Medium- to Long-Term Forecasts of Sea Surface Height Anomalies Using a Spatiotemporal Empirical Orthogonal Function Method

YUXIN ZHAO, a DEQUAN YANG, a WEI LI, b CHANG LIU, a XIONG DENG, a RIXU HAO, a and ZHONGJIE HE a

a College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin, China
b School of Marine Science and Technology, Tianjin University, Tianjin, China

(Manuscript received 19 February 2020, in final form 9 September 2020)

ABSTRACT: A spatiotemporal empirical orthogonal function (STEOF) forecast method is proposed and used in medium- to long-term sea surface height anomaly (SSHA) forecast. This method embeds temporal information in empirical orthogonal function spatial patterns, effectively capturing the evolving spatial distribution of variables and avoiding the typical rapid accumulation of forecast errors. The forecast experiments are carried out for SSHA in the South China Sea to evaluate the proposed model. Experimental results demonstrate that the STEOF forecast method consistently outperforms the autoregressive integrated moving average (ARIMA), optimal climatic normal (OCN), and persistence prediction. The model accurately forecasts the intensity and location of ocean eddies, indicating its great potential for practical applications in medium- to long-term ocean forecasts.

KEYWORDS: Altimetry; Empirical orthogonal functions; Forecasting techniques; Statistical forecasting

1. Introduction

Sea surface height anomalies (SSHAs) result from ocean processes and are important variables in climate research (Gill and Niller 1973; Meyssignac et al. 2017; Zhuang et al. 2010). Changes in regional sea surface height are caused by both physical ocean characteristics, such as temperature and salinity, and dynamic ocean processes, resulting in complex spatiotemporal variations (Talley 2011). Investigations of SSHA contribute to improving forecasts of the changing oceans and to revealing the dynamic mechanisms that underpin these changes.

Numerical models have played a dominant role in ocean forecasts for decades. However, there is still room for improvement in these models. Due to numerical ocean models require significant computational resources, high-precision global numerical forecasts are typically carried out on supercomputers to produce forecasts. Numerical ocean models also rely on atmospheric forcing fields as inputs. The lead time of ocean forecasts is limited by that of weather forecasts, so long-term ocean forecasts remain difficult. Finally, the construction of the initial fields and the parameterizations of physical processes used in numerical models need improvement. The long-term ocean forecasts take into account the coupled ocean–atmosphere model (Jin et al. 2008). Parameterization methods describing meteorological and oceanographic processes at different scales lead to difficulties in their coupling.

Statistical forecasts apply probability theory and mathematical statistics to analyze large amounts of historical data and to identify trends that are useful for forecasting. The advantage of the statistical forecast method is that it does not prescribe physical conditions. As long as there is sufficient correlation between data series, statistical methods can be used for forecasts. Statistical models are better able to forecast the results of physical processes that are not clearly defined. Barnston et al. (1994) compared the forecast performance of a dynamic model, a statistical model, and a dynamical–statistical hybrid model for El Niño–Southern Oscillation (ENSO). Results revealed that there was no significant difference in forecast skill among these models. In the early years of statistical forecasts, limited observation methods and insufficient observational data meant that statistical forecasts underperformed other methods. With the application of satellite observations and the availability of high-quality reanalysis products, the amount of data available for ocean regions has increased dramatically, providing support for the application of statistical forecast methods in ocean regions.

The empirical orthogonal function (EOF) analysis method, also known as principal component analysis (PCA), is used to extract dominant feature vectors from feature structures in analysis matrix data (Jolliffe 2002; Pearson 1901). Lorenz (1963) introduced the EOF into meteorological research for the first time in the 1950s, and since then it has been widely applied in the geosciences (Chen et al. 2010; Oh and Suh 2018). In recent years, it has become common practice to combine EOF techniques with statistical methods to improve forecast performance in ocean statistical forecasts (Chowdhury and Chu 2015; Lee et al. 2016; Sharma et al. 2010). Xue and Leetmaa (2000) used a Markov model to forecast sea surface temperature (SST) and sea level in the tropical Pacific. Landman and Mason (2001) forecasted SST after an EOF dimensionality reduction based on linear forecast technology, and generated results with high accuracy for ENSO amplitude. Niedzielski and Kosek (2005) proposed a P-order multiple regression model for SSHA forecasts based on the monthly SSHA observed by the TOPEX/Poseidon satellite and NOAA SST data. Niedzielski and Kosek (2009) adapted a polynomial-harmonic deterministic least squares model and its combination with an autoregressive model (LS + AR) to predict sea level anomalies.
and found that the LS+AR algorithm has outstanding forecast performance for ocean gridded data. Ubilava and Helmers (2013) applied a smooth transition autoregression (STAR) model to SST anomaly forecasts to improve the accuracy of ENSO forecasts. Results showed that the STAR model provided more accurate short-term forecasts than the linear autoregressive model. Imani et al. (2014) used an EOF to reduce the complexity of long-term series, and applied an autoregressive integrated moving average (ARIMA) model for time series forecasts to achieve the forecast of sea level in the Caspian Sea.

Ocean forecasts are not only about forecasting the temporal change of the oceans, but also about estimating their spatial distribution. From the statistical forecasts perspective, regional ocean forecast is a spatiotemporal sequence forecast problem. In the aforementioned study, the EOF was used to reduce the dimensionality, thus transforming spatiotemporal forecast problems into multiple steps time series forecast problems. The recursive approach is the most intuitive strategy for forecasting multiple steps time series (Taieb and Bontempi 2011). A well-known drawback of the recursive method is its sensitivity to estimation errors (Bontempi et al. 2012). For medium- to long-term forecasts, dozens of model iterations are typically required. The recursive method is prone to accumulate forecast errors, limiting their applicability to long-term forecasts. Therefore, medium- to long-term spatiotemporal ocean forecasts remain challenging, and further improvements are needed.

In this paper, a statistical forecast model is proposed and applied to medium- to long-term forecasts of SSHA in the South China Sea. We design the spatiotemporal empirical orthogonal function (STEOF) forecast method to incorporate temporal information into the EOF. The method applies historical data fitting to tackle spatiotemporal forecast problems, effectively avoiding error accumulation.

The remainder of this paper is organized as follows. Section 2 presents the SSHA forecast problem and EOF method. In section 3, we describe the STEOF method and the statistical forecast model. The experimental data and the results of medium- to long-term SSHA forecast experiments are given in section 4. Finally, section 5 summarizes the conclusions of this work and discusses future research.

2. Preliminaries

a. Sea surface height anomaly forecast

The purpose of the SSHA forecast is to estimate future changes in the SSHA field using regional historical data. The SSHA forecast is performed on certain spatial and temporal resolutions, so the SSHA forecast can be regarded as a spatiotemporal forecast problem.

The SSHA field is generally divided into grid points by latitude and longitude. The SSHA for each grid point interacts with its neighboring grid points. It is critical to preserve spatial correlation within the gridded data. Point-by-point forecasts without consideration of spatial correlation can lead to inaccurate forecast performance. Therefore, the SSHA forecast problem is addressed in three steps: 1) process the historical data using a dimension-reduction method; 2) use statistical methods to forecast the processed data; and 3) reconstruct the spatiotemporal forecast results.

b. Empirical orthogonal function method

Empirical orthogonal function analysis is a method used to decompose data in terms of a set of orthogonal basis functions, which are based on the data. The EOF method decomposes a spatiotemporal variable field into spatial patterns that do not change with time and time coefficients that only change with time. The spatial patterns represent the spatial characteristics of the field, and the time coefficient is a linear combination of spatial point variables in the field, called the principal component (PC) (Hannachi 2004). The first few PCs account for a large part of the total variance of the original field, and thus the dominant characteristics of a system can be captured with just a few PCs. Therefore, it is reasonable to study the temporal variations of the major components rather than the fields themselves to reveal the evolution of the field. The advantage of the EOF method is that the spatial pattern is determined by the characteristics of the variable field sequence, rather than being artificially specified in advance, and thus better represents the basic structure of the field. This method has rapid expansion and convergence, and efficiently accounts for a large amount of information.

EOF analysis is widely used in meteorology, climate research, and the geosciences. The eigenvectors correspond to spatial patterns, and the PCs correspond to temporal variations. Therefore, EOF analysis is also called spatiotemporal decomposition in the geosciences. The principle of EOF analysis is as follows.

First, the data are preprocessed into an anomaly matrix $X_{m \times n}$,

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix},$$  

where $m$ denotes the number of spatial grid points and $n$ represents the size of the time series. Second, the covariance matrix $C_{m \times m}$ of matrix $X_{m \times n}$ is calculated as

$$C_{m \times m} = \frac{1}{n}X_{m \times n}X_{m \times n}^T.$$  

Then, the eigenvalues ($\lambda_1, \ldots, \lambda_m$) and eigenvectors $V_{m \times m}$ of $C_{m \times m}$ are calculated as

$$C_{m \times m} = \lambda_{m \times m} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{bmatrix}.$$  

The eigenvalues $\lambda$ are arranged in descending order (i.e., $\lambda_1 > \lambda_2 > \ldots > \lambda_m$). Each nonzero eigenvalue corresponds to a column of eigenvectors, also referred to as the spatial pattern. For example, the eigenvector corresponding to $\lambda_1$ is called the first spatial pattern (i.e., the first column of $V_{m \times m}$) and so on.
The spatial patterns are projected onto the matrix \( X_{m \times n} \) to obtain the PC corresponding to the eigenvector:

\[
\text{PC}_{m \times n} = V_{m \times n}^T \times X_{m \times n}.
\]

(5)

The data of each row in the \( \text{PC}_{m \times n} \) correspond to the PCs of each column of eigenvectors. The PC of the first spatial pattern corresponds to the first row of \( \text{PC}_{m \times n} \), and so on.

3. Methodology

a. Spatiotemporal empirical orthogonal function

Here, we develop the STEOF method, which integrates temporal information into EOF spatial patterns. We propose

\[
X = \begin{bmatrix}
x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\
\vdots & \ddots & \ddots & \vdots \\
x_{m,1} & x_{m,2} & \cdots & x_{m,n}
\end{bmatrix},
\]

\[
X = \begin{bmatrix}
x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\
\vdots & \ddots & \ddots & \vdots \\
x_{m,1} & x_{m,2} & \cdots & x_{m,n}
\end{bmatrix}^T
\]

(6)

where \( X \) denotes the sample matrix, \( n \) represents the number of spatial grid points, \( m \) denotes the number of time series, and \( k \) is the number of periods.

We usually employ the Jacobi method to solve the eigenvalues and eigenvectors of the covariance matrix of \( X \). When the rank of the matrix is large, the Jocobi method has a huge amount of computation. The number of spatiotemporal grid points \( m \times n \) is much larger than the number of periods \( k \). Hence, matrix transformation is performed to reduce the computational cost. It is evident that \( C = X \cdot X^T \) and \( C^* = X^T \cdot X \) have the same nonzero eigenvalues, but their eigenvectors differ. Therefore, after the eigenvectors of the \( C^* \) matrix are obtained by matrix transformation, the eigenvectors of the \( C \) matrix can be calculated as

\[
C^* \times V^* = C^* \times \Lambda
\]

(7)

where \( V^* \) denotes the eigenvector of \( C^* \), \( \Lambda \) represents the diagonal matrix of the eigenvalues, and

\[
V_k = \frac{1}{\sqrt{\lambda_k}} X \times V^*,
\]

(8)

where \( V_k \) is the first \( k \) eigenvectors of \( C \).

The eigenvector is a time series of a spatial pattern, which contains both spatial information and temporal information, which we name the spatiotemporal base. There are many common orthogonal spatiotemporal bases. Each spatiotemporal base represents the variation of a common spatial pattern with time. Therefore, the STEOF method extracts the primary characteristics of the temporal variations of a spatial pattern based on historical data.

b. Spatiotemporal empirical orthogonal function forecast method

A spatiotemporal base can be used as a basis vector. The larger the data sample is, the more spatiotemporal bases can be decomposed. A set of spatiotemporal bases is sufficient to construct a complete space when the set is large enough. The linear combination of spatiotemporal bases in complete space has the capacity for constructing any vector in the space.

Using the spatiotemporal bases from STEOF decomposition, a spatiotemporal forecast can be transformed from a time extrapolation problem into a spatiotemporal base fitting problem. The decomposition results of multiple spatiotemporal sequences establish a set of spatiotemporal bases. Then, spatiotemporal observations and spatiotemporal bases can be used to forecast the spatiotemporal series:

\[
O = \begin{bmatrix}
o_{1,1} & o_{2,1} & \cdots & o_{n,1} \\
o_{1,2} & o_{2,2} & \cdots & o_{n,2} \\
o_{1,k} & o_{2,k} & \cdots & o_{n,k}
\end{bmatrix}^T.
\]

(9)

The spatiotemporal bases \( M \) are divided into two parts: the fitting spatiotemporal bases \( M_i \) that are of the same period as the spatiotemporal observations, and the forecast spatiotemporal bases \( M_{ij} \).
where \( t \) denotes the forecast start time, \( n \) denotes the number of spatial grid points, \( l \) represents the number of observation times, \( p \) is the number of forecast time steps, and \( k \) is the \( k \)th spatiotemporal base.

The eigenvectors are orthogonal to each other; i.e., the spatiotemporal bases are linearly independent. For linearly independent bases, least squares estimation (LSE) is the optimal fitting method. We employ the LSE method to solve fitting coefficients of spatiotemporal observations and fitting spatiotemporal bases. The fitting coefficients are the projection of the spatiotemporal observations on each spatiotemporal base, and describe the similarity between a set of observations and a spatiotemporal base:

\[
O = M_f \cdot S,
\]

where \( S \) denotes the fitting coefficients and \( k \) represents the \( k \)th pattern.

Each spatiotemporal base can be regarded as a change rule of spatiotemporal sequence. When the spatiotemporal sequence is similar to a rule in the fitting stage, it is reasonable that the change of the spatiotemporal sequence conforms to the rule in the forecasting stage. We forecast the future spatiotemporal sequence by reconstructing fitting coefficients and the forecast spatiotemporal bases. Therefore, the STEOF forecast model combines the STEOF method with LSE to forecast a spatiotemporal sequence.

\[
Y = M_p \cdot S = [y_{1,t+1} \ y_{2,t+1} \ \cdots \ y_{n,t+1} \ y_{1,t+2} \ y_{2,t+2} \ \cdots \ y_{n,t+2} \ y_{1,t+p} \ y_{2,t+p} \ \cdots \ y_{n,t+p}]^T,
\]

where \( Y \) denotes the forecast result, \( n \) denotes the number of spatial grid points, \( t \) represents the forecast start time, and \( p \) is the number of forecast time steps.

4. Model design and performance test results

a. Experimental setup

The data used are the ocean reanalysis products of the northwest Pacific Ocean in China Ocean Reanalysis (CORA) developed and released by the National Marine Data and Information Service Center (http://cora.nmdis.org.cn/). The ocean dynamic model used to generate the data is the Princeton Ocean Model with Generated Coordination System (POMgcs), which considers the impact of wave breaking and tide mixing on the vertical distribution of sea temperature (Han et al. 2013).
Table 1. CORA regional ocean reanalysis.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporal range</td>
<td>1 Jan 1958–31 Dec 2016</td>
</tr>
<tr>
<td>Spatial range</td>
<td>10°S–52°N, 99°–150°E</td>
</tr>
<tr>
<td>Temporal resolution</td>
<td>Daily average</td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>1/2°–1/8° (variable grid)</td>
</tr>
</tbody>
</table>

generally high in the northwest and low in the southeast. There are two distinct ocean processes at work here: the South China Sea western boundary current and the Luzon cyclone cold vortex (Wang et al. 2012). The study area is the region 0°–24°N, 99°–125°E, which covers the majority of the South China Sea (Fig. 1).

The analysis focuses on daily SSHA variations. The SSHA data of the study area is extracted from the CORA data. A 0.5° × 0.5° resolution grid is used for the experiments. We run the experiments on a work station of Intel Xeon W-2133 CPU @3.60 Hz, 32 G RAM with Windows 10 Education 64-bit operating system.

The interannual variations of SSHA show significant seasonality (Cheng et al. 2016). For sea surface height forecasts, it is reasonable to select 1 year as the period for the STEOF. Therefore, when constructing sample matrices, each row of the matrix is a daily time series of different spatial grid points in the same year, and each column of data is a series of different years at the same spatiotemporal position.

We use reanalysis data to complete STEOF decomposition and coefficient fitting. During the model evaluation, we refer to the reanalysis data corresponding to the forecast period as the actual value. To verify the effectiveness of the STEOF model, we evaluate the method using several commonly used metrics, namely, the root-mean-square error (RMSE) and the correlation coefficient (r). The error formulation and the performance metrics are expressed as follows:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_{i,j} - X_{f,i,j})^2},
\]

where \(X_{i,j}\) denotes the actual values and \(X_{f,i,j}\) denotes the forecast values at position \(i\), \(i\) denotes an element index in row–column order, and \(n\) denotes the total number of grid points, and

\[
r = \frac{\sum_{i=1}^{n} (X_{i,j} - \bar{X}_j)(X_{f,i,j} - \bar{X}'_j)}{\sqrt{\sum_{i=1}^{n} (X_{i,j} - \bar{X}_j)^2 \sum_{i=1}^{n} (X_{f,i,j} - \bar{X}'_j)^2}},
\]

where \(\bar{X}_j\) denotes the mean of \(X_n\), and \(\bar{X}'_j\) denotes the mean of \(X_f\).

The RMSE indicates the overall difference between the forecast results and the actual values, and \(r\) indicates the similarity between the forecast results and the actual values in the spatiotemporal domain. The smaller the RMSE, the more accurate the forecast results. The closer \(r\) is to 1, the more similar a forecast distribution to the actual distribution.

b. Algorithm validation

First, we carry out SSHA forecast experiments to demonstrate the effectiveness of the algorithm. The experiment is a forecast of SSHA for April 2015. The spatiotemporal bases are STEOF-decomposed from historical SSHA data (1958–2014) from March to April. We use the data for March 2015 and the corresponding fitting spatiotemporal bases to estimate the fitting coefficient. The fitting coefficient and the forecast spatiotemporal bases are used to conduct the forecast. Because the number of spatiotemporal bases affect the fit, we selected historical data for different years in the experiment.

The 30- (1958–87), 40- (1958–97), 50- (1958–2007), and 57-yr (1958–2014) data are individually decomposed to explore the influence of spatiotemporal base number on the forecast performance. In general, the amount of historical data is insufficient to obtain complete spatiotemporal bases. To explore conditions where the spatiotemporal bases are complete, we include data for the forecast year in the STEOF decomposition in one experiment; i.e., the data for all 58 years are used for the STEOF decomposition. Figure 2 shows the spatial distribution of actual values and forecast results for the different sets of spatiotemporal bases.

Results indicate that the number of spatiotemporal bases significantly affects the forecast performance. Specifically, the general distribution of SSHA is roughly accurate for the 30-yr dataset, but the warm eddy in the central South China Sea is not well captured. The eddy in the northern South China Sea is captured using the 40-yr dataset, but the intensity of the eddy in the central South China Sea is weak. The warm eddy from the 50-yr dataset is weaker than the actual eddy and east of the actual location. The 57-yr dataset captures the SSHA rise in

![Fig. 1. The study area.](Image)
the Gulf of Thailand. The intensity and location of the eddies are close to the actual values, but the radius of the warm eddy is too large. As the number of spatiotemporal bases increases, the similarity of the spatial distributions of the forecast and actual values increases, and the forecast mesoscale eddy positions and intensities become more accurate. When a complete set of spatiotemporal bases are used, the forecast results are completely consistent with the actual values. The RMSE and correlation coefficient between the forecast results and actual values for the different base sets are shown in Figs. 3 and 4.

The trends of forecast error for 30, 40, and 50 years are similar, indicating that the spatiotemporal bases that dominate the forecasting process are analogous. However, the trend in forecast error of 17–23 days for 57 years is different from the others. The principal component of the first spatiotemporal base in 30, 40, and 50 years is 16.44%, 16.35%, and 16.55%, respectively. The principal component of the first spatiotemporal base in 57 years is 28.78%. It is shown that the first spatiotemporal base from the decomposition of the 57-yr data is significantly different from the base of the other data. The variation in the 57-yr spatiotemporal base during the forecast period differs from the other bases, resulting in a different forecast error trend than the others. The spatiotemporal base set from the decomposition of the 57-yr data contain a spatiotemporal base that is more consistent with the evolution of the forecast year. The general trend is that with an increase in the number of spatiotemporal bases, the forecast error noticeably decreases, and the stability of the correlation coefficient improves significantly. For the 57-yr spatiotemporal base set forecast results, the correlation coefficient remains >0.95 with little variation. Therefore, with the accumulation of historical samples, the applicability of the spatiotemporal base set to forecasts increases, as it is better able to capture observed spatiotemporal variations. In the 58-yr spatiotemporal base set forecast experiment (the complete spatiotemporal base set scenario), the STEOF decomposition includes the spatiotemporal variations of the forecast period. Results from this experiment show that the STEOF forecasts are perfect for 30 consecutive days. It is important to note, however, that this experiment evaluates the

**FIG. 2.** Forecast performance of the 15-day forecast based on different spatiotemporal bases. (a) The forecast result based on the spatiotemporal bases for 30-yr historical data. (b)–(e) As in (a), but for the results of 40, 50, 57, and 58 years (including forecast information), respectively. (f) The actual value.

**FIG. 3.** RMSE between different spatiotemporal bases.
effectiveness of the method in the idealized case of a complete set of spatiotemporal bases. To evaluate more realistic scenarios, the forecast period data are not included in the STEOF decompositions in subsequent experiments. The forecast performance in these experiments is expected to be inferior to the idealized case.

The running time of the algorithm for different numbers of historical samples are presented in Table 2. The experiments show that the running time becomes longer as the sample number increases. The running time of the 58-yr historical sample is 269.351 s, which is much smaller than the running time of the numerical forecasts. We conclude that STEOF is an extremely efficient forecasting algorithm.

c. Parameter determination

Experiments are carried out with different numbers of fitting and forecast days, and assessed their effects on forecast performance. As discussed above, the amount of historical data affects the forecast performance. Accordingly, data from 1958 to 2014 (57 years of historical data in total) are used in the STEOF decomposition. We estimate the fitting coefficients and evaluate the forecast performance based on the 2015 data. We design experiments to explore the predictability of the forecast model with different fitting periods. Taking the 30-day forecast as an example, data from different periods are used for the fitting. The experiments are conducted with data for 10, 20, 30, 60, and 90 days. The RMSE and correlation coefficients of the forecasts are listed in Table 3.

In this experiment, the forecast performance improves as the number of fitting days increases. The RMSE is the smallest for 90 fitting days. Forecast performance is good enough when the number of fitting days is 30. There is no visible difference between the correlation coefficients for more than 30 fitting days. Figure 5 shows the spatial distribution of the forecast results for different numbers of fitting days. With 10 or 20 fitting days, the warm eddy radius in the center of the South China Sea is larger than the actual radius. The 30–90 fitting-day results accurately capture the distribution of SSHA, with negligible differences among the forecasts.

We conduct experiments to investigate the performance of different forecast days. The 57-yr historical dataset is decomposed to construct the spatiotemporal bases. Data for 2015 are used for coefficient fitting and forecasting. In each experiment, the number of fitting days is the same as the number of forecast days. The forecast performance for 7, 15, 30, 60, 90, and 120 days are compared. The RMSE and correlation coefficients for the forecast results using various numbers of forecast days are listed in Table 4.

The RMSE of the 90-day forecast is 4.4 cm, and the correlation coefficient is 0.967. A greater number of forecast days means a greater number of forecast grid points. It becomes increasingly difficult to forecast each grid point accurately, and forecast error increases slightly. However, even when the number of forecast days increases, the model still provides accurate forecasts. There is no dramatic increase in error for long-term forecasts.

As is evident in Fig. 6, the fluctuation of RMSE is affected by the variation of SSHA. The ratio of the standard deviation to RMSE represents the signal-to-noise ratio. Results reveal that the forecast algorithm has high precision and the performance is acceptable. As the number of forecast days increases, the error increases slightly before becoming stable. This indicates that the STEOF forecast method can be used to provide accurate medium- to long-term SSHA forecasts, with errors that do not increase significantly with the number of forecast days.


d. Seasonal performance

We perform experiments to compare the forecast performance of the STEOF method across seasons. To evaluate the performance of the method for different seasons, forecast experiments are conducted for four different periods: spring (March–May), summer (June–August), autumn (September–November), and winter (December–February). As there is a cross-year forecast for winter, we use 56 years of the historical data for the STEOF decomposition. We forecast the SSHA of the next season based on the data of the previous season (Table 5). The RMSE and correlation coefficients of the forecasts for different seasons are provided in Table 6.

Results indicate that the highest correlation coefficient is found for summer forecast, implying that its spatial distribution

<table>
<thead>
<tr>
<th>Metrics</th>
<th>10 days</th>
<th>20 days</th>
<th>30 days</th>
<th>60 days</th>
<th>90 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE(m)</td>
<td>0.041</td>
<td>0.039</td>
<td>0.038</td>
<td>0.038</td>
<td>0.037</td>
</tr>
<tr>
<td>r</td>
<td>0.951</td>
<td>0.957</td>
<td>0.959</td>
<td>0.959</td>
<td>0.959</td>
</tr>
</tbody>
</table>

FIG. 4. The correlation coefficient between different spatiotemporal bases.
is most consistent with the actual value. The spring forecast has the smallest RMSE, meaning it has the smallest overall difference from the actual value. The overall assessment is that the summer forecast is more accurate than all others, and the winter forecast performs relatively poorly and could be further improved. The annual RMSE is 5.3 cm and the correlation coefficient is 0.943.

Figure 7 shows the STEOF forecast results for different numbers of forecast days in different seasons. For the same number of forecast days, the winter forecast has the largest errors among the seasons. The forecast overall distribution of SSHA, eddy positions and radii are consistent with the actual values. In addition, short-term (7 days), medium-term (30 days), and long-term (60 days) seasonal forecasts demonstrate remarkable performance. Results accurately capture the SSHA trend, eddy movement, and changes in eddy intensity in each season. Hence, the forecast model demonstrates good generalization ability for each season.

e. Forecast performance comparison

To evaluate the forecast performance of the STEOF method, it is compared with three other statistical forecast methods: persistence prediction, optimal climatic normal (OCN) and ARIMA methods. Persistence prediction is the most convenient method of statistical forecasting. The simplest form of persistence prediction is to use the value of the previous moment as the forecast for the next moment. When the variable at time \( t \) is \( X_t \), then the forecast for the variable at time \( t + 1 \) is

\[
X_{t+1} = X_t. \tag{18}
\]

The method commonly used for medium- to long-term ocean forecasts is the OCN, which uses the average of \( k \)-year historical data as the forecast for the next year (Huang et al. 1996). Suppose \( X_s, s = 1, 2, \ldots, n \), is a time series of data at yearly intervals. The OCN \( X_{s,k} \) is constructed as follows:

\[
\bar{X}_{s,k} = \frac{1}{k} \sum_{j=1}^{k} X_{s-j}, \tag{19}
\]

where \( n \) denotes the total number of years, and the average of \( k \) years is used as the forecast.

The ARIMA method is a time series forecast method, which is usually combined with the EOF method for ocean forecasts (Box et al. 2015). The ARIMA model is one of the most popular time series forecast models (Kumar et al. 2009). In this method, the forecast is assumed to be a linear function of historical data and random errors. The ARIMA model is a function of \( p, d, \) and \( q \), and is calculated as

\[
\left( 1 - \sum_{i=1}^{p} \phi_i L^i \right) (1 - L)^d X_t = \left( 1 + \sum_{j=1}^{q} \theta_j L^j \right) \epsilon_t, \tag{20}
\]

| TABLE 4. Forecast results for different forecast days. |
|---|---|---|---|---|---|---|---|
| Metrics | 7 days | 15 days | 30 days | 60 days | 90 days | 120 days |
| RMSE(m) | 0.035 | 0.040 | 0.043 | 0.044 | 0.044 | 0.0457 |
| \( r \) | 0.977 | 0.968 | 0.964 | 0.965 | 0.967 | 0.9648 |
where $p$ denotes the order of the autoregressive part of the model, $q$ denotes the order of the moving average part of the model, $d$ denotes the degree of regular differencing, $\phi_i$ denotes the autoregressive parameters, $\theta_j$ denotes the moving average parameter, $L$ denotes the lag operator, and $\varepsilon_t$ denotes the noise component of the stochastic model assumed to be NID $\left(0, \sigma^2 \right)$. The orders of the model are determined based on autocorrelation functions and partial autocorrelation functions. The set of parameters that minimize the Akaike information criterion and the Bayesian information criterion of the model is selected as the optimal parameter set.

We adopt the above methods to forecast the SSHA from April to May 2015. Each method selects the best parameter setting through experiments. For the STEOF forecasts, the STEOF decomposition is carried out using data from February to May for 57 years. We use data from February to March 2015 for coefficient fitting and forecasting. The fitting days are 60 days and the forecast days are 60 days. For the persistence prediction, the actual value on 31 March 2015 is taken as the forecast result. For the OCN, the average of April–May data from 1985 to 2014 is used as the forecast result. For the ARIMA model, data from 1958 to 2014 are used for the EOF decomposition to obtain spatial patterns and principal components. The first three modes with more than 70% of principal components are selected for the forecast experiment. We employ an ARIMA model to estimate the variation of the principal component with time. The autoregressive order is 4, the difference order is 1, and the moving average order is 3. The principal component forecast result and the corresponding spatial pattern are used to reconstruct the spatiotemporal forecast for a certain mode, and the reconstruction results of different modes are superimposed to achieve the SSHA forecast. The forecast results from the different methods are shown in Fig. 8.

Experimental results indicate that the STEOF model performs better than the persistence prediction, OCN and ARIMA models. The sea surface height in the coastal areas of Guangdong decrease, the eddy in the central South China Sea gradually forms, and the SSHA in the Sulu Sea near the Philippines noticeably increases, in agreement with the actual value. The persistence prediction has the same forecast results at different forecast days. The results are only indicative of the initial state of the ocean and do not reflect changes in the ocean. The OCN method can only forecast the overall trend of sea surface height, and its forecast results are significantly different from the actual values, which indicates that the OCN method insufficiently captures ocean phenomena. In short-term forecasts, the spatial distribution of SSHA forecasted by ARIMA is similar to the actual values. In long-term (60-day) forecasts, the ARIMA forecast results near Philippine Island are noticeably higher than actual values, and the eddy positions are not accurate. Compared with ARIMA, the STEOF forecast results show more small-scale marine variations, and more accurately forecast the radii and locations of eddies, especially in coastal areas. The STEOF method shows remarkable performance for long-term forecasts. As shown in Fig. 8, there is an eddy at 7°N, 112°E for the 60th day in actual value. The eddy originates from the currents near the Malaysian Islands, which pass along the Rossby wave with time. Comparing the forecast results of different methods, only the STEOF method can accurately forecast the location, radius and intensity of the eddy. Persistence prediction and OCN cannot forecast the location of the eddy. The spatial patterns of ARIMA are fixed and it is difficult to show the continuous movement of the current. The spatiotemporal bases of STEOF method contain long-term evolving spatiotemporal information with the capability of capturing Rossby wave. As a result, the STEOF method effectively forecasts the dynamical processes in the South China Sea, a capability that is not available with other statistical methods.

The RMSE between actual values and forecast results of persistence prediction, OCN, ARIMA, and STEOF models is shown in Fig. 9. The error of the STEOF method is stable in the range 2–5 cm within 60 days, with an average RMSE of 3.78 cm. The persistence prediction has very small errors in the initial stages of forecasting, indicating its suitability for a very short-term forecast. However, it does not work well for long-term forecasting.

### Table 5. Forecast data for different seasons.

<table>
<thead>
<tr>
<th>Season</th>
<th>Fitting days</th>
<th>Forecast days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autumn</td>
<td>Jun–Aug 2015</td>
<td>Sep–Nov 2015</td>
</tr>
</tbody>
</table>

### Table 6. Forecast results for different seasons.

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE(m)</td>
<td>0.064</td>
<td>0.047</td>
<td>0.048</td>
<td>0.052</td>
</tr>
<tr>
<td>$r$</td>
<td>0.934</td>
<td>0.935</td>
<td>0.953</td>
<td>0.948</td>
</tr>
</tbody>
</table>
FIG. 7. Forecast performance of different seasons on SSHA in the South China Sea. (a) (top) The forecast result of SSHA on day 7, and (bottom) the actual value of SSHA on the day 7 for (from left to right) winter, spring, summer, and autumn. (b),(c) As in (a), but for the result of day 30 and day 60, respectively.
forecasts, with an average error of 10.13 cm. The average RMSE of the OCN method is 11.22 cm, which is much larger than that of the other methods. The average RMSE of the ARIMA method is 7.4 cm, and the forecast error fluctuates dramatically, increasing over time.

Figure 10 shows the correlation coefficient between actual values and forecasts of persistence prediction, OCN, ARIMA and STEOF models. Because the STEOF method preserves spatial patterns in the data, it can accurately forecast the spatial distribution of SSHA, with an average correlation coefficient of 0.96. Compared with the STEOF method, the correlation coefficient of the ARIMA method is less stable and fluctuates significantly for long-term forecasts. The average correlation coefficient of the ARIMA forecast method is 0.86. The highest correlation coefficient of the OCN method is 0.81 and the average correlation coefficient is 0.72. The correlation coefficients for the persistence prediction show a significant downward trend with time, with an average correlation coefficient of 0.67. This indicates that the STEOF method accurately forecasts the spatiotemporal variations of the ocean, and retains a high correlation with the actual value, even for long-term forecasts.

5. Conclusions

In this paper, the STEOF forecast method is proposed for medium- to long-term statistical forecasts of SSHA. The STEOF improves forecasts of the spatial distributions of SSHA by incorporating temporal information in the spatiotemporal base. We reformulate SSHA forecasts as a spatiotemporal forecast problem, and convert it into a base fitting problem using the STEOF forecast method. Hence, the method avoids forecast

![Figure 8](image_url)  
**Figure 8.** Forecast performance of different methods on SSHA in the South China Sea. (a) The forecast result and actual value for day 7 for (from left to right) persistence prediction method, OCN method, ARIMA method, STEOF method, and actual values. (b)–(d) As in (a), but for the result of day 15, day 30, and day 60, respectively.

![Figure 9](image_url)  
**Figure 9.** RMSE of STEOF method (blue line), ARIMA method (red line), OCN method (yellow line), and persistence prediction method (purple line) for SSHA forecast result.
The STEOF forecast method has demonstrated remarkable performance for medium- to long-term forecasts. Results show that the RMSE within 60 days is 3.78 cm and the correlation coefficient is as high as 0.96. Thus, the proposed STEOF forecast method has higher forecast accuracy and consistently outperforms persistence prediction, OCN and ARIMA algorithms. The STEOF forecast method has demonstrated remarkable performance for medium- to long-term forecasts. Moreover, compared with numerical forecasts, the STEOF method requires less computational resources and has the potential for application to real-time forecast systems.

This work has focused on the development of a spatiotemporal decomposition method and using linear spatiotemporal bases to forecast SSHA. Because the STEOF prediction method relies on historical data, it has insufficient forecasting capacity for extreme phenomena. The accuracy of the STEOF method for forecasting SSHA in winter needs to be further improved. The ocean is a complex dynamic system, in which there are many nonlinear processes. The STEOF model with linear bases is poorly able to effectively forecast small-scale ocean phenomena. In future work, we are about to study nonlinear forecast methods and their performance in forecasts of finescale ocean phenomena. Besides, we will attempt to conduct forecast experiments with observational data. Furthermore, we will investigate the forecast performance of the STEOF method on ocean multivariate in the South China Sea, such as sea surface temperature, sea surface salinity, and sea surface velocity.

**Acknowledgments.** This research was funded by National CMOSt Key Research and Development projects (2017YFC1404100, 2017YFC1404104 and 2018YFC1406202), the NSFc (51379049, 41676088, 41775100, and 41830964), and the HEU and CSC (which supported Dequan Yang studying abroad at Technical University of Munich for 2 years).

**Data availability statement.** The sea surface height anomaly data that support the findings of this study are available from China Ocean ReAnalysis (http://cora.nmdis.org.cn/).

**REFERENCES**


