Acta Oceanol. Sin., 2022, Vol. 41, No. 2, P. 41–50
https://doi.org/10.1007/s13131-021-1880-5
http://www.aosocean.com
E-mail: ocean2@hyxb.org.cn

Bias correction of sea surface temperature retrospective forecasts in the South China Sea

Guijun Han[1], Jianfeng Zhou[1], Qi Shao[1, 2], Wei Li[1, 2*], Chaoliang Li[1*], Xiaobo Wu[1], Lige Cao[1], Haowen Wu[1], Yundong Li[1], Gongfu Zhou[1]

[1] School of Marine Science and Technology, Tianjin University, Tianjin 300072, China

[2] Tianjin Key Laboratory for Oceanic Meteorology, Tianjin 300074, China

Received 21 May 2021; accepted 12 June 2021

Abstract

Offline bias correction of numerical marine forecast products is an effective post-processing means to improve forecast accuracy. Two offline bias correction methods for sea surface temperature (SST) forecasts have been developed in this study: a backpropagation neural network (BPNN) algorithm, and a hybrid algorithm of empirical orthogonal function (EOF) analysis and BPNN (named EOF-BPNN). The performances of these two methods are validated using bias correction experiments implemented in the South China Sea (SCS), in which the target dataset is a six-year (2003–2008) daily mean time series of SST retrospective forecasts for one-day in advance, obtained from a regional ocean forecast and analysis system called the China Ocean Reanalysis (CORA), and the reference time series is the gridded satellite-based SST. The bias-correction results show that the two methods have similar good skills; however, the EOF-BPNN method is more than five times faster than the BPNN method. Before applying the bias correction, the basin-wide climatological error of the daily mean CORA SST retrospective forecasts in the SCS is up to −3°C; now, it is minimized substantially, falling within the error range (±0.5°C) of the satellite SST data.

Key words: sea surface temperature retrospective forecasts, bias correction, backpropagation neural network, empirical orthogonal function analysis, South China Sea

Citation: Han Guijun, Zhou Jianfeng, Shao Qi, Li Wei, Li Chaoliang, Wu Xiaobo, Cao Lige, Wu Haowen, Li Yundong, Zhou Gongfu. 2022. Bias correction of sea surface temperature retrospective forecasts in the South China Sea. Acta Oceanologica Sinica, 41(2): 41–50, doi: 10.1007/s13131-021-1880-5

1 Introduction

Like numerical weather forecasts, numerical ocean forecasts are becoming more and more common. This is in part due to significant advances in the development of real-time ocean observing system, the ability of modeling the ocean and assimilating ocean data, and availability of high performance computing resources. A common problem, exactly as facing weather forecasters today, is the presence of systematic model errors, referred to as the biases (Dalcher and Kalnay, 1987), resulted from deficiencies of ocean models themselves. As a result, two kinds of bias correction techniques used to deal with the numerical weather forecasting errors are considered by marine forecasters: one is offline bias correction, and the other is online bias correction (Danforth et al., 2007). Offline bias correction is a typical postprocedure for the output of an operational numerical weather prediction model by accounting for the model bias derived from the statistics of short-term forecast errors with respect to a reference time series. This technique can trace its roots all the way back to the approaches of "perfect prog" (PP) (Klein et al., 1959; Klein, 1971) and model output statistics (MOS) (Glahn and Lowry, 1972; Carter et al., 1989). Note that such offline bias correction is empirical, and has no dynamic effect on the forecasts; in contrast, online bias correction is more robust for being state-dependent and made along with model integration. For instance, data assimilation in the presence of forecast bias (Dee and Da Silva, 1998) belongs to the category of online bias correction. However, online bias correction algorithms are usually computationally complicated and expensive. Offline bias correction is particularly valuable for being independent of the model and easy to implement so that up to now it is still of great interest in both research and practice (Bhargava et al., 2018; Chang et al., 2019).

Sea surface temperature (SST) is not only used as the main index describing the thermal condition of the ocean surface but also an important indicator of ocean primary productivity, and plays a key role in the interaction between the ocean and the atmosphere. Therefore, it is of great scientific significance and practical application prospect to correct SST forecast bias in numerical weather prediction, ocean forecasting, and climate research (Kug et al., 2008; Ashfaq et al., 2011; LaRow, 2013; Abhilash et al., 2014; Narapusetty et al., 2014; Vitart and Balmaseda, 2018; Hernández-Díaz et al., 2019; Voldoire et al., 2019). This study is motivated by recent extensive applications of various types of artificial neural networks in areas where more reliable reference data from meteorologic and oceanographic observations and reanalyses are available. Two methods employing artificial neural network trained by the traditional backpropagation (BP) algorithm are proposed as post-processing solutions for

Foundation item: The National Key Research and Development Program of China under contract No. 2018YFC1406206; the National Natural Science Foundation of China under contract No. 41876014.
*Corresponding author, E-mail: liwei1978@tju.edu.cn; chaoliang_li0608@163.com

minimizing biases of SST forecasts in the South China Sea (SCS). As a data-driven method, the accessibility of the time series of gridded satellite-based SST dataset is of great favor in the implementation of these two methods for bias correction of SST forecasts. The target dataset of SST forecasts in the SCS to be corrected is taken from an archived retrospective forecast made by a regional ocean forecasting system for the coastal waters of China and its adjacent seas.

The rest of the paper is organized as follows. In Section 2, the data and methods used in this study are introduced. The results of the bias correction experiments are presented in Section 3, followed by a brief summary and discussion in Section 4.

2 Data and methods

2.1 *Data collation*

The study area is the SCS (2°–24°N, 100°–122°E; Fig. 1). Note that since the technique of empirical orthogonal function (EOF) analysis used in building the following bias-correction algorithm has good performance for enclosed basins, the Sulu Sea, which may degrade the results to some extent for the SCS region as a whole, is excluded in this study.

2.1.1 *Forecasts*

Daily mean SST forecasts in the SCS with a spatial resolution of (1/4)° used in this paper for bias verification and correction in the bias correction experiments are retrospective forecasts for one-day in advance produced by the regional China Ocean Reanalysis (CORA) system (Han et al., 2011, 2013). The CORA system covering the coastal waters of China and its adjacent seas is developed based on the parallelized Princeton Ocean Model (POM; Mellor et al., 2002); it uses a generalized coordinate system and a sequential three-dimensional variational scheme implemented within a multigrid framework (Li et al., 2008, 2010). The ocean observations being incorporated into the CORA system consist of the NOAA (1/4)° daily optimum interpolation SST (OISST) v2.0 (Reynolds et al., 2007) from the advanced very high resolution radiometer (AVHRR), which is obtained from the National Centers for Environmental Information (NCEI) (https://www.ncei.noaa.gov/data/sea-surface-temperature-optimum-interpolation), altimetry sea-surface height anomaly (SSHA) from the Archiving Validation and Interpretation of Satellite Oceanographic (AVISO) (http://www.jason.oceanobs.com), and temperature/salinity profiles from the World Ocean Database 2009 (WOD09) (https://www.nodc.noaa.gov/OC5/WOD09/pr_wod09.html). Evaluations of the CORA product (Wu et al., 2013; Zhang et al., 2016; Chao et al., 2020), especially in the SCS region (Xing et al., 2012; Chen et al., 2013; Fan et al., 2020), show its capability in reconstructing historical ocean states in the coastal waters of China and its adjacent seas. Note that the data assimilation scheme in the CORA system has been changed to produce retrospective forecasts for the year of 2009 and beyond (Chao et al., 2020). Therefore, for data consistency and considering the time span of the gridded satellite SST dataset described below, the daily mean SST retrospective forecasts employed in this study cover a period of six years from January 1, 2003 through December 31, 2008. This daily mean SST retrospective forecast dataset will be called the CORA SST from now on.

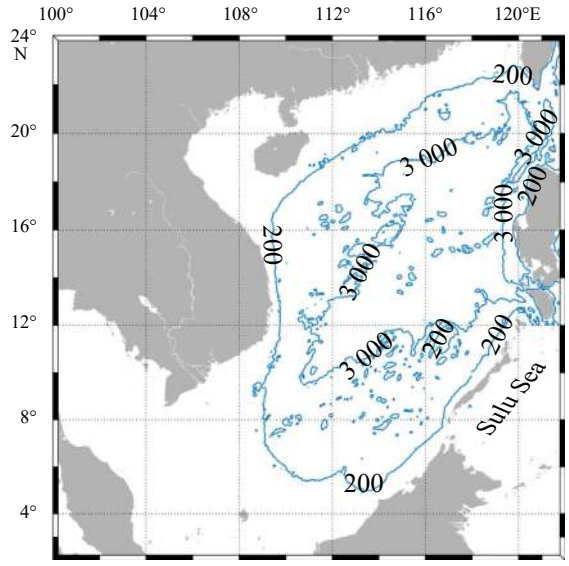

Fig. 1. Study area: the South China Sea (2°–24°N, 100°–122°E) with the Sulu Sea excluded. The blue contours are isobaths (unit: m).

2.1.2 *Observations*

The SST observational dataset, employed as a baseline for bias correction of the CORA SST, is the NOAA (1/4)° daily OISST v2.0 (Reynolds et al., 2007) from the AVHRR-AMSR with observations near land coming from the AVHRR, and measurements over the open ocean from the advanced microwave scanning radiometer (AMSR) on the Earth Observing System. The AVHRR-AMSR SST time series, available during 1 June 2002 through 4 October 2011 from the NCEI (https://www.ncei.noaa.gov/data/sea-surface-temperature-optimum-interpolation), is an analysis constructed by adjusting biases of satellite and ship observations referenced to buoys.

The *in-situ* SST observations used in this research come from the EN4.2.1 dataset of the Met Office Hadley Centre (https://www.metoffice.gov.uk/hadobs/en4/, hereinafter EN4). The World Ocean Database (WOD) dataset of the NCEI is the main source used for constructing the EN4 (Good et al., 2013). According to Good et al. (2013), quality control has been applied in constructing the EN4 data, especially for bias correction in expendable bathythermograph and mechanical bathythermograph profiles.

2.2 *Methodology*

The bias correction technique applied in this work is based on the widely used neural network (NN) trained by employing the traditional BP algorithm (Rumelhart et al., 1986), called the backpropagation neural network (BPNN). The CORA SST bias is corrected by the BPNN method on each (1/4)° grid in the SCS region. Thus, there are 4 534 grid points in total, which need to be done one by one. In the meantime, hybridization of the BPNN algorithm with the EOF analysis, called the EOF-BPNN, is introduced to consider the relationship between spatial grids of the data ignored by the BPNN method when implementing the bias correction on each (1/4)° grid alone.

The BPNN method, consisting of an input layer, a hidden layer, and an output layer, and with all these layers connected, is trained in the supervised learning manner. In this study, the original time series of six-year dataset of biases in the CORA SST with respect to the AVHRR-AMSR SST in the SCS is split into training and testing subsets: the first five-year (from January 1, 2003 to December 31, 2007) dataset is used as the samples for training the neural networks in both methods, and the last year (2008) dataset is used as the testing samples to validate effects of

the bias correction on the CORA SST. Using the training samples, the nonlinear relationship between inputs and outputs can be established by minimizing the cost function in the BPNN algorithm, i.e., mean squared error measuring the distance between the prediction and target output. The Adam optimization algorithm (Kingma and Ba, 2015) and the Leaky ReLU activation function (Maas et al., 2013) are put to use respectively. To avoid overfitting and improving the learning speed of the neural networks, batch normalization algorithm is used to normalize the training samples (Boutilier et al., 2006).

The EOF analysis in the EOF-BPNN method is first used to transform the spatial-temporal dataset of biases between the CORA and AVHRR-AMSR SSTs to spatial patterns of variability (called EOFs) and temporal projections of these patterns (aka, principle components, or PCs); then, the same algorithm in the BPNN method is used to deal with these PCs. It is anticipated that the direct benefit of doing so is to make the EOF-BPNN method, a time-saving technique compared to performing the BPNN method alone on every (1/4)° grid over the entire study region. The North significance test method (North et al., 1982) is adopted to identify whether these EOFs are statistically significant.

It is worth noting that the reason for using five-year long dataset to train the neural networks mentioned above is that the EOF analysis employed in the EOF-BPNN method assumes stationarity of the data. Under the assumption of stationarity, the length of time series data used in the EOF analysis should be as long as possible. Through trial-and-error, it is found that the employment of five-year long time series data in the EOF-BPNN method is appropriate for this research. The same strategy of data partition is thus used in both BPNN and EOF-BPNN methods to facilitate the comparison of the two methods.

3 Results

By applying the EOF analysis to the first five-year dataset of biases in the CORA SST, the EOFs and their corresponding PCs are obtained. Table 1 shows the cumulative fractions of the total variance accounted for by the EOFs.

The first 803 modes, accounting for approximately 99.9% of the total variance, are chosen to represent the overall signal in the bias correction experiment of applying the EOF-BPNN method. By doing so, the results of implementing bias correction experiments corresponding to these two methods show that the EOF-BPNN method has a significant advantage in terms of computational efficiency since it performs more than five times faster than the BPNN method.

Both methods of BPNN and EOF-BPNN yield results that are close to each other (Tables 2 and 3). Hence, in what follows, only those figures revealing the performance of the EOF-BPNN method are demonstrated. Bias, mean absolute error (MAE), and root mean squared error (RMSE) are used as metrics to measure retrospective forecast accuracy of the daily mean CORA SST before and after bias correction.

Table 1. Cumulative proportions of the EOFs

EOFs	Cumulative proportion/%
1	18
6	50
10	60
18	70
35	80
77	90
1 826	100

Table 2. Mean SSTs of the AVHRR-AMSR and CORA, and mean differences between the two datasets before and after applying the bias correction

	Uncorrected SST (ΔT)/°C	Bias-corrected SST (ΔT)/°C	
		BPNN	EOF-BPNN
AVHRR-AMSR	27.79	–	–
CORA	27.73 (–0.06)	27.78 (–0.01)	27.78 (–0.01)

Note: – represents no data.

Table 3. MAE and RMSE of the CORA SST with and without bias correction

		MAE/°C	RMSE/°C
Uncorrected		0.63	0.84
Bias-corrected	BPNN	0.23	0.32
	EOF-BPNN	0.22	0.30

3.1 *Bias correction of climatological field*

Figure 2a illustrates the spatial bias map obtained by time averaging the errors over the six-year daily mean retrospective SST forecasts taken from the CORA system. The climatological bias of the CORA SST with respect to the AVHRR-AMSR SST in the SCS is within the range of –3.02–1.04°C, with the mean basin-wide bias being –0.06°C (Table 2); and the biases around 82% of the basin are within the intervals of ±0.50°C (see the black contours in Fig. 2a), which is assumed to be the RMSE range of the satellite dataset according to previous studies in the East Asian marginal seas (0.5–0.7°C; Sakaida et al., 2000; Lee et al., 2005; Qiu et al., 2009). The region of the climatological cold or warm bias of the CORA SST is roughly bounded by the 200-m isobath (Fig. 1). In general, the cold bias is found along the coastal waters shallower than 200 m, while the warm bias is observed in the central SCS. Such spatial distribution characteristics of climatological bias in the CORA SST may be due to the following two facts. First, the AVHRR SST is cold with respect to the AVHRR-AMSR SST in the global ocean (Balmaseda et al., 2013), and the same is true in the SCS (Figs 3a and b). There is a cold bias of the high-resolution (1/4)° OISST product relative to the low-resolution (1°) product (Balmaseda et al., 2013). As noted by Qiu et al. (2009), the AVHRR SST has been validated to show a regional bias of about –0.4°C compared with independent *in-situ* SST observations from the drifting buoys in a two-year period from January 1, 2004 to December 31, 2005 in the northern SCS, which is confirmed by highly accurate *in-situ* SSTs from research vessels. Second, the initial fields of the CORA forecasts are formed by assimilating the (1/4)°-resolution AVHRR SST, the AVISO altimetry SSHA measurements in waters deeper than 200 m and *in-situ* temperature/salinity profiles (Han et al., 2013). Under such circumstances, the cold bias in the initial SST field in the central SCS (not the coastal regions) caused by assimilating the AVHRR SST could be compensated by assimilating the altimetry SSHA and *in-situ* temperature/salinity profiles (Figs 2a and 3a); and this leads to an initialization of SST in the right direction to give improved SST forecasts. However, the cold biases in the coastal waters shallower than 200 m remain. This statement can be confirmed by comparing the CORA SST with the EN4 data shown in Fig. 2c. The scatter plot in Fig. 2c displays monthly mean SST from the CORA data versus that from satellite (blue triangles) and EN4 (green plus signs) data. Although the EN4 data have an uneven temporal and spatial distribution (not shown), an agreement between the CORA and EN4 SSTs exists, and the mean basin-wide bias is –0.02°C. Figure 2c clearly shows that the majority of the blue triangles fall below the 45-degree reference line, which further stresses the fact that generally speak-

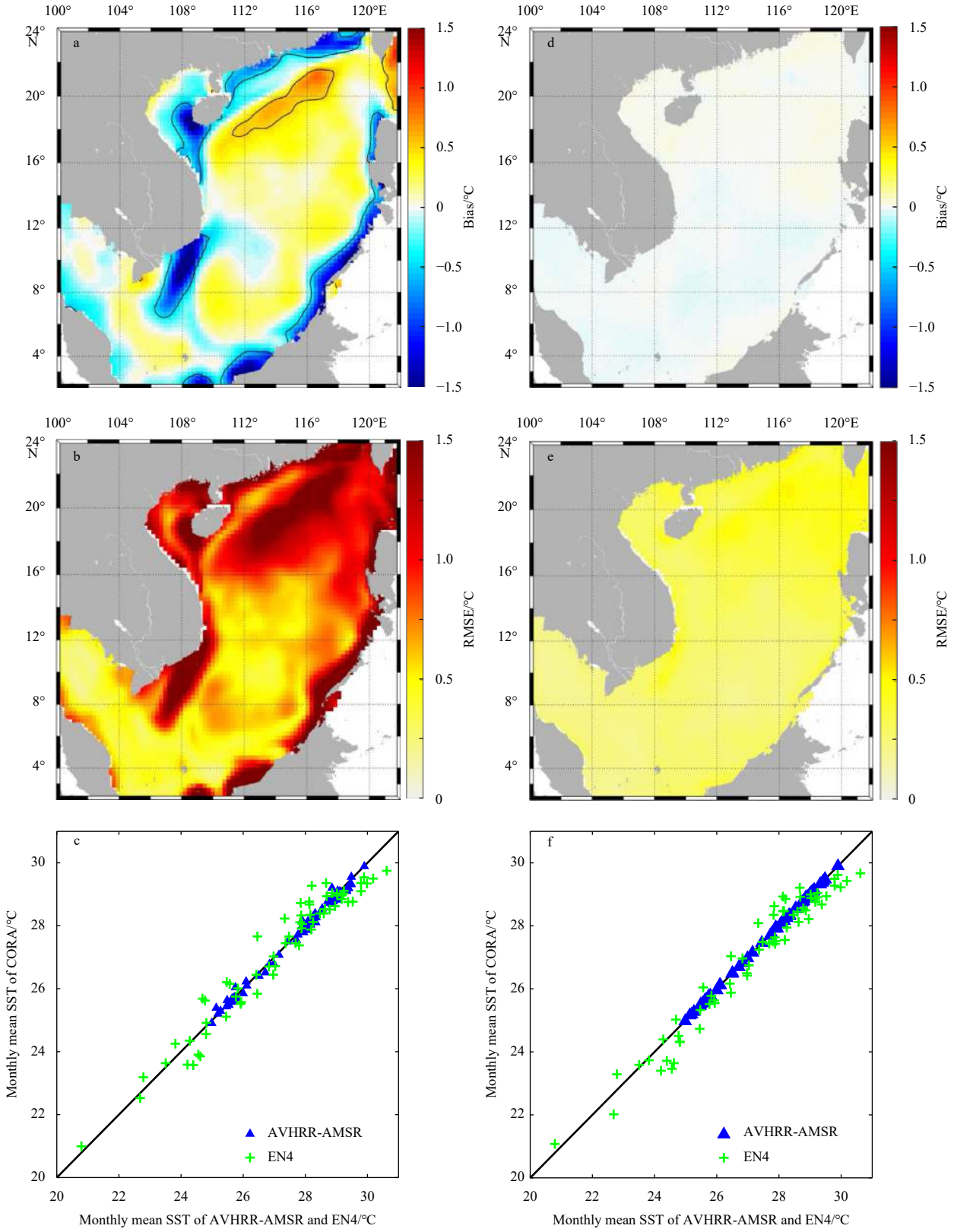

Fig. 2. Maps of climatological bias (a) and RMSE (b) of the uncorrected CORA SST with respect to the AVHRR-AMSR SST during 2003–2008; and scatter plot of monthly mean SST from the AVHRR-AMSR (blue triangles) and EN4 (green plus signs) dataset versus monthly mean SST from the uncorrected CORA dataset (c). Panels (d, e and f) are the corresponding ones for the bias-adjusted results. The black contours in a indicate the RMSE range (±0.50°C) of the satellite data. The 45-degree black line in c and f is a reference.

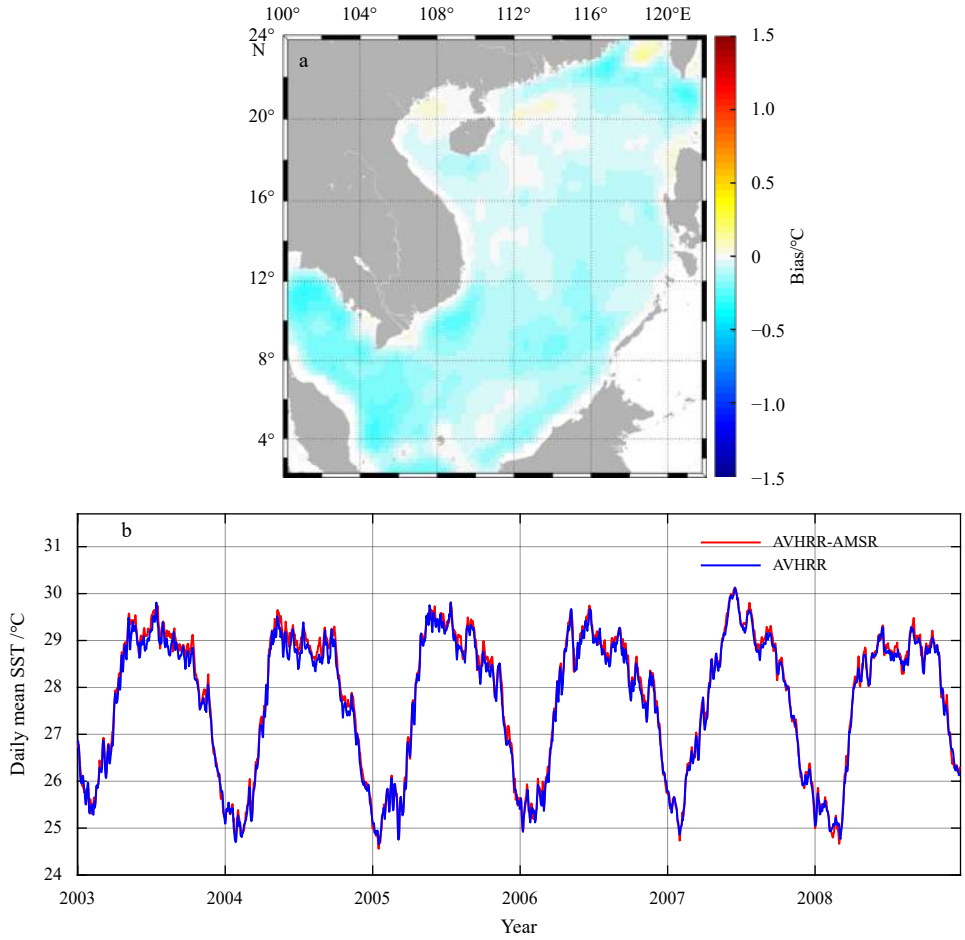

Fig. 3. Map of differences between the AVHRR and AVHRR-AMSR SSTs (a) and time series of daily averaged SST from these two products (b) during 2003–2008 in the SCS.

ing the CORA SST is lower than the AVHRR-AMSR SST in the SCS. Nevertheless, the gridded AVHRR-AMSR SST dataset, thanks to its availability, is used as the baseline for correcting the CORA SST bias to demonstrate performances of both BPNN and EOF-BPNN methods. After applying the bias correction, consistency between the CORA and AVHRR-AMSR SSTs is achieved, as shown by the blue triangles in Fig. 2f. In contrast, the basin-wide deviation (mean bias) between the CORA and EN4 SSTs has increased from –0.02°C to –0.19°C. It therefore reminds us that the quality of the satellite SST in the SCS could be improved by applying the regional algorithm, as suggested by Qiu et al. (2009).

In addition, the coastal zones to the southwest of Hainan Island and to the south of Vietnam show relatively large cold biases while the northern central SCS region and the Luzon Strait area show apparent warm biases outside the error limits of the satellite data. Further studies are required to demonstrate the causes of these large deviations by exploring, for example, details of ocean model behaviors and data assimilation schemes.

The globally averaged forecast RMSE of the uncorrected CORA SST is 0.84°C (Table 3), and it presents the same spatial pattern (Fig. 2b) with the climatological bias by showing relatively large errors in the northern SCS. After the bias correction, the spatial climatological bias (Fig. 2d) and RMSE (Fig. 2e) are globally reduced by falling inside the intervals of ±0.02°C and being less than 0.54°C, respectively. Tables 2 and 3 clearly indicate significant adjustments with respect to the mean basin-wide bias (decrease to –0.01°C) and RMSE (decrease to ~0.30°C) after applying the BPNN and EOF-BPNN methods, respectively. In the meantime, the basin-wide MAE is reduced from 0.63°C to ~0.20°C (Table 3).

3.2 *Bias correction for climatological mean annual and seasonal cycles*

Figures 4a and b show the climatological mean annual SST for the entire SCS using the observed and CORA data before and after the bias adjustment. Similar interannual variations are observed between the CORA SST (black dotted curve) and satellite SST (red solid curve), except for the abnormal values of the CORA SST in 2007 (Fig. 4a), and the CORA SST was negatively biased with a maximum of –0.11°C in 2006. Due to the sparse and uneven distribution of *in-situ* observations, no such simultaneous changes on interannual timescale are detected when comparing the CORA SST (black dotted curve in Fig. 4b) with the EN4 data (black solid curve in Fig. 4b); and comparatively speaking, the mean annual SST values of the CORA are closer to those of the EN4 than to those of the AVHRR-AMSR. The result of the EOF-BPNN method in Fig. 4b shows, as expected, the remarkable convergence of corrected CORA SST (blue solid curve) toward the satellite SST data during both the first five-year training period and the last year's testing period.

The climatological SST seasonal cycles from the observed and CORA SST data before and after the bias adjustment are presented in Figs 4c and d. The seasons of spring, summer, autumn, and winter in this research refer to those in the northern hemisphere;

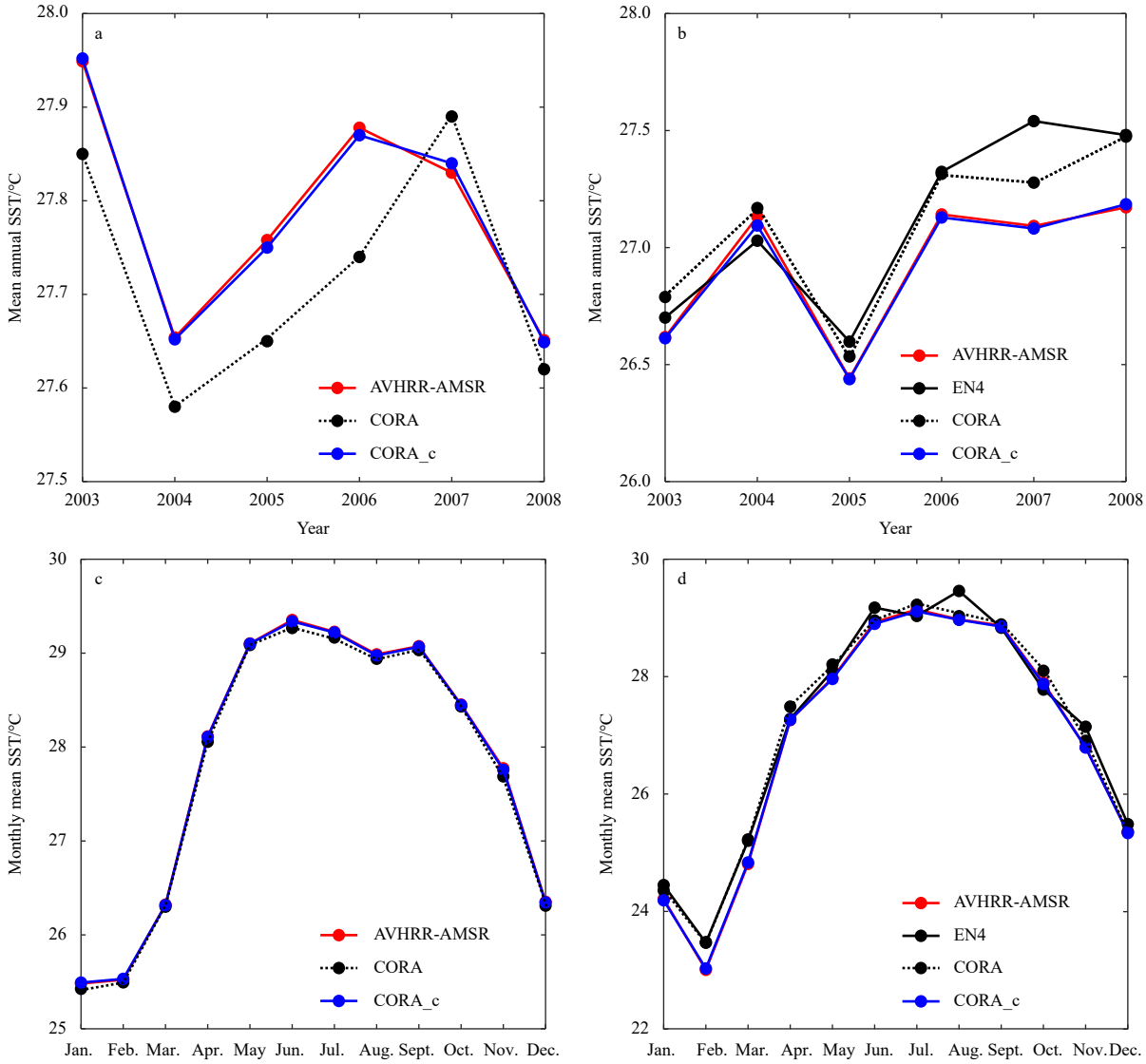

Fig. 4. Mean annual SST time series (a) and seasonal SST cycles (c) from the AVHRR-AMSR (red solid curve) and CORA data with (black dotted curve) and without (blue solid curve) bias correction. Panels b and d are the counterparts of a and c with the gridded AVHRR-AMSR and CORA SST being interpolated onto the observed points of the EN4 data (black solid curve).

summer, for example, refers to June, July and August (Xie et al., 2003). The CORA SST (black dotted curve in Fig. 4c) shows a cold bias (–0.20°C) against the satellite data (red solid curve in Fig. 4c) during the summer months of June and July. The deviations between the CORA SST (black dotted curve in Fig. 4d) and EN4 SST (black solid curve in Fig. 4d) are smaller during winter and early-spring (December, January, February, and March) than during the other months, with the largest warm bias in August. When interpolating the gridded CORA and AVHRR-AMSR data onto the observed points of the EN4 data, the climatological seasonal cycle (Fig. 4d) is roughly the same as those of the gridded CORA and satellite data (Fig. 4c), except that the seasonal minimum temperature occurs in February, instead in January. This may be caused by the spatiotemporal inhomogeneous distribution of the EN4 data. After applying the EOF-BPNN method to reduce the CORA SST bias, the adjusted CORA SST (see blue solid curves in Figs 4c and d) is in good agreement with the satellite data in both the training stage (2003–2007) and testing stage (2008).

Figures 5a–d show that the uncorrected CORA SST bias has a clear seasonality and spatial structure. With the same features in the spatial structure of cold and warm biases revealed in climatology of Fig. 2a, the forecasting SST bias evolves with season in strength and extent. The most striking feature is the strong cold and warm biases with values outside the error range of satellite SST indicated by black contours in the figures; and it is a research focus hereinafter. First, the summer season in Fig. 5b is examined since it is characterized by simultaneously demonstrating strong cold and warm biases both in strength and extent along the coastal regions and in the central northern SCS, respectively, which are the largest among four seasons. Another prominent cold bias area throughout the seasons and consistent with that in the northern SCS is the southern coastal region off Vietnam (see Figs 5a–d); in summer, it is strongest and extends northward along the coast all the way up to the Beibu Gulf. During autumn (Fig. 5c) the warm bias in the central northern SCS recedes in both strength and extent to reach the minimum of all four seasons. In winter (Fig. 5d) the warm bias shows a signific-

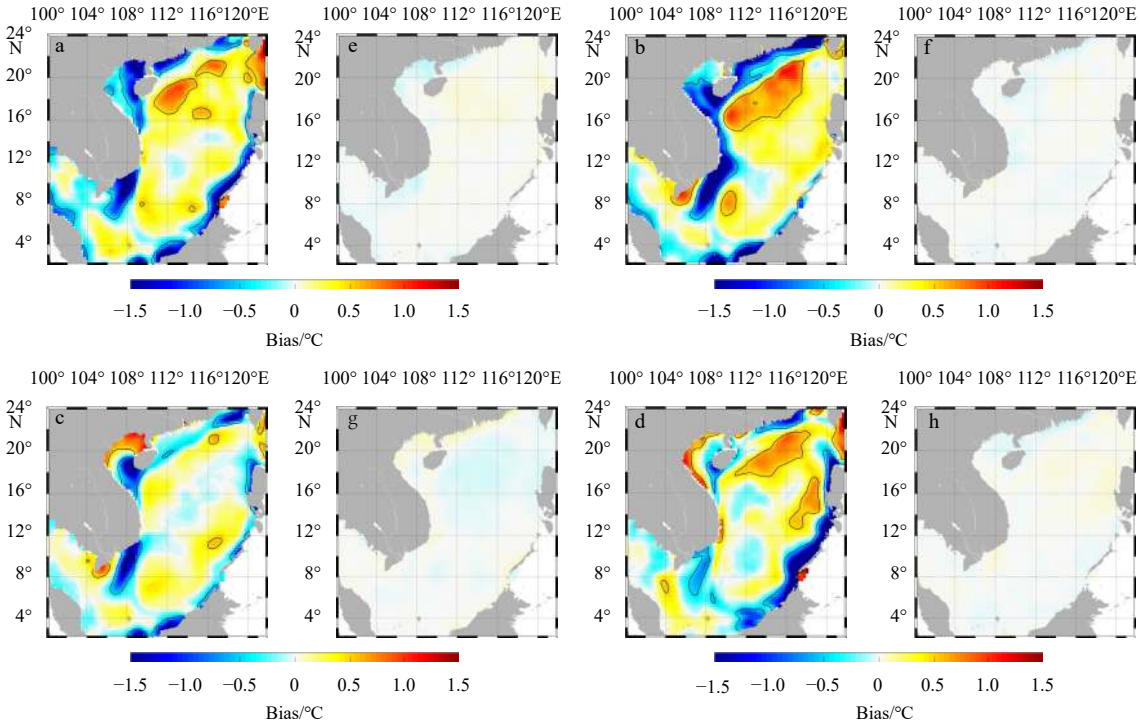

Fig. 5. Maps of seasonal bias during spring (a), summer (b), autumn (c), and winter (d) in the CORA SST compared with the AVHRR-AMSR SST. Panels e to h are the same as a to d, except with bias correction. The black contours in the left panels indicate the RMSE range (±0.50°C) of the satellite data.

ant development, whereas the cold biases drop to the lowest. The aforementioned cold and warm biases grow in both strength and extent during spring (Fig. 5a). The strong cold bias in the coastal area of the southeastern SCS exhibits seasonal fluctuation, being opposite in phase with that in the northern SCS. It is the strongest in winter (Fig. 5d), begins to weaken in spring (Fig. 5a) and through summer (Fig. 5b), and reaches the minimum in autumn (Fig. 5c). The seasonal variation of the CORA SST bias in the head area of the Beibu Gulf is also obvious: the coldest bias occurs during spring (Fig. 5a), and is weakened in summer (Fig. 5b), with a weak warm bias appearing at the right-most area of the head. This weak warm bias develops into the maximum in autumn (Fig. 5c), and starts to shrink during winter (Fig. 5d). A large warm bias is present all the time in the Luzon Strait, making an appearance of the maximum in spring (Fig. 5a) and minimum in summer (Fig. 5b).

Figures 5e–h illustrate the spatial distributions of the corrected seasonal bias in the CORA SST by applying the EOF-BPNN method. Just as expected in the bias-corrected climatology (Fig. 2d), the convergence of the adjusted CORA SST to the satellite data is quite remarkable, with deviation values being 100% inside the satellite error range. The seasonality and spatial structure of the corrected bias still hold with relatively small basin-wide deviation in summer (Fig. 5f).

3.3 *Bias correction for monthly and daily mean SST fields*

Figure 6a shows time series of the basin-wide monthly and daily mean bias of the uncorrected CORA SST from 2003 to 2008. The cold bias (blue bars) on daily time scale in the CORA SST generally occurs more often, and is also stronger than the warm bias (red bars) with the exceptions of 2007 and early-2008, during which the warm bias came into prominence. Such features are shared by the monthly mean bias (black curve). At the same time, the daily mean warm and cold biases mostly remain inside the satellite error range (±0.50°C), with the warmest of 0.71°C appearing in the spring of 2007 and the coldest of –0.58°C in the fall of 2006.

After applying the EOF-BPNN method, as shown in Fig. 6b, the corrected daily mean bias in the CORA SST shows no extremum of being either cold or warm no matter which stage the output of the algorithm is in, namely, the training stage (2003–2007) or testing stage (2008). The corrected daily mean warm (red bars) and cold (blue bars) biases approximately fall within the range of ±0.20°C. During both the first five-year training period (2003–2007) and the last one-year testing period (2008), the corrected monthly mean deviation (see black curve in Fig. 6b) converges to zero, which means that the basin-wide CORA SST is in good agreement with the satellite SST on monthly time scale.

An interesting feature in Fig. 6b is that, unlike the uncorrected bias being warm or cold in some specific periods (Fig. 6a), the corrected basin-wide daily mean warm and cold biases are distributed evenly on both sides of the zero line. The very same situation arises when examining the performance of the BPNN method applied to each individual grid in the bias-correction experiment (not shown). It seems that this is the strategy adopted by the BPNN algorithm to minimize the mean-squared error cost function. This is an appealing topic for future research involving choice of different cost functions and comparative studies of other popular machine learning techniques to uncover the essence of such feature.

4 Conclusions and discussion

Accurate marine forecasts are becoming more and more crucial to satisfy needs arising from the development of modern so-

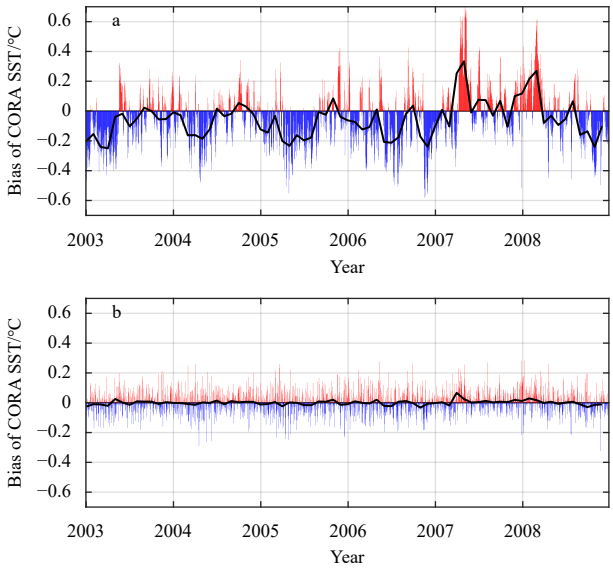

Fig. 6. Basin-wide monthly mean (black curve) and daily mean (red and blue bars) time series of the difference between the AVHRR-AMSR and CORA SSTs before (a) and after (b) applying the bias-correction method of EOF-BPNN.

ciety. The simplest way to reduce forecast uncertainties and improve the skill of numerical ocean forecasts is to empirically correct model outputs. With this in mind, two methods for offline bias correction of SST forecasts are proposed and analyzed in this study: a BPNN algorithm, and a hybrid algorithm of EOF analysis and BPNN (called EOF-BPNN). With the availability of time series of daily mean SST retrospective forecasts for a lead time of one-day in the SCS, an archived dataset produced by the regional CORA forecast and analysis system in the coastal waters of China and its adjacent seas, the feasibility of these two established schemes is validated in the bias correction experiments with the SST retrospective forecast time series spanning a six-year period from January 1, 2003 to December 31, 2008, during which the satellite SST dataset of the AVHRR-AMSR from the NOAA daily OISST v2.0 is used as the benchmark. The performance evaluation is demonstrated by metrics of bias, MAE, and RMSE measuring accuracy of the CORA daily mean SST retrospective forecasts before and after bias correction. Significant error reduction in the CORA SST forecasts has been achieved using both BPNN and EOF-BPNN methods. Although successful basin-wide bias correction acquired by these two methods is very similar with respect to various time scales ranging from daily, monthly, seasonal, interannual, to climatology, the EOF-BPNN method performs more than five times faster than the BPNN method. From an operational point of view, the EOF-BPNN algorithm would be an appropriate candidate due to less computational burden. The metric used to measure the quality of the bias correction is that the basin-wide climatological errors of the bias-corrected forecasts fall within the error range (±0.5°C) of the satellite SST data. Note that the error range (±0.5°C) is given according to previous validation studies of the satellite SST data in the East Asian marginal seas (0.5–0.7°C; Sakaida et al., 2000; Lee et al., 2005; Qiu et al., 2009). In this sense, the performance of the BPNN method can be judged as good and satisfactory.

The results of the bias-correction experiments presented above include outputs of the BPNN and EOF-BPNN methods during both the first five-year training stage (2003–2007) and the last-year testing stage (2008). Actually, when addressing and directing attention to the outputs of the two algorithms for the last-year testing stage only, the same conclusion can be drawn by looking into the details, say the spatial distributions of the seasonal bias and RMSE for the corrected CORA SST. It thus provides us with confidence in the capabilities of these two algorithms for correcting bias of the one-day-ahead SST forecasts. Once bias correction in daily forecasted SST is done, by using the established model such as the Modular Ocean Data Assimilation System (Fox et al., 2002), it can be projected onto the mixed layer where the correction has clear impact.

While the results from implementing the BPNN algorithm to bias correction of one-day-ahead retrospective forecasted SST from the CORA system are very encouraging, there are issues that need to be addressed before its application in an operational context. The satellite SST is known to have cold bias against *in-situ* SST (see Section 3.1); thus, improving satellite SST (used as the reference in this research) should be given priority. Only then, data-driven approach such as the one described in this paper can be implemented. Unfortunately, *in-situ* observations remain insufficient for large regions, particularly along the coastal region where large SST forecast bias occurs just like the situation encountered in the SCS in this study. Therefore, it will be a great challenge for enhancing the satellite SST quality in the SCS. What is more, impacts of these two bias-correction algorithms on the CORA SST retrospective forecasts for longer than one day in advance need to be investigated in future. The time series of the CORA SST retrospective forecast in one-day lead time is used for the bias-correction experiments in this research. The biases in the retrospective forecasted SST in this case are therefore strongly model- and assimilation-dependent. However, from the analysis above, it can be concluded that the relatively large cold biases along the coastal waters shallower than 200 m in the SCS are more likely attributed to the deficiencies of the ocean model in the CORA system, although the effect of data assimilation during the distribution of observational increments by the multigrid three-dimensional variational scheme cannot be ruled out until the numerical model and data assimilation scheme in the CORA system are thoroughly and separately calibrated. Therefore, it is difficult to tell which type of biases can be corrected at this stage. Future efforts should aim not only to reduce SST forecast errors but to simultaneously assess and estimate uncertainties in the ocean model for longer lead time forecasts (greater than five days, for example) by eliminating the impact of data assimilation as much as possible in order to enhance the skill of the CORA forecasting system in terms of ocean model development. All in all, it is definitely a topic for further investigation by calibrating the numerical model and data assimilation scheme, respectively, to achieve the ultimate goal of continuously upgrading the CORA system, which has been developed for more than a decade. It is anticipated that the continuous improvements in the CORA system being able to produce the reanalysis of historical ocean state with reliable error estimates will provide us reference data for bias correction to the subsurface temperature and salinity forecasts using the data-driven approach developed in this research.

Acknowledgements

We thank the NOAA's National Centers for Environmental Information for providing the gridded AVHRR-AMSR data of the daily OISST v2.0 (https://www.ncei.noaa.gov/data/sea-surface-temperature-optimum-interpolation), and the Met Office Hadley Centre for providing the EN4.2.1 dataset (https://www.metoffice.gov.uk/hadobs/en4/download-en4-2-1.html). We also thank

the editor and two anonymous reviewers for their helpful comments to improve the quality of this paper.

References

Abhilash S, Sahai A K, Borah N, et al. 2014. Does bias correction in the forecasted SST improve the extended range prediction skill of active-break spells of Indian summer monsoon rainfall?. Atmospheric Science Letters, 15(2): 114–119, doi: 10.1002/asl2.477

Ashfaq M, Skinner C B, Diffenbaugh N S. 2011. Influence of SST biases on future climate change projections. Climate Dynamics, 36(7–8): 1303–1319, doi: 10.1007/s00382-010-0875-2

Balmaseda M A, Mogensen K, Weaver A T. 2013. Evaluation of the ECMWF ocean reanalysis system ORAS4. Quarterly Journal of the Royal Meteorological Society, 139(674): 1132–1161, doi: 10.1002/qj.2063

Bhargava K, Kalnay E, Carton J A, et al. 2018. Estimation of systematic errors in the GFS using analysis increments. Journal of Geophysical Research: Atmospheres, 123(3): 1626–1637, doi: 10.1002/2017JD027423

Boutilier C, Patrascu R, Poupart P, et al. 2006. Constraint-based optimization and utility elicitation using the minimax decision criterion. Artificial Intelligence, 170(8–9): 686–713, doi: 10.1016/j.artint.2006.02.003

Carter G M, Dallavalle J P, Glahn H R. 1989. Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. Weather and Forecasting, 4(3): 401–412, doi: 10.1175/1520-0434(1989)004<0401:SFBOTN>2.0.CO;2

Chang Y, Schubert S D, Koster R D, et al. 2019. Tendency bias correction in coupled and uncoupled global climate models with a focus on impacts over North America. Journal of Climate, 32(2): 639–661, doi: 10.1175/JCLI-D-18-0598.1

Chao Guofang, Wu Xinrong, Zhang Lianxin, et al. 2020. China Ocean ReAnalysis (CORA) version 1.0 products and validation for 2009–18. Atmospheric and Oceanic Science Letters, 14(5): 100023, doi: 10.1016/j.aosl.2020.100023

Chen Xiao, Yan Youfang, Cheng Xuhua, et al. 2013. Performances of seven datasets in presenting the upper ocean heat content in the South China Sea. Advances in Atmospheric Sciences, 30(5): 1331–1342, doi: 10.1007/s00376-013-2132-1

Dalcher A, Kalnay E. 1987. Error growth and predictability in operational ECMWF forecasts. Tellus A, 39A(5): 474–491, doi: 10.1111/j.1600-0870.1987.tb00322.x

Danforth C M, Kalnay E, Miyoshi T. 2007. Estimating and correcting global weather model error. Monthly Weather Review, 135(2): 281–299, doi: 10.1175/MWR3289.1

Dee D P, Da Silva A M. 1998. Data assimilation in the presence of forecast bias. Quarterly Journal of the Royal Meteorological Society, 124(545): 269–295, doi: 10.1002/qj.49712454512

Fan Maoting, Wang Huizan, Zhang Weimin, et al. 2020. Evaluation of the China ocean reanalysis (CORA) in the South China Sea. Journal of Oceanology and Limnology, 38(6): 1640–1653, doi: 10.1007/s00343-019-9146-1

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

Glahn H R, Lowry D A. 1972. The use of model output statistics (MOS) in objective weather forecasting. Journal of Applied Meteorology and Climatology, 11(8): 1203–1211, doi: 10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2

Good S A, Martin M J, Rayner N A. 2013. EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. Journal of Geophysical Research: Oceans, 118(12): 6704–6716, doi: 10.1002/2013JC009067

Han Guijun, Li Wei, Zhang Xuefeng, et al. 2011. A regional ocean reanalysis system for coastal waters of China and adjacent seas. Advances in Atmospheric Sciences, 28(3): 682–690, doi: 10.1007/s00376-010-9184-2

Han Guijun, Li Wei, Zhang Xuefeng, et al. 2013. A new version of regional ocean reanalysis for coastal waters of China and adjacent seas. Advances in Atmospheric Sciences, 30(4): 974–982, doi: 10.1007/s00376-012-2195-4

Hernández-Díaz L, Nikiéma O, Laprise R, et al. 2019. Effect of empirical correction of sea-surface temperature biases on the CRCM5-simulated climate and projected climate changes over North America. Climate Dynamics, 53(1–2): 453–476, doi: 10.1007/s00382-018-4596-2

Kingma D P, Ba J. 2015. Adam: a method for stochastic optimization. In: Proceedings of the 3rd International Conference for Learning Representations. San Diego, CA, USA: ICLR

Klein W H. 1971. Computer prediction of precipitation probability in the United States. Journal of Applied Meteorology and Climatology, 10(5): 903–915, doi: 10.1175/1520-0450(1971)010<0903:CPOPPI>2.0.CO;2

Klein W H, Lewis B M, Enger I. 1959. Objective prediction of five-day mean temperatures during winter. Journal of the Atmospheric Sciences, 16(6): 672–682, doi: 10.1175/1520-0469(1959)016<0672:OPOFDM>2.0.CO;2

Kug J S, Lee J Y, Kang I S. 2008. Systematic error correction of dynamical seasonal prediction of sea surface temperature using a stepwise pattern project method. Monthly Weather Review, 136(9): 3501–3512, doi: 10.1175/2008MWR2272.1

LaRow T E. 2013. The impact of SST bias correction on North Atlantic hurricane retrospective forecasts. Monthly Weather Review, 141(2): 490–498, doi: 10.1175/MWR-D-12-00152.1

Lee M A, Chang Yi, Sakaida F, et al. 2005. Validation of satellite-derived sea surface temperatures for waters around Taiwan. Terrestrial, Atmospheric and Oceanic Sciences, 16(5): 1189–1204, doi: 10.3319/TAO.2005.16.5.1189(O)

Li Wei, Xie Yuanfu, Deng S M, et al. 2010. Application of the multigrid method to the two-dimensional Doppler radar radial velocity data assimilation. Journal of Atmospheric and Oceanic Technology, 27(2): 319–332, doi: 10.1175/2009JTECHA1271.1

Li Wei, Xie Yuanfu, He Zhongjie, et al. 2008. Application of the multigrid data assimilation scheme to the China seas' temperature forecast. Journal of Atmospheric and Oceanic Technology, 25(11): 2106–2116, doi: 10.1175/2008JTECHO510.1

Maas A L, Hannun A Y, Ng A Y. 2013. Rectifier nonlinearities improve neural network acoustic models. In: Proceedings of the 30th International Conference on Machine Learning. Atlanta, GA, USA: ICLR

Mellor G L, Häkkinen S M, Ezer T, et al. 2002. A generalization of a sigma coordinate ocean model and an intercomparison of model vertical grids. In: Pinardi N, Woods J, eds. Ocean Forecasting: Conceptual Basis and Applications. Berlin, Heidelberg: Springer, 55–72

Narapusetty B, Stan C, Kumar A. 2014. Bias correction methods for decadal sea-surface temperature forecasts. Tellus A: Dynamic Meteorology and Oceanography, 66(1): 23681, doi: 10.3402/tellusa.v66.23681

North G R, Bell T L, Cahalan R F, et al. 1982. Sampling errors in the estimation of empirical orthogonal functions. Monthly Weather Review, 110(7): 699–706, doi: 10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2

Qiu Chunhua, Wang Dongxiao, Kawamura H, et al. 2009. Validation of AVHRR and TMI-derived sea surface temperature in the northern South China Sea. Continental Shelf Research, 29(20): 2358–2366, doi: 10.1016/j.csr.2009.10.009

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

Rumelhart D E, Hinton G E, Williams R J. 1986. Learning representations by back-propagating errors. Nature, 323(6088): 533–536, doi: 10.1038/323533a0

Sakaida F, Kudoh J I, Kawamura H. 2000. A-HIGHERS—the system to produce the high spatial resolution sea surface temperature maps of the western North Pacific using the AVHRR/NOAA. Journal of Oceanography, 56(6): 707–716, doi: 10.1023/A:1011181918048

Vitart F, Balmaseda M. 2018. Impact of sea surface temperature bi-

ases on extended-range forecasts. ECMWF Technical Memoranda. Reading, UK: ECMWF

Voldoire A, Exarchou E, Sanchez-Gomez E, et al. 2019. Role of wind stress in driving SST biases in the Tropical Atlantic. Climate Dynamics, 53(5–6): 3481–3504, doi: 10.1007/s00382-019-04717-0

Wu Yang, Cheng Guosheng, Han Guijun, et al. 2013. Analysis of seasonal and interannual variability of sea surface temperature for China Seas based on CORA dataset. Haiyang Xuebao (in Chinese), 35(1): 44–54, doi: 10.3969/j.issn.0253-4193.2013.01.006

Xie Shangping, Xie Qiang, Wang Dongxiao, et al. 2003. Summer upwelling in the South China Sea and its role in regional climate variations. Journal of Geophysical Research: Oceans, 108(C8): 3261, doi: 10.1029/2003JC001867

Xing Yansong, Cheng Guosheng, Shu Yeqiang, et al. 2012. Anomalous characteristics of the ocean circulation in South China Sea during the El Niño events. Oceanologia et Limnologia Sinica, 43(2): 201–209

Zhang Min, Zhou Lei, Fu Hongli, et al. 2016. Assessment of intraseasonal variabilities in China Ocean Reanalysis (CORA). Acta Oceanologica Sinica, 35(3): 90–101, doi: 10.1007/s13131-016-0820-2