

Article

El Niño Index Prediction Based on Deep Learning with STL Decomposition

Ningmeng Chen ^{1,2}, Cheng Su ^{1,3} , Sensen Wu ^{1,2}  and Yuanyuan Wang ^{2,4,*}

¹ School of Earth Sciences, Zhejiang University, Hangzhou 310027, China

² Zhejiang Provincial Key Laboratory of Geographic Information Science, Hangzhou 310028, China

³ Zhejiang Provincial Key Laboratory of Geoscience Big Data and Deep Resource, Hangzhou 310028, China

⁴ Ocean Academy, Zhejiang University, Zhoushan 316022, China

* Correspondence: wangyuanyuanxy@zju.edu.cn

Abstract: ENSO is an important climate phenomenon that often causes widespread climate anomalies and triggers various meteorological disasters. Accurately predicting the ENSO variation trend is of great significance for global ecosystems and socio-economic aspects. In scientific practice, researchers predominantly employ associated indices, such as Niño 3.4, to quantitatively characterize the onset, intensity, duration, and type of ENSO events. In this study, we propose the STL-TCN model, which combines seasonal-trend decomposition using locally weighted scatterplot smoothing (LOESS) (STL) and temporal convolutional networks (TCN). This method uses STL to decompose the original time series into trend, seasonal, and residual components. Each subsequence is then individually predicted by different TCN models for multi-step forecasting, and the predictions from all models are combined to obtain the final result. During the verification period from 1992 to 2022, the STL-TCN model effectively captures index features and improves the accuracy of multi-step forecasting. In historical event simulation experiments, the model demonstrates advantages in capturing the trend and peak intensity of ENSO events.

Keywords: ENSO; El Niño index; deep learning; TCN; STL



Citation: Chen, N.; Su, C.; Wu, S.; Wang, Y. El Niño Index Prediction Based on Deep Learning with STL Decomposition. *J. Mar. Sci. Eng.* **2023**, *11*, 1529. <https://doi.org/10.3390/jmse11081529>

Academic Editor: Jean-Louis Pinault

Received: 29 June 2023

Revised: 22 July 2023

Accepted: 29 July 2023

Published: 31 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

El Niño–Southern Oscillation (ENSO) is one of the most intense ocean–atmosphere coupling phenomena worldwide [1], characterized by a clear periodicity with a cycle of 2–7 years [2]. Although originating in the tropical Pacific region, ENSO affects various parts of the world through atmospheric teleconnections. It often causes widespread climate anomalies, triggers various meteorological disasters, and has impacts on ecosystems and socio-economics [3–5]. In the context of global warming, climate change may aggravate the impact of various disasters caused by it [6]. ENSO has gained significant attention and widespread research interest worldwide [7], seeking to monitor the current status of ENSO and predict its future evolution to prepare for possible abnormal impacts and mitigate the negative effects of ENSO.

To provide a more intuitive representation and effective monitoring of ENSO, previous researchers have summarized a set of indices to characterize the occurrence, intensity, and type of ENSO events. Different indices have been proposed and adopted by research institutions in different countries. For example, the Japan Meteorological Agency uses the JMA index [8]; the Australian Bureau of Meteorology uses the BOM index; the NECP uses the Niño 3.4 index for event monitoring; and the latest standard of the China Meteorological Administration also uses the Niño 3.4 index to define events [9].

Traditional ENSO prediction methods are divided into two main categories: dynamical and statistical models [10]. The dynamical models model and numerically simulate the phenomenon based on the physical laws related to the formation and development of ENSO. While dynamical models effectively capture the underlying physical laws [11], they face

challenges such as significant discrepancies in prediction results between different models and high computational complexity [12]. Statistical models use the historical observation data to predict ENSO [13]. Compared with dynamical models, statistical models are less costly and easier to develop. Statistical models include linear and nonlinear statistical models. Traditional statistical methods are mainly constructed with linear statistical models, such as Canonical Correlation Analysis (CCA) [14], Principal Oscillation Patterns (POP) [15], multiple linear regression (MLR), such as ENSO climatology and persistence (CLIPER) [16], the Linear Inverse Model (LIM) [17] and Markov chains [18]. Jiang et al. [19] used a generalized typical mixed model with principal component correlation analysis (PC-CCA) to make forecasts of sea surface temperature in the Niño region, which could only achieve effective forecasts up to three seasons in advance. Rosmiati et al. [20] used the ARIMA method to predict the ENSO region SST Niño 3.4 and found that the ARIMA model stage was well suited to predict short-term ENSO events. However, traditional statistical models may struggle to extract the nonlinear features of ENSO, resulting in limited long-term predictive ability.

Recently, with the rise of machine learning, deep learning methods have been recognized as having significant potential for ENSO prediction [21]. Artificial neural networks (ANN) and convolutional neural networks (CNN) are powerful prediction models that have been applied to El Niño forecasting [22,23]. Yuan et al. [12] achieved an effective prediction with an 11-month lead time for the Niño 3.4 index using the CNN model [12,24]. Zhou et al. [24] used the long short-term memory networks (LSTM) model to forecast the Niño 3.4 index and analyzed the seasonal forecast errors of the model. Despite the dominance of periodicity in ENSO events, their observed characteristics exhibit irregularity and chaotic behavior due to the influence of many complex atmospheric and oceanic processes. Time series data of various indices used to characterize these phenomena also possess non-linear and non-stationary features. However, traditional statistical models have not been able to effectively simulate the non-linear characteristics, resulting in limited long-term predictability. When using deep learning methods for ENSO prediction, a single model is often used for index forecasting [23]. However, a single model may not fully capture significant periodicity in time series data. This limitation can be encountered when faced with significant periodicity in the data. Chen et al. [25] suggest that combining decomposition models with prediction models is an effective approach for time series data. Among these, Seasonal and Trend decomposition using LOESS (STL) has been widely used for time series pre-processing [26]. Jiao et al. [27] combined it with an LSTM model to predict public transportation passenger flow during the COVID-19 pandemic, effectively improving accuracy.

Taking into account the non-stationarity, periodicity, and seasonality of the ENSO index time series, this study proposes a Temporal Convolutional Networks model based on Seasonal and Trend decomposition using LOESS (STL), with the Niño 3.4 index as the research object. Firstly, STL decomposed the original index data into trend, seasonal, and residual components. Subsequently, the TCN model was developed to predict the three components separately, and the predictions were then summed to obtain the final index. We used the root mean square error (RMSE), mean absolute error (MAE), and Pearson correlation coefficient (PCC) to evaluate the model's performance.

The structure of the rest of this manuscript is as follows: Section 2 describes the data, models used in the study and evaluation metrics. Section 3 discusses and analyzes the results. In Section 4, we summarize the strengths and limitations of the STL-TCN model and outline our future research directions.

2. Materials and Methods

2.1. Data

The data used in this study are from the Niño 3.4 index, provided by the Physical Sciences Laboratory (PSL) of the National Oceanic and Atmospheric Administration (NOAA).

This dataset is derived from the monthly mean sea-surface temperature data (HadISST1) provided by the Hadley Centre of the UK Met Office, using the 1981–2010 period as the climate baseline. The Niño 3.4 index data used in this study cover the period from January 1871 to December 2022, with a total of 1824 data points. The visualization of the data is shown in Figure 1.

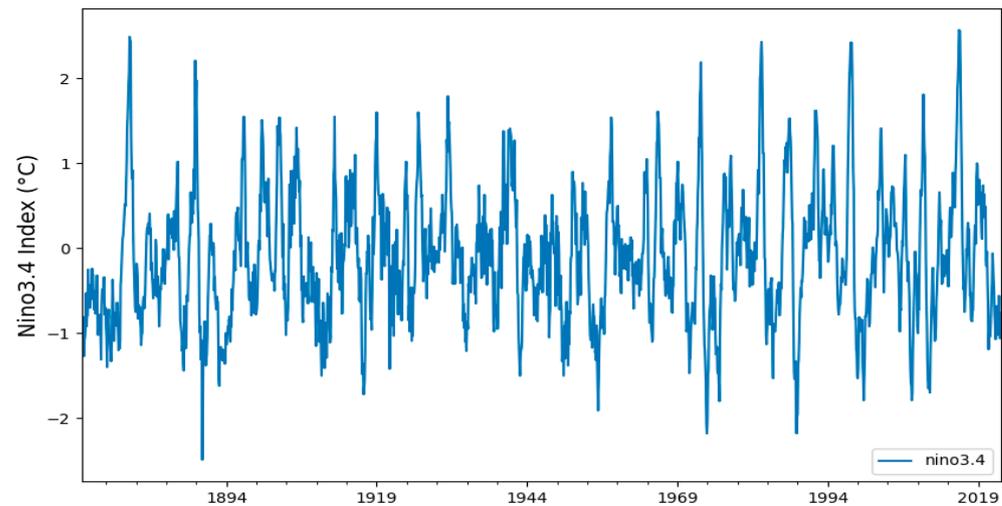


Figure 1. Niño 3.4 Index.

2.2. Seasonal and Trend Decomposition Using LOESS

STL is a general and robust method for decomposing time series, whereas LOESS is a method for estimating non-linear relationships [28]. The STL decomposition algorithm was originally proposed by Cleveland [29], and its basic idea is to decompose the original time series (X_v) into three components: trend (T_v), seasonal (S_v), and remainder (R_v). The trend component is the low-frequency component of the data, representing the trend and direction of change; the seasonal component is the high-frequency component of the data, representing the regular change of the data over time, usually with a fixed period and amplitude; and the residual component is the remaining component of the original series after subtracting the trend component and the seasonal component, containing the noise in the series as follow:

$$X_v = T_v + S_v + R_v \quad (1)$$

LOESS in STL allows the smoothing of data while preserving the essential features of the data. It smooths the time series by assigning weights to the neighborhood of each data point based on distance, and then performs a polynomial regression fit at each data point, using the points closest to them as explanatory variables. This decomposition is additive so that summing the components yields the original series. Compared to traditional linear regression models, locally-weighted regression can better accommodate nonlinear data relationships and is more robust to outliers and noise.

STL is mainly divided into two procedures: inner loop and outer loop. The inner loop is nested in the outer loop, and the specific process of k th is as follows:

1. Detrending. Subtract the trend component from the original series: $X_v - T_v^{(k)}$.
2. Cycle-subseries smoothing. In Step (1) the detrended time series is broken into cycle-subseries. Each subseries is smoothed using LOESS, and the smoothed subseries are combined into a new series, denoted as $C_v^{(k+1)}$.
3. Low-Pass filtering of smoothed cycle subseries. The series obtained from Step (2) is processed using low-pass filtering, and then the regression operation is performed using LOESS, denoted as $L_v^{(k+1)}$.
4. Detrending of smoothed cycle subseries. $S_v^{(k+1)} = C_v^{(k+1)} - L_v^{(k+1)}$.
5. Deseasonalizing. $X_v - S_v^{(k+1)}$.
6. Trend smoothing. The series obtained in step (5) is smoothed using LOESS to obtain a new trend component $T_v^{(k+1)}$.

In the outer loop, the seasonal and trend components obtained in the inner loop are used to calculate the remainder component $R_v^{(k+1)} = X_v - S_v^{(k+1)} - T_v^{(k+1)}$.

2.3. Temporal Convolutional Networks

Temporal Convolutional Networks (TCN) is a sequence modeling approach based on convolutional neural networks, specifically designed for modeling and predicting time series data. It was originally proposed by Bai et al. [30] and consists primarily of causal convolutions, dilated convolutions, and residual connections.

Causal convolution is a convolutional operation that preserves the temporal order of the input sequence. The convolution kernel only operates on the past portion of the input sequence and cannot access the future portion, thereby ensuring the causality of the convolution. This causality property makes the model more interpretable and stable, as it helps to avoid issues of information leakage or spurious correlations.

Dilated convolution, also known as atrous convolution, overcomes the limitation of standard causal convolution when dealing with long temporal sequences. Standard causal convolution requires an increase in the number of network layers or larger filters to capture longer dependencies. However, by introducing dilated convolution, we can enlarge the receptive field of neurons without increasing the model parameters and computational complexity. This allows the model to effectively capture features from farther distances. Refer to Figure 2 for an illustration of this concept.

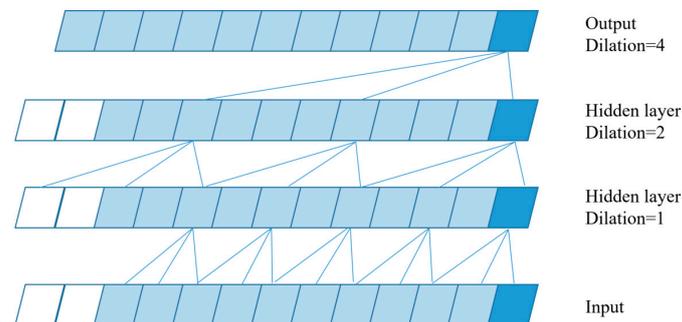


Figure 2. A stack of dilated convolutional layers for dilations 1, 2, and 4.

A residual block consists of two causal dilated convolutional layers, each followed by an optional batch normalization and activation function, as illustrated in Figure 3. By utilizing causal and dilated convolutions, these convolutional layers can pass the output of the previous layer to the next layer for processing. Each residual block also includes a skip connection, which aims to address the gradient vanishing problem by allowing information to bypass one or more layers in the network. This ensures the preservation of the original input data even in very deep networks.

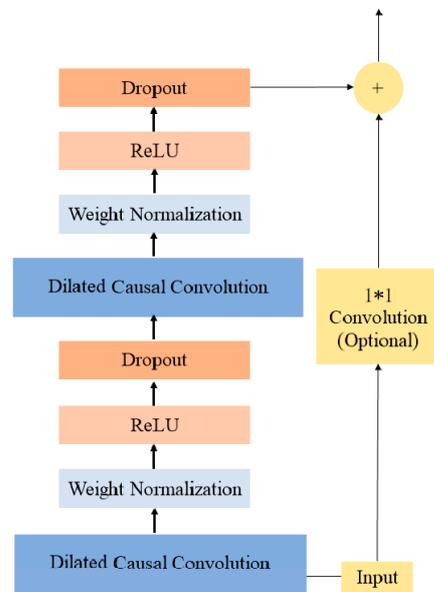


Figure 3. Residual block. * means multiplication sign.

2.4. A Multi-Step El Niño Index Forecasting Strategy

Early time series forecasting uses existing historical data to predict a single data point in the future, e.g., to predict tomorrow’s temperature, i.e., single-step forecasting, but forecasting a single value provides more limited information. In many cases, multi-step forecasting methods are needed to better predict future trends and changes, i.e., historical time series $[y_1, \dots, y_N]$ are used to predict H values $[y_{N+1}, \dots, y_{N+H}]$, where $H > 1$ indicates the period of the forecast. At present, there are four main methods for multi-step prediction: direct strategy, recursive strategy, direct and recursive fusion strategy, and multiple outputs.

In this paper, we used a direct multi-step forecasting strategy, also known as an independent strategy, to train multiple models to predict multiple values, with each model predicting one value independently of the other. The specific calculation process is shown as follows:

$$\hat{y}_{t+1} = model_1(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-n})$$

$$\hat{y}_{t+2} = model_2(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-n})$$

$$\hat{y}_{t+3} = model_3(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-n})$$

2.5. Proposed Model

The STL-TCN model combines STL and TCN. The Niño 3.4 index is decomposed into simpler and meaningful components via STL. This helps the TCN to capture the trend, cyclical and seasonal features in the series, which in turn improves the forecast accuracy. The specific steps are as follows:

1. The Niño 3.4 index is decomposed into the trend component T_v , seasonal component S_v and remainder component R_v by STL.
2. Normalize the three components.
3. Prediction of each of these three components using TCN neural network.
4. The trend component forecast, seasonal component forecast, and residual component forecast of Niño 3.4 index will be inverse normalized and summed to obtain the final Niño 3.4 forecast, The detailed process can be seen in Figure 4.

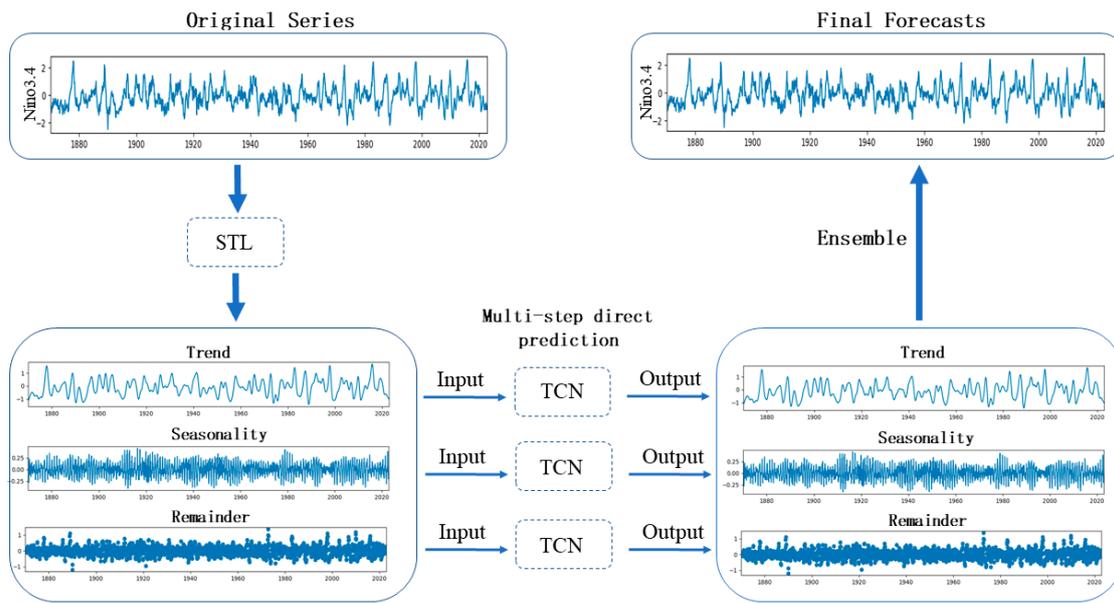


Figure 4. STL-TCN model framework.

In this experiment the whole data set is divided into training and test sets according to the ratio of 8:2. The specific division is shown in Table 1.

Table 1. Data Segmentation.

Data	Data Type	Time Range
Niño 3.4	Training set	1 January 1871 to 31 December 1991
	Test set	1 January 1992 to 31 December 2022

The model consists of STL, TCN layers, and a fully connected layer. The main hyperparameters of TCN include the size of the convolutional kernel, dilation factors, and the number of convolutional kernels. In this study, the hyperparameters for TCN are set as follows: the one-dimensional convolutional kernel size is set to 7; the dilation factors are sequentially set to 1, 2, and 4; the residual modules consist of three layers, and the number of convolutional kernels in each layer is set to 128, 64, and 32, respectively; and the dropout parameter is set to 0.2. The optimizer used is the Adam algorithm, with a learning rate of 0.001. The maximum number of training epochs is set to 20, the batch size is set to 4, and the random seed is fixed.

2.6. Evaluation Metrics

To evaluate the model’s performance, the following three evaluation metrics were selected in this study. (1) Root mean square error (RMSE). (2) Mean absolute error (MAE). (3) Pearson correlation coefficient (PCC). The calculation formulae are as follows:

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

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \tag{3}$$

$$PCC = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \tag{4}$$

y_i represents the observed value (true value); \hat{y}_i represents the predicted value; \bar{y} and $\bar{\hat{y}}$ are the averages of y_i and \hat{y}_i , respectively.

3. Results and Discussion

3.1. Models Are Trained Using Different Time Windows

The time window refers to the duration of past observed values considered when predicting future data points. Selecting an appropriate time window is crucial for the performance of prediction models in many forecasting problems. If the time window is too short, the model might fail to capture long-term trends and patterns, thus limiting its predictive capabilities. On the other hand, if the time window is too long, the model may capture excessive historical information, including noise and irrelevant details, that do not hold representativeness for future predictions, thereby affecting the generalization of the model's performance. Therefore, it is essential to carefully choose the time window, considering the characteristics of the data and the forecasting objectives, to achieve optimal predictive performance.

Predicting the Niño 3.4 index is a typical task in time series forecasting. The time window within a time series model has a significant impact on its performance, as it directly affects the structure of the temporal data and consequently influences the model's training and performance. To determine the optimal time window, this study varied the time window of the model and evaluated its performance. The time windows were increased by intervals of 3 months, ranging from 3 to 36. Specifically, the following time windows were tested: {3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36}. These time windows were used to train the model, and performance metrics such as PCC, RMSE, and MAE were computed. Table 2 presents the predictive performance of different time windows for a 12-month lead time. It can be observed that when the time window is 12 months, the PCC reaches its highest value of 0.62, while the RMSE and MAE reach their lowest values of 0.70 °C and 0.55 °C, respectively. In contrast, the predictive performance is poorer when the time window is 3 months. This indicates that shorter time windows may not provide sufficient information for forecasting long-term sequences, making it challenging for the model to capture effective periodic and seasonal features. From Figure 5, it can be seen that as the time window exceeds 12 months, the errors increase while the correlation coefficients decrease, resulting in a deterioration of the model's predictive performance. This suggests that a larger time window is not necessarily better. Therefore, when predicting the Niño 3.4 index, a time window of 12 months is most suitable.

Table 2. Predicted results with different time windows at 12-months lead.

Time Window	PCC	RMSE	MAE
3	0.48	0.80	0.61
6	0.56	0.74	0.57
9	0.60	0.71	0.55
12	0.62	0.70	0.55
15	0.61	0.72	0.56
18	0.60	0.71	0.55
21	0.59	0.72	0.55
24	0.57	0.72	0.55
27	0.53	0.77	0.59
30	0.53	0.75	0.57
36	0.51	0.76	0.59

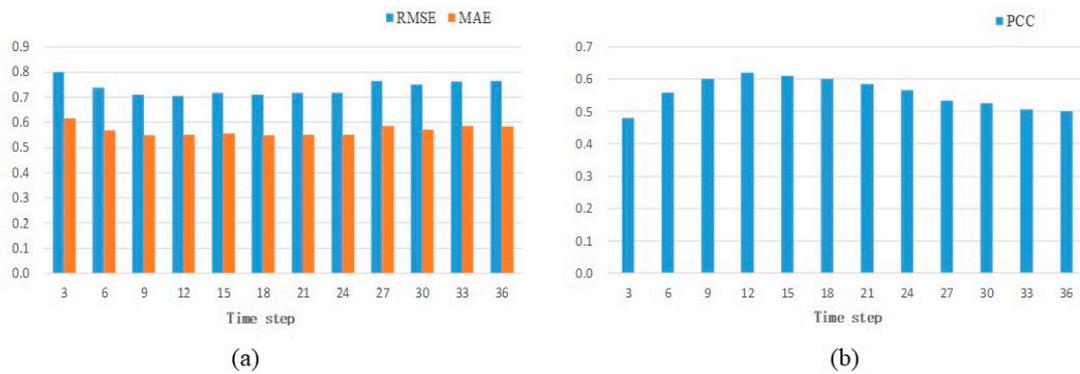


Figure 5. Performance comparison of different time windows for a 12-month lead time: (a) RMSE and MAE; (b) PCC.

3.2. Comparison and Analysis of Different Models

To verify the accuracy and validity of the prediction model proposed in this paper, other more commonly used models were constructed for comparison experiments, specifically the gate recurrent unit (GRU), multiple layer perceptron (MLP), LSTM, TCN and STL-LSTM. Figure 6 illustrates the variations in PCC, RMSE, and MAE metrics for different models when forecasting the Niño 3.4 index ahead by 1–24 months, while Table 3 presents detailed results for selected months. From the figure, it can be observed that as the lead time increases, the predictive performance of all models tends to decline. At a lead time of one month, the correlation coefficient metrics for all models are quite similar, with STL-TCN slightly higher. However, except for the STL-TCN and STL-LSTM models, the predictive performance of the remaining models rapidly deteriorates, providing effective forecasts for only about six months. The STL-TCN model achieves the longest effective forecasting period, followed by the STL-LSTM model.

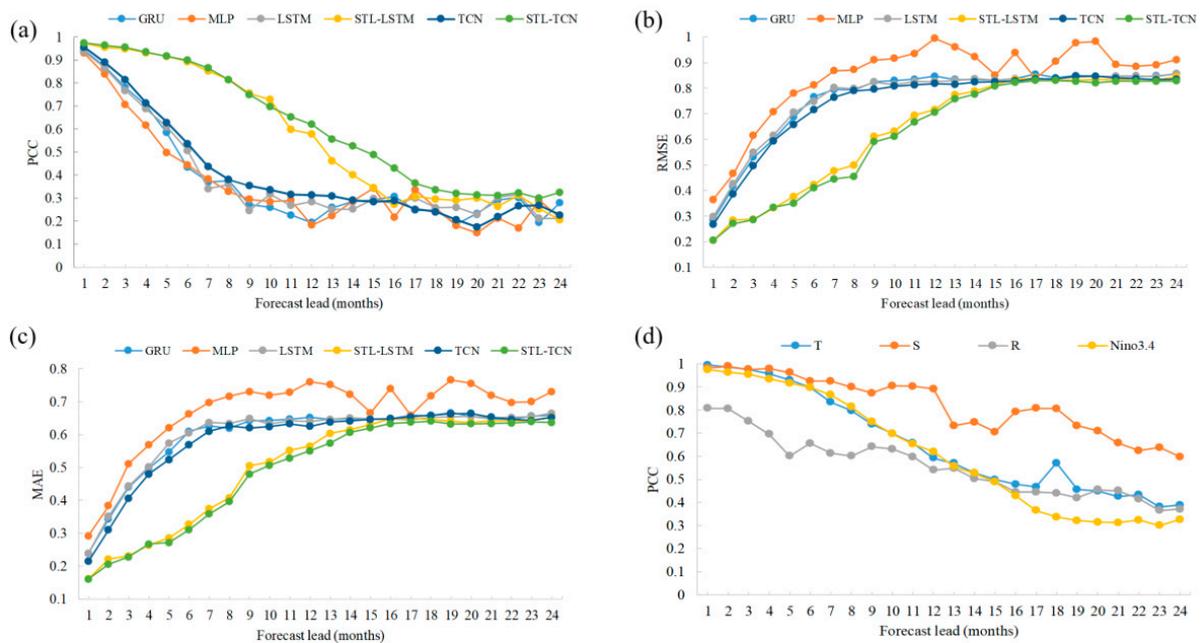


Figure 6. Comparison results of different models: (a) PCC of different models; (b) RMSE of different models; (c) MAE of different models; (d) The PCC results for each component of the STL-TCN model. T, S and R represent the forecasting performance of the decomposed trend component, seasonal component, and residual component, respectively.

Table 3. Comparison results of different models with different lead months.

Metric	Model	1	3	6	9	12
PCC	GRU	0.94	0.78	0.43	0.27	0.19
	MLP	0.93	0.71	0.44	0.30	0.18
	LSTM	0.94	0.77	0.51	0.25	0.28
	TCN	0.96	0.81	0.54	0.35	0.31
	STL-LSTM	0.97	0.95	0.89	0.75	0.58
	STL-TCN	0.97	0.96	0.90	0.75	0.62
RMSE	GRU	0.28	0.53	0.77	0.82	0.85
	MLP	0.36	0.61	0.81	0.91	0.99
	LSTM	0.30	0.55	0.75	0.82	0.83
	TCN	0.27	0.50	0.71	0.80	0.82
	STL-LSTM	0.20	0.29	0.42	0.61	0.72
	STL-TCN	0.20	0.29	0.41	0.59	0.70
MAE	GRU	0.24	0.44	0.61	0.64	0.65
	MLP	0.29	0.51	0.66	0.73	0.76
	LSTM	0.24	0.44	0.61	0.65	0.64
	TCN	0.22	0.41	0.57	0.62	0.63
	STL-LSTM	0.16	0.23	0.33	0.51	0.56
	STL-TCN	0.16	0.23	0.31	0.48	0.55

Overall, the STL-TCN model exhibits the best predictive performance with the lowest being the RMSE and MAE. The MLP model performs relatively poorly, which demonstrates the weaker ability of traditional feed-forward neural networks, such as MLP, in handling temporal relationships. TCN, on the other hand, excels at capturing long-term dependencies in the data, leading to better performance. Both LSTM and GRU demonstrate similar predictive performance, although GRU has fewer parameters. However, based on correlation coefficient indicators, the basic GRU model slightly underperforms compared to the basic LSTM model. The STL-TCN model shows a significant improvement in predictive performance compared to the TCN model, and, similarly, the STL-LSTM model exhibits notable enhancements over the LSTM model. These findings indicate that applying STL to extract components such as trend, seasonality, and residuals effectively reduces the inter-component interactions and contributes to improving the accuracy of multi-step predictions.

To further explore the impact of different decomposed sequences on predictive performance, we conducted predictions for each sequence and present the results in Figure 7d, illustrating the decomposed trend component (T), seasonal component (S) and residual component (R). By comparing the prediction results of each component, we can better understand their contributions to the final prediction of the Niño 3.4 index. Based on the results, we can observe that the seasonal component exhibits the best prediction performance. This may be attributed to its regular and periodic nature, which makes it easier to predict. The results of the trend component exhibit the closest similarity to the final forecasting results of the Niño 3.4. The remainder component is also known as the noise or the residuals of the decomposition. It contains the random fluctuations, measurement errors, or other factors that cannot be attributed to the underlying trend or seasonal effects. These unpredictable elements make it challenging to capture their future behavior accurately.

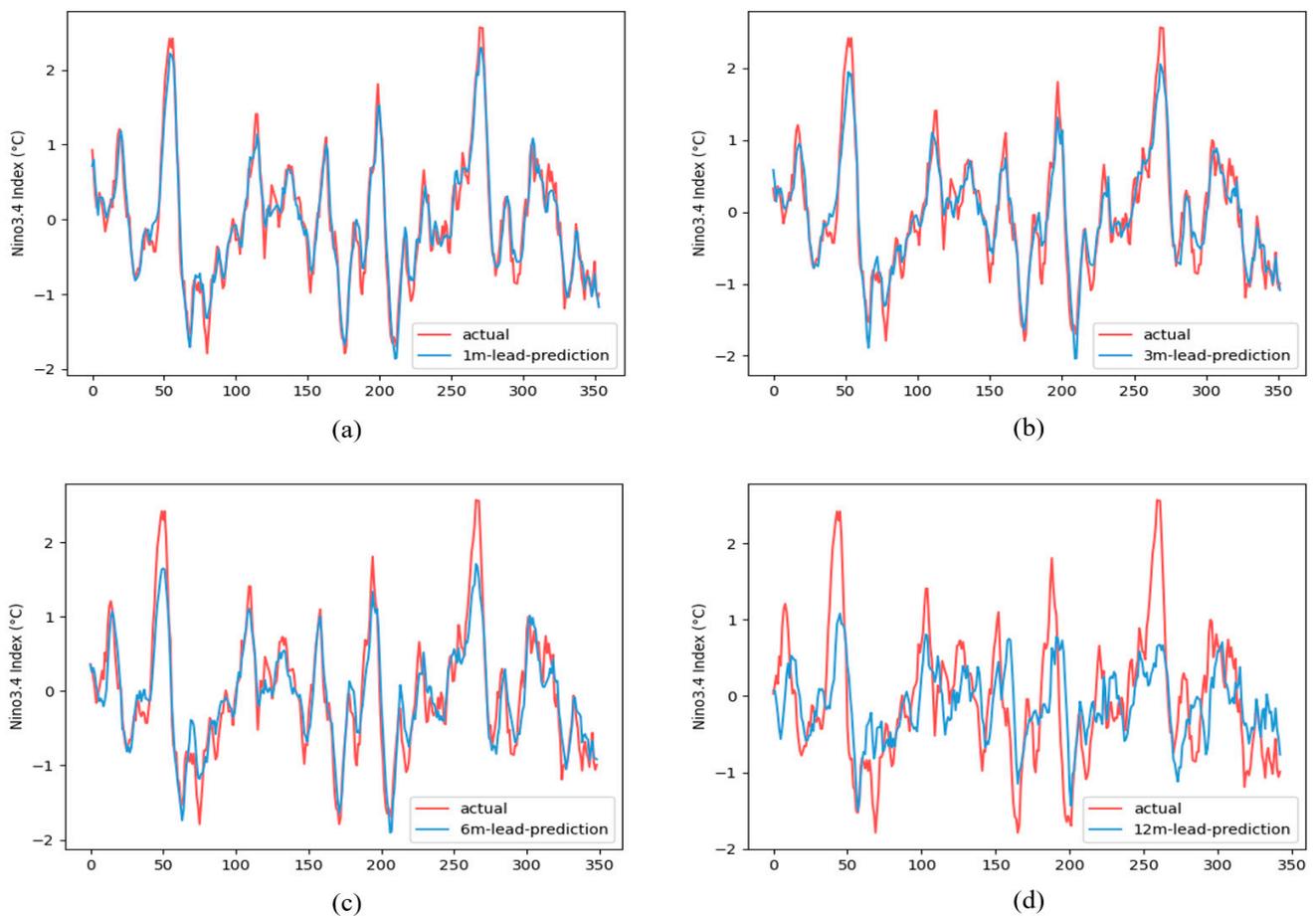


Figure 7. Forecast results based on STL-TCN Niño 3.4 index with different lead months: (a) 1 month lead; (b) 3 month lead; (c) 6 month lead; (d) 12 month lead.

Figure 7a–d show the predicted Niño 3.4 index results of the STL-TCN model 1, 3, 6, and 12 months in advance, respectively. As can be seen from the figures, all the predicted Niño 3.4 index curves are basically consistent with the actual curves in terms of growth trends and turning points, indicating that the model can effectively capture the time series variation characteristics of the Niño 3.4 index. The curve obtained by forecasting one month in advance has the highest degree of overlap with the actual Niño 3.4 index curve, with a PCC of 0.97, indicating a strong linear correlation between the predicted and actual values; and an RMSE of 0.20 °C and MAE of 0.16 °C indicate relatively small forecasting errors. However, with the increase in the prediction time, the fluctuation of the predicted value gradually decreases and a lag phenomenon appears. This phenomenon may stem from the increased uncertainty due to the increase in forecast length. The lowest match between the forecast results and the actual value curve is found 12 months ahead of time, with the PCC dropping to 0.62, RMSE increasing to 0.70 °C and MAE increasing to 0.55 °C, indicating an increase in forecast error.

3.3. ENSO Event Prediction and Analysis

To further validate the prediction effectiveness of the STL-TCN model for ENSO events, the model was used to predict ENSO events in 1998/1999, 2009/2010, 2018/2019, and 2020/2021. As can be seen in Figure 8 the STL-TCN model more accurately predicts ENSO events of significant scale in 1997/1998, and ENSO events of moderate scale in 2018/2019 and 2020/2021. In all three events, the model can better simulate the development process, intensity and duration of the events, proving its strength in capturing the dynamic characteristics of ENSO. For the 2009/2010 ENSO event, the STL-TCN model performs

better in predicting the peak of the El Niño event; however, the prediction decreases in the subsequent La Niña phase. The model fails to accurately simulate the development of the La Niña event after 14 months, showing a trend opposite to the actual situation. In summary, the STL-TCN model shows better results in predicting ENSO events of different intensities, especially in capturing event trends and peak intensities, and provides an effective tool for ENSO event studies.

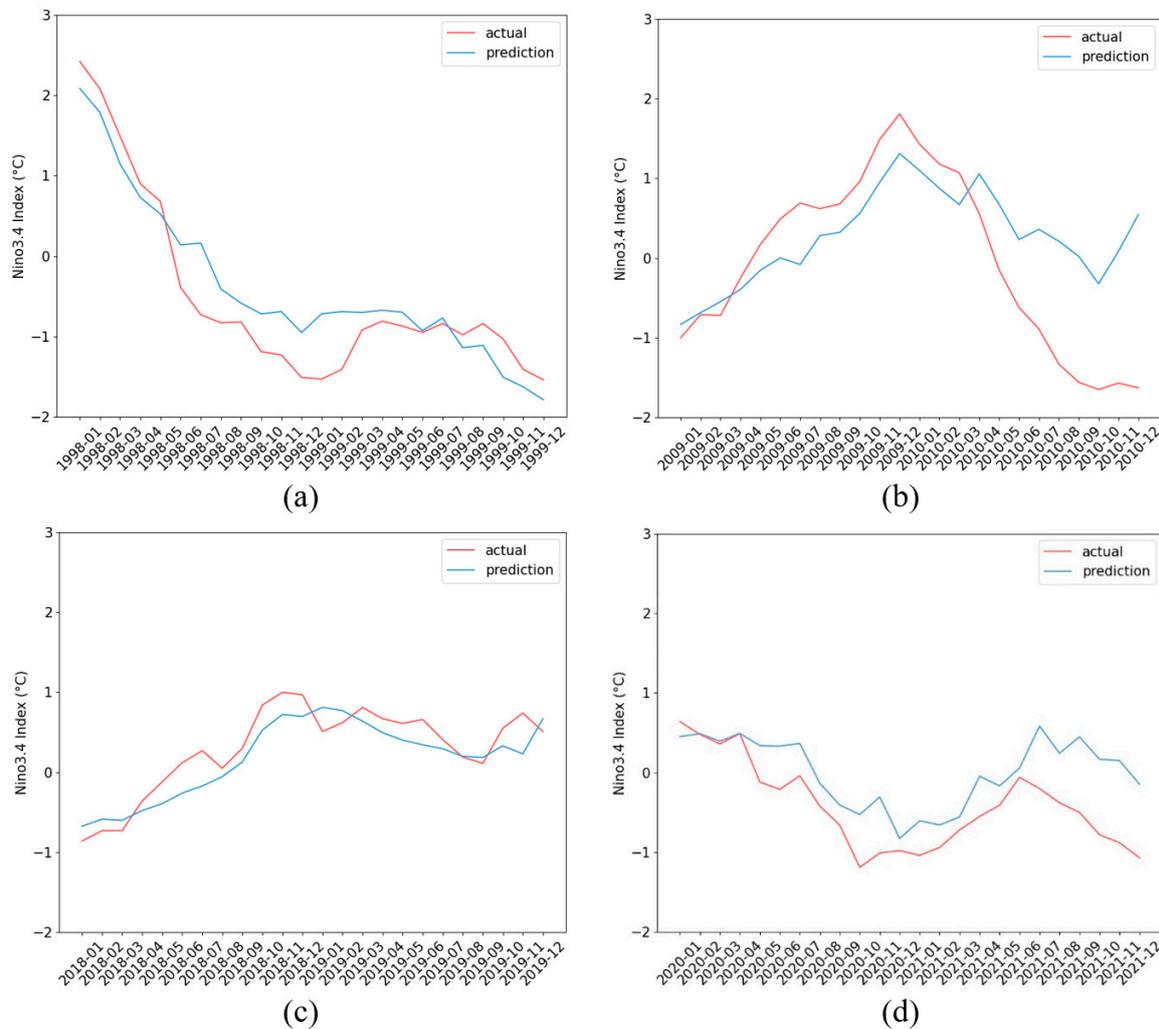


Figure 8. The predicted results of the STL-TCN model for the four ENSO events and the measured values: (a) ENSO event of 1998/1999; (b) ENSO event of 2009/2010; (c) ENSO event of 2018/2019; (d) ENSO event of 2020/2021.

4. Conclusions

This paper presents a novel hybrid model that combines the STL decomposition algorithm and the TCN model for Niño 3.4 index prediction. Taking RMSE, MAE, and R as the evaluation indicators of prediction accuracy, through experimental verification and comparative analysis, the following conclusions are drawn:

1. Combining the STL time series decomposition method and the TCN model, the cumulative error of long time series forecasting can be significantly reduced, and the forecasting accuracy can be greatly improved at the same time.
2. Compared with the popular LSTM, the STL-TCN model performs better and has better prediction results.

3. The STL-TCN model can effectively forecast the ENSO events in 1998/1999, 2009/2010, 2018/2019 and 2020/2021, and the prediction results of these events can fit the fluctuations and trends of their changes well.

In this study, the Niño 3.4 index was selected as the object for time series prediction, but ENSO also contains other indices. Therefore, applying model migration learning to other relevant ENSO indices for prediction can be considered in the future.

Author Contributions: Conceptualization, N.C., Y.W., C.S. and S.W.; investigation, Y.W.; Methodology, N.C., Y.W. and C.S.; Writing—original draft preparation, N.C.; Writing—review and editing, N.C., C.S., S.W. and Y.W.; Supervision, Y.W., S.W. and C.S.; Visualization, N.C. Funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Key Research and Development Program of China (grant 2021YFB3900900), the Provincial Key R&D Program of Zhejiang (grant 2021C01031), and the National Natural Science Foundation of China (grant 42001323). This research was also supported by the Deep-time Digital Earth (DDE) Big Science Program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data used in this paper can be downloaded from: https://psl.noaa.gov/gcos_wgsp/Timeseries/Nino34/ (accessed on 1 February 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ashok, K.; Yamagata, T. The El Niño with a difference. *Nature* **2009**, *461*, 481–484. [\[CrossRef\]](#)
2. Timmermann, A.; An, S.; Kug, J.; Jin, F.; Cai, W.; Capotondi, A.; Cobb, K.M.; Lengaigne, M.; McPhaden, M.J.; Stuecker, M.F.; et al. El Niño—Southern Oscillation complexity. *Nature* **2018**, *559*, 535–545. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Klein, S.A.; Brian, J.S.; Lau, N.C. Remote Sea Surface Temperature Variations during ENSO: Evidence for a Tropical Atmospheric Bridge. *J. Clim.* **1999**, *12*, 917–932. [\[CrossRef\]](#)
4. Taschetto, A.S.; Ummenhofer, C.C.; Stuecker, M.F.; Dommenges, D.; Ashok, K.; Rodrigues, R.R.; Yeh, S. ENSO Atmospheric Teleconnections. In *El Niño Southern Oscillation in a Changing Climate*; AGU: Washington, DC, USA, 2020; pp. 309–335, ISBN 9781119548164.
5. de Guenni, L.B.; García, M.; Muñoz, Á.G.; Santos, J.L.; Cedeño, A.; Perugachi, C.; Castillo, J. Predicting monthly precipitation along coastal Ecuador: ENSO and transfer function models. *Theor. Appl. Climatol.* **2017**, *129*, 1059–1073. [\[CrossRef\]](#)
6. Stevenson, S.; Fox-Kemper, B.; Jochum, M.; Neale, R.; Deser, C.; Meehl, G. Will There Be a Significant Change to El Niño in the Twenty-First Century? *J. Clim.* **2012**, *25*, 2129–2145. [\[CrossRef\]](#)
7. Luo, J.; Yuan, C.; Sasaki, W.; Behera, S.; Masumoto, Y.; Yamagata, T.; Lee, J.; Masson, S. Current status of intraseasonal-seasonal-to-interannual prediction of the Indo-Pacific climate. In *Indo-Pacific Climate Variability and Predictability*; World Scientific: Singapore, 2016.
8. Trenberth, K.E. The definition of el nino. *Bull. Am. Meteorol. Soc.* **1997**, *78*, 2771–2778. [\[CrossRef\]](#)
9. Ren, H.; Lu, B.; Wan, J.; Tian, B.; Zhang, P. Identification Standard for ENSO Events and Its Application to Climate Monitoring and Prediction in China. *J. Meteorol. Res.* **2018**, *32*, 923–936. [\[CrossRef\]](#)
10. Ren, H.; Zheng, F.; Luo, J.; Wang, R.; Liu, M.; Zhang, W.; Zhou, T.; Zhou, G. A Review of Research on Tropical Air-Sea Interaction, ENSO Dynamics, and ENSO Prediction in China. *J. Meteorol. Res.* **2020**, *34*, 43–62. [\[CrossRef\]](#)
11. Barnston, A.; Tippett, M.; L'Heureux, M.; Li, S.; Dewitt, D. Skill of Real-Time Seasonal ENSO Model Predictions During 2002–11: Is Our Capability Increasing? *Bull. Am. Meteorol. Soc.* **2012**, *93*, 48. [\[CrossRef\]](#)
12. Guo, Y.; Cao, X.; Liu, B.; Peng, K. El Niño Index Prediction Using Deep Learning with Ensemble Empirical Mode Decomposition. *Symmetry* **2020**, *12*, 893. [\[CrossRef\]](#)
13. Barnston, A.G.; Tippett, M.K.; Ranganathan, M.; Heureux, M.L.L. Deterministic skill of ENSO predictions from the North American Multimodel Ensemble. *Clim. Dyn.* **2019**, *53*, 7215–7234. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Collins, D.C.; Reason, C.J.C.; Tangang, F. Predictability of Indian Ocean sea surface temperature using canonical correlation analysis. *Clim. Dyn.* **2004**, *22*, 481–497. [\[CrossRef\]](#)
15. Zhou, L.; Zhang, R. A Hybrid Neural Network Model for ENSO Prediction in Combination with Principal Oscillation Pattern Analyses. *Adv. Atmos. Sci.* **2022**, *39*, 889–902. [\[CrossRef\]](#)
16. Lloyd-Hughes, B.; Saunders, M.A.; Rockett, P. A Consolidated CLIPER Model for Improved August–September ENSO Prediction Skill. *Weather Forecast.* **2004**, *19*, 1089–1105. [\[CrossRef\]](#)

17. Lou, J.; Kane, T.J.O.; Holbrook, N.J. A Linear Inverse Model of Tropical and South Pacific Seasonal Predictability. *J. Clim.* **2020**, *33*, 4537–4554. [[CrossRef](#)]
18. Xue, Y.; Leetmaa, A. Forecasts of tropical Pacific SST and sea level using a Markov model. *Geophys. Res. Lett.* **2000**, *27*, 2701–2704. [[CrossRef](#)]
19. Jiang, Z.; Ding, Y.; Zhou, Q. A generalized model using canonical mixed regression and forecasting test to ENSO. *J. Trop. Meteorol.* **1999**, *4*, 322–329.
20. Rosmiati, R.; Liliyasi, S.; Tjasyono, B.; Ramalis, T.R. Development of ARIMA technique in determining the ocean climate prediction skills for pre-service teacher. *J. Phys. Conf. Ser.* **2021**, *1731*, 12072. [[CrossRef](#)]
21. Ye, M.; Nie, J.; Liu, A.; Wang, Z.; Huang, L.; Tian, H.; Song, D.; Wei, Z. Multi-Year ENSO Forecasts Using Parallel Convolutional Neural Networks With Heterogeneous Architecture. *Front. Mar. Sci.* **2021**, *8*, 717184. [[CrossRef](#)]
22. Ham, Y.; Kim, J.; Luo, J. Deep learning for multi-year ENSO forecasts. *Nature* **2019**, *573*, 568–572. [[CrossRef](#)]
23. Nootboom, P.D.; Feng, Q.Y.; López, C.; Hernández-García, E.; Dijkstra, H.A. Using network theory and machine learning to predict El Niño. *Earth Syst. Dynam.* **2018**, *9*, 969–983. [[CrossRef](#)]
24. Pei, Z.; Yingjie, H.; Bingyi, H.; Jun, W. Spring predictability barrier phenomenon in ENSO prediction model based on LSTM deep learning algorithm. *Beijing Da Xue Xue Bao* **2021**, *57*, 1071–1078.
25. Chen, D.; Zhang, J.; Jiang, S. Forecasting the Short-Term Metro Ridership With Seasonal and Trend Decomposition Using Loess and LSTM Neural Networks. *IEEE Access* **2020**, *8*, 91181–91187. [[CrossRef](#)]
26. Li, Y.; Bao, T.; Gong, J.; Shu, X.; Zhang, K. The Prediction of Dam Displacement Time Series Using STL, Extra-Trees, and Stacked LSTM Neural Network. *IEEE Access* **2020**, *8*, 94440–94452. [[CrossRef](#)]
27. Jiao, F.; Huang, L.; Song, R.; Huang, H. An Improved STL-LSTM Model for Daily Bus Passenger Flow Prediction during the COVID-19 Pandemic. *Sensors* **2021**, *21*, 5950. [[CrossRef](#)]
28. Yang, S.; Deng, Z.; Li, X.; Zheng, C.; Xi, L.; Zhuang, J.; Zhang, Z.; Zhang, Z. A novel hybrid model based on STL decomposition and one-dimensional convolutional neural networks with positional encoding for significant wave height forecast. *Renew. Energy* **2021**, *173*, 531–543. [[CrossRef](#)]
29. Cleveland, R.B.; Cleveland, W.S.; Mcrae, J.E.; Terpenning, I. STL: A seasonal-trend decomposition. *J. Off. Stat* **1990**, *6*, 3–73.
30. Bai, S.; Kolter, J.Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv* **2018**, arXiv:1803.01271.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.