Research Article

Estimation of Sea Level Change in the South China Sea from Satellite Altimetry Data

Shanwei Liu,1 Yue Jiao,1 Qinting Sun,2 and Jinghui Jiang1

1College of Oceanography and Space Informatics, China University of Petroleum Qingdao, Qingdao 266580, China
2School of Geosciences, China University of Petroleum Qingdao, Qingdao 266580, China

Correspondence should be addressed to Shanwei Liu; shanweiliu@163.com and Qinting Sun; 1136613517@qq.com

Received 19 November 2020; Revised 11 January 2021; Accepted 27 January 2021; Published 16 February 2021

Academic Editor: Shaukat Ali

Copyright © 2021 Shanwei Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The South China Sea is China’s largest marginal sea area, and it is rich in oil and gas mineral resources; thus, estimating its sea level changes is of practical significance. Based on linear and nonlinear sea level change characteristics, this paper decomposes 1992–2019 monthly mean sea level anomaly time series in the South China Sea into trend, seasonal, and random terms. This paper compares the seasonal autoregressive integrated moving average (SARIMA) and Prophet models for estimating the trend and seasonal terms and the long short-term memory (LSTM) and radial basis function (RBF) models for estimating random terms, and the more suitable models were selected. A Prophet-LSTM combined model was developed based on the accuracy results. This paper uses the combined model to study the effect of known data length on the experimental results and determines the best prediction duration. The results show that the combined model is suitable for short-term and medium-term estimations of 12–36 months. The accuracy at 36 months is 0.962 cm, which proves that the combined model has high application value for estimating sea level changes in the South China Sea.

1. Introduction

The South China Sea is China’s largest marginal sea and is a transportation hub for maritime energy transportation; it also has abundant reserves of oil and gas resources [1]. In recent years, the sea level of the South China Sea has continued to rise [2], and, over the next 30 years, the sea level along the coast will also rise by 50–180 mm [3], which not only will have serious impacts on the natural environment, ecosystems, and social economy of the coastal areas [4] but will also pose challenges to maritime transport and energy extraction. Research on these changes can assist in measuring regional climate change, contribute disaster warning information, and provide a scientific basis for the coordination of maritime traffic and the rational exploitation of energy. Therefore, research on the estimation of sea level change trends in the South China Sea is necessary.

In terms of sea level estimation, predecessors tried different methods based on mathematical statistics, physical mechanisms, or combined model predictions according to different regions and data types. Mathematical statistics methods are mainly based on the mathematical law of sea level change time series, which is used to fit and extrapolate data [5]. Early mathematical statistical methods included simple linear regression [6], multivariate stepwise regression, maximum entropy spectrum analysis [7], and Kalman filtering [8]. The physical mechanism method considers climate change and sea temperature and salinity. Chen and others used the CCSM3 climate system model to simulate sea level changes, and the results showed that the global sea level will rise by 30 cm in the 21st century [9]. Zhang and others used the sea-atmosphere coupled model to predict the trends and spatial distribution of sea level in the South China Sea in the 21st century [10]. The use of these methods is relatively simple and does not take into account the non-stationary and uncertain characteristics of sea level changes. With the emergence of various neural network models, the estimation of nonstationary characteristics of sea level changes has become more possible; moreover, the combination of mathematical methods and neural networks can
improve accuracy. Zhao and others used a model that combined least squares and radial basis function neural network to predict sea level anomaly series in offshore China, and the reliability of the model for short-term predictions was demonstrated, and the accuracy reached 0.65 cm [11]. Among the methods selected in this article, SARIMA is widely used in epidemiological prediction [12] and the Prophet model has good performance in user traffic prediction [13]. As neural network models, LSTM and RBF have certain application value in rainfall and river flow predictions [14, 15].

Currently, few studies have focused on the estimation of sea level changes in the South China Sea. To more accurately estimate the change trends of the South China Sea, this paper divides the monthly average sea level abnormal time series according to the sea level change characteristics in the South China Sea as follows:

(i) Trend term and seasonal term combination series
(ii) Random term series

In terms of model selection, the SARIMA [12] and Prophet [16] models are suitable for fitting stationary and trending series and the LSTM [17] and RBF [18] models perform well when fitting nonlinear and random series. Therefore, the SARIMA and Prophet models are selected to fit the combination series, the RBF and LSTM models are selected to fit the random term series, and a combination of estimation models which is more suitable for the study area is determined by accuracy evaluation standards, such as RMSE and $R^2$. The experimental steps are shown in Figure 1.

2. Material and Methods

2.1. Data and Study Area. The satellite altimetry data used in the experiment are from the GDR data sets of T/P, Jason-1, Jason-2, and Jason-3, which were all released by the French Space Centre (CNES). This paper processes the data in accordance with the steps in [19]. The data obtained after processing are the mean sea level anomaly (SLA) time series data from October 1992 to December 2019 in units of months, as shown in Figure 2.

The research area of this paper is the South China Sea and it is shown in Figure 3. The coordinates are 110°–119°E, 14°–23°N, and the total area is approximately 1.19 million square kilometers. The area is located between the Pacific Ocean and the Indian Ocean, and it includes many important shipping lanes for material transportation. In addition, there are abundant oil and gas resources.

2.2. Seasonal Autoregressive Integrated Moving Average Model. The SARIMA model evolved based on the ARIMA model, which takes into account the seasonal factors of the time series [20–22]. It adopts the method of seasonal difference to estimate parameters and can effectively predict time series with seasonality, trend, and periodicity. The SARIMA model has performed well in industrial and medical research in recent decades [23]. The general form of the SARIMA model is $\text{SARIMA} (p, d, q) (P, D, Q) [S]$, where $p$ is the autoregressive order, $P$ is the seasonal autoregressive order, $q$ is the moving average order, $Q$ is the seasonal moving average order, $d$ is the difference order, $D$ is the seasonal difference order, and $S$ is the seasonal period. The SARIMA model is expressed in the following equation:

$$\phi_p(B)\Phi_P(B^S)(1 - B)^d(1 - B^S)^D y_t = \theta_q(B)\Theta_Q(B^S)\mu_t,$$

(1)

where $y_t$ is the time series, $\mu_t$ is the random term, $\Phi_p(B)$ is the nonseason AR($p$) part, $\Phi_P(B^S)$ is the season AR($P$) part, $(1 - B)^d$ is the $d$-order progressive difference, $\theta_q(B)$ is the nonseason MA($q$) part, and $\Theta_Q(B^S)$ is the season MA ($Q$) part.
2.3. Prophet Model. The Prophet model is a time series curve-fitting tool developed by Taylor and Letham [24]. This model is suitable for fitting time series with strong seasonal effects and has strong robustness against missing values, abnormal values, and trend changes. The basic form of the model is shown in the following equation:

\[ P(t) = g(t) + s(t) + h(t) + \varepsilon_t, \]  

where \( g(t) \) is the trend term used to fit aperiodic changes in the time series, \( s(t) \) is the periodic term and it uses a Fourier series to approximate the periodic component, \( h(t) \) is the holiday term, and \( \varepsilon_t \) is the error term. Prophet is robust to missing data and trend changes and usually handles abnormal values well.

2.4. Long Short-Term Memory Model. Hochreiter and Schmidhuber proposed the LSTM model [25–28], which is considered a special recurrent neural network. It can solve the problems of gradient disappearance and gradient explosion and automatically learn the sequence features; therefore, it performs better in the prediction of longer time series and is more suitable for the prediction of sea level changes. The historical update information in LSTM model is controlled by the input gate, forget gate, and output gate as shown in the following equation:

\[
\begin{align*}
\bar{c}_t &= \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c), \\
i_t &= \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i), \\
f_t &= \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f), \\
c_t &= f_t \odot c_{t-1} + i_t \odot \bar{c}_t, \\
o_t &= \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o), \\
h_t &= o_t \odot \tanh(c_t).
\end{align*}
\]  

The steps are given as follows:

1. Calculate the candidate memory unit value \( \bar{c}_t \) at the current moment, where \( W_{xc} \) and \( W_{hc} \) represent the input data and the unit output weight at the previous moment, respectively.
2. Calculate the input gate value, where \( x_t \) represents the current input data, \( h_{t-1} \) represents the unit output value at the previous time, and \( C_{t-1} \) represents the memory unit value at the previous time.
3. Calculate the forget gate value \( f_t \) and control the influence of historical information on the current state.
4. Calculate the state value \( h_t \) at the current moment, where \( \odot \) represents the point-by-point product.
5. Calculate the output gate value \( o_t \).
6. Calculate the output of the final LSTM unit.

2.5. Radial Basis Function Neural Network Model. The RBF neural network model is a feedforward neural network, and it has strong nonlinear mapping ability and can approximate a nonlinear function with arbitrary precision [29, 30]. Therefore, it is more suitable for the prediction of random terms in time series. The network consists of an input layer, a hidden layer, and an output layer. The radial basis function acts on the high-dimensional mapping between the input layer and the hidden layer, and the linear least square method is used to calculate the weight between the hidden layer and the output layer. The model is generally expressed as shown in the following equation:

\[ y_t = \tilde{y}_t + \varepsilon_t = \sum_{i=1}^{M} \varphi_i(x_t) \omega_i(t) + \varepsilon_t, \]

where \( M \) is the number of hidden layers, \( \varphi_i(x_t) \) is the radial basis function of the \( i \)th node, \( \omega_i(t) \) is the corresponding regression coefficient, and \( i = 1, 2, \ldots, M. \)

3. Results and Discussion

The curve in Figure 2 approximately reflects the characteristics of sea level changes in the study area. The peak value corresponds to summer and autumn, and the valley value corresponds to winter and spring. The sea level changes in this area are seasonal and cyclical. To consider both linear and nonlinear features when predicting sea level changes, the monthly mean SLA time series in the study area is decomposed into “season,” “trend,” and “random” terms based on the principle of addition (Figure 4).

In Figure 4(a), the trend term indicates that as the global climate warms, the sea level of the South China Sea will gradually rise, and the year corresponding to the falling part of the curve will correspond to time periods with strong El Niño phenomena, which fully reflects the trend of a slow rise and occasional decline in sea level observed from 1992–2019. In Figure 2(b), the seasonal term presents the same changes every year, thus reflecting the seasonal characteristics of sea level changes.
level changes. In Figure 2(c), the noise generated by uncertain factors during sea level change is presented. This paper chooses 1992–2017 SLA data as the training set and 2018–2019 SLA data as the test set. Then, according to both the stationary and nonstationary characteristics of sea level change, the above decomposition results are divided into two groups: trend-seasonal series and random term series. The SARIMA and Prophet methods are used to fit the seasonal trend series, and the LSTM and RBF models are used to fit random term series. The root mean square error and the coefficient of determination $R^2$ were used as the criteria for evaluating the estimation results.

3.1. Estimation of Trend-Seasonal Series

3.1.1. Trend-Seasonal Series Estimated Using the SARIMA Model. The selection and fitting of the SARIMA model mainly include the following steps:

(i) Determine the main structure of SARIMA $(p, d, q)$ $(P, D, Q)$ through experience and autocorrelation and partial autocorrelation function graphs.

(ii) Experiment to obtain other unknown parameters.

(iii) Evaluate the degree of fit through a residual test.

(iv) Perform predictions based on known data [23].

From the above steps, $D = 1$, $d = 0$, $P = 0$, and $Q = 2$. To determine other parameters of the SARIMA model, this paper adopts an experimental method that takes the autoregressive order $p$ and the moving average order $q$ from 0 to 5. The Akaike information criterion (AIC), RMSE, and $R^2$ were used to evaluate the degree of model fit (Table 1).

The AIC is a measure of the fitting effect of a statistical model. The smaller the AIC value, the better the model fitting effect of the following equation:

$$\text{AIC} = 2k - 2 \ln(L),$$

where $k$ is the number of parameters and $L$ is the likelihood function.

The RMSE can reflect the deviation between the predicted value and the true value and is shown in the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2},$$

where $m$ is the number of predicted values, $y_i$ is the actual value, and $\hat{y}_i$ is the predictive value.

The coefficient of determination $R^2$ reflects the degree to which the model fits the test data. If the $R^2$ value is close to 1, the model fitting effect result will be better as shown in the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2},$$

where $ar{y}$ is the mean of the actual values.

### Table 1: SARIMA model parameters and fitting results.

<table>
<thead>
<tr>
<th>Groups</th>
<th>SARIMA model</th>
<th>AIC</th>
<th>RMSE (cm)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SARIMA $(2, 0, 2) (0, 1, 2)$</td>
<td>$[-526.94]$</td>
<td>1.681</td>
<td>0.396</td>
</tr>
<tr>
<td>2</td>
<td>SARIMA $(3, 0, 1) (0, 1, 2)$</td>
<td>$[-527.96]$</td>
<td>1.809</td>
<td>0.301</td>
</tr>
<tr>
<td>3</td>
<td>SARIMA $(4, 0, 1) (0, 1, 2)$</td>
<td>$[-526.27]$</td>
<td>1.758</td>
<td>0.339</td>
</tr>
<tr>
<td>4</td>
<td>SARIMA $(5, 0, 3) (0, 1, 2)$</td>
<td>$[-529.05]$</td>
<td>1.723</td>
<td>0.365</td>
</tr>
</tbody>
</table>

Figure 4: Decomposition results of the monthly mean sea level anomaly time series: (a) trend term, (b) seasonal term, and (c) random term.
where \( n \) is the number of predicted values, \( y_i \) is the actual value, \( \bar{y} \) is the mean of true values, and \( \hat{y}_i \) is the predictive value.

The \( R^2 \) values of the 4 groups in Table 1 are all small, indicating that the fitting effect is poor and the SARIMA model is not applicable, which may be related to the unobvious trend of sea level rise in 2018 and 2019. Thus, the advantages of the SARIMA model cannot be used.

3.1.2. Trend-Seasonal Series Estimated Using the Prophet Model. The Prophet model transforms a time series into a combination pattern of different time dimensions and then adds the overall trend. The model has a high degree of packaging and few adjustable parameters. According to the sea level change time series characteristics, the estimation frequency is set to “month,” the holiday term is empty, and the seasonal mode is set to “Multiplicative” and “Additive.” The accuracy of the final estimation results is shown in Table 2.

Obviously, the RMSE of the Additive model is small and the degree of fit is relatively satisfactory; thus, it is more suitable for estimating the time series of the sea level changes in this area.

3.2. Estimation of Random Term Series

3.2.1. Random Term Series Estimated Using the LSTM Model. The LSTM model can learn and remember long-term series information and perform selective forgetting. It is suitable for fitting and estimating the random terms of the time series in our study area. This paper uses 327 months of time series data from 1992 to 2019 as input data, and the number of output data sets is 24. The optimal LSTM estimation model was selected by adjusting the number of hidden layers. According to the RMSE and \( R^2 \) results in Table 3, when the hidden layer is 2, the LSTM model fits the best, and the RMSE value is 0.937 cm.

3.2.2. Random Term Series Estimated Using the RBF Model. The RBF neural network model has a strong nonlinear mapping ability and is suitable for fitting of random terms. It has 3 main parameters: the radial basis expansion speed \( S \), the maximum number of neurons MN, and the network parameter DF, which are added each time. Generally, these parameters are determined based on experience. This paper attempted multiple parameter combinations and achieved good results. Although the RMSE of the model in the table is small, \( R^2 \) is also too small, indicating that the fitting effect is not good. Therefore, the RBF model is not suitable for the estimation of random terms in this time series (Table 4).

3.3. Prediction by the Combination of Prophet and LSTM Models. According to the results of the model selection experiment, the Prophet model predicts the trend-seasonal series better and the LSTM model predicts the random term series better. This paper chooses the combination of Prophet and LSTM models to predict sea level changes in this area. To explore the influence of the known series length on the results and determine the best prediction duration, this paper sets up training samples and test samples of different lengths to test the prediction effect of the combined model. The experimental results are shown in Table 5. The mean absolute error (MAE) represents the average value of the absolute error between the predicted value and the observed value, which can avoid the problem of mutual cancellation of errors and accurately reflect the actual prediction error. The calculation formula is as follows:

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - y_i|.
\]

where \( n \) is the number of predicted values, \( \hat{y}_i \) is the predictive value, and \( y_i \) is the actual value.

According to the results in Table 5, when the prediction duration is 12–36 months, the accuracy indicators, such as the RMSE, are relatively ideal, indicating that the combined Prophet-LSTM model is suitable for medium- and short-term estimations of the study area for 12–36 months, and the best RMSE of 0.962 cm is obtained. Therefore, this paper uses all known monthly time series of sea level anomalies in the South China Sea from 1992 to 2019 as training samples to estimate sea level changes from 2020 to 2022. The results are shown in Figure 5.
4. Conclusions

The current research aims to evaluate the capability of different models in estimating sea level variability in the South China Sea. Based on satellite altimetry data from 1992 to 2019, this paper compares the estimation effects of the SARIMA, Prophet, LSTM, and RBF models by grouping, and the combination of Prophet and LSTM models was selected. The detailed results are as follows:

(1) A comparison of the estimation accuracy of the SARIMA and Prophet models shows that the Prophet model can better predict the trend-seasonal series.

(2) A comparison of the estimation accuracy of the LSTM and RBF models shows that the LSTM model can better predict random term series.

(3) The combined model has high accuracy and good performance for 12–36-month short- and medium-term sea level change predictions.

The estimation of the time series in this paper simply considers the changing characteristics of the time series. If the temperature, salinity, tides, ocean currents, and climate anomalies of the sea can be included as reference parameters, then the accuracy will be further improved.

Data Availability

The excel data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest in this paper.

Acknowledgments

The authors acknowledge the AVISO website of the French Space Center (CNES) for providing the satellite data. This research has been supported by the Fundamental Research Funds for the Central Universities (17CX02071), NSFC (61571009), and the Key R&D Program of Shandong Province (2018GHY115046).

References


Table 5: Combination model fitting results of training samples of different lengths.

<table>
<thead>
<tr>
<th>Type of model</th>
<th>Prediction duration (months)</th>
<th>RMSE (cm)</th>
<th>MAE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prophet-LSTM</td>
<td>12</td>
<td>1.243</td>
<td>0.744</td>
<td>0.831</td>
</tr>
<tr>
<td>Prophet-LSTM</td>
<td>24</td>
<td>1.334</td>
<td>1.083</td>
<td>0.814</td>
</tr>
<tr>
<td>Prophet-LSTM</td>
<td>36</td>
<td>0.962</td>
<td>0.718</td>
<td>0.824</td>
</tr>
<tr>
<td>Prophet-LSTM</td>
<td>48</td>
<td>2.261</td>
<td>1.849</td>
<td>0.567</td>
</tr>
<tr>
<td>Prophet-LSTM</td>
<td>60</td>
<td>4.241</td>
<td>3.967</td>
<td>−0.623</td>
</tr>
</tbody>
</table>

Figure 5: Estimation results of the combined Prophet-LSTM model (2020–2022).


