2(z) = \lambda(z)\langle \delta\mathrm{SLA}, \delta T(z)\rangle + \theta(z)\langle \delta\mathrm{SST}, \delta T(z)\rangle. \quad (14)$$

The essence of $\sigma_b^{d^*}/\sigma_b^d$ is the EV for the vertical regression, i.e., $\mathrm{EV} = \sigma_r^2(z)/\sigma_f^2(z)$. Specifically, (1) σ_b^d characterizes the signal variance: $\sigma_f^2(z) = \langle \delta T(z), \delta T(z)\rangle$ (related to the climate field); (2) $\sigma_b^{d^*}$

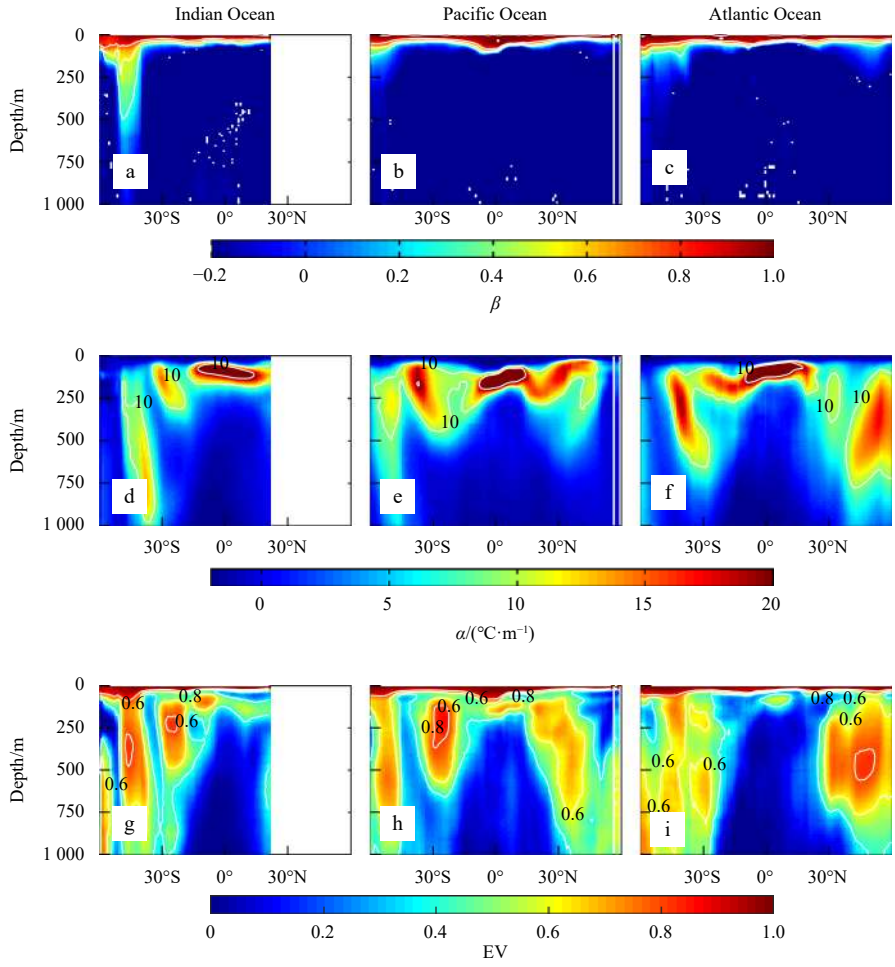

Fig. 8. Depth-latitude sections of annual β (a–c), α (d–f), and EV (g–i) along 70°E in the Indian Ocean (left column), 160°W in the Pacific Ocean (central column) , and 30°W in the Atlantic Ocean (right column), as shown by the three straight lines in Fig. 5a.

characterizes the regression variance: $\sigma_r^2(z) = \lambda(z)\langle \delta\mathrm{SLA}, \delta T(z)\rangle + \theta(z)\langle \delta\mathrm{SST}, \delta T(z)\rangle$ (related to the synthetic field); (3) thus, $\sigma_\mathrm{b}^{\mathrm{d}^*}/\sigma_\mathrm{b}^\mathrm{d}$ can account for the synthetic field RMSD relative to the climate field RMSD, i.e., p_synt.

The depth-latitude sections of the three regression parameters, i.e., β, α, EV, in each ocean are shown in Fig. 8. At the surface, the distribution of p_synt is determined by the accuracy in satellite SST data itself, which is relatively high at lower latitudes (Fig. 5a). In the interior ocean (Figs 5b–e), the magnitude of p_synt in each layer is determined by the magnitude of EV, with a smaller p_synt reflecting a larger EV (Figs 8c, f and i). On the other hand, the magnitude of EV is determined by the magnitudes of β (Figs 8a, d and g) and α (Figs 8b, e and h). Specifically, from the surface to the mixed layer depth, the magnitude of β rapidly decreases in the vertical direction (Figs 8a, d and g) such that the low-p_synt areas along the equator rapidly shrink (from Figs 5a and b). Conversely, below the mixed layer α stays greater at middle and high latitudes and smaller in the tropics (Figs 8b, e and h), which results in a shift of low-p_synt areas to extratropical oceans (from Figs 5b–e).

(2) Rationale behind the *in situ*-only field RMSD feature

For the *in situ*-only field, its background and observation errors can only be distinguished in observational space, as described in Section 3.1. The comparison between the predefined background and observation errors ($\sigma_\mathrm{b}^\mathrm{p}$, $\sigma_\mathrm{o}^\mathrm{p}$; diagonal parts of left matrices of Eqs (6) and (7)) and the diagnostic background and observation errors ($\sigma_\mathrm{b}^\mathrm{d}$, $\sigma_\mathrm{o}^\mathrm{d}$; diagonal parts of right matrices of Eqs (6) and (7)) can be used to analyze the impacts of Argo observations.

The essence of $\sigma_\mathrm{o}^\mathrm{d}/\sigma_\mathrm{b}^\mathrm{d}$ (i.e., p_ob^d) is the NSR. Specifically, $\sigma_\mathrm{b}^\mathrm{d}$ characterizes the large-scale and mesoscale signals that *in situ* observations can resolve (related to the climate field RMSD); $\sigma_\mathrm{o}^\mathrm{d}$ characterizes the small-scale signals that the *in situ* observations cannot resolve (related to the *in situ*-only field RMSD); thus, p_ob^d can account for the *in situ*-only field RMSD relative to the climate field RMSD, i.e., p_ins.

At hindcast positions, the profiles of predefined ($\sigma_\mathrm{o}^\mathrm{p}$, $\sigma_\mathrm{b}^\mathrm{p}$) and diagnostic errors ($\sigma_\mathrm{o}^\mathrm{d}$, $\sigma_\mathrm{b}^\mathrm{d}$), as well as corresponding ratio quantities (p_ob^p=$\sigma_\mathrm{o}^\mathrm{p}/\sigma_\mathrm{b}^\mathrm{p}$, p_ob^d=$\sigma_\mathrm{o}^\mathrm{d}/\sigma_\mathrm{b}^\mathrm{d}$), are shown in Fig. 9 (for the tropics) and Fig. 10 (for the extratropics). Under the premise that the instrument error is ignorable, i.e., only the representative error is contained in $\sigma_\mathrm{o}^\mathrm{p}$, it is known from Section 2 that $(\sigma_\mathrm{o}^\mathrm{p})^2$ and $(\sigma_\mathrm{b}^\mathrm{p})^2$ are proportional to the ocean variance (σ_ocean^2). Therefore, at any latitude and depth, p_ob^p as a predefined quantity is approximately constant. For the static scheme, p_ob^p is roughly $\sqrt{W_3/(W_1+W_2)} = 0.86$. In contrast, p_ob^d as an a posteriori quantity varies with latitude and depth.

In the tropics (Fig. 9), $\sigma_\mathrm{o}^\mathrm{d}$ is smaller (greater) than $\sigma_\mathrm{o}^\mathrm{p}$ above (below) 500 m and σ_b^d is greater (smaller) than $\sigma_\mathrm{b}^\mathrm{p}$ above (below) 300 m (Fig. 9a); as a result, p_ob^d is smaller (greater) than p_ob^p above (below) 400 m (Fig. 9b); in this way, p_insi is small in the upper lay-

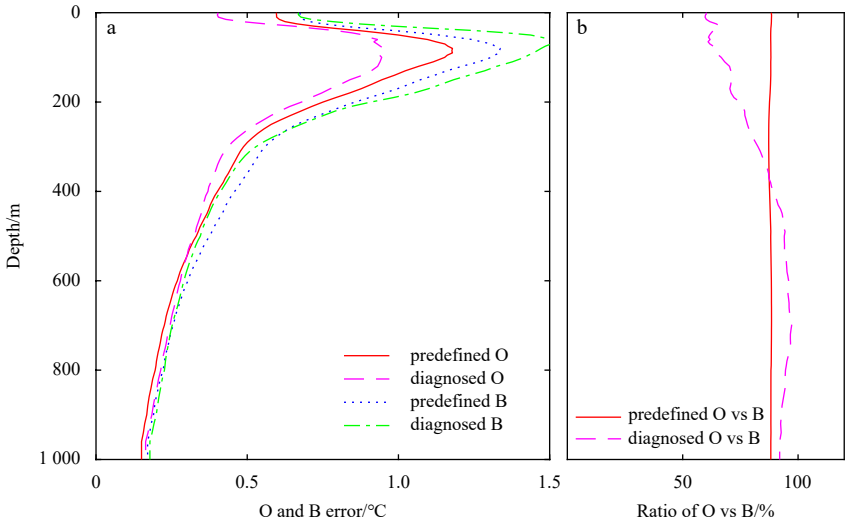

Fig. 9. Profiles of σ_o^p, σ_o^d, σ_b^p, and σ_b^d (a); p_{ob}^p, p_{ob}^d in the tropical ocean (30°S–30°N) (b). B denotes background error and O denotes observation error.

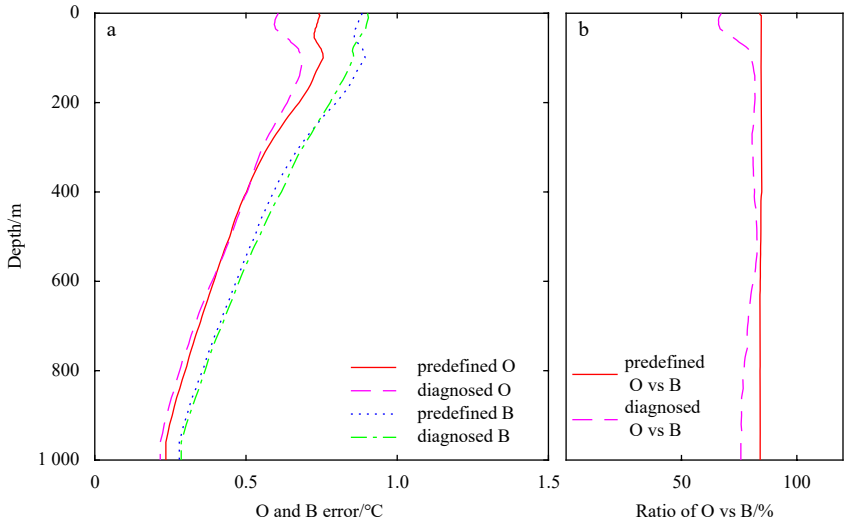

Fig. 10. Profiles of σ_o^p, σ_o^d, σ_b^p, and σ_b^d (a); p_{ob}^p, p_{ob}^din the extratropical ocean (30°–60°S, 30°–60°N) (b). B denotes background error and O denotes observation error.

er and large in the deep layer (Figs 5f–j). It can be inferred that in the tropical upper oceans, the actual large-scale and mesoscale signals intend to be stronger than the predefined signals, and the actual small-scale signals intend to be weaker, perhaps due to some significant large-scale phenomenon (e.g., ENSO) arising here.

In the extratropics (Fig. 10), σ_o^d (σ_b^d) is roughly equal to σ_o^p (σ_b^p), except above 300 m, where σ_o^d is notably less than σ_o^p (Fig. 10a); as a result, the profiles of p_{ob}^d and p_{ob}^p are basically parallel, except between the surface and 100 m, where p_{ob}^d deviates from p_{ob}^p (Fig. 10b); in this way, p_{insi} remains roughly constant throughout all layers (Figs 5f–j). It can be inferred that in the extratropics, the predefined NSRs are reasonable and do not vary with depth.

(3) Rationale behind the combined-field RMSD feature

The p_{comb} (right column of Fig. 5) and p_{insi} (central column of Fig. 5) have similar patterns due to the dominant influence of *in situ* Argo observations, and thus, both can be explained by the noise-to-signal ratio theory above. Conversely, the contribution of satellite measurements to analysis improvement can only be significant for independent observations (see Section 4.2).

4 Evaluation of the analysis of independent observations

4.1 *Patterns of synthetic, in situ-only, and combined field RMSD*

As an obvious disadvantage, the diagnostics in Section 3 cannot be directly applied in a model space. Therefore, a simplification is made in that error information in the model space can be reflected by a combination of information on the assimilated observations and information on the independent observations. Thus, as an extension of Section 3.2, an additional RMSD analysis is performed for independent validation. Some independent observations were randomly selected for validation, which accounted for 10% of the total number of profiles, versus the hindcast validations used for OI calculation, which accounted for 90% of the total number of profiles in Section 3.2. The counterparts (analysis values vs. independent observations) were taken by interpolating the four fields onto the positions of the independent validations within the monthly time window. Similar to the defin-

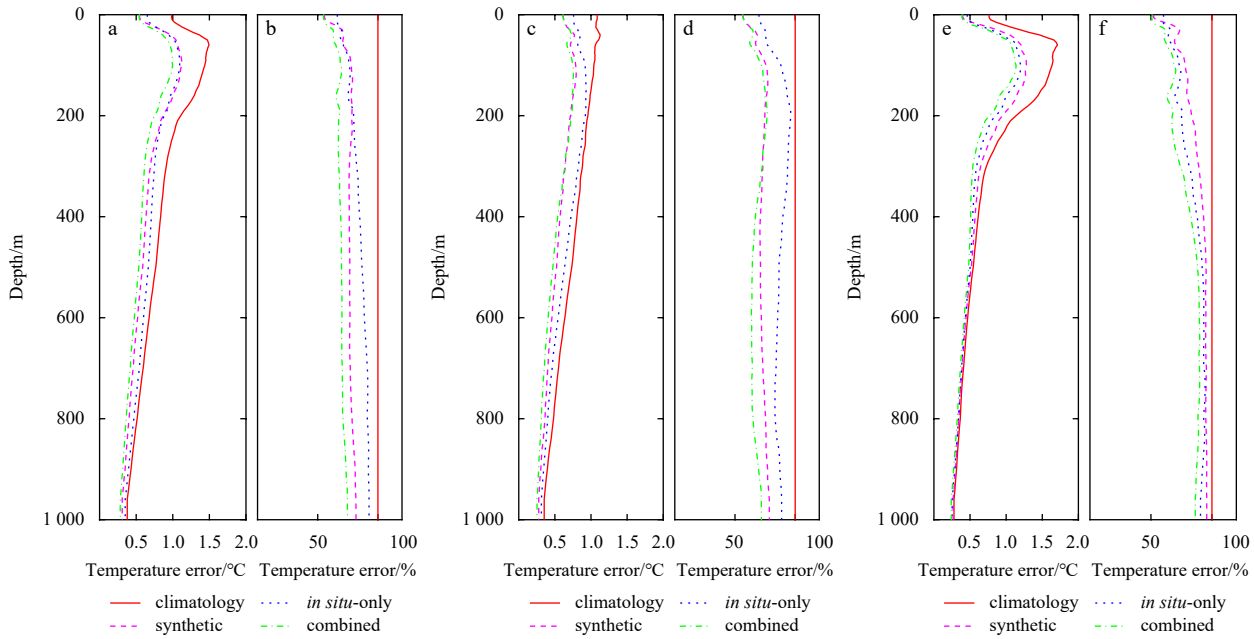

Fig. 11. Profiles of temperature error (RMSD) toward independent observations over the global ocean (60°S–60°N): σ_{clim}, σ_{synt}, σ_{insi}, and σ_{comb} (a); $p_{synt}(=\sigma_{synt}/\sigma_{clim})$, $p_{insi}(=\sigma_{insi}/\sigma_{clim})$, and $p_{comb}(=\sigma_{comb}/\sigma_{clim})$ (b). c–d. The same as a and b but over the extratropical ocean (30°–60°S, 30°–60°N). e–f. The same as a and b but over the tropical ocean (30°S–30°N).

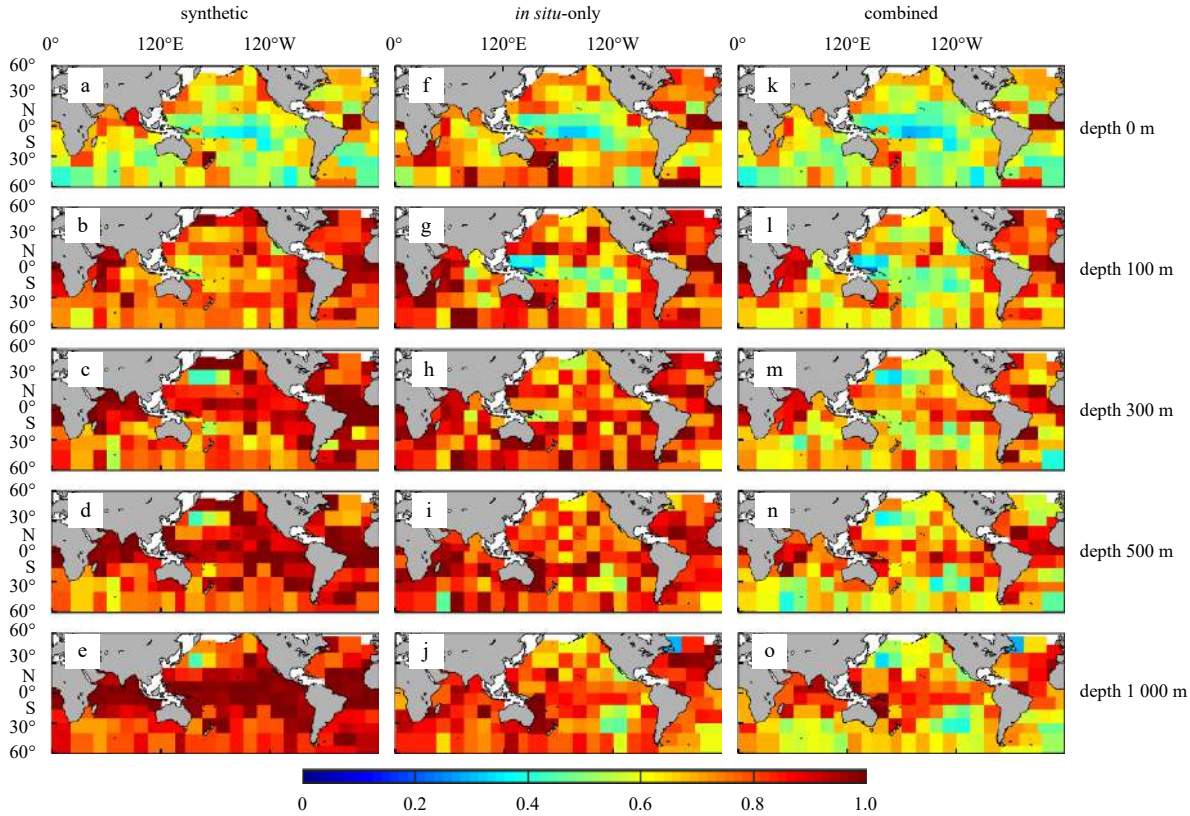

Fig. 12. Global maps of p_{synt} $(=\sigma_{synt}/\sigma_{clim})$ (a–e), p_{insi} $(=\sigma_{insi}/\sigma_{clim})$ (f–j), and p_{comb} $(=\sigma_{comb}/\sigma_{clim})$ (k–o) toward independent observations at each horizontal layer.

ition in Section 3.2.1, the latter three RMSDs are sometimes expressed in the "noise-signal" ratio form, i.e., p_{synt}, p_{insi}, and p_{comb}.

Firstly, the RMSD profiles of different sea areas (i.e., globally, in the extratropics, and in the tropics) are shown in Fig. 11. Globally (Figs 11a and b), the synthetic field has an even smaller RMSD by 0–15% than the *in situ*-only field (except on the 50–250 m layer where they are roughly equivalent). For the *in situ*-only field and the combined field, their RMSD both increased com-

pared to those on assimilated observations but to completely different degrees. At 1 000 m, for example, approximately 10% and 25% of the temperature signals are reconstructed from the two fields, respectively. As a result, the combined field reduces the RMSD by 5%–20% relative to the *in situ*-only field. This reduction indicates that at independent validation points and on a global scale, using only satellite data is generally advantageous over using only *in situ* data. In extratropical oceans (Figs 11c and d), the synthetic RMSD is smaller than the *in situ*-only RMSD by 10%–25% and even close to the combined RMSD above 300 m. This difference indicates that at independent validation positions and in extratropical oceans, using satellite data only is clearly advantageous over using *in situ* data only, and merging both datasets is closer to using satellite data only. In tropical oceans (Figs 11e and f), opposite to the case in the extratropics, the synthetic RMSD is larger than the *in situ* only RMSD by 0%–10%, but both of them are not much larger than the combined RMSD. On independent validation positions and in tropical oceans, using only *in situ* data is slightly advantageous compared to satellite only data, and combining both data is closer to using only *in situ* data.

Secondly, the horizontal maps of RMSD toward independent observations are shown in Fig. 12. For the synthetic field (Figs 12a–e), the RMSD is the same as the RMSD on assimilated observations (Figs 5a–e) since no *in situ* observations have been involved. For the *in situ*-only field (Figs 12f–j), at all layers, the RMSD is much larger than the RMSD on assimilated observations (Figs 5f–j), suggesting that the observations become less influential on the analysis toward independent observations; at the surface and 100 m, low-RMSD areas are still concentrated in the tropics, but at 300 m and below there are no obvious low-RMSD areas, particularly in extratropical oceans where, notably, *in situ* observations fail to improve the analysis. For the combined field (Figs 12k–o), the RMSD has a larger magnitude but similar distribution (i.e., varying with latitude) compared to that using assimilated observations (Figs 5k–o); on the other hand, the RMSD is roughly the same in tropical oceans but significantly smaller in extratropical oceans than the *in situ*-only field RMSD (Figs 12f–j), which is consistent with the results shown in Figs 11c–f.

4.2 *Explaining patterns of RMSD toward independent observations*

For the synthetic field, the rationale behind its RMSD feature is similar to the case of assimilated observations since it is synthesized in the model space, not in the observational space.

The *in situ*-only field has an increased RMSD on independent observations (Figs 12f–j) than on assimilated observations (Figs 5f–j) since few independent observations fed into the OI. In addition, the differentiation between the amplitudes of increased RMSD in the tropical oceans and in the extratropical oceans (i.e., the RMSD increases moderately in the tropics but considerably for the extratropics) can be explained by the global feature of covariance correlation scales. Specifically, in the tropics (extratropics), the correlation scales are larger (smaller), and thus, more (fewer) observations are fed into the OI calculation. This differentiation, however, is attenuated by the dominant influence of observations (rather than that of backgrounds or correlation scales) on assimilated observations.

For the combined field, the p_{comb} (Figs 12k–o) is significantly smaller than p_{insi} (Figs 12f–j) only in the extratropics, suggesting that the synthetic field, as a background, is much more advantageous over the climate field in the extratropics where covariance scales are relatively small; thus, the analysis is affected by fewer *in situ* observations.

5 Conclusions and discussion

Most algorithms for data assimilation are based on the theory of least-variance in linear statistical estimation. This theory requires the a priori specification of the expectation and covariance statistics of the errors that affect the data to be assimilated. However, those statistical quantities are generally poorly known, and determining them is one of the major problems in data assimilation.

In oceanography applications, the data to be assimilated usually consist of both satellite and Argo data. They feed into the assimilation system as background estimates and additional observations, respectively, in a “two-step” framework. This framework first performs vertical projection with satellite measurements and then a correction with *in situ* observations. Therefore, it is important to evaluate the contribution of each type of data to reconstruction analysis to better configure the assimilation parameters in the future. This is accomplished by finding the true background and observational errors using some diagnostics.

To achieve these goals, two comparative schemes were designed, with the background in OI set to the climate field and the synthetic field (derived from only satellite measurements). The resulting analysis fields are referred to as the *in situ*-only field (derived from only *in situ* observations) and the combined field (derived from both *in situ* observations and satellite measurements), respectively.

Firstly, consistency diagnostics were fulfilled in the observational space (Desroziers et al., 2005). This a posterior diagnosis method was performed iteratively to optimize the background error and observation error covariances in the observation space. For the static scheme, the predefined ocean variance was expanded by 1.8 times to compensate for the impact of insufficient *in situ* observations. For the dynamic scheme, the predefined ocean variance was flattened slightly to weaken the impact of the discrepancies in the time spans of *in situ* observations and the *synthetic* field.

Then, a comparison was made between the reconstructed temperature analyses (synthetic field, *in situ*-only field, and combined field) toward assimilated/independent validation to show a detailed RMSD pattern. The pattern includes horizontal maps and vertical profiles of RMSD concerning different latitudes (globally, tropics, extratropics) and depths (shallow, deep). The main conclusions are as follows:

(1) In all scenarios, the synthetic field can always reconstruct considerable temperature signals. The addition of satellite data can reduce the RMSD by up to 25%, and thus, the combined field has minimal RMSD among all the fields. Intuitively, satellite data improve the analysis mainly around positions where *in situ* observations are sparse.

(2) The advantage of satellite data over *in situ* data varies with ocean area. Generally, in extratropical oceans, analysis using only satellite data performs better than analysis using only *in situ* data, with RMSD reduced by 10%–25%. By contrast, in tropical oceans analysis using only satellite data performs worse than using only *in situ* data, with RMSD increasing by 0%–10%. Thus, satellite data improve the analysis mainly in extratropical oceans.

(3) The synthetic field and *in situ*-only field have respective error features on each layer. At the sea surface and in the mixed layer, both have smaller errors in tropical oceans. In deep waters, the former has smaller errors around the Northwest Pacific and other oceans where western boundary flows intend to prevail; in contrast, the latter has smaller errors around the North Pacific

and other oceans at slightly higher latitudes.

Finally, the rationale behind the distribution of RMSD was discussed using quantities based on the diagnostic background and observation errors in the observation space, along with some diagnostics in the model space. Some physical oceanography concepts were introduced to explain these error characteristics.

(1) The synthetic field has a smaller RMSD within the global mixed layer and midlatitude deep waters, as in the northwest Pacific Ocean. This is controlled by the explained variance of the vertical projection, i.e., the correlation between the sea surface elements and the interior ocean temperature. This is associated with two marine phenomena. First, the mixed layer, with a shape of "shallow in tropics and deep in extratropics", is mostly confined to the upper 100 m in the tropics (An et al., 2012.). Therefore, within this layer, the internal temperature anomaly (δT) corresponding to a certain SST anomaly (δSST), i.e., β, remains uniformly high. Second, the ocean baroclinic or steric variably is stronger through the water columns of some subtropical and midlatitude oceans where interior circulations prevail. They include the North Pacific interior gyre circulation, the South Pacific and South Indian gyres, and the "Tasman leakage" connection south of Australia between the subtropical South Pacific and Indian gyres (Roemmich and Gilson, 2009). Therefore, over these ocean areas, the internal temperature anomaly (δT) corresponding to a certain SLA anomaly (δSLA), i.e., α remains high.

(2) The *in situ*-only field (on assimilated observations) has a smaller RMSD in the tropical upper layer and at midlatitudes. This is determined by the actual NSR of ocean temperature, i.e., the diagnostic observation error divided by background error, which is consistent with previous studies. For example, von Schuckmann et al. (2009) proposed that generally, the signal-to-noise ratio (SNR) is high in the upper layer of tropical basins and

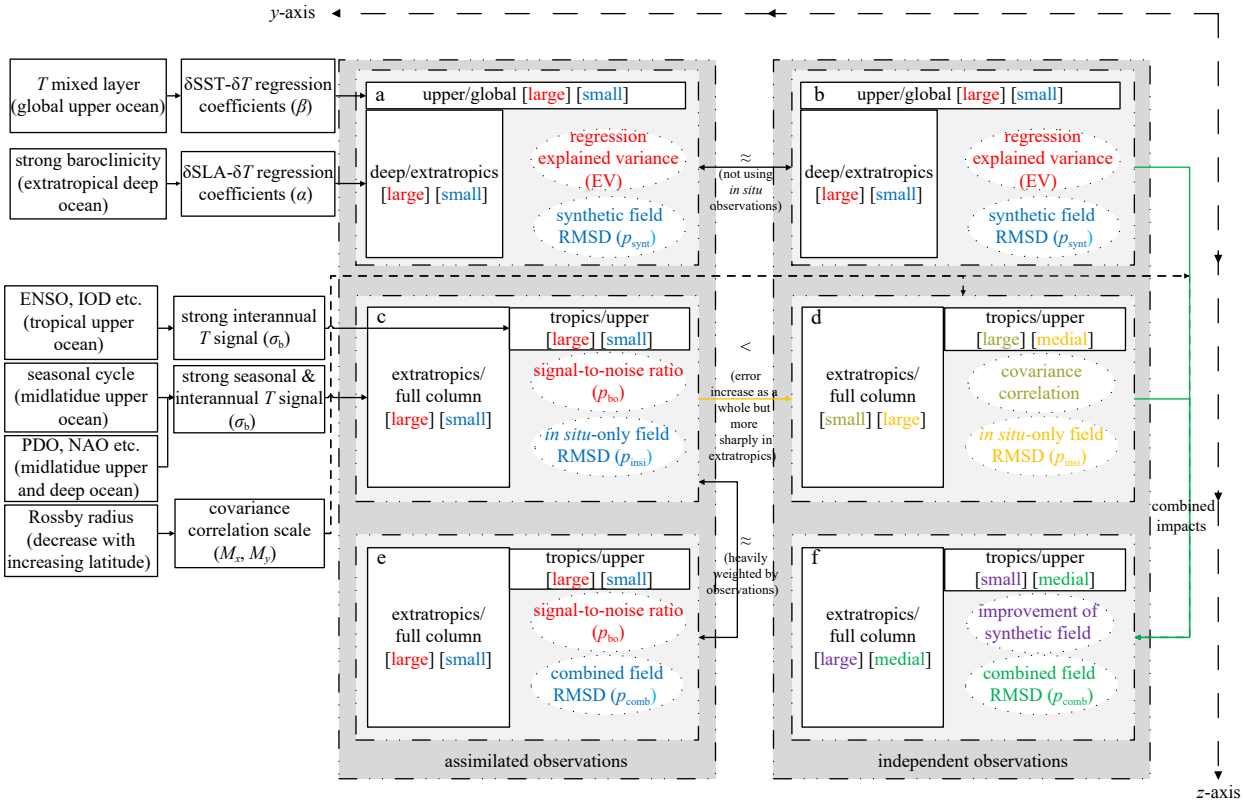

Fig. 13. A summary of the distribution and rationale of RMSD for the synthetic field (a and b), *in situ*-only field (c and d), and combined field (e and f) toward the assimilated observations (left column) and independent observations (right column). The *y*-axis denotes the increase in latitude (from the tropics to the extratropics); the *z*-axis denotes the increase in depth (from the upper ocean to the deep ocean). The blue (red) text denotes that the quantities are smaller (larger) in the corresponding regions (i.e., latitude or depth) than in other regions; the brown text denotes the differences between the RMSD toward independent observations and the RMSD toward assimilated observations due to factors denoted by the yellow text; the green text denotes the differences between the RMSD of the combined field and the RMSD of the *in situ*-only field toward independent observations due to factors denoted by the purple text. Presented in the left boxes are some ocean phenomena possibly accounting for these RMSD features. a, b. The synthetic field toward both assimilated and independent observations: the RMSD is smaller within the global mixed layer and the midlatitude deep waters, as in the northwest Pacific Ocean, due to the larger explained variance of the vertical regression there. c. *In situ*-only field toward assimilated observations: the RMSD is smaller in the tropical upper layer and at midlatitudes due to the larger actual NSR of ocean temperature there. d. The *in situ*-only field toward independent observations: the increase amplitudes of RMSD (relative to that toward assimilated observations) is moderate in the tropics but considerable in the extratropics due to larger covariance correlation scales in the tropics. e. The combined field toward assimilated observations: the RMSD is roughly equal to that of the *in situ*-only field toward assimilated observations due to the dominant role of observations at these points. f. The combined field toward independent observations: the decrease amplitudes of RMSD (relative to that of the *in situ*-only field toward independent observations) is small in the tropics but large in the extratropics due to more notable improvement from the synthetic field in the extratropics.

at midlatitudes where seasonal to interannual ocean variability appears to be large. The Reynolds L4 SST product (Reynolds et al., 2007) applies a similar OI scheme, so its SST error distribution also reflects these characteristics. However, the NSR issue is much more complicated than discussed here. The variances calculated herein include ocean temperature variability on different timescales, which involve different latitudes and depths. For example, amplitudes of the annual harmonic of temperature, mainly triggered by seasonal variations in surface heating, are maximal in the surface layer at midlatitudes, where amplitudes decrease with depth but still exceed 1°C above 100 m and the explained variance accounts for 80% of the total variance (Antonov et al., 2004). Interannual fluctuations occur in several latitudinal bands of the global ocean, i.e., at midlatitudes between 30°N/S and 60°N/S, in the subtropics between 20°N/S and 35°N/S, and in the tropics between 20°S and 20°N. At midlatitudes, temperature is surface intensified and reaches down to more than 1 000 m; deep reaching interannual fluctuations also characterize the subtropical basins with maxima at the surface layer and ~500 m; in the tropical basins, dominant anomalies are mostly confined to the upper 500 m. These ocean fluctuations are often controlled by changes in modes of the atmospheric circulation. They include, e.g., the North Atlantic Oscillation (NAO) (Eden and Willebrand, 2001), the Pacific Decadal Oscillation (PDO) (Harrison and Carson, 2007), the El Niño Southern Oscillation (ENSO) in the upper layer of the tropical basin, the Indian Ocean Dipole (IOD) (Saji et al., 1999), and the two primary modes in the tropical Atlantic (Servain et al., 1999).

(3) On the one hand, the covariance correlation scale decreases as the latitude increases. This inversion results in differentiation between the tropics (increased moderately) and the extratropics (increased considerably) in terms of increased amplitudes of *in situ*-only field RMSD on independent observations than on assimilated observations. On the other hand, the synthetic background has a smaller RMSD than the climate background. This results in differentiation between the combined field (smaller RMSD) and the *in situ*-only field (larger RMSD) in terms of the RMSD toward extratropical independent observations. Both differentiations are visible by comparing Figs 4 and 11 or Figs 5 and 12 and explain why satellite measurements mainly improve the analysis toward independent observations in the extratropics. That is, only when the *in situ* observations do not play a dominant role (on independent observations and with smaller covariance correlation scales) can the synthetic field significantly influence the analysis. This influence is different at different latitudes since the synthetic field is more (less) accurate in the extratropics (tropics).

All the rationales mentioned and theories involved are summarized in Fig. 13 (Note: SNR is used instead of the NSR, to be consistent with EV in the increasing direction).

Therefore, in global 3-D ocean temperature reconstruction, satellite measurements and *in situ* observations have their respective dominant areas. This means that some of the preprocessing methods described herein are limited. For example, covariance matrix optimization was carried out by adjusting the averaged weight coefficient and expansion coefficient at all levels and over global oceans but did not consider the differences in NSR at different latitudes and depths. Neither did the setting of covariance correlation scales. In addition, it must be stressed that the analyses above were constrained to the OI framework here. There are other dynamical methods (e.g., 3DVAR) that can also consider satellites and Argo simultaneously through dynamical propagation. Additional work is required to investigate whether the conclusions here apply to other methods as in the OI framework. The aim of this study is to refine the characteristics of satellite and *in situ* measurements in terms of the baroclinicity, NSR, and correlation scale, from the surface to the depth and from low to high latitudes. This is the first step in making reasonable adjustments to assimilation systems. For example, from the conclusion above, it's known that the predefined NSR should be reduced in the tropical upper oceans and midlatitude oceans but increased in the tropical interior oceans; vertical regression methods should be further optimized, particularly at low latitudes. Due to length limitations, this article did not discuss other issues, such as the effective resolution of the analysis. On the one hand, the use of high-resolution satellite measurements can reduce the aliasing induced by unresolved mesoscale variability. This is a serious problem in assimilating Argo observations by reconstructing the large-scale and low-frequency part of the fields (Guinehut et al., 2004). On the other hand, the ability of satellite measurements to reconstruct mesoscale or even small-scale fields is also worth discussing (Nardelli et al., 2012).

References

Agarwal N, Sharma R, Basu S, et al. 2007. Derivation of salinity profiles in the Indian Ocean from satellite surface observations. IEEE Transactions on Geoscience and Remote Sensing, 4(2): 322–325, doi: 10.1109/LGRS.2007.894163

Alves J O S, Haines K, Anderson D L T. 2001. Sea level assimilation experiments in the tropical Pacific. Journal of Physical Oceanography, 31(2): 305–323, doi: 10.1175/1520-0485(2001)031<0305:SLAEIT>2.0.CO;2

An Yuzhu, Zhang Ren, Wang Huizan, et al. 2012. Study on calculation and spatio-temporal variations of global ocean mixed layer depth. Chinese Journal of Geophysics, 55(7): 2249–2258

Antonov J I, Levitus S, Boyer T P. 2004. Climatological annual cycle of ocean heat content. Geophysical Research Letters, 31(4): L04304, doi: 10.1029/2003GL018851

Ballabrera-Poy J, Mourre B, Garcia-Ladona E, et al. 2009. Linear and non-linear T–S models for the eastern North Atlantic from Argo data: Role of surface salinity observations. Deep-Sea Research I, 56: 1605–1614, doi: 10.1016/j.dsr.2009.05.017

Carnes M R, Teague W J, Mitchell J L, 1994. Inference of subsurface thermohaline structure from fields measurable by satellite. Journal of Atmospheric and Oceanic Technology, 11(2): 551–566

Chapnik B, Desroziers G, Rabier F, et al. 2004. Properties and first application of an error-statistics tuning method in variational assimilation. Quarterly Journal of the Royal Meteorological Society, 130: 2253–2275, doi: 10.1256/qj.03.26

Chapnik B, Desroziers G, Rabier F, et al. 2006. Diagnosis and tuning of observational error in a quasi-operational data assimilation setting. Quarterly Journal of the Royal Meteorological Society, 132(615): 543–565, doi: 10.1256/qj.04.102

Cooper M, Haines K. 1996. Altimetric assimilation with water property conservation. Journal of Geophysical Research: Oceans, 101(C1): 1059–1077, doi: 10.1029/95JC02902

Dee D P, da Silva A M. 1999. Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. Monthly Weather Review, 127(8): 1822–1834, doi: 10.1175/1520-0493(1999)127<1822:MLEOFA>2.0.CO;2

Desroziers G, Berre L, Chapnik B, et al. 2005. Diagnosis of observation, background and analysis-error statistics in observation space. Quarterly Journal of the Royal Meteorological Society, 131(613): 3385–3396, doi: 10.1256/qj.05.108

Desroziers G, Ivanov S. 2001. Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. Quarterly Journal of the Royal Meteorological Society, 127(574): 1433–1452, doi: 10.1002/qj.49712757417

Eden C, Willebrand J. 2001. Mechanism of interannual to decadal variability of the North Atlantic circulation. Journal of Climate,

14(10): 2266–2280, doi: 10.1175/1520-0442(2001)014<2266:MOITDV>2.0.CO;2

Fox D N, Teague W J, Barron C N, et al. 2002. The modular ocean data assimilation system (MODAS). Journal of Atmospheric and Oceanic Technology, 19(2): 240–252, doi: 10.1175/1520-0426(2002)019<0240:TMODAS>2.0.CO;2

Gaillard F. 2012. ISAS-Tool version 6: Method and Configuration. Plouzané, France: Ifremer Cnrs Ird Ubo

Gaillard F, Autret E, Thierry V, et al. 2009. Quality control of large Argo datasets. Journal of Atmospheric and Oceanic Technology, 26(2): 337–351, doi: 10.1175/2008JTECHO552.1

Gaillard F, Reynaud T, Thierry V, et al. 2016. In situ-based reanalysis of the global ocean temperature and salinity with ISAS: variability of the heat content and steric height. Journal of Climate, 29(4): 1305–1323, doi: 10.1175/JCLI-D-15-0028.1

Guinehut S, Dhomps A L, Larnicol G, et al. 2012. High resolution 3-D temperature and salinity fields derived from *in situ* and satellite observations. Ocean Science, 8(5): 845–857, doi: 10.5194/os-8-845-2012

Guinehut S, Le Traon P Y, Larnicol G, et al. 2004. Combining Argo and remote-sensing data to estimate the ocean three-dimensional temperature fields—A first approach based on simulated observations. Journal of Marine Systems, 46(1–4): 85–98

Guinehut S, Le Traon P Y, Larnicol G, et al. 2006. What can we learn from Global Altimetry/Hydrography comparisons?. Geophysical Research Letters, 33(10): L10604

Harrison D E, Carson M. 2007. Is the world ocean warming? Upper-ocean temperature trends: 1950–2000. Journal of Physical Oceanography, 37(2): 174–187, doi: 10.1175/JPO3005.1

Hollingsworth A, Lönnberg P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus A, 38(2): 111–136, doi: 10.3402/tellusa.v38i2.11707

Larnicol G, Guinehut S, Rio M H, et al. 2006. The global observed ocean products of the French Mercator project. In: Proceedings of the Symposium on 15 Years of Progress in Radar Altimetry. Noordwijk, Netherlands: ESA

Lellouche J M, Greiner E, Le Galloudec O, et al. 2018. Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1/12° high-resolution system. Ocean Science, 14(5): 1093–1126, doi: 10.5194/os-14-1093-2018

Lellouche J M, Le Galloudec O, Drévillon M, et al. 2013. Evaluation of global monitoring and forecasting systems at Mercator Océan. Ocean Science, 9(1): 57–81, doi: 10.5194/os-9-57-2013

Maes C, Behringer D. 2000. Using satellite–derived sea level and temperature profiles for determining the salinity variability: A new approach. Journal of Geophysical Research: Oceans, 105(C4): 8537–8547, doi: 10.1029/1999JC900279

Meinen C S, Watts D R. 2000. Vertical structure and transport on a transect across the North Atlantic Current near 42 N: Time series and mean. Journal of Geophysical Research: Oceans, 105(C9): 21869–21891, doi: 10.1029/2000JC900097

Mitchell D, Wimbush M, Watts D, et al. 2004. The residual GEM technique and its application to the southwestern Japan East Sea. Journal of Atmospheric and Oceanic Technology, 21(12): 1895–1909, doi: 10.1175/JTECH-1668.1

Nardelli B B, Guinehut S, Pascual A, et al. 2012. Towards high resolution mapping of 3-D mesoscale dynamics from observations. Ocean Science, 8(5): 885–901, doi: 10.5194/os-8-885-2012

Nardelli B B, Santoleri R. 2004. Reconstructing synthetic profiles from surface data. Journal of Atmospheric and Oceanic Technology, 21(4): 693–703, doi: 10.1175/1520-0426(2004)021<0693:RSPFSD>2.0.CO;2

Nardelli B B, Santoleri R. 2005. Methods for the reconstruction of vertical profiles from surface data: Multivariate analyses, residual GEM, and variable temporal signals in the North Pacific Ocean. Journal of Atmospheric and Oceanic Technology, 22(11): 1762–1781, doi: 10.1175/JTECH1792.1

Reynolds R W, Smith T M, Liu Chunying, et al. 2007. Daily high-resolution-blended analyses for sea surface temperature. Journal of Climate, 20(22): 5473–5496, doi: 10.1175/2007JCLI1824.1

Roemmich D, Gilson J. 2009. The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. Progress in Oceanography, 82(2): 81–100, doi: 10.1016/j.pocean.2009.03.004

Saji N H, Goswami B N, Vinayachandran P N, et al. 1999. A dipole mode in the tropical Indian Ocean. Nature, 401(6751): 360–363

Servain J, Wainer I, McCreary J P, et al. 1999. Relationship between the equatorial and meridional modes of climatic variability in the tropical Atlantic. Geophysical Research Letters, 26(4): 485–488, doi: 10.1029/1999GL900014

Talagrand O. 1999. A posteriori verification of analysis and assimilation algorithms. In: Proceedings of the ECMWF Workshop on Diagnosis of Data Assimilation Systems. Reading, UK: European Centre for Medium-Range Weather Forecasts

Von Schuckmann K, Gaillard F, Le Traon P Y. 2009. Global hydrographic variability patterns during 2003–2008. Journal of Geophysical Research, 114(C9): C09007

Wu X, Yan X H, Jo Y H, et al. 2012. Estimation of Subsurface Temperature Anomaly in the North Atlantic Using a Self-Organizing Map Neural Network. Journal of Atmospheric and Oceanic Technology, 29(11): 1675–1688, doi: 10.1175/JTECH-D-12-00013.1