



Long-Term Niño 3.4 Index Prediction Using BiLSTM with CEEMDAN

YUJIAO LIU¹, JING LI¹, YINLONG LI³, SHUHE LEI¹ and XIAOLING DOU^{2*}

¹*Ocean University of China, Qingdao, China*

²*International Christian University, Tokyo, Japan*

³*Geovis Wisdom Technology Co., Ltd., Qingdao, China*

(Received 28 November 2025, Accepted 20 April 2026)

*Corresponding author's email: dou.xiaoling@icu.ac.jp

सार – एल नीनो-दक्षिणी दोलन (El Niño-Southern Oscillation–ENSO) सूचकांक के पूर्वानुमान की पारंपरिक विधियों के समक्ष कई महत्वपूर्ण चुनौतियाँ हैं, जिनमें दीर्घकालिक पूर्वानुमान की सीमित व्यवहार्यता, पर्याप्त पूर्वानुमान अनिश्चितताएँ तथा बसंत पूर्वानुमान बाधा (Spring Predictability Barrier–SPB) शामिल हैं। हाल के वर्षों में डीप लर्निंग आधारित विधियों में उल्लेखनीय प्रगति हुई है, जिससे ENSO के अधिक सटीक एवं दक्ष पूर्वानुमान संभव हुए हैं। इस अध्ययन में सीमडैन-बिल्स्टैम (CEEMDAN-BiLSTM) नामक एक नवीन हाइब्रिड मॉडलिंग फ्रेमवर्क प्रस्तुत किया गया है, जिसमें बाइ-डायरेक्शनल लॉन्ग शॉर्ट-टर्म मेमोरी (Bidirectional Long Short-Term Memory–BiLSTM) नेटवर्क तथा अडैप्टिव नॉइज़ के साथ कम्प्लीट एन्सेम्बल एम्पिरिकल मोड डिकम्पोज़िशन (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise–CEEMDAN) को सम्मिलित किया गया है। इसका उद्देश्य अस्थिरता और अस्थिरता वाली अनुक्रमिक संकेत-श्रृंखलाओं का विश्लेषण करना तथा अनुक्रम की समग्र सूचना को ग्रहण करना है। इसके अतिरिक्त, मॉडल प्रत्येक पूर्वानुमान अवधि के लिए इष्टतम इनपुट लैग निर्धारित करने हेतु पूर्वानुमान-अवधि-विशिष्ट लैग चयन रणनीति का उपयोग करता है, जिससे अल्पकालिक और दीर्घकालिक दोनों पूर्वानुमानों के प्रदर्शन में सुधार होता है। जहाँ अधिकांश मौजूदा अध्ययनों में मासिक Niño3.4 सूचकांक का दो वर्ष तक का सटीक पूर्वानुमान करना कठिन रहा है, वहीं प्रस्तावित विधि इस दीर्घकालिक पूर्वानुमान को सफलतापूर्वक प्राप्त करती है। 24 माह पूर्व के पूर्वानुमान में मॉडल का मूल माध्य वर्ग त्रुटि (RMSE) 0.2819 तथा पियर्सन सहसंबंध गुणांक (PCC) 0.9450 रहा, जो पूर्वानुमान सटीकता और सहसंबंध दोनों के संदर्भ में अन्य मॉडलिंग विधियों से उल्लेखनीय रूप से बेहतर है। यह विधि सामान्यतः प्रबल एल नीनो घटनाओं के दौरान भी दो वर्ष पूर्व Niño3.4 सूचकांक का पूर्वानुमान करने में सक्षम है। इसके अतिरिक्त, 24 माह पूर्व वसंत ऋतु के पूर्वानुमानों के लिए औसत PCC 0.8872 प्राप्त हुआ, जिससे SPB को प्रभावी रूप से कम किया जा सका। परिणामों से यह भी पता चला कि BiLSTM की तुलना में CEEMDAN-BiLSTM ने अल्पकालिक तथा दीर्घकालिक दोनों पूर्वानुमानों में सुधार किया; एक माह पूर्व के पूर्वानुमान में PCC में 8.11% की वृद्धि तथा 24 माह पूर्व के पूर्वानुमान में RMSE में 33.63% की कमी हुई। प्रयोगों के परिणाम दर्शाते हैं कि CEEMDAN-BiLSTM पारंपरिक न्यूरल नेटवर्क मॉडलों की तुलना में अधिक सटीक, स्थिर और विश्वसनीय पूर्वानुमान प्रदान करता है।

ABSTRACT. Traditional methods for forecasting the El Niño-Southern Oscillation (ENSO) index face significant challenges, including limited long-term forecast feasibility, substantial forecast uncertainties, and the spring predictability barrier (SPB). Recently, deep learning approaches have achieved remarkable advancements, enabling more precise and efficient ENSO forecasting results. In this study, we introduce CEEMDAN-BiLSTM, a novel hybrid modeling framework incorporating Bidirectional Long Short-Term Memory (BiLSTM) networks and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) to analyze signals characterized by non-linearity and non-stationarity sequences and capture overall information of the sequence. Moreover, the model employs a horizon-specific lag selection strategy to determine the optimal input lags for each prediction horizon, enhancing both short- and long-term forecast performance. Although most existing studies struggle to precisely predict the monthly Nino3.4 index extending up to two years horizon, our method achieves this long-term prediction. For 24-month-ahead predictions, our model has a root mean square error (RMSE) of 0.2819 and a Pearson correlation coefficient (PCC) of 0.9450, significantly outperforming alternative modeling approaches in the aspects of prediction accuracy and correlation. The method is also able to forecast the Nino3.4 index during typical strong El Niño events with two years ahead. Additionally, it achieves an average PCC of 0.8872 for 24-month-ahead spring predictions, effectively mitigating the SPB. Results also show that the CEEMDAN-

BiLSTM improves both short-term (1-month-ahead PCC increased by 8.11%) and long-term (24-month-ahead RMSE decreased by 33.63%) predictions over BiLSTM. The results of the experiments indicate that CEEMDAN-BiLSTM delivers prediction results that are both more accurate and more stable and dependable in comparison to conventional neural network models.

Key words – ENSO, Niño 3.4, BiLSTM, CEEMDAN, Prediction.

1. Introduction

The El Niño-Southern Oscillation (ENSO) reflects the coupling of ocean-atmosphere dynamics at low latitudes, manifested through oceanic transitions involving alternating El Niño and La Niña phases, alongside the atmospheric Southern Oscillation (SO). El Niño denotes episodic warming in the tropical Pacific Ocean occurring every 2 to 7 years, while La Niña represents its cooling counterpart. SO involves fluctuations in atmospheric pressure contrasts across the eastern–western Pacific basin regions close to the equator (Li *et al.*, 2019). Although ENSO's physical processes originate in the tropical Pacific Ocean, its climatic impacts are far-reaching, affecting regions well beyond the equatorial Pacific. ENSO not only alters the conditions of the Pacific Ocean (Yeh *et al.*, 2009) but also exerts substantial impacts on the global climate system (Yang *et al.*, 2018; Piao *et al.*, 2021), agriculture (Anderson *et al.*, 2018), and fisheries (Sharp & McLain, 2015). Therefore, analyzing and predicting ENSO is crucial for understanding the initiation and development of ENSO events, allowing us to take proactive measures to mitigate the potential disasters associated with ENSO.

Currently, ENSO forecasting approaches are generally grouped into two primary types: dynamical and statistical techniques (McPhaden, 2018; Tang *et al.*, 2018; Ren *et al.*, 2020). Meteorologists and oceanographers worldwide emphasize the significance of grasping the dynamics and operational mechanisms of ENSO to provide a solid theoretical foundation (Timmermann *et al.*, 1998; Chen & Cane, 2008; Timmermann *et al.*, 2018). Dynamical methods incorporate the fully nonlinear physical equations that regulate the intricate ocean-atmosphere coupled dynamics (Wang & Fang, 1996; Tseng *et al.*, 2017). Since the 1980s, various researchers have developed dynamic methods for ENSO forecasting. They can offer physical interpretability by analyzing key processes, such as equatorial wave propagation (Chen & Cane, 2008). Modern numerical models with high-resolution and coupling features improve simulation accuracy (Tseng *et al.*, 2017). While dynamical methods entail substantial computational costs and are not accessible beyond the atmospheric research community.

Statistical methods are typically divided into linear and nonlinear methodologies. Linear methods use observations to establish simplified relationships between predictors and ENSO dynamics (Wang *et al.*, 2020; Chen

et al., 2023). While linear methods are widely employed in ENSO prediction (L'Heureux & Coauthors, 2017), these nonlinear patterns present a challenge for conventional statistical approaches, which frequently struggle to account for these intricate patterns or identify the underlying features (Liu *et al.*, 2023). On the other hand, nonlinear methods have often employed deep learning techniques, like end-to-end neural networks, which have exhibited good predictive performance for El Niño events (Hsieh *et al.*, 2009; Seleznev & Mukhin, 2023).

The El Niño index serves as crucial data for investigating El Niño events (Hanley *et al.*, 2003; Wolter & Timlin, 2011). As diverse El Niño indices represent typical time-series data with an inherent quasi-periodic characteristic, time series prediction models are applicable for forecasting the El Niño index. ENSO indexes exhibit periodic interannual variability and are influenced by high-frequency random noise resulting from seasonal fluctuations, which can significantly impact prediction accuracy (Ham *et al.*, 2019; Wang *et al.*, 2023). In the past few years, although deep learning-based ENSO prediction methods have significantly outperformed conventional linear models in aspects of both prediction accuracy and timeliness, accurately forecasting ENSO events more than one year ahead remains a significant challenge. Furthermore, ENSO prediction is highly uncertain (Chen *et al.*, 2020). A classic illustration is the El Niño event spanning 2014–2016, which imposed considerable challenges on conventional theory-based ENSO forecasting models. Specifically, contrary to the projections of most models, no El Niño event materialized in 2014. By contrast, the 2015 El Niño, ranked among the most intense episodes in observational records, presented notable difficulties for most models when conducting predictions with a 12-month lead time (Ineson *et al.*, 2018; Tang *et al.*, 2018; Xie & Fang, 2020). Moreover, the 2015 El Niño varied greatly relative to the extreme 1997–1998 episodes regarding its formation and warming characteristics (Lian *et al.*, 2017; Broni-Bedaiko *et al.*, 2019). Therefore, selecting an appropriate method that can accurately forecast ENSO events over extended periods is crucial.

Given the above challenges, the application of methods based on climate networks has become increasingly mainstream (Patil *et al.* 2023). Commonly used models encompass convolutional neural networks (CNN), recurrent neural networks (RNN), and their derivative models, which can extract latent patterns from

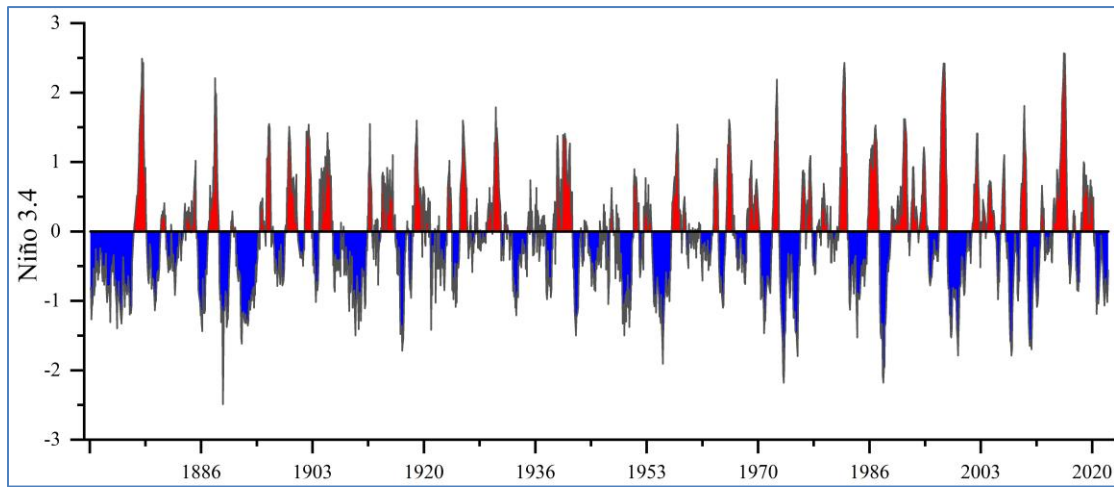


Fig. 1 Original Niño 3.4 data. Red and blue represent positive and negative values, respectively

extensive observational data to enhance the forecasting of future ENSO tendencies. For instance, Ham *et al.* (2019) employed a CNN for ENSO prediction. They successfully extended the forecast horizon to one and a half years, outperforming prevailing dynamical forecasting systems by effectively capturing spatiotemporal patterns in climate data. Furthermore, Chen *et al.* (2023) introduced a multi-factor deep learning architecture built on residual neural networks and observational datasets, demonstrating superior predictive performance compared to traditional operational forecasting models for an effective forecast horizon of up to 8 months. Kim *et al.* (2017) applied Convolutional Long Short-Term Memory (ConvLSTM), a variant of Long Short-Term Memory (LSTM) for rainfall prediction. As demonstrated in a wide range of research works, neural network (NN) models have proven to constitute a practical and effective approach for climate prediction, with their predictive capabilities generally exceeding those of conventional statistical approaches. (Yuan *et al.*, 2019; Zhang *et al.*, 2017).

Furthermore, research has demonstrated that signal-processing techniques can substantially enhance the predictive performance of time-series models. Yan *et al.* (2020) proposed an approach that combines ensemble empirical mode decomposition (EEMD) with temporal convolutional networks (TCN). By decomposing ENSO signals into subcomponents using EEMD and forecasting each component with TCN, their method achieved accurate ENSO predictions with a maximum forecast horizon of 12 months. Guo *et al.* (2020) developed a novel integrated NN framework, EEMD-CNN-LSTM, integrating EEMD, CNN, and LSTM. However, the most existing decomposition-based deep learning studies for ENSO prediction mainly rely on Empirical Mode Decomposition (EMD) or EEMD-type decomposition schemes and unidirectional sequence

models. These approaches may still suffer from residual noise and mode mixing issues, and they often lack sufficient empirical evidence on prediction robustness for long lead times (such as 18–24 months) and across the Spring Predictability Barrier (SPB). Thus, there remains a critical need for a more noise-robust decomposition strategy and a temporal modeling approach capable of capturing bidirectional sequential dependencies to improve ENSO predictability at long lead times.

To address the above limitations, we propose an integrated CEEMDAN–BiLSTM framework for ENSO index prediction, where CEEMDAN is employed to provide a noise-robust decomposition with reduced mode mixing, and BiLSTM is used to capture bidirectional temporal dependencies for ENSO forecasting at long lead times.

We select the Niño 3.4 index to represent oceanic ENSO events in this study, providing the basis for validating the introduced method. The results further demonstrate that our method effectively identifies both high- and low-frequency features for prediction, achieving better results in both predictive accuracy and long-term robustness. The main contributions of this work are threefold: (i) introducing CEEMDAN to enhance noise robustness and alleviate mode mixing in ENSO index decomposition; (ii) proposing a horizon-specific lag configuration strategy to determine the optimal lag order for each prediction horizon and leveraging BiLSTM to model bidirectional temporal dependencies for extended lead-time prediction (up to 24 months); and (iii) systematically evaluating long-horizon performance with particular attention to forecasts across the Spring Predictability Barrier.

The subsequent sections of this paper are organized as follows: Section 2 offers a detailed outline of the methodology. Section 3 elaborates on the parameter configurations and three evaluation metrics for assessing the prediction outcomes. Section 4 showcases the predicted effectiveness and analyzes the results. Section 5 summarizes the key findings and concludes the study.

2. Data and methodology

2.1. Data set and preprocessing

The dataset employed in this research is obtained from the ENSO index (Niño 3.4) supplied by the National Oceanic and Atmospheric Administration (NOAA) (<https://psl.noaa.gov/enso/dashboard.html>). Fig. 1 illustrates the original El Niño index series used in this study.

It comprises monthly data from March 1870 to July 2022, totaling 1,829 observations. In this paper, we use the first 1,477 observations (March 1870 to March 1993) as the training set and the remaining 352 observations (April 1993 to July 2022) as the test set, corresponding to an approximately 8:2 split. Table 1 provides an overview of the dataset's descriptive statistical indicators.

2.2. Complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN)

CEEMDAN is built upon the methodologies of EMD, EEMD, and Complementary Ensemble Empirical Mode Decomposition (CEEMD). EMD-based techniques excel in dynamically breaking down datasets into multiple intrinsic mode functions (IMFs) and residual components, each featuring unique frequency and scale characteristics. These methods are widely used across diverse domains owing to their strong capability in managing non-linear and non-stationary time-series datasets. However, the EMD approach tends to suffer from multimodal distortion due to random phenomena caused by anomalies (Huang *et al.*, 1998). To address this issue, Wu and Huang (2009) introduced the EEMD method, which adds white noise to the original data. The uniformly distributed frequencies of white noise ensure that the data remains continuous and focused, slightly reducing the signal-to-noise ratio (SNR) metric and effectively addressing mode mixing issues. EEMD constitutes a notable advancement in the development of EMD, yet it has a limitation: the introduced white noise cannot be fully removed, potentially affecting feature extraction.

In response to this limitation, Yeh *et al.* (2010) proposed CEEMD, which uses pairs of standardized white noise with identical magnitudes but opposite signs. This approach effectively eliminates the added noise and

TABLE 1

Statistical information of the dataset

Cases	Time	Sample size	Mean	Std.	Min.	Max.	Skewness	Kurtosis
Niño 3.4	1870.3-2022.7	1829	-0.1045	0.7689	-2.4900	2.5700	0.4673	0.4284

reduces errors associated with ensemble-based averaging in the process of data reconstruction. Torres *et al.* (2011) further refined this approach, leading to the development of CEEMDAN, which optimizes the decomposition process by incorporating adaptive noise.

The complete CEEMDAN decomposition procedure applied in this study is systematically described as follows:

(i) Signal Standardization

The original signal x is first normalized to eliminate the effect of amplitude differences:

$$x_{norm} = \frac{x}{\sigma_x},$$

where σ_x denotes the standard deviation of the original signal.

(ii) Noise Injection

A total of Number of Realizations (NR) independent Gaussian white noise sequences is generated, each noise sequence is decomposed using EMD to obtain its intrinsic mode functions (IMFs), denoted as MWN_i .

(iii) First IMF Extraction

For each noise sequence, the original signal is combined with the first IMF of the noise sequence, scaled to a specified noise intensity N_{std} :

$$\xi_i = x_{norm} + N_{std} \cdot \frac{MWN_i}{std(MWN_i)}.$$

Each augmented signal ξ_i is decomposed by EMD with a maximum of 1 mode:

$$temp = emd(\xi_i).$$

The IMFs from all noise realizations are averaged to obtain the IMF of the original signal, and the residual signal is updated iteratively for subsequent IMF extraction.

(iv) Subsequent IMF Extraction

For the extraction of subsequent IMFs, the current residual signal $medias(k)$, where k denotes the index of the IMF to be extracted in the current iteration, is combined with a scaled noise sequence:

$$\xi_i = medias(k) + N_{std} \cdot noise,$$

where noise is the $(k+1)$ -th IMF of the pre-decomposed white noise. The IMFs from all noise realizations are averaged to obtain the IMF of the original signal, and the residual signal is updated iteratively for subsequent IMF extraction.

(v) Stopping Criterion

The decomposition process terminates when the residual signal can no longer be decomposed into a valid IMF, which is determined by:

$$es_{imf} = \min(size(emd(medias))),$$

if $es_{imf} \leq 1$, the residual signal is a monotonic or constant sequence, and no further valid IMFs can be extracted, thus ending the iteration.

(vi) Final IMF Assembly

All extracted IMFs are concatenated into a single matrix, with the final residual signal appended as the last component:

$$mod = [mod; medias].$$

The decomposed components are then rescaled to the original amplitude of the input signal:

$$mod := mod \cdot \sigma_x.$$

CEEMDAN offers several advantages: (i) It introduces an adjustable SNR in each decomposition to control noise levels. (ii) It allows for accurate reconstruction of IMFs without significant noise interference. (iii) It is more computationally efficient than EMD, EEMD, and CEEMD, requiring fewer iterations. By including adaptive noise, CEEMDAN effectively mitigates mode aliasing in EMD and overcomes challenges posed by nonlinear and non-stationary data, enabling the decomposition of the original dataset into IMFs with distinct frequencies.

2.3. Bi-directional long short-term memory (BiLSTM)

LSTM is a deep learning framework constructed to capture long-term dependencies in data by utilizing gate structures and internal loop units (Hochreiter, 1997).

However, LSTM cannot fully capture contextual relationships in a sequence, which BiLSTM addresses by incorporating bidirectional information flow. Depending on the direction of the input sequence, the LSTM can operate as either a forward or backward model. The forward and backward models can be combined into a bidirectional LSTM, which enables the learning of long-term dependencies in both directions (Graves & Jaitly, 2014).

The forward LSTM processes data sequentially from step 1 to step s , whereas the backward LSTM processes in reverse order. For inputting x_t , the hidden state h_t and cell state c_t are the outputs of LSTM. The operation of the LSTM can be summarized as follows.

Step 1: Forget Gate Computation f_t . This step decides which data from the prior cell state c_{t-1} ($c_0 = 0$) should be retained, ensuring that essential information is preserved.

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f),$$

where h_{t-1} ($h_0 = 0$) denotes the hidden layer output at time $t-1$, and x_t is the input at time t , and w_f, b_f stand for the weight matrix and bias of the forget gate, respectively. The activation function σ adopts the sigmoid function, which produces values in $[0,1]$.

Step 2: The input gate updates the cell state as follows:

$$\begin{aligned} \hat{c}_t &= \tanh(w_c[h_{t-1}, x_t] + b_c), \\ i_t &= \sigma(w_i[h_{t-1}, x_t] + b_i), \\ c_t &= f_t \odot c_{t-1} + i_t \odot \hat{c}_t, \end{aligned}$$

where $\hat{c}_t$ denotes the candidate cell state, which stores the data contained in x_t and h_{t-1} . The activation function $\tanh(\cdot)$ produces $\hat{c}_t$ with values between -1 and 1 . The variable i_t is employed to preserve the data in $\hat{c}_t$, and $\hat{c}_t$ is merged with c_{t-1} to form a new cell state c_t . The $\odot$ represents matrix dot product operation, while w_i, w_c as well as b_i, b_c correspond to the weight matrices and biases, respectively.

Step 3: The output gate calculates the value o_t to determine what portion of the cell state c_t ought to be output and passed to the hidden state h_t . The output o_t is then obtained by Eq. (1).

$$\begin{aligned} o_t &= \sigma(w_o[h_{t-1}, x_t] + b_o), \\ h_t &= o_t \times \tanh(c_t). \end{aligned} \quad (1)$$

where o_t denotes the output-gate vector, and h_t is the hidden state at time step t , which serves as the output of the LSTM unit.

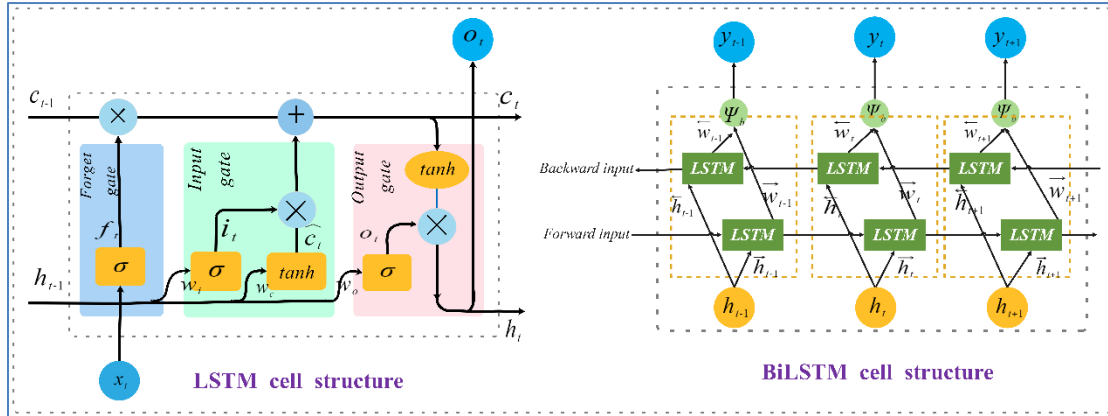


Fig. 2 Network structures of LSTM and BiLSTM models

Building upon the LSTM framework, the computation procedure of BiLSTM is outlined as below:

$$\begin{aligned}\vec{h}_t &= LSTM(\vec{h}_{t-1}, x_t), \\ \tilde{h}_t &= LSTM(\tilde{h}_{t-1}, x_t), \\ y_t &= \psi_b(\vec{w}_t \times \vec{h}_t + \tilde{w}_t \times \tilde{h}_t + b_t),\end{aligned}$$

where $\vec{h}_t$ and $\tilde{h}_t$ denote the forward and backward hidden states, respectively; $\vec{w}_t$ and $\tilde{w}_t$ are the weight matrices for the forward and backward layers, respectively, and b_t refers to the bias of the output layer. The activation function $\psi(\cdot)$ combines $\vec{h}_t$ and $\tilde{h}_t$ to produce the predicted output y_t .

The structures of LSTM and BiLSTM are shown on Fig. 2.

2.4. Model design and evaluation approaches

Before inputting the original time-series data into the BiLSTM, the data must be normalized. We select min-max normalization. This data preprocessing approach linearly converts the original dataset. The detailed calculation formula is presented below:

$$\tilde{y} = (y - y_{min}) / (y_{max} - y_{min}).$$

Additionally, a sensitivity analysis is conducted to investigate how the lag order l affects the Niño 3.4 index forecasting performance of the BiLSTM framework. The evaluation metric is the RMSE, as defined in Eq. (2), computed for BiLSTM forecasts with different lag orders across multiple prediction horizons. The lag orders considered range from 3 to 30 months, while the forecast horizons span 1 to 24 months. This analysis offers a quantitative evaluation of the function of lag order in forecasting sea surface temperature (SST) fluctuations in the eastern equatorial Pacific. The findings demonstrate

that the optimal lag order differs by forecast horizon. Specifically, shorter horizons (1 and 24 months ahead) achieve optimal accuracy with a lag order of 3, intermediate horizons (3–6 months ahead) with a lag order of 6, and longer horizons (12–18 months ahead) with a lag order of 9. These findings highlight that increasing lag length incorporates more historical information into the model, thereby improving its ability to constrain and predict future variability, albeit at the cost of higher forecast errors with excessively long lags. Consequently, subsequent analyses adopt the optimal lag order determined for each forecast horizon.

To summarize, the core conclusions of the lag order analysis are as follows: The optimal lag order for ENSO prediction is not a fixed value but varies with the forecast horizon. A targeted, horizon-specific optimal lag configuration (3 for short horizons, 6 for intermediate horizons, and 9 for longer horizons) can balance the utilization of historical information and prediction accuracy, laying a reliable foundation for subsequent model comparisons and performance evaluations.

2.4.1. Parameter details

The BiLSTM model is composed of a BiLSTM layer, a dropout layer, and two dense layers, and a regularization mechanism. The dropout layer randomly drops a fraction of the neuron outputs to prevent overfitting, while the regularization mechanism adds constraints to the weights to improve the generalization ability. Batch normalization is utilized on the output of the recurrent layer for boosting training stability and overall performance. The CEEMDAN algorithm addresses the modal aliasing inherent in traditional EMD through a novel integration of staged noise perturbation and adaptive residual decomposition, with optimized parameterization (as shown in Table 2) ensuring stable multiscale signal resolution. The proposed model generates predicted Niño 3.4 index values, which are

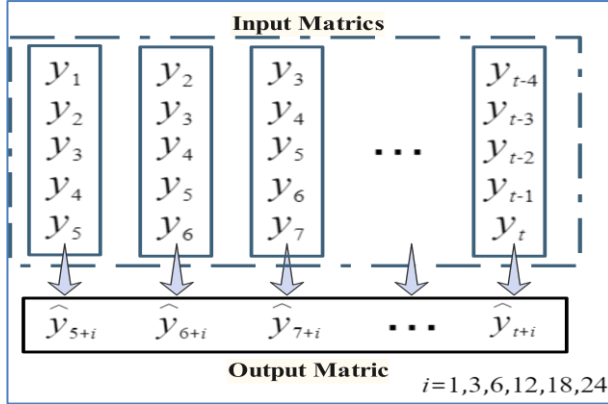


Fig. 3. Schematic illustration of the direct multi-step prediction architecture: lagged time series data as input matrix for predicting future values with i representing the forecast horizon

compared with actual observations to evaluate its effectiveness. All methods are conducted using MATLAB 2022a, with the detailed parameters provided in TABLE .

2.4.2. Evaluation metrics

Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Pearson Correlation Coefficient (PCC) are used in this study to evaluate the errors, which are calculated as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^{pre} - y_i^{Act})^2}, \quad (2)$$

where n represents the length of the test set, y_i^{Act} is the observations, y_i^{pre} is the predicted value.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i^{pre} - y_i^{Act}|,$$

$$PCC = \frac{\sum_{i=1}^n (y_i^{Act} - \bar{y}^{Act})(y_i^{pre} - \bar{y}^{pre})}{\sqrt{\sum_{i=1}^n (y_i^{Act} - \bar{y}^{Act})^2 \sum_{i=1}^n (y_i^{pre} - \bar{y}^{pre})^2}},$$

where $\bar{y}^{Act}$ is the average of observations from the test set, and $\bar{y}^{pre}$ is the average of predicted value from the test set.

2.4.3. The optimal lag order selection

This study adopts a fundamental concept for El Niño index forecasting that aligns with conventional time series prediction methods, using historical datasets to forecast future conditions. Moreover, the prediction process of the CEEMDAN - BiLSTM adopted a direct multistep prediction method. Specifically, consider the given time

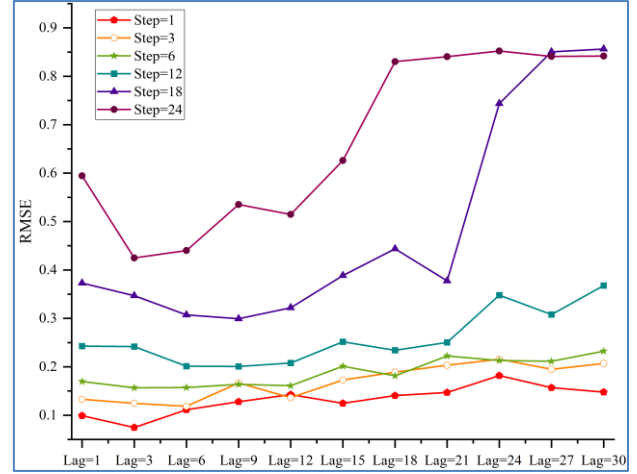


Fig. 4. RMSE of ENSO predictions across different lead times and lag orders.

series $\{y_t\}$, our goal is to predict the values of the i -forecast horizon $\hat{y}_{t+1}, \hat{y}_{t+2}, \dots, \hat{y}_{t+i}$. In this context, we set the lag order to l , and then we utilize the l historical data points of the time series $Y_{t-l+1}, \dots, Y_{t-1}, Y_t$ as the input for BiLSTM model. Consequently, the model output is the predicted value for future i -time steps. To illustrate, we consider lag order of $l = 5$. Fig. 3 illustrates the model's input and output data. Specifically, when the lag order is configured as 5, we use five historical data points as the input array, which serves as the basis for generating forecast outputs for the subsequent i ($i = 1, 3, 6, 12, 18, 24$) months. Take $i = 3$ as an example. We use $y_1, y_2, \dots, y_5$ to predict y_8 , $y_2, y_3, \dots, y_6$ to predict y_9 , and $y_3, y_4, \dots, y_7$ to predict y_{10} , and so on. That is, for $i=3$, we use $y_{t-4}, y_{t-3}, \dots, y_t$ to predict y_{t+3} .

Based on the sensitivity analysis in Fig. 4, we select the lag order that minimizes RMSE for each forecast horizon and apply this horizon-specific setting consistently in both the BiLSTM baseline and the proposed CEEMDAN-BiLSTM model.

To summarize, the key findings of the lag-order selection are as follows: The optimal lag order for ENSO prediction is not fixed but varies with the forecast horizon. We implement a horizon-specific lag configuration to ensure fair and consistent model comparisons, which strengthens the reliability of our performance evaluation.

2.5. The structure of the method for forecasting Niño 3.4 index

To conduct long-horizon forecasting of the Niño 3.4 index over long horizons, a structured methodology is designed that combines data preprocessing, optimal lag order selection, and BiLSTM modeling. The overall workflow is presented in Fig. 5.

TABLE 2

Hyperparameter settings of the deep learning models employed in this study

Parameter category	Parameter name	Value or description
BiLSTM		
Network structure parameters	Number of hidden units of BiLSTM	70
	Number of fully connected layer output units	70
	Dropout rate of dropout layer	0.25
Training option parameters	Optimization algorithm	Adaptive Moment Estimation (Adam)
	Maximum number of iterations	600
	Initial learning rate	0.005
	Learning rate decay factor	0.1
CEEMDAN		
Nstd	Noise standard deviation ratio set	1
NR	Number of noise perturbation experiments for improved decomposition stability	20
MaxIter	Maximum number of iterations for EMD set	2000
SNRFlag	Noise is not standardized during processing (set to 2 for standardization)	1
IMFs	Number of intrinsic mode functions	10

Step 1: Decomposition. Input the preprocessed Niño 3.4 index into the CEEMDAN module, decomposing it into several IMFs (IMF1, IMF2, ..., IMF_n) and a residual component (R_n). These components correspond to the fluctuation characteristics of the original series at different time scales, respectively.

Step 2: Candidate lag configuration. A set of candidate lag orders is predefined for subsequent hyperparameter searching. For each decomposed component, input matrices are constructed under each candidate lag setting for model training.

Step 3: Parallel Training with BiLSTM Subnetworks and optimal lag selection. Feed each IMF component (IMF1, IMF2, ..., IMF_n) and the residual component (R_n) obtained from CEEMDAN decomposition into their corresponding BiLSTM sub-networks (BiLSTM1, BiLSTM2, ..., BiLSTM_{n+1}) for parallel training. Based on the outputs of these BiLSTM sub-networks under different candidate lags (marked by the dashed ellipse in Fig. 5), the optimal lag for each forecast horizon is determined by prediction performance metrics. BiLSTM can capture temporal dependencies by leveraging both past and future information in the sequence.

Step 4: Fusion of Output Matrices. Sum (Σ) the outputs of all BiLSTM sub-networks to generate an output matrix that contains predicted values for multiple forecast horizons, including 1, 3, 6, 12, 18, and 24 months. Finally, the predicted Niño 3.4 index is obtained.

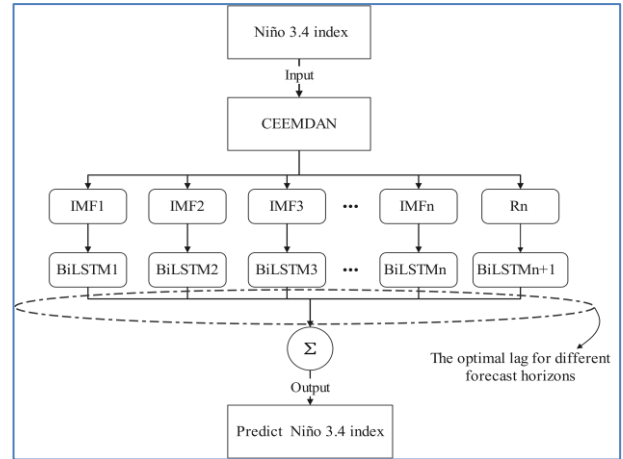


Fig. 5. Flowchart of the proposed method

Step 5: Evaluation. Finally, the model performance is evaluated through contrasting the forecast outcomes with the actual measured values.

In summary, this methodology involves selecting the optimal lag order, constructing input matrices, employing a BiLSTM model to process the data, and evaluating the forecast results, thereby facilitating the long horizon forecasting of the Niño 3.4 index.

3. Results and discussion

We assess the performance of the proposed CEEMDAN-BiLSTM model through a series of results in this section. All subsequent result analyses are based on

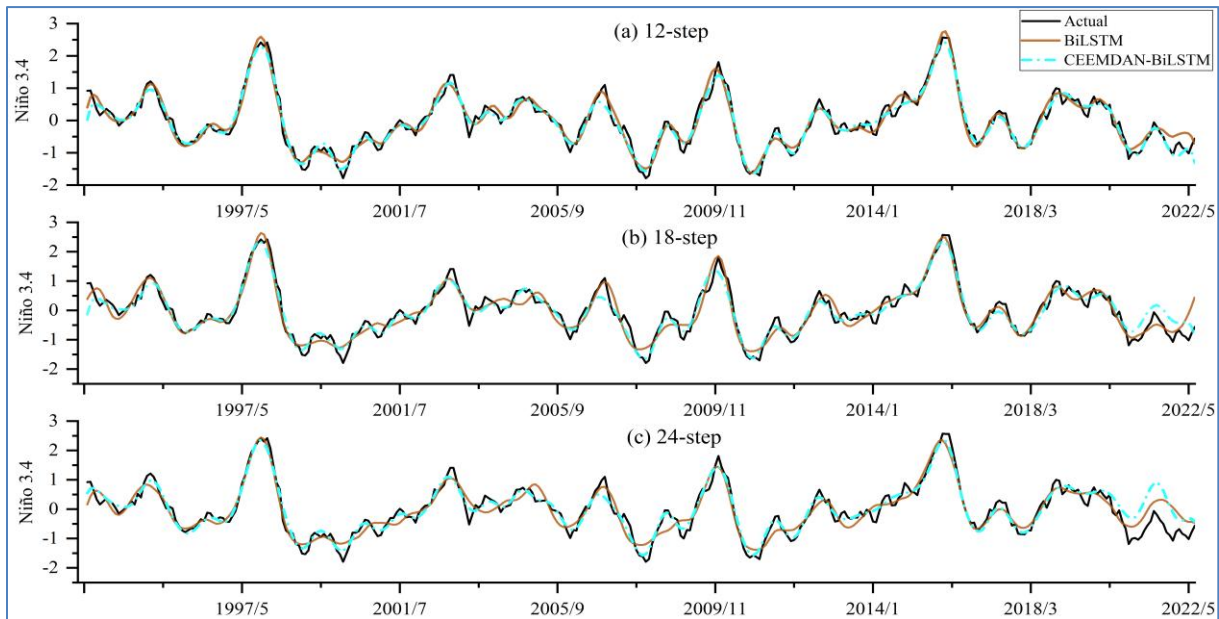


Fig. 6. Niño 3.4 index observations and predictions of the proposed method from 1993 to 2022 at (a) 12-, (b) 18-, and (c) 24-month lead times

the prediction results under the optimal lag order. First, the effectiveness of CEEMDAN-BiLSTM in long-term prediction is comprehensively evaluated. Next, we select two strong El Niño events, 1997-1998 and 2014-2016, as case studies to visualize the forecasting ability of CEEMDAN-BiLSTM. Finally, we investigate the SPB by examining the monthly correlation between real Niño 3.4 index values and multi-step-ahead forecast results.

3.1. Long-term predictability of the CEEMDAN-BiLSTM

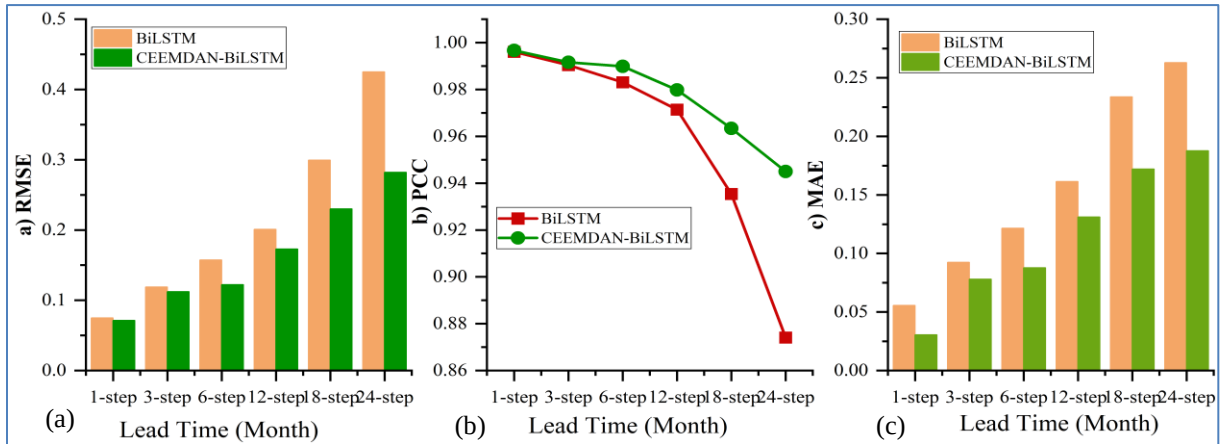
The evaluation process of CEEMDAN-BiLSTM is as follows. First, the original Niño 3.4 time series is decomposed using CEEMDAN, resulting in nine IMFs and one residual term. Next, each component is predicted individually using the BiLSTM models. After that, the results are conducted over various prediction months ahead, including 1-, 3-, 6-, 12-, 18-, and 24-month. Finally, to facilitate performance visualization, graphs are plotted to compare predicted values with actual observations.

Fig. 6 presents the forecasted and actual values for the period from April 1993 to July 2022, providing a visual representation of the long-term predictive performance. It is obvious that the prediction curves of BiLSTM and CEEMDAN-BiLSTM follow the actual values across long-term forecasting horizons of (a) 12 steps, (b) 18 steps, and (c) 24 steps. Notably, CEEMDAN-BiLSTM outperforms BiLSTM in terms of fitting data fluctuations. Under various long-term prediction steps, especially the 24-step ahead prediction, its prediction curves adhere more closely to the

actual values' undulations, precisely capturing the amplitude and rhythm of data fluctuations. This showcases its superior ability to fit the fluctuating characteristics of data, which is crucial for enhancing the accuracy of long-term predictions.

In Fig. 7, panels a) and c) display bar charts representing RMSE and MAE, respectively, while Fig. 7 b) presents line graphs illustrating the variation of PCC with prediction step ahead. It is evident that CEEMDAN-BiLSTM outperforms BiLSTM across all three metrics, RMSE, PCC, and MAE. Notably, at long term prediction, BiLSTM exhibits a sharp increase in error and a significant drop in correlation, whereas CEEMDAN-BiLSTM maintains superior performance with greater stability.

The RMSE values for BiLSTM predictions at 1, 3, 6, 12, 18, and 24 months ahead are [0.0747, 0.1186, 0.1572, 0.2009, 0.2994, 0.4248]. While the RMSE values for CEEMDAN-BiLSTM are [0.0711, 0.1121, 0.1219, 0.1726, 0.2300, 0.2819]. In the 24-step-ahead prediction, CEEMDAN-BiLSTM achieves a 33.63% reduction in RMSE compared to BiLSTM. Furthermore, CEEMDAN-BiLSTM maintains a PCC above 0.9, ensuring strong predictive correlation even at long forecasting horizons. The results highlight the superior performance of the CEEMDAN-BiLSTM model in multi-step forecasting, particularly for long-term predictions. Additionally, we can still observe that BiLSTM performs well, achieving correlations above 0.85 for 24-month-ahead predictions, which also demonstrates the effectiveness of our chosen baseline model.



Figs. 7(a-c). Performance evaluation results from 1993 to 2022.4: (a) RMSE bar charts; (b) Pearson correlation coefficient (PCC) curves; (c) RMSE bar charts for different prediction lead times (months)

3.2. Examples for strong El Niño events

To assess the forecasting performance of the CEEMDAN-BiLSTM for ENSO, two typical ENSO events (in 1997-1998 and 2014-2016) are selected as test cases.

Figs. presents line plots comparing the predicted Niño 3.4 index against the observed values for the periods (a) from May 1997 to April 1998 and (b) from October 2014 to May 2016, covering forecast lead times of 1, 3, 6, 12, 18, and 24 months. For short-term prediction (1, 3, and 6 months), the model exhibits high precision, where the forecasted values closely aligned with the observed data, particularly in capturing the peak and magnitudes of both ENSO events. The framework successfully grasps the general variation trend of the Niño 3.4 index, indicating robust performance in the short-term prediction. However, as the forecast lead time increases (12, 18, and 24 months), the accuracy declines, with increasing deviations between predicted and observed values. Despite this, the CEEMDAN-BiLSTM continues to capture the general trend of ENSO events over longer lead times, although its ability to accurately predict the timing and intensity of peaks diminishes. This reflects the growing uncertainty and complexity of ENSO dynamics at extended prediction horizons, highlighting the challenges associated with accurate long-term prediction stemming from the nonlinear and inherently unpredictable characteristics of ENSO phenomena.

The performance of the CEEMDAN-BiLSTM is evaluated by comparing the predicted values with the observed data over various lead times (1, 3, 6, 12, 18, and 24 months). As shown in

Figs. 9, the model demonstrates high accuracy for short-term forecasts (1, 3, and 6 months), with correlation coefficients close to 1 and low RMSE. For the 1997-1998 event, the model achieved a PCC of 0.9967 at a 1-month

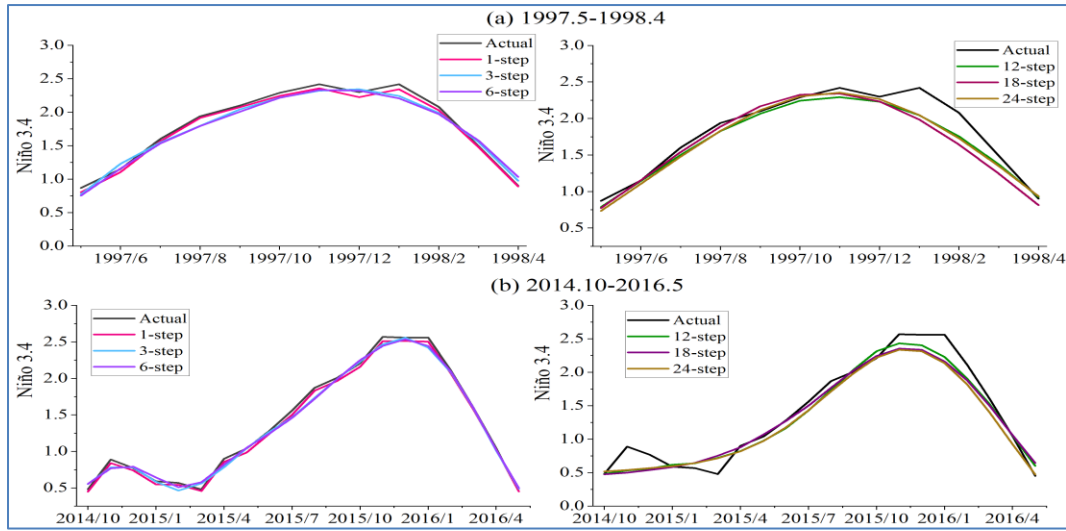
lead time and an RMSE of 0.0711, reflecting strong predictive performance. However, as the forecast lead time increases to 12 the performance of the model shows a decline, with a decrease in PCC and an increase in RMSE. Despite this, the model maintains a strong correlation for longer-term predictions, with PCC values greater than 0.90 for the 1997-1998 event even at the 24-month ahead prediction. For the 2014-2016 case, the PCC drops after 12 months but remains above 0.95, indicating the robustness of this framework in capturing both short-term variations and long-term tendencies. These findings highlight the difficulties in forecasting ENSO phenomena over prolonged lead times due to the inherent complexity and nonlinear dynamics of the system, while underscoring the effectiveness of BiLSTM effectiveness in forecasting across varying time scales.

In conclusion, for the results of two typical El Niño events, in 1997-1998 and 2014-2016, the CEEMDAN-BiLSTM delivers impressive 24-month-ahead correlation coefficients, significantly surpassing the performance of existing prediction methods. These results highlight the exceptional adaptability of the model in forecasting extreme events, further underscoring its ability to handle complex and variable climatic phenomena.

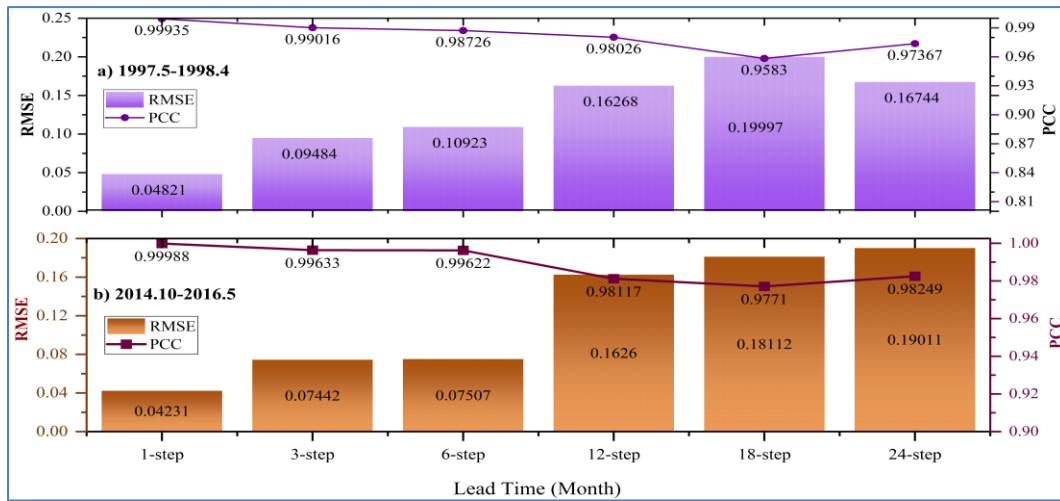
3.3. Spring forecast barrier (SPB)

To investigate the potential presence of SPB in the forecasting process of the introduced model, we calculated the correlation coefficients of the Niño 3.4 index for each month of the different lead times from April 1993 to July 2022. The results are shown in Fig. 10.

We can see in Fig. 10, our framework remains affected by the SPB, as indicated by the correlation drop in boreal spring (April – May), particularly at longer lead times (18–24 months). It is important to note that the



Figs. 8(a&b). Niño 3.4 index predictions (proposed model) and observations for the (a) 1997–1998 and (b) 2014–2016 periods



Figs. 9(a&b). PCC and RMSE values versus prediction lead time for the (a) 1997–1998 and (b) 2014–2016 periods



Fig. 10. Monthly correlations between observed and forecasted Niño 3.4 index at 1-, 3-, 6-, 12-, 18-, and 24-month lead times. All correlations are statistically significant at the 5% significance level, based on the t-distribution p-values employed to evaluate the significance of the correlation coefficients

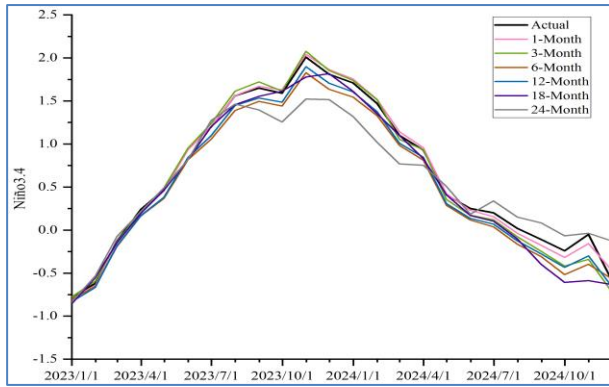


Fig. 11. Comparison of Observed and Predicted Niño 3.4 Index (2023–2024) Across Different Prediction Lead Times

TABLE 3

Results of statistical tests across different predictions lead time

Lead Time (Month)	DM	P values of DM	P values of Wilcoxon
1	0.1625	0.8709	0.0000***
3	0.6285	0.5297	0.0000***
6	4.7267	0.0000***	0.0000***
12	2.7581	0.0058***	0.0000***
18	6.1930	0.0000***	0.0000***
24	4.5589	0.0000***	0.0000***

Note: Significance levels are clearly annotated using *, **, and *** to denote the 10%, 5%, and 1% levels, respectively

proposed model does not fully overcome the SPB but partially alleviates its impact by maintaining relatively stable correlations during spring. Specifically, the average correlation coefficient of our model during northern hemisphere spring reaches 0.8272, and prediction results from March to May significantly outperform those reported in other studies, further verifying the merits of our framework in improving ENSO forecasting during spring (Yeh *et al.*, 2009; Li *et al.*, 2019; Ye *et al.*, 2021).

Additionally, the SPB affects the prediction accuracy across different prediction horizons (1, 3, 6, 12, 18, and 24 months), considering both the training period (March 1870 to March 1993) and the testing period (April 1993 to July 2022) of the model. In short-term predictions (1–6 months ahead), the model exhibits high correlations throughout the entire year, with a PCC close to 1.0, further validating the capability of the CEEMDAN-BiLSTM model to predict ENSO events accurately. However, as the forecast horizon increases (beyond 12 months), the correlation during spring (March to May) begins to decrease, particularly for the 24-step forecast in May, where the PCC decreases to 0.8272. This decline is more pronounced in May than in other

months, suggesting the influence of SPB. Despite this, the model still mitigates the SPB, maintaining a PCC above 0.80, which demonstrates its capacity to reduce the impact of SPB on long-term predictions.

The observed mitigation of the SPB may stem from the fact that CEEMDAN decomposes the Niño 3.4 index into multi-scale IMFs and a residual term, thereby suppressing high-frequency stochastic noise and reducing mode mixing. This decomposition enhances the signal-to-noise ratio of the low-frequency interannual components that are more predictable and dynamically relevant to ENSO evolution. By forecasting each component separately and then aggregating them, the BiLSTM sub-networks can focus on learning cleaner and scale-specific temporal dependencies, which is particularly beneficial during boreal spring when ENSO predictability is typically weakened.

In summary, the CEEMDAN-BiLSTM framework demonstrates clear advantages in alleviating SPB-related skill degradation, leading to more stable springtime forecasts, especially at long lead times.

3.4. Statistical test of the proposed model

In this subsection, we employ the Diebold–Mariano (DM) test (Diebold *et al.*, 1995) and Wilcoxon signed-rank test to statistically validate the superiority of the proposed CEEMDAN-BiLSTM model over the benchmark BiLSTM model, with results presented in Table 3. The DM test, which has been widely applied in time series forecasting validation (Wang *et al.*, 2018; Sun *et al.*, 2019; Niu & Wang, 2019), assesses predictive accuracy by examining the statistical significance of differences in mean squared error (MSE) between the two models. A positive DM statistic indicates that CEEMDAN-BiLSTM yields lower prediction error, while a p-value < 0.05 rejects the null hypothesis of equal predictive performance.

For short lead times (1–3 months), the DM test yields p-values > 0.05, suggesting no statistically significant difference between the models-consistent with the observation that both models perform well in short-term forecasting. In contrast, for long lead times (6–24 months), the DM test returns p-values < 0.01 (***) alongside positive DM statistics ranging from 2.7581 to 6.1930. This confirms that CEEMDAN-BiLSTM delivers statistically significant improvements in long-term prediction accuracy, with the most pronounced enhancement observed at the 18-month lead time (DM = 6.1930, p < 0.0001).

The Wilcoxon signed-rank test, a non-parametric alternative robust to non-normal error distributions, evaluates differences in mean absolute error (MAE) across all lead times (1–24 months). The results show p-values

TABLE 4

Multi-Step Prediction Performance Comparison in 2023–2024 Niño3.4 Independent Test Set

Lead Time (Month)	BiLSTM			CEEMDAN-BiLSTM		
	RMSE	MAE	PCC	RMSE	MAE	PCC
1	0.0679	0.0439	0.9973	0.0487	0.0392	0.9988
3	0.1177	0.1092	0.9972	0.0912	0.0644	0.9986
6	0.1681	0.1408	0.9886	0.1583	0.1113	0.9962
12	0.2815	0.2113	0.9734	0.2387	0.1886	0.9802
18	0.5951	0.5029	0.9001	0.4383	0.3661	0.9117
24	0.7837	0.6649	0.8514	0.6507	0.5371	0.8946

TABLE 5

Comparison of predictive performance between the proposed CEEMDAN-BiLSTM model & relevant studies on Niño 3.4 index prediction

Authors	Methods	Data set	Multi-step forecast	RMSE	PCC
Nooteboom et al. (2018)	ARIMA-ANN	Niño 3.4	12-step	NRMSE(0.17)	-----
Broni-Bedaiko et al. (2019)	LSTM	Niño 3.4	12-step	RMSE(0.90)	-----
Ham et al. (2019)	CNN	Niño 3.4	11-step	-----	PCC(0.73)
Yan et al. (2020)	EEMD-TCN	Niño 3.4/SOI	12-step	RMSE(0.51)	PCC(0.83)
Guo et al. (2020)	EEMD-CNN-LSTM	Niño 3.4	12-step	RMSE(0.29)	PCC(0.52)
Gupta et al. (2022)	ConvLSTM	Niño 3.4	12-step	RMSE(0.72)	-----
Ibebuchi & Richman (2024)	AE-LSTM	Niño 3.4	18-step	RMSE(0.55)	PCC(0.91)
The proposed	CEEMDAN-BiLSTM	Niño 3.4	18-step	RMSE(0.23)	PCC(0.96)

Note: NRMSE = Normalized Root Mean Square Error; SOI = Southern Oscillation Index; RMSE = Root Mean Square Error; PCC = Pearson Correlation Coefficient; “—” indicates the metric was not reported in the original study; Bold values represent the optimal performance in this comparison

< 0.0001 (***) for all horizons, confirming that CEEMDAN-BiLSTM consistently and significantly outperforms the BiLSTM, with the improvement not attributable to randomness in model training.

Taken together, these statistical tests validate that the performance improvement of the CEEMDAN-BiLSTM model is stable and reliable, with value for long-term El Niño prediction.

3.5. Out-of-sample Test Case

To verify the model’s real-time predictive capability and generalization performance, we conducted an independent validation using the 2023–2024 Niño3.4 dataset as an out-of-sample test case. The model parameters were kept consistent with those calibrated on historical data to simulate real-time application scenarios.

As shown in Table 4, the evaluation metrics (RMSE, PCC, MAE) for the 2023–2024 validation set are of the same magnitude as those from the original test set, confirming the stability of the model’s performance. Across all prediction lead times (3-step to 24-step), the CEEMDAN-BiLSTM model consistently outperforms the standalone BiLSTM model, with lower RMSE and MAE values, as well as higher PCC values. For example, at the 12-step lead time, CEEMDAN-BiLSTM achieves an

RMSE of 0.2387 and a PCC of 0.9802, compared to 0.2815 and 0.9734 for BiLSTM.

Fig. 11 illustrates the Niño 3.4 index predictions across lead times from 1 to 24 months, alongside the observed values. Shorter lead times (1–3 months) demonstrate near-perfect alignment with observations, while even longer lead times (18–24 months) maintain consistent trend tracking, confirming the model’s ability to provide reliable early warnings of El Niño onset. These results highlight the model’s potential for real-time climate prediction and disaster risk reduction.

This independent validation confirms that the CEEMDAN-BiLSTM model can effectively capture the multi-scale oscillatory features of ENSO signals, providing more accurate and stable predictions for real-world events.

3.6. Further discussion: comparison with existing studies

In this subsection, we provide a comprehensive discussion and analysis to further confirm the advantages of the proposed framework by contrasting its performance against that of existing ENSO index forecasting models.

As can be seen from Table 5, the predictive performance of various models differs remarkably, even for the same ENSO index. It should be additionally explained

that NRMSE (Normalized Root Mean Square Error) is a normalization of RMSE, so its value is smaller relative to RMSE, where

$$NRMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^{pre} - y_i^{Act})^2 / (y_{max}^{Act} - y_{min}^{Act})}.$$

Notably, the proposed method demonstrates consistently strong performance, attaining an 18-month-ahead forecast with an RMSE of 0.23 and a PCC of 0.96 for the Niño 3.4 index. These long-term forecasts outperform the short-term results of several existing methods, underscoring the reliability and effectiveness of our method. In addition to its strong predictive accuracy, the method aligns well with current research trends by introducing advanced deep learning techniques. Overall, the experimental findings indicate that the proposed approach serves as a promising solution for ENSO forecasting, especially in long-horizon prediction. Beyond these statistical gains, the operational implications are also noteworthy.

From an operational forecasting perspective, the improved long-lead predictability (18–24 months) may provide earlier guidance for multi-season planning, while the enhanced springtime stability suggests that SPB-related skill degradation is alleviated rather than fully eliminated. Moreover, the decomposition–prediction–reconstruction pipeline offers a transparent way to separate low-frequency predictable signals from high-frequency variability, thereby supporting more reliable and interpretable ENSO forecasts. The relatively low computational burden further makes the framework suitable for routine real-time updates as new observations become available.

4. Conclusions

Boosting the predictability of ENSO is critical for improving the precision and practical value of global seasonal climate forecasts. However, the inherent complexity, nonlinearity, and temporal variability of ENSO events pose persistent challenges for traditional prediction methods, particularly in achieving long lead-time accuracy and alleviating the SPB. In response, we develop a decomposition-assisted deep learning framework that explicitly targets long-lead ENSO predictability and SPB-related uncertainty.

In the present research, we propose an integrated framework combining CEEMDAN, BiLSTM, and a horizon-specific lag configuration strategy, which effectively captures the intricate dynamics of ENSO. CEEMDAN excels at decomposing complex, nonlinear, and non-stationary time series into straightforward IMFs,

preserving critical features while reducing noise and mode mixing, which is a key factor in mitigating SPB. This decomposition enhances the learning capability of BiLSTM, which is adept at capturing long-term dependencies and bidirectional temporal relationships in sequential data. Furthermore, the introduced horizon-specific lag configuration systematically identifies the optimal lag order for each prediction lead time, ensuring consistent model settings and fair performance comparisons. By integrating these components, our approach significantly improves prediction accuracy and robustness. To the best of our knowledge, this work represents an early attempt to integrate CEEMDAN with BiLSTM for ENSO index prediction, extending existing decomposition-based deep learning frameworks toward improved long-lead forecasting and enhanced robustness across the SPB.

Experimental results confirm the effectiveness of the proposed framework. The standalone BiLSTM model can achieve reasonable predictive performance for long-horizon forecasting. Compared with the single BiLSTM model, the CEEMDAN-BiLSTM hybrid framework yields further performance improvements across different forecast horizons. Notably, the proposed model exhibits prominent advantages for spring forecasting tasks, which constitutes a key highlight of this work. These findings demonstrate the complementary strengths of CEEMDAN and BiLSTM: CEEMDAN decomposes complex climate time-series signals, while BiLSTM effectively captures long-range temporal dependencies.

The proposed approach not only advances ENSO forecasting capabilities but also provides a promising framework for improving long-term climate prediction models. Looking ahead, this framework presents a solid foundation for further research. Future work could extend the method to multivariate input, regionalized ENSO impacts, or hybrid models combining physical and data-driven techniques. Ultimately, this study contributes to advancing ENSO prediction technology and supports broader efforts toward more reliable long-term climate forecasting.

Acknowledgements

We thank the editor and anonymous reviewers for their constructive comments on this work. This work was supported by the National Natural Science Foundation of China [11871444].

Authors' Contributions

Yujiao Liu: Visualization, Methodology, Writing-original draft. (email: liuyujiao0508@163.com).

Jing Li: Data curation, Formal analysis, Methodology. (email: lijing6813@163.com).

Yinlong Li: Visualization. (email: onth05@163.com).

Shuhe Lei: Conceptualization, Supervision, Funding acquisition. (email: shuhelei@ouc.edu.cn).

Xiaoling Dou: Conceptualization, Supervision, Editing.

Disclaimer: The contents and views expressed in this research article are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

- Anderson, W., Seager, R., Baethgen, W., Cane, M., 2018, "Trans-Pacific ENSO teleconnections pose a correlated risk to agriculture", *Agricultural and Forest Meteorology*, **262**, 298–309, doi: <https://doi.org/10.1016/j.agrformet.2018.07.023>.
- Broni-Bedaiko, C., Katsriku, F. A., Unemi, T., et al., 2019, "El Niño-Southern Oscillation forecasting using complex networks analysis of LSTM neural networks", *Artificial Life and Robotics*, **24**, 445–451, doi: <https://doi.org/10.1007/s10015-019-00540-2>.
- Chen, D., Cane, M. A., 2008, "ENSO Prediction and Predictability", *Journal of Computational Physics*, **227**, 7, 3625–3640, doi: <https://doi.org/10.1016/j.jcp.2007.05.014>.
- Chen, H. C., Tseng, Y. H., Hu, Z., Ding, R., 2020, "Enhancing the ENSO Predictability beyond the Spring Barrier", *Scientific Reports*, **10**, 1–12, doi: <https://doi.org/10.1038/s41598-020-57853-7>.
- Chen, N., Fang, X., 2023, "A simple multiscale intermediate coupled stochastic model for El Niño diversity and complexity", *Journal of Advances in Modeling Earth Systems*, doi: https://doi.org/10.48550/arXiv.2206.06649?utm_source=chatgpt.com.
- Chen, Y., Huang, X. M., Luo, J. J., et al., 2023, "Prediction of ENSO using multivariable deep learning", *Atmospheric and Oceanic Science Letters*, **16**, 04, 53–58, doi: <https://doi.org/10.1016/j.aosl.2023.100350>.
- Diebold, F. X., Mariano, R. S., 1995, "Comparing predictive accuracy", *Journal of Business & Economic Statistics*, **13**, 3, 253–263, doi: <https://doi.org/10.1080/07350015.1995.10524599>.
- Guo, Y., Cao, X., Liu, B., Peng, K., 2020, "El Niño Index Prediction Using Deep Learning with Ensemble Empirical Mode Decomposition", *Symmetry*, **12**, 6, 893, doi: <https://doi.org/10.3390/sym12060893>.
- Gupta, M., Kodamana, H., Sandeep, S., 2022, "Prediction of ENSO Beyond Spring Predictability Barrier Using Deep Convolutional LSTM Networks", *IEEE Geoscience and Remote Sensing Letters*, **19**, 1–5, doi: <https://doi.org/10.1109/LGRS.2020.3032353>.
- Graves, A., Jaitly, N., 2014, "Towards end-to-end speech recognition with recurrent neural networks", *Proceedings of the 31st International Conference on Machine Learning*, **32**, 2, 1764–1772, doi: <https://proceedings.mlr.press/v32/graves14.html>.
- Ham, Y. G., Kim, J. H., Luo, J. J., 2019, "Deep learning for multi-year ENSO forecasts", *Nature*, **573**, 7775, 568–572, doi: <https://doi.org/10.1038/s41586-019-1559-7>.
- Hanley, D. E., Bourassa, M. A., O'Brien, J. J., Smith, S. R., Spade, E. R., 2003, "A quantitative evaluation of ENSO indices", *Journal of Climate*, **16**, 1249–1258, doi: [https://doi.org/10.1175/1520-0442\(2003\)16%3C1249:AQOEI%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)16%3C1249:AQOEI%3E2.0.CO;2).
- Hochreiter, S., Schmidhuber, J., 1997, "Long short-term memory", *Neural Computation*, **9**, 8, 1735–1780, doi: <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Hsieh, W. W., 2009, "Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels", Cambridge University Press, doi: <https://doi.org/10.1017/CBO9780511627217>.
- Huang, N., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N., Tung, C., Liu, H. H., 1998, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, **454**, 903–995, doi: <https://doi.org/10.1098/rspa.1998.0193>.
- Ibebuchi, C. C., Richman, M. B., 2024, "Deep learning with autoencoders and LSTM for ENSO forecasting", *Climate Dynamics*, **62**, 5683–5697, doi: <https://doi.org/10.1007/s00382-024-07180-8>.
- Ineson, S., Balmaseda, M. A., Davey, M. K., et al., 2018, "Predicting El Niño in 2014 and 2015", *Scientific Reports*, **8**, 10733, doi: <https://doi.org/10.1038/s41598-018-29130-1>.
- Kim, S., Hong, S., Joh, M., Song, S.-K., 2017, "Deeptrain: ConvLstm network for precipitation prediction using multichannel radar data", *arXiv Preprint arXiv:1711.02316*, doi: <https://doi.org/10.48550/arXiv.1711.02316>.
- L'Heureux, M. L., Coauthors., 2017, "Observing and predicting the 2015/16 El Niño", *Bulletin of the American Meteorological Society*, **98**, 1363–1382, doi: <https://doi.org/10.1175/BAMS-D-16-0009.1>.
- Li, B., Zhang, L., Zhang, Q., Yang, S., 2019, "An EEMD-based denoising method for seismic signal of high arch dam combining wavelet with singular spectrum analysis", *Shock and Vibration*, 2019, **1**, 1–9, doi: <https://doi.org/10.1155/2019/4937595>.
- Lian, T., Chen, D. K., Tang, Y., 2017, "Genesis of the 2014–2016 El Niño events", *Science China Earth Sciences*, **60**, 9, 1589–1600, doi: <https://doi.org/10.1007/s11430-016-8315-5>.
- Liu, Y., Donat, M. G., England, M. H., et al., 2023, "Enhanced multi-year predictability after El Niño and La Niña events", *Nature Communications*, **14**, 6387, doi: <https://doi.org/10.1038/s41467-023-42113-9>.
- MathWorks., n.d., "ranksum - Rank sum test", MathWorks Help Center, doi: <https://www.mathworks.cn/help/stats/ranksum.html>.
- McPhaden, M. J., 2018, "Understanding and Predicting El Niño and the Southern Oscillation", *New Frontiers In Operational Oceanography*, doi: <https://doi.org/10.17125/gov2018.ch23>.
- Niu, X., Wang, J., 2019, "A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting", *Applied Energy*, **241**, 519–539, doi: <https://doi.org/10.1016/j.apenergy.2019.01.122>.
- Nooteboom, P. D., Feng, Q. Y., López, C., Hernández-García, E., Dijkstra, H. A., 2018, "Using network theory and machine learning to predict El Niño", *Earth System Dynamics*, **9**, 969–983, doi: <https://doi.org/10.3389/fpsy.2019.00153>.
- Patil, K. R., Doi, T., Jayanthi, V. R., Behera, S., 2023, "Deep learning for skillful long lead ENSO forecasts", *Frontiers in Climate*, **4**, 1058677, doi: <https://doi.org/10.3389/fclim.2022.1058677>.
- Piao, J. L., Chen, W., Chen, S. F., Gong, H. N., Wang, L., 2021, "Mean states and future projections of precipitation over the monsoon transitional zone in China in CMIP5 and CMIP6 models", *Climatic Change*, **169**, 35, doi: <https://doi.org/10.1007/s10584-021-03286-8>.

- Ren, H. L., Zheng, F., Luo, J. J., Wang, R., Liu, M., Zhang, W., Zhou, T., Zhou, G., 2020, "A Review of Research on Tropical Air-Sea Interaction, ENSO Dynamics, and ENSO Prediction in China", *Journal of Meteorological Research*, **34**, 43–62, doi: <https://doi.org/10.1007/s13351-020-9155-1>.
- Seleznev, A., Mukhin, D., 2023, "Improving statistical prediction and revealing nonlinearity of ENSO using observations of ocean heat content in the tropical Pacific", *Climate Dynamics*, **60**, 1–15, doi: <https://doi.org/10.1007/s00382-022-06298-x>.
- Sharp, G. D., McLain, D. R., 2015, "Fisheries, El Niño-Southern Oscillation and upper ocean temperature records: an eastern Pacific example", *Oceanography*, **6**, 1, 13–22, doi: <https://doi.org/10.5670/oceanog.1993.18>.
- Sun, S., Wei, Y., Tsui, K. L., Wang, S., 2019, "Forecasting tourist arrivals with machine learning and internet search index", *Tourism Management*, **70**, 1–10, doi: <https://doi.org/10.1016/j.tourman.2018.09.013>.
- Tang, Y. M., Zhang, R. H., Liu, T., Duan, W. S., Yang, D. J., Zheng, F., Ren, H. L., Lian, T., Gao, C., Chen, D. K., Mu, M., 2018, "Progress in ENSO prediction and predictability study", *National Science Review*, **5**, 6, 826–839, doi: <https://doi.org/10.1093/nsr/nwy105>.
- Timmermann, A., Latif, M., Voss, R., Grötzner, A., 1998, "Northern Hemispheric interdecadal variability: A coupled air–sea mode", *Journal of Climate*, **11**, 1906–1931, doi: [https://doi.org/10.1175/1520-0442\(1998\)011%3c1906:NHIVAC%3e2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011%3c1906:NHIVAC%3e2.0.CO;2).
- Timmermann, A., An, S. I., Kug, J. S., et al., 2018, "El Niño–southern oscillation complexity", *Nature*, **559**, 535–545, doi: <https://doi.org/10.1038/s41586-018-0252-6>.
- Tseng, Y. H., Hu, Z. Z., Ding, R., et al., 2017, "An ENSO prediction approach based on ocean conditions and ocean–atmosphere coupling", *Climate Dynamics*, **48**, 2025–2044, doi: <https://doi.org/10.1007/s00382-016-3188-2>.
- Torres, M. E., Colominas, M. A., Schlotthauer, G., Flandrin, P., 2011, "A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise", *Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing*, 4144–4147, doi: <https://doi.org/10.1109/ICASSP.2011.5947265>.
- Wang, B., Fang, Z., 1996, "Chaotic oscillations of tropical climate: A dynamic system theory for ENSO", *Journal of the Atmospheric Sciences*, **53**, 19, 2786–2802, doi: [https://doi.org/10.1175/1520-0469\(1996\)053<2786:COOTCA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<2786:COOTCA>2.0.CO;2).
- Wang, J., Du, P., Lu, H., Yang, W., Niu, T., 2018, "An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting", *Applied Soft Computing*, **72**, 321–337, doi: <https://doi.org/10.1016/j.asoc.2018.07.025>.
- Wang, X., Slawinska, J., Giannakis, D., 2020, "Extended-range statistical ENSO prediction through operator-theoretic techniques for nonlinear dynamics", *Scientific Reports*, **10**, 2636, doi: <https://doi.org/10.1038/s41598-020-59128-7>.
- Wang, G. G., Cheng, H. L., Zhang, Y. M., Yu, H., 2023, "ENSO analysis and prediction using deep learning: A review", *Neurocomputing*, **520**, 216–229, doi: <https://doi.org/10.1016/j.neucom.2022.11.078>.
- Wang, H. S., Zhang, Y. P., Liang, J., Liu, L. L., 2023, "DAFA-BiLSTM: Deep autoregression feature augmented bidirectional LSTM network for time series prediction", *Neural Networks*, **157**, 240–256, doi: <https://doi.org/10.1016/j.neucom.2022.10.009>.
- Wolter, K., Timlin, M. S., 2011, "El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEIext)", *International Journal of Climatology*, **31**, 1074–1087, doi: <https://doi.org/10.1002/joc.2336>.
- Wu, Z., Huang, N. E., 2009, "Ensemble Empirical Mode Decomposition: a Noise-Assisted Data Analysis Method", *Adv. Data Sci. Adapt. Anal.*, **1**, 01, 1–41, doi: <https://doi.org/10.1142/S1793536909000047>.
- Xie, R., Fang, X., 2020, "The unusual 2014–2016 El Niño events: Dynamics, prediction and enlightenments", *Science China Earth Sciences*, **63**, 626–633, doi: <https://doi.org/10.1007/s11430-019-9561-2>.
- Yan, J., Mu, L., Wang, L., et al., 2020, "Temporal convolutional networks for the advance prediction of ENSO", *Scientific Reports*, **10**, 1, 8055, doi: <https://doi.org/10.1038/s41598-020-65070-5>.
- Yang, S., Li, Z. N., Yu, J. Y., Hu, X. M., Dong, W. J., He, S., 2018, "El Niño–Southern Oscillation and its impact in the changing climate", *National Science Review*, **5**, 6, 840–857, doi: <https://doi.org/10.1093/nsr/nwy046>.
- Yeh, S. W., Kug, J. S., Dewitte, B., Kwon, M. H., Kirtman, B. P., Jin, F. F., 2009, "El Niño in a changing climate", *Nature*, **461**, 511–514, doi: <https://doi.org/10.1038/nature08316>.
- Yeh, J., Shieh, J. S., Huang, N. E., 2010, "Complementary Ensemble Empirical Mode Decomposition: a Novel Noise Enhanced Data Analysis Method", *Adv. Data Sci. Adapt. Anal.*, **2**, 135–156, doi: <https://doi.org/10.1142/S1793536910000422>.
- Ye, M., Nie, J., Liu, A., Wang, Z., Huang, L., Tian, H., Wei, Z., 2021, "Multi-year ENSO forecasts using parallel convolutional neural networks with heterogeneous architecture", *Frontiers in Marine Science*, **8**, 717184, doi: <https://doi.org/10.3389/fmars.2021.717184>.
- Yuan, S., Luo, X., Mu, B., Li, J., Dai, G., 2019, "Prediction of North Atlantic Oscillation index with convolutional LSTM based on ensemble empirical mode decomposition", *Atmosphere*, **10**, 5, 252, doi: <https://doi.org/10.3390/atmos10050252>.
- Zhang, Q., Wang, H., Dong, J., Zhong, G., Sun, X., 2017, "Prediction of sea surface temperature using long short-term memory", *IEEE Geoscience and Remote Sensing Letters*, **14**, 10, 1745–1749, doi: <https://doi.org/10.1109/LGRS.2017.2733548>.

