1. Introduction

Study on the variation of the sea surface height anomaly (SSHA) is an important issue in physical oceanography and meteorological science. Changes in SSHA will influence the frequency and impact of extreme sea level events which engender lots of negative impact [1, 2]. Previous studies found that the regional SSHA variability at the interannual time scale dominated by ocean variability is much larger than the global SSHA [3–5]. As the largest marginal sea in Western Pacific, South China Sea (SCS) is a crucial link between Pacific and Indian Ocean, and it has significant impacts on human activities and the sustainable development of coastal economy and society.

Analysis and forecast are two crucial aspects of SCS SSHA. As far, most attentions have been paid to the analysis of the SSHA in SCS, like the annual variability [6, 7], interannual variability [8, 9], seasonal variability [10], the rising rate [11, 12], the forcing of the variations [13–16], and the variability associated with El Niño and Southern Oscillation (ENSO) [8, 9, 16, 17]. For the forecast of SCS SSHA, the forecast model of SCS SSHA, mainly, includes the dynamical models and statistical methods. Now the 3D dynamical prediction system of the SCS SSHA has been improved [18, 19]. For example, Wei et al. [18] use a finite-grid dynamical model covering SCS to produce monthly and annual mean SSHA of the SCS from 1992 to 2000, and the model-produced data is in good agreement with altimeter measurements. While the dynamical approaches have been studied widely [1, 20, 21], little attentions have been paid to the statistical forecast of SSHA in SCS, which is simple and feasible in practice. The Pacific ENSO Applications Climate (PEAC) Centre at the National Oceanographic and Atmospheric Administration (NOAA) uses a statistical model to calculate site-specific seasonal sea level outlooks, and the results indicated that the statistical model is potentially useful in predicting seasonal sea level variations in the U.S.-affiliated Pacific Islands (USAPI) [22]. Imani et al. [23] used the Holt-Winters exponential smoothing technique to analyze and forecast Caspian Sea level anomalies from 15-year altimetry...
data from 1993 to 2008 and found that the modeling results of a 3-year forecasting time span (2005–2008) agree well with the observed time series. As far as we know, no literature on the statistical prediction of SCS SSHA has been documented. Also it is feasible to use the statistical models to implement the forecast of the SCS SSHA compared with the complicated dynamical models. In this study, based on the long-term monthly time series of SCS SSHA derived from the Simple Ocean Data Analysis (SODA), we first use a time series decomposition method to analyze the variability of monthly SCS SSHA. Then, a statistical model is constructed to fit the monthly SCS SSHA and used to perform an array of forecast experiments. Finally, the statistical prediction skill of SCS SSHA is investigated.

The remainder of this study is organized as follows: the data, the time series decomposition method, and the statistical model construction are introduced in Section 2. Section 3 presents the analysis results of SSHA in SCS while the forecast results are shown in Section 4. Summary and discussion are given in Section 5.

2. Methodology

2.1. Data

2.1.1. SODA. An ocean reanalysis product, namely, simple ocean data assimilation (SODA), is used in this study, which is based on the Parallel Ocean Program ocean model with an average horizontal resolution of 0.5° × 0.5° and with 40 vertical levels during January 1948 to December 2007 [24]. SODA data is downloaded from web site at http://dsrs.atmos.umd.edu/DATA/soda_2.2.A/. For our study, we first calculate the monthly mean values of SSH so as to derive the monthly SSHA. Then, we average the monthly SSHA over the South China Sea (5°N–25°N, 105°E–121°E) to form the basic datasets used in this study.

2.1.2. AVISO. A merged gridded product of Maps of Sea Level Anomaly produced by AVISO (Archivage Validation et Interprétation des données des Satellites Oceanographiques) based on TOPEN/Poseidon, Janson 1, ERS-1, and ERS-2 satellite data is used for evaluating the correctness of SODA data. This product provides SSHA as from January 1993 to December 2007, which consists of maps produced every day on a 1/8° × 1/8° Cartesian grid. The monthly SSHA in AVISO is first computed and then used to derive the monthly SCS SSHA.

2.2. Statistical Methods for Modeling and Forecasting

2.2.1. Decomposition Method. In this section, we briefly introduce the time series decomposition method, that is, the centralized moving average scheme. This method partitions a monthly time series into the following three components: the interannual component $T_t$, the seasonal component $S_t$, and the residual component $R_t$ as follows:

$$x_t = T_t + S_t + R_t,$$

where $x_t$ represents the monthly value at time $t$. The computational processes of $T_t$, $S_t$, and $R_t$ are listed as follows.

(a) Interannual Component. A simple centralized moving average technique with 12-month time scale is used to obtain the interannual term as

$$T_t = \frac{0.5x_{t-6} + x_{t-5} + \cdots + x_{t+5} + 0.5x_{t+6}}{12},$$

where $N$ represents the length of $x_t$. Apparently, with the smoothing-scale being set to 12 months, $\{T_t\}$ filters the short time scale (less than 12 months) information of $\{x_t\}$ and keeps the interannual variations. Thus, $\{T_t\}$ describes the interannual variabilities of $\{x_t\}$. Note that the valid period of $\{T_t\}$ is from July 1948 to June 2007. The following decomposition and analysis as well as the forecast experiments will be also based on $\{x_t\}$ during the same period.

(b) Seasonal Component. After the interannual term of $\{x_t\}$ has been computed with the above method, we first subtract $T_t$ from $\{x_t\}$ to get the remaining term. In this study, the seasonal term ($S_t$) corresponds to the climatology of the remaining term. Thus, although the length of $S_t$ is the same as $T_t$, the period of $S_t$ is 12 months.

(c) Residual Component. With the interannual term ($T_t$) and the seasonal term ($S_t$) being determined, the residual term $R_t$ of $x_t$ is directly computed by

$$R_t = x_t - T_t - S_t.$$

According to the time scales of $T_t$ and $S_t$, it is easy to derive that $R_t$ reflects the variability of $x_t$ whose time scale is interseasonal or smaller. As the resolution of $x_t$ is monthly, it is expected that $R_t$ is noise-dominant. In addition, the summation of $R_t$ and $T_t$ represents the anomaly of $x_t$. Due to the static nature of $S_t$, the forecasting of $x_t$ mainly depends on the prediction of the anomaly of $x_t$, which consists of $R_t$ and $T_t$. Because of the different time scales of $R_t$ and $T_t$, different statistical models are used to fit and forecast $R_t$ and $T_t$.

2.2.2. Models for Fitting and Forecasting. To investigate the statistical predictability of SSHA in SCS, the statistical forecast model is first constructed in this section.

(a) The Model for Interannual Term. Due to the simplicity and robustness [23, 25], Holt’s linear exponential smoothing technique is widely used to handle nonseasonal time series by introducing smoothing parameters [26], like the precipitation prediction [27] and the maxima and minima air temperature prediction [28]. In this study, we also use Holt’s linear exponential smoothing technique to fit and forecast $T_t$. The fit formula of $T_t$ is

$$\hat{T}_{t+1} = a_t + b_t,$$

Advances in Meteorology
\[ a_t = \alpha \hat{T}_t + (1 - \alpha) (a_{t-1} + b_{t-1}), \]
\[ b_t = \beta (a_t - a_{t-1}) + (1 - \beta) b_{t-1}. \]

Here \( \alpha \) and \( \beta \) denote the smoothing parameters. The initial values of \( a_0 \) and \( b_0 \) are usually set to 0. Given preset values (i.e., first guesses) of \( \alpha \) and \( \beta \), the fitness of \( T_i \) (denoted as \( \hat{T}_i \), \( t = 8 \), \( N - 6 \)) can be calculated sequentially with known \( \hat{T}_{t-1} \). Note that \( 7 \) is the starting index of \( T_i \) (see (2)) and \( \hat{T}_7 = T_7 \). Then a cost function between \( T_i \) and \( \hat{T}_i \) is established. Taking \( \alpha \) and \( \beta \) as the control variables to be optimized, an optimization algorithm, like the L-BFGS-B algorithm [29], is used to obtain the optimum \( \alpha \) and \( \beta \). In addition, to avoid the overflow of \( \alpha \) and \( \beta \), the [0, 1] bound is applied to these two parameters during the optimizing process.

For the forecast of \( T_{t+k} \), a linear function of the lead time is constructed as follows:

\[ T_{t+k} = a_t + b_t k. \]

Here \( k \) is the lead time in month.

(b) The Model for Residual Term. Due to the noise-dominant nature of \( R_t \), we use an auto-regression integrated moving average model (denoted as ARIMA\((p,d,q)\)) [30, 31]) to construct the fit model of \( R_t \). Here \( p \) represents the order of the autoregression model (i.e., AR\((p)\)); \( d \) is the order of time differential of \( R_t \); \( q \) indicates the order of the moving average model. The model can be formulated as

\[ \hat{W}_t = \varphi_1 \hat{W}_{t-1} + \varphi_2 \hat{W}_{t-2} + \cdots + \varphi_p \hat{W}_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}, \]

\[ W_t = \nabla^d R_t, \]

where \( W_t \) represents the time series of \( d \)-order time differentiate of \( R_t \), \( t \) is the time, \( \varphi_1, \varphi_2, \ldots, \varphi_p \), are coefficients of the autoregressive integrated model, \( \theta_1, \theta_2, \ldots, \theta_q \) are coefficients of the moving average model, and \( \varepsilon_t \) is a white noise time series [32].

Figure 1 shows the time series of monthly SCS SSHA (cm) derived from AVISO (black) and SODA (red).

3. Modeling Construction

Affected by various factors, such as solar radiation, evaporation and precipitation, monsoon, and El Niño and Southern Oscillation (ENSO), the SCS SSHA has significant characteristics on different time scales. In this section, we roughly analyze the time variability of SCS SSHA with the time series decomposition method (Section 2.2.1) based on the SODA product.

3.1. Verification of SODA SSHA. Before applying the time series decomposition to the SCS SSHA derived from the SODA product, the correctness of this dataset should be first verified. Figure 1 shows the time series of the monthly SCS SSHA derived from SODA (red) and AVISO (black). It can be seen that the most prominent signal is the seasonality in both products. The correlation coefficient between two time series reaches 0.89 with the significance level above 95%. The root-mean-square error between two time series is 2.56 cm. The SODA product can well capture the temporal variability of SCS SSHA.

To detect the significant periods of these two datasets, we perform the power spectrum analysis. Figure 2 presents the power spectrum of SCS SSHA for AVISO (blue) and SODA (black), where the dashed curves represent the 95% confidence upper limits. Obviously, both time series have 1-year and half-year significant periods.

Due to the limited length of AVISO SSHA data, we cannot systematically verify the correctness of the whole time series of SCS SSHA derived from SODA data. However, based on the above analysis, it is reasonable to assume that the quality of the time series of SCS SSHA from SODA data is reliable.

3.2. Results of Time Series Decomposition. Given that the correctness of the SODA SSHA data has been validated by AVISO data, we apply the time series composition to the monthly SCS SSHA from SODA with a 60-year length. Figure 3 shows the time series of three decomposed terms of SSHA in SCS, where the interannual term, residual term, and seasonal term are represented by black, red, and blue curves. For the interannual term, it has significant interannual variability. During most of El Niño years, SCS SSHA is smaller than that during other years. Moreover, there are some inflexions most of which correspond to El Niño or
3.3. Model Construction. Before performing the forecast experiments of SSHA, the forecast model of SCS SSHA should be first established. Note that the SSHAs from now on indicate the SSHA anomalies. Due to the invariant property of the seasonal term, the forecast model is divided into two parts: the Holt-Winters model for the interannual term and the ARIMA model for the residual term. Given the time series of SCS SSHA, the fitness models for the interannual term and the residual term are first, respectively, constructed and then used to forecast these two quantities. Finally the forecasted interannual term and the residual term are added together to form the forecasted anomaly of SCS SSHA. In this section, we examine the qualities of fitness for the interannual and residual terms.

Figures 4(a) and 4(b) show the fitted values (red curve) and SODA values (blue curve) of the interannual and residual terms of SSHA in SCS. Note that as the time series composition needs to remove the first and last 6 months’ SSHAs, the actual fitting model uses the time series between July 1948 and June 2007. For the interannual term, the Holt-Winters model can exactly track the trajectory of the SODA data with little biases in the points of minimum and maximum. For the residual term, the ARIMA model can well simulate the phase of the SODA data, although the simulation of the amplitude is not so good, which may be attributed to the nearly white noise nature of the residual term. Figure 4(c) plots the fitted and SODA values of total SSHA in SCS. The fitted curve is strongly correlated with the SODA time series. Figure 4(d) presents the time series of the biases (computed as the difference between the fitted value and the observed value) of the fitted interannual term (black curve), residual term (blue dashed curve), and total SCS SSHA (red curve). The bias of the fitted total SCS SSHA oscillates around 0 and mainly consists of residual biases. The analysis indicates that the added statistical model is effective and flexible, and then we use the added model to do forecast experiments as follows.

4. Forecast Results

The last 30 years’ analysis results are used to perform the forecast experiment which is similar to the dynamical forecast of physical processes (like ENSO [34, 35]). For each month from December 1977 to December 2007, we first decompose the time series (from January 1948 to the current month) of SSHA in SCS into three parts and then use Holt-Winters and ARIMA models to fit the interannual and residual terms, respectively. Then, both models are integrated forward up to 2 years to forecast the interannual and residual terms of SSHA in SCS. Thus, 361 forecast experiments are conducted. Note that as the time series composition needs to remove the SSHAs during the first and last 6 months, the actual initial forecast months for 361 forecast experiments correspond to June 1977 and June 2007. The forecast SSHA anomaly (interannual term plus residual term) is compared to the SODA data with seasonal cycle removed. Figure 5 gives a sketch map of the forecast experiments.

Based on the forecast results, we investigate the predictability of SCS SSHA from the following three aspects.
4.1. Forecast Skill. We first totally examine the prediction skill of SSHA in SCS with the anomaly correlation coefficient (ACC) and the root-mean-square error (RMSE) relative to the SODA data. The ACC and RMSE are defined as follows:

\[
\text{ACC}_s = \frac{\sum_{i=1}^{361} (x'_{i,s} - \bar{x}_s) (x^0_{i,s} - \bar{x}_s)}{\sqrt{\sum_{i=1}^{361} (x'_{i,s} - \bar{x}_s)^2} \sqrt{\sum_{i=1}^{361} (x^0_{i,s} - \bar{x}_s)^2}},
\]

\[
\text{RMSE}_s = \sqrt{\frac{1}{361} \sum_{i=1}^{361} \left( \frac{x'_{i,s} - x^0_{i,s}}{\sigma_c} \right)^2},
\]

where \(s\) and \(i\) index the lead time and the forecast experiment; \(x'\) and \(x^0\) represent the forecasted and SODA anomaly of SCS SSHA while \(\bar{x}\) and \(\bar{x}^0\) denote the averages of \(x'\) and \(x^0\) over all forecast experiments; \(\sigma_c\) indicates the climatological standard deviation of anomaly of SCS SSHA.

Figures 6(a) and 6(b) show the variation of RMSE (Figure 6(a)) and ACC (Figure 6(b)) of forecasted SSHA anomaly in SCS with respect to lead time (in months) and start month. The solid curves in Figures 6(a) and 6(b) indicate the 1.0-contour and 0.6-contour, respectively. According to the results of RMSE, we find that the RMSEs forecasted from Spring (like April and May) and Autumn (like October and November) are smaller than the other seasons. The minima of the RMSE happen in May and November, and the maxima occur in June. If an ad hoc 0.6 value of ACC is used to define the valid lead time, the valid forecast time of SCS SSHA anomaly is 9 (7) months in Spring (other seasons).

To further investigate the contributions of the interannual term and residual term to the total quantity, we separately analyze the predictabilities of the interannual term and the residual term. Figures 6(c), 6(d), 6(e), and 6(f) present the same results as Figures 6(a) and 6(b) but for the interannual term and the residual term, respectively. We can see that the valid forecast time of the interannual term is 12 (9) months in Summer (other seasons). In contrast, the prediction skill of the residual term is much worse than the total quantity, especially for the ACC and the RMSEs forecasted from Spring and Autumn. This may be caused by the white noise in the forecast model (i.e., the ARIMA model) of the residual term, which increases the uncertainty of the forecasted residual
Figure 6: Variations of RMSE (a, c, and e) and ACC (b, d, and f) of forecasted SCS SSHA anomaly (a, b), forecasted interannual term (c, d), and forecasted residual term (e, f) with respect to lead time (in month) and start month. The solid curve in (a, c, and e) [b, d, f] is (1.0-contour) [0.6-contour]. Note that RMSEs in (a, c, and e) have been normalized by the climatological standard deviations of SCS SSHA anomaly, interannual term, and residual term, respectively.
term. Therefore, the forecast skill of the anomaly of SCS SSHA is mainly contributed by the counterpart of the interannual term and more or less reduced by the bad prediction skill of the residual term.

To figure out the possible reason causing the prediction skill of the interannual term, we calculate the number of extreme (including maximum and minimum) values of the interannual term (decomposed from the SODA SCS SSHA time series between July 1977 and June 2007) that happens in each month (Figure 7). Due to the linear property of the forecast model (i.e., the Holter-Winter model) of the interannual term, the linear trend between the maximum and the minimum can enhance the prediction skill of the interannual term. Thus, if a forecast is started from a month where the extreme value happens, the prediction skill is likely to be improved. From Figure 7, we can see that most extreme values happen in July which is consistent with the results in Figure 6(d) with an exception of the case of June.

4.2. Seasonal Forecast Skill. According to the analysis in last section, the valid forecast time of SCS SSHA is about 7 months. We now assess the seasonal forecasts in this section. Figure 8 plots the SODA data (blue) and 6-month (red in Figure 8(a)) or 9-month (red in Figure 8(b)) lead SCS SSHA anomalies. We can see that, in the seasonal forecast, the statistical model can capture the main variability of SSHA anomaly in SCS, especially for the 6-month lead. The correlation coefficient between the SODA data and the 6-month (9-month) lead SCS SSHA anomaly is 0.77 (0.45) with 95% significance level.

The SSHA in SCS can be forecasted in 6-month lead for all three decades. As the previous studies indicated, the SCS is modulated by PDO, and the SSH variability in SCS is coherent with PDO [16]. We also find that the ACCs for the last decade and the total thirty years with the 7 lead months are higher than those for other two decades, when the two decades (1977–1997) are in warm Pacific Decadal Oscillation (PDO) phase [36]. In addition, when we construct the forecast model, the SODA SSHA time series is not long enough. Thus, the forecast model mainly covers both the cold and the warm phases, and the forecast in cold phase decades is better than in warm. Conversely, the RMSEs for the last decade are the worst, which may be related to the quality of SODA. From the results of the total years, the valid forecast time of SCS SSHA is about 7 months. Obviously, the prediction skill of SCS SSHA anomaly has apparent decadal variability.

5. Summary and Discussion

In this study, we first simply contrast the time series of monthly SSSHAs calculated from SODA and AVISO datasets in SCS to verify the correctness of the SODA SCS SSHA which is used to construct the forecast model of SCS SSHA. Results show that SODA well correlates with AVISO data. Afterwards, we use the long-term time series from 1948 to 2007 of SODA SCS SSHA to establish the statistical forecast model. A time series decomposition method is used to decompose the monthly time series of SCS SSHA into the following three parts: interannual term, seasonal term, and
residual term. Then we, respectively, use the Holt-Winters and ARIMA models to fit the interannual and residual terms of SSHA in SCS, aiming to construct the forecast models of these two parts. Results of fitting demonstrate that the Holt-Winters model can exactly track the trajectory of SODA counterpart while the ARIMA can well simulate the phase of the residual term. Finally, we perform an array of forecast experiments based on the above two models. Results show that the valid forecast time of SCS SSHA is about 7 months. The predictability of SCS SSHA in Spring and Autumn is stronger than that in Summer and Winter, and the forecast skill of the anomaly of SCS SSHA is mainly contributed by the counterpart of the interannual term and more or less reduced by the bad prediction skill of the residual term. In addition, the prediction skill of SCS SSHA has remarkable decadal variability.

In the future studies, we will focus on the following three aspects. First, the statistical method used in this study can be applied to each grid to investigate the spatial distribution of the predictability of SSHA in SCS. Second, the results produced by this statistical model should be compared to other statistical models or dynamical models. Third, the physical mechanism causing the decadal variability of the prediction skill of SCS SSHA should be further investigated.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgments**

The authors would like to express their gratitude to two reviewers for their helpful comments and suggestions, which contributed to greatly improve the original manuscript. This research was jointly supported by grants of National Basic Research Program (2013CB430304), National Natural Science Foundation (41176003, 41206178, 41376013, 41376015, and 41306006), National High-Tech R&D Program (2013AA09A505), and Global Change and Air-Sea Interaction (GASI-01-01-12) of China.

**References**


Advances in Meteorology


Submit your manuscripts at http://www.hindawi.com