



# Deep learning with autoencoders and LSTM for ENSO forecasting

Chibuikie Chiedozie Ibebuchi<sup>1,2</sup> · Michael B. Richman<sup>3</sup>

Received: 8 November 2023 / Accepted: 28 February 2024  
© The Author(s) 2024

## Abstract

El Niño Southern Oscillation (ENSO) is the prominent recurrent climatic pattern in the tropical Pacific Ocean with global impacts on regional climates. This study utilizes deep learning to predict the Niño 3.4 index by encoding non-linear sea surface temperature patterns in the tropical Pacific using an autoencoder neural network. The resulting encoded patterns identify crucial centers of action in the Pacific that serve as predictors of the ENSO mode. These patterns are utilized as predictors for forecasting the Niño 3.4 index with a lead time of at least 6 months using the Long Short-Term Memory (LSTM) deep learning model. The analysis uncovers multiple non-linear dipole patterns in the tropical Pacific, with anomalies that are both regionalized and latitudinally oriented that should support a single inter-tropical convergence zone for modeling efforts. Leveraging these encoded patterns as predictors, the LSTM - trained on monthly data from 1950 to 2007 and tested from 2008 to 2022 - shows fidelity in predicting the Niño 3.4 index. The encoded patterns captured the annual cycle of ENSO with a 0.94 correlation between the actual and predicted Niño 3.4 index for lag 12 and 0.91 for lags 6 and 18. Additionally, the 6-month lag predictions excel in detecting extreme ENSO events, achieving an 85% hit rate, outperforming the 70% hit rate at lag 12 and 55% hit rate at lag 18. The prediction accuracy peaks from November to March, with correlations ranging from 0.94 to 0.96. The average correlations in the boreal spring were as large as 0.84, indicating the method has the capability to decrease the spring predictability barrier.

**Keywords** Deep learning · Machine learning · Autoencoders · LSTM · ENSO · Tropical pacific ocean · Predictions · Non-linear · Spring predictability barrier

## 1 Introduction

El Niño Southern Oscillation (ENSO) is a dominant recurrent climate pattern influencing global weather patterns (Trenberth and Hoar 1996). Centered in the tropical Pacific Ocean, its influence extends across the globe, affecting weather patterns, ocean conditions, atmospheric circulation, and even the global economy (Hsiang et al. 2011; Odériz et al. 2020; Hrudya et al. 2021; Dufrenot et al. 2023). ENSO encompasses two primary states or phases: El Niño and La Niña. Both arise from ocean-atmosphere interactions (Dawson and O'Hare 2000). During El Niño events, the central

and eastern tropical Pacific waters become anomalously warm, leading to a weakening of the trade winds (Cai et al. 2021). Moreover, the upwelling off the west coast of South America diminishes, reducing the cold tongue of oxygen-rich water in the far eastern tropical Pacific. Conversely, La Niña events are defined by cooler-than-average sea surface temperatures (SSTs) in the central and eastern tropical Pacific, strengthening of trade winds, and an extensive cold tongue off the west coast of South America. These shifts impact not only SSTs but also generate significant atmospheric, biologic and economic anomalies (Cai et al. 2020; Hu et al. 2021; Reddy et al. 2022). Owing to the widespread global impacts (Timmermann et al. 2018), enhanced characterization and forecast of ENSO at multiple lead times is crucial.

ENSO events are characterized by different locational manifestations, often referred to as classical and non-classical ENSO. Classical, canonical or Eastern Pacific ENSO is the most well-known form of El Niño. The Niño 3.4 index is the most often applied indicator that describes the classical ENSO pattern (Lee and McPhaden 2010). It measures sea

✉ Chibuikie Chiedozie Ibebuchi  
cibebuch@kent.edu

<sup>1</sup> Department of Geography, Kent State University, Kent, Ohio, USA

<sup>2</sup> ClimRISE lab, Kent State University, Kent, Ohio, USA

<sup>3</sup> School of Meteorology, University of Oklahoma, Norman, OK, USA

surface temperature anomalies in the east-central tropical Pacific and serves as a primary arbiter of ENSO's strength and phase. Over the years, research has uncovered the existence of multiple non-classical ENSO patterns, suggesting its variability and non-linear nature (Geng et al. 2020; Dasgupta et al. 2021). These ENSO patterns diverge from the classical ENSO based on their specific geographical location, spatial patterns, duration, and global impact, reinforcing the complexity of the ENSO climate phenomenon (Ashok et al. 2007; Kao and Yu 2009; Yeh et al. 2009; Timmermann et al. 2018).

Historically, two main types of models have been used to predict ENSO: dynamic and statistical (Ham et al. 2019). Dynamic models, grounded in the physical laws governing the atmosphere and oceans, simulate the interactions that give rise to ENSO. Such models have struggled to depict the observed tropical patterns, including ENSO (Hidalgo and Alfaro 2015; Ortega et al. 2021; Zhao and Sun 2022; Zhang et al. 2022). In contrast, statistical models rely on observed relationships between selected variables known to modulate ENSO to make predictions. Like their dynamic counterparts, current statistical models provide imperfect representations of ENSO in its prediction. Consequently, both approaches have their merits and challenges in achieving consistent, long-term predictive success; therefore, it is becoming increasingly popular to couple the dynamical model with a statistical model to reduce dynamic model biases (L'Heureux et al. 2020; Zhang et al. 2022; Cho et al. 2022).

Owing to the chaotic nature of the oceans and atmosphere, ENSO predictability varies with the lead time. Generally, ENSO is most predictable at shorter term forecasts, such as  $\leq 6$ -months (Jin et al. 2008). The inherent non-linearity in the development within each ENSO phase, combined with other factors, e.g., external noise, model inaccuracies and climate change, make its long-term prediction challenging (L'Heureux et al. 2020). Inaccurate ENSO predictions are exacerbated during the spring months, and at any lead-time that includes the spring (e.g., 12-months and longer forecasts), likely because of seasonal transition, use of linear models, ocean-atmosphere feedback, and numerical model limitations, among others (Barnston et al. 2012; Chen et al. 2023a, b). Additionally, capturing the onset and termination of ENSO events remains problematic (Duan and Wei 2013; Wu et al. 2021).

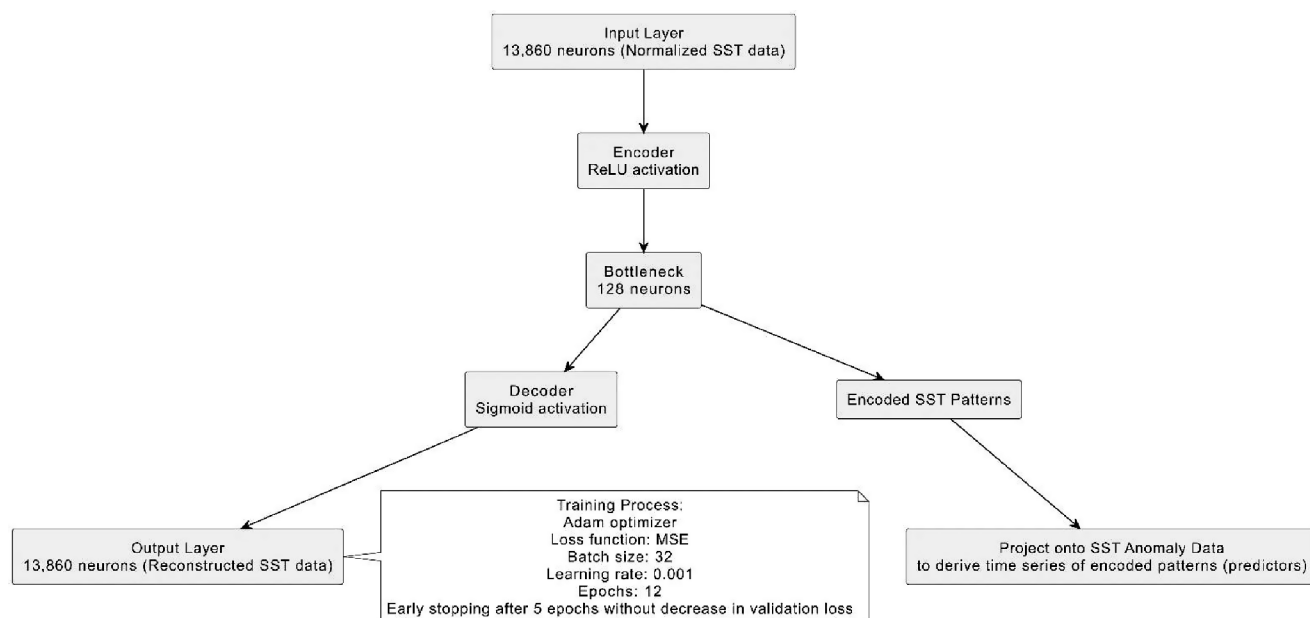
In recent years there has been a burgeoning in the application of Artificial Neural Networks (ANN) in climate science, including for ENSO prediction (Kim et al. 2022; Liu et al. 2022; Zhou and Zhang 2023; Wang et al. 2023; Jonnalagadda and Hashemi 2023). ANNs have shown potential in capturing non-linear relationships intrinsic to ENSO (Ham et al. 2019; Zhao and Sun 2022; Zhang et al. 2022). Some

studies employing recurrent neural networks and Convolutional Neural Networks (CNN) have achieved modest success in lead times ranging from 6 to 12 months and beyond, often outperforming traditional models (Mu et al. 2021, 2022; Liu et al. 2023; Patil et al. 2023; Wang and Huang 2023; Chen et al. 2023a, b). Compared to traditional ANN and CNN, the combination of Autoencoders (AE) (Saha et al., 2020; Ibebuchi and Richman 2024) and Long Short-Term Memory Networks (LSTM) represent a viable novel alternative for ENSO prediction. AEs can compress vast amounts of climatic data into meaningful patterns, potentially unveiling novel ENSO predictors. LSTMs, with their ability to "remember" long-term dependencies (Mu et al. 2020), are ideal for predicting a phenomenon with the temporal complexity of ENSO. Therefore, this study complements the existing ANN and CNN ENSO forecast literature, by presenting a novel approach combining AE and LSTM deep learning to forecast ENSO in 6-, 12- and 18-month lead times.

## 2 Data and methods

SST data in the tropical Pacific Ocean were obtained from the extended reconstructed sea surface temperature, version 5 (Huang et al. 2017), at monthly temporal resolution from 1950 to 2022. The horizontal resolution of the SST data is  $2^\circ$  longitude and latitude. This provided 13,860 grid points and 876 data (time) series. The monthly SST data were pre-processed, i.e., deseasonalized by subtracting the long-term mean from the corresponding monthly SST values. The monthly Niño 3.4 index data were obtained from <https://psl.noaa.gov/data/climateindices/list/>.

The steps outlined below were followed in developing the deep learning model for predicting the Niño 3.4 index. First, the SST patterns in the tropical Pacific Ocean were encoded using AE (Hinton and Salakhutdinov 2006) applied to the deseasonalized  $876 \times 13,860$  SST data matrix. AE performs a task similar to the unsupervised learning technique, principal component analysis (PCA), which includes reducing the dimensionality of the SST anomaly data, denoising the data and extracting the most crucial inherent patterns in the data. However, unlike linear PCA, AE is a neural network architecture that can extract non-linear patterns from the data. AE compresses the SST anomaly data into a compact latent space representation using an encoder and then reconstructs the original SST data from this representation using a decoder (Hinton and Salakhutdinov 2006). Figure 1 shows the schematic diagram of the AE model architecture for deriving the predictors of SST anomalies in the tropical Pacific, which are utilized subsequently for ENSO prediction.



**Fig. 1** Schematic diagram of the AE model architecture applied to obtain predictors (spatiotemporal patterns) in the tropical Pacific for ENSO prediction

The SST anomaly data were further preprocessed by normalizing between [0,1], ensuring uniformity of the data values, which is important in training neural networks (improves accuracy, lessens overfitting and improves interpretability) and helps in faster convergence during training (Pal and Sudeep 2016; Phan et al. 2021). The input and output neurons correspond to the dimension of the input data, equating to 13,860 grids for the data used in this study. The hidden neurons are selected by iterating over several configurations to find the best trade-off between data representation and model complexity. Thus, different neurons were experimented with, alongside the optimal epoch number, resulting in a solution that minimizes the reconstruction error between the original data and the reconstructed data output by the autoencoder. By using batches of 5 epochs (Forouzesh and Thiran 2021), training is halted if there is not a significant decrease in validation loss after a given batch, thereby preventing potential overfitting and unnecessary computation. Figure A1 depicts both training and validation loss curves across epochs for each neuron configuration. For configurations with 128 and 256 neurons, both the training and validation loss patterns show a consistent downward trend, eventually asymptoting to a minimum loss. Such patterns indicate that the models are learning effectively without showing obvious signs of overfitting, as an overfitting signature would show a simultaneous decreasing training loss coupled with an increasing validation loss. Although, both the 128- and 256-neuron configurations yielded comparable performance with the least losses (Figure A1), the

more parsimonious model of the 128-neuron configuration was selected for its computational efficiency.

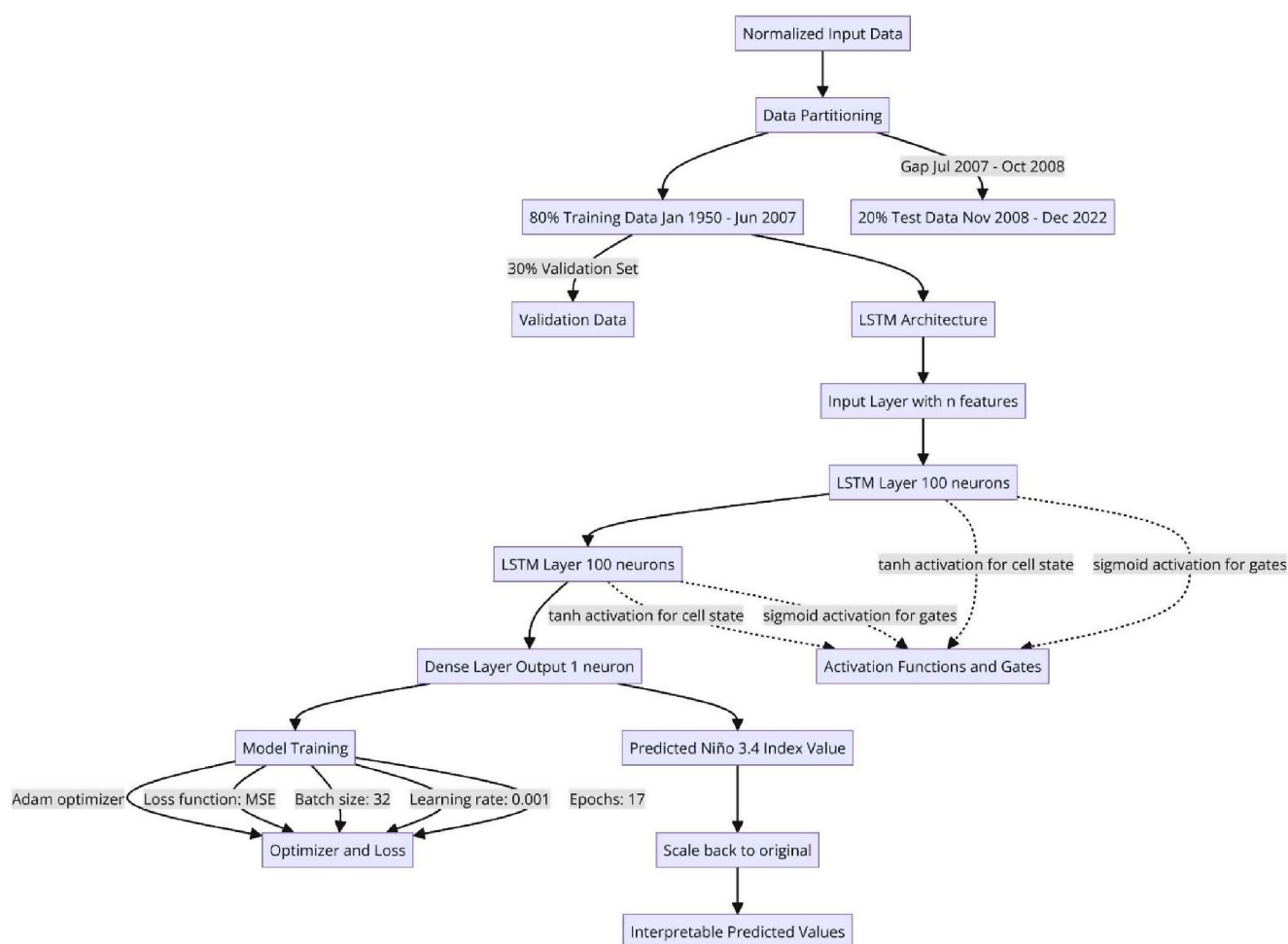
The encoder utilized the rectified linear unit activation function (Glorot et al. 2011) that introduces non-linearity in the model. The decoder used sigmoid to produce outputs between [0,1]. The Adam optimizer (Kingma and Ba 2014) was used in training the AE because of its adaptive learning rates and efficiency. The compiled architectures were used to train the AE, and the encoder part was applied to reduce the dimension of the SST anomaly data, preserving crucial SST anomaly patterns that were used as predictors of the Niño 3.4 index. Identifying and encoding regions or patterns that act as predictors of the ENSO mode is crucial as not all patterns or regions in the SST data might have predictive power. By focusing on these key patterns, the model may provide more specific insights into the dynamics and evolution of ENSO. The associated time series of the encoded spatial patterns, necessary for predicting the time series of the Niño 3.4 index, were derived by projecting the encoded spatial patterns onto the monthly global SST anomaly data in the tropical Pacific. Moreover, the final predictors were selected only if they had a Pearson correlation of at  $\geq 0.8$  with the Niño 3.4 index. We used 0.8 since it was suitable as a trade-off between a tractable number of patterns and forecast accuracy.

Second, the predictive deep learning model was defined using a dataset of both the input predictors (i.e., time series of encoded SST patterns) and the predictand output (Niño 3.4 index). Both the predictors and predictand, which have different scales, were normalized using Min-Max

normalization (Patro and Sahu 2015). This normalization was applied to the training data to derive the scaling factors. The same factors were then used to scale both the training, validation and testing data. This practice ensures consistency and avoids information from the test set leaking into the model training process, leading to seemingly inflated performance of the predictions. The same scaling factors derived from the training data were also used to transform the predicted values back to the original Niño 3.4 index scale, ensuring interpretability. Figure 2 shows the flow chart in applying the LSTM model.

For model training and testing, the dataset was portioned into an 80% training subset, from January 1950 to June 2007. Because of the inclusion of an 18-month lag, the test data subset starts from November 2008 and continues until December 2022, thereby omitting the period between July 2007 and October 2008. This adjustment ensures that we can validate forecasts for all considered lags within a consistent and continuous timeframe. From the 80% training subset, 30% of it was used for validation during the hyperparameter tuning.

Following Ham et al. 2019, who applied deep learning to predict the Niño 3.4 index up to 18 months lag, predictors were lagged by 6, 12 and 18 months, which means, for instance, that the model would predict the Niño 3.4 index for a specific month using encoded patterns from at least 6 months prior. LSTM neurons are recurrent neural network structures adept at remembering patterns over long sequences, making them suitable for time series forecasting tasks. The input layer is defined by the shape of the input data; in this case, we used one input layer and, within that input layer, we have  $n$  input features, where  $n$  is the number of predictors. One dense layer with a single neuron was used as the output layer. The initial layer of the model consists of an LSTM layer with 100 neurons. The choice of 100 neurons was determined through experimental evaluations to maximize the predictive accuracy of the validation data. The LSTM model utilized in this study employs Keras's default implementation (Mohan et al. 2018), which incorporates the constant error carousel mechanism as described by Hochreiter and Schmidhuber (1997). This mechanism preserves the internal cell state across time steps by a recurrent



**Fig. 2** Flow chart in applying the LSTM for ENSO predictions

connection with a fixed weight of 1.0, ensuring that the signal's strength doesn't vanish or approach infinity over long sequences. The model's forget gate is also leveraged to reset the cell state when necessary, providing the network with the ability to learn and decide when to clear previous information. A vanilla LSTM architecture (Van Houdt et al. 2020) was applied. The Keras LSTM mechanism employs the tanh activation for the cell state and the sigmoid activation for its gates. This LSTM layer is configured to return sequences, allowing for subsequent LSTM layers to receive input sequences, thereby enhancing the depth and complexity the model can capture. Consequently, a second LSTM layer was added, also with 100 units. The model terminates with a dense layer containing a single neuron, responsible for outputting the forecast Niño 3.4 index value. For compiling the model, the Adam optimizer was selected (Choi et al. 2019), and the mean squared error was employed as the loss function.

Third, several metrics were utilized to evaluate the monthly predicted Niño 3.4 index against the actual monthly Niño 3.4 index during the test period. These metrics include the Pearson correlation, which is efficient in comparing the distance between two standardized vectors, the mean absolute error and the root mean square error. However, since the primary concern of stakeholders (e.g., policymakers, climatologists) is to predict extreme events (i.e., strong or very strong ENSO events), owing to their significant climatic impacts, another metric has been added, introducing a threshold to examine if the predicted values are consistent with actual values exceeding that threshold. Three non-overlapping thresholds were applied after z-score standardization is employed to convert the Niño 3.4 index values into anomalies that represent deviations from the climatological normal of 1991–2020, facilitating the identification of unusual or extreme events: (1) when the actual index is  $|z| \geq 1.5$  standard deviations above the mean for the test period to define very strong positive and negative ENSO events; (2) when the actual index is  $|z| \geq 1.0$  standard deviations above the mean for the test period to define strong positive and negative ENSO events; (3) when the actual index is  $|z| \geq 0.5$  standard deviations above the mean for the validation period to define moderate positive and negative ENSO events. Consequently, a true positive (hit) is defined as when both actual and predicted values for a date exceed the threshold in the same direction. For example, for a strong El Niño event (implying index  $> +1.5$  threshold) if the actual index is greater than  $+1.5$  at a given date, the predicted should be greater than  $+1.5$  on that same date to be a hit. False negative (miss) is defined as when the actual value for a date exceeds the threshold but the predicted value doesn't. False positive (false alarm) is when the predicted value for a date exceeds the threshold but the

actual value doesn't. A false positive might imply a water manager preparing for an ENSO event that doesn't occur (which could have economic implications), whereas a false negative could simply be that manager caught off guard by an unpredicted event and could also lead to unpreparedness for potential adverse impacts, which might be devastating in some cases. Further, false positives and negatives lead to mistrust of the models.

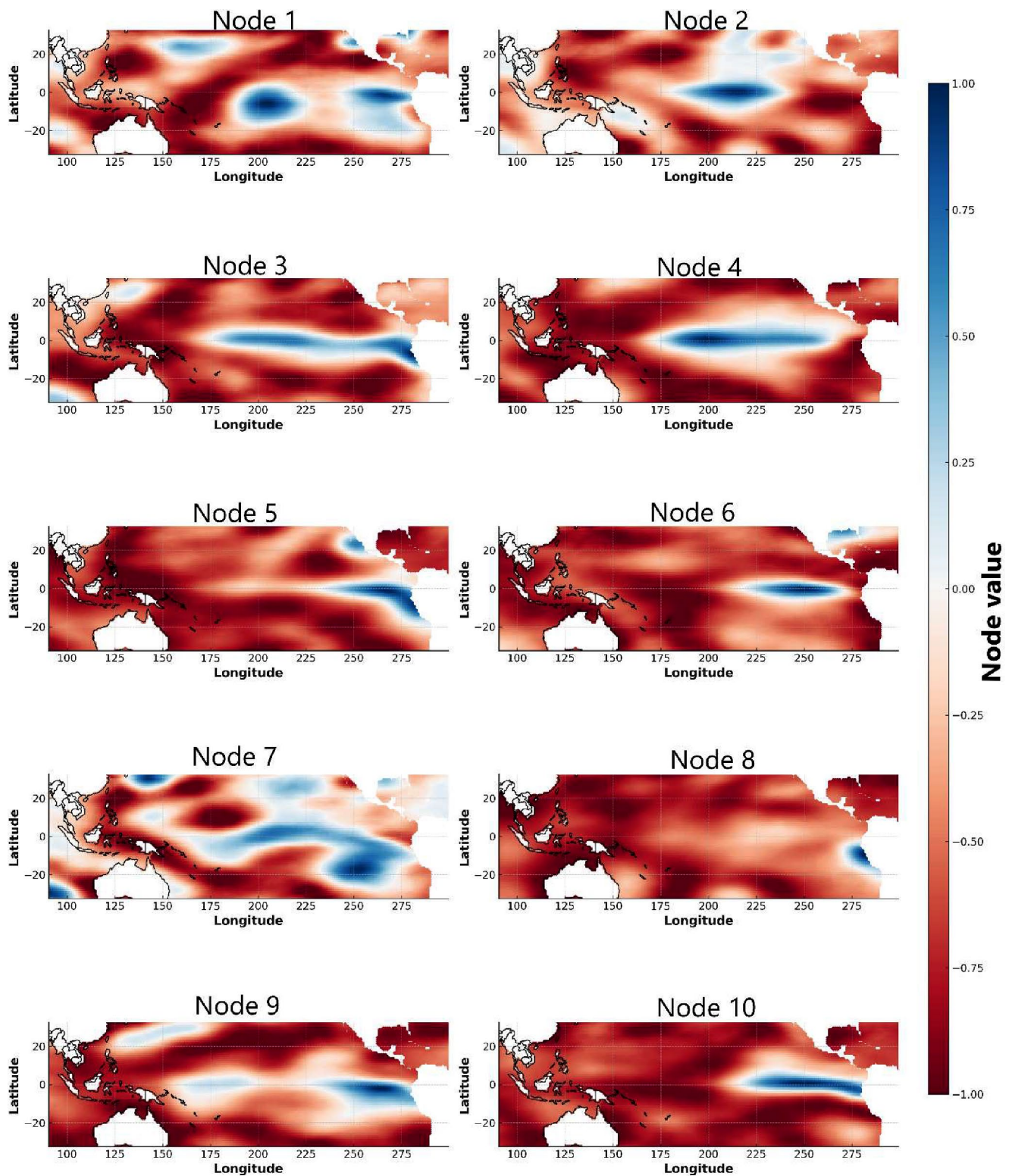
Further evaluation metrics are (1) Hit Rate (Sensitivity or Recall), defined as the ratio of hits to the total of hits and misses; (2) False Alarm Ratio, defined as the ratio of false alarms to the sum of hits and false alarms; (3) Critical Success Index (CSI) Threat Score), defined as the ratio of hits to the sum of hits, misses, and false alarms (Schaefer 1990); (4) mean absolute error of duration: defined as the average of the absolute difference in months or days between the actual and predicted event durations.

### 3 Results and discussion

The correlation criterion of at least 0.8 between the time series of the AE patterns and the Niño 3.4 index revealed that 10 patterns would be investigated. Figure A2 also shows the correlation between the 10 encoded patterns and the Niño 3.4 index at lags 0, 6, and 12. All correlations were at least 0.8 at lag 0. Nodes 3 and 4 have the highest correlations with the Niño 3.4 index. Figure 3 shows the spatial pattern of the encoded patterns representing distinct non-linear variability patterns in the tropical Pacific Ocean. By comparing the largest positive and negative excursions of the time series in the panels of Figure A3, the temporal variability of the encoded time series is consistent with those from the Niño 3.4 index, indicating that the encoded patterns using the AE possess sufficient spatiotemporal characteristics of ENSO. Table A1 shows that a substantial portion of the variance in each node's time series is explained by their linear association with the Niño 3.4 index with the variance overlap ranging from approximately 63.75–85.68%. Thus, there is still a portion of the variance difference (ranging from about 14.32–36.25%) that cannot be accounted for using a linear association. This remaining unexplained variance implies nonlinear methods might offer additional insights. Hence, we utilize the encoded patterns as predictors for ENSO prediction using the deep LSTM predictive model.

Node 4, which has the highest correlation ( $R=0.93$ ) to the Niño 3.4 index, shows a typical Niño 3.4 pattern associated with the SST anomalies in the central to eastern tropical Pacific. Other nodes reveal the distinct zonal dipole patterns in the tropical Pacific. This contrasts with the traditional use of linear PCA to depict ENSO that does not discriminate the encoded patterns (Ibebuchi, 2024). From Fig. 4, nodes 5, 6,



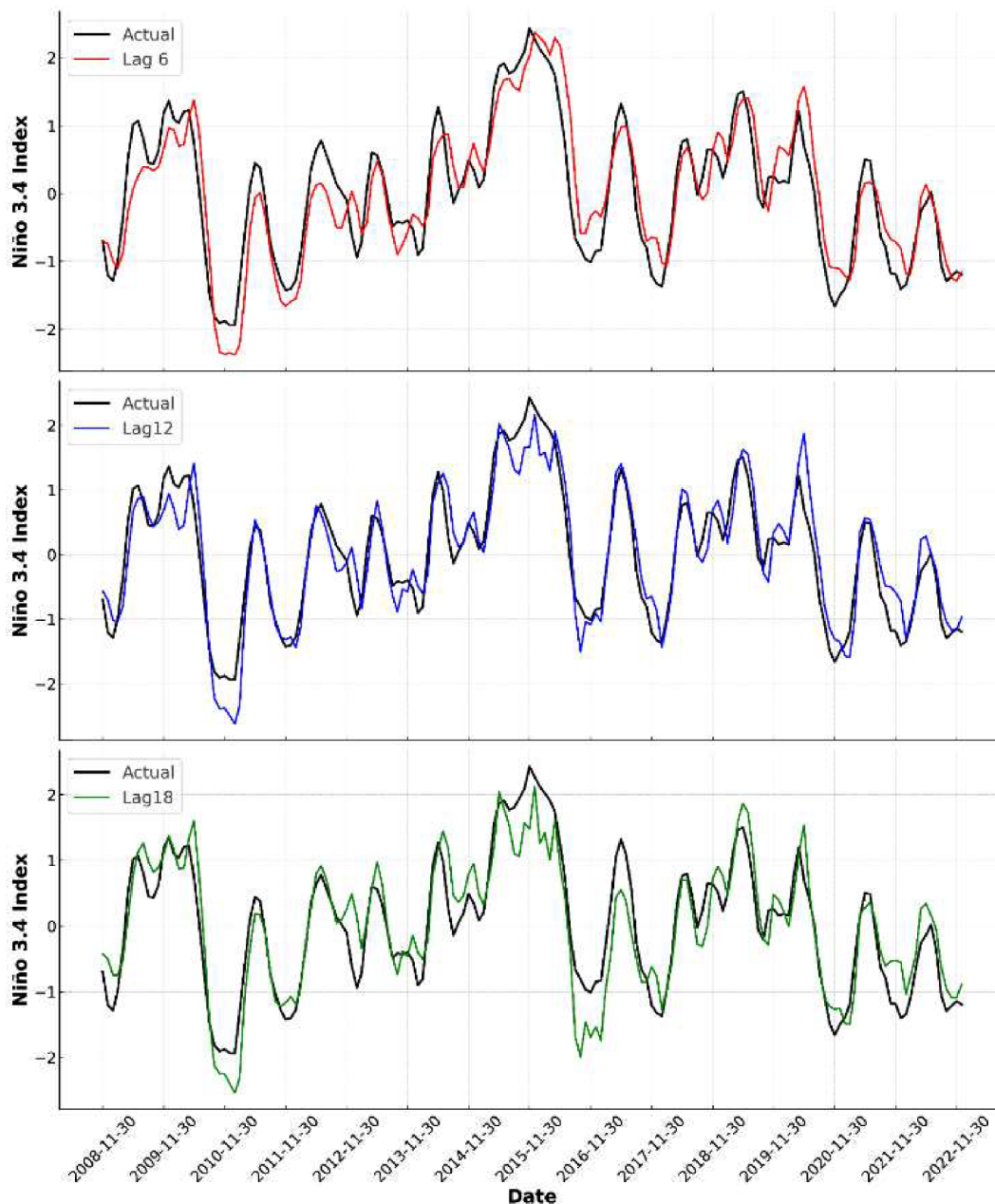


**Fig. 3** Encoded SST pattern in the tropical Pacific Ocean using the autoencoder neural network. The patterns were z-score standardized to obtain anomalies, which aids their interpretability.

8, 9, and 10 can be grouped as East Pacific ENSO patterns, whereas nodes 2 and 4 can be grouped as Central Pacific ENSO patterns. Nodes 1, 3, and 7 are associated with non-canonical patterns over the tropical Pacific involving both the central and eastern tropical Pacific Ocean. Such non-canonical El Niño patterns and their impact on Atlantic tropical activity have been reported by Larson et al. (2012).

At different lag times, these non-linear patterns act as predictors of ENSO variability. Interestingly, Figure A2 shows that the encoded patterns in Fig. 4 have larger correlations with the Niño 3.4 index at lag 12 compared to lag

6, which implies these nonlinear patterns have a reasonable linear predictive relationship with the Niño 3.4 index; and that they can predict ENSO up to a year ahead. Whereas the slightly smaller correlation at lag 6 might be linked to the spring predictability barrier (SPB), the fact that the lag 12 correlation increases over lag 6 suggests this method lowers the predictability barrier compared to traditional linear models (Lopez and Kirtman, 2014; Chen et al. 2020). Other studies have also documented the existence of several non-linear types of ENSO consistent with the encoded patterns in Fig. 4 (An et al., 2004; Levine and McPhaden 2015).



**Fig. 4** Monthly time series (2008-11-30 to 2022-12-31) of the actual and predicted Niño 3.4 index at 6-, 12-, and 18-month lead times. Above-average (below-average) values indicate El Niño (La Niña) events, respectively

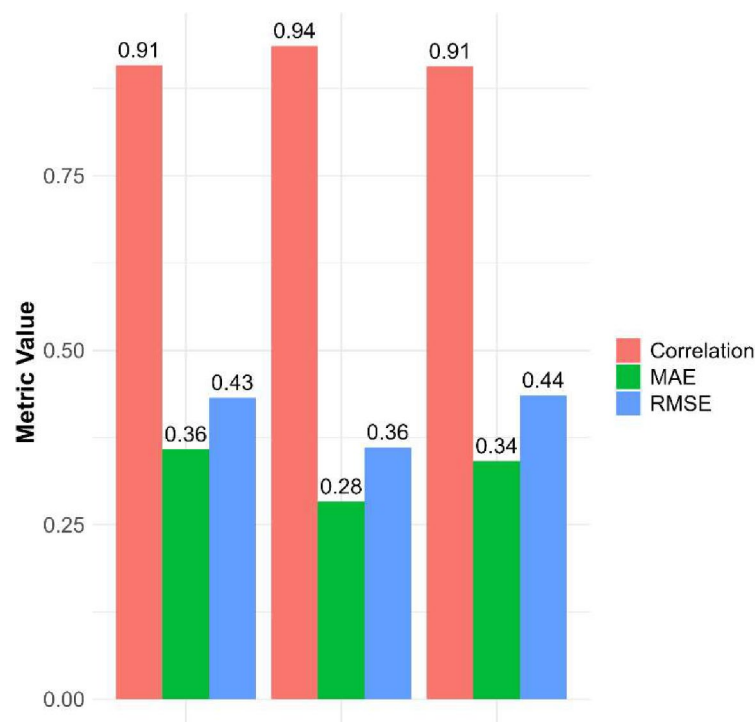
From Fig. 3, node 3, which is associated with Niño 3.4, and nodes 2 and 4, which are associated with Niño 4 (i.e., SST anomalies in the central Pacific) are the most common ENSO patterns. Node 8 reproduces the cold tongue index with a correlation of 0.70. As the reproduction of the cold tongue is part of the ENSO phenomenon (Hu et al. 2019) and its reproduction has been problematic in climate models (e.g., see discussion in Ying et al. 2019), node 8, and perhaps several other AE patterns, may be useful in validating ENSO in such models.

Further, the time series of the spatially encoded patterns in Fig. 3 are applied as predictors of the ENSO mode in at least a 6-month lead time. Figure 4 shows the actual and predicted Niño 3.4 index and Fig. 5 shows the validation metrics. From Fig. 4, the monthly variability of the actual Niño 3.4 index was captured by the predicted index, though in some months there are visual dissimilarities in the magnitude of the actual and predicted values. Also, the accuracy of the predictions for the distinct strength of ENSO events appears to be dependent on the lag considered. For example, Fig. 4 indicates that the 2015 El Niño event was better represented at lag 6 compared to lags 12 and 18. Considering the evaluation metrics in Fig. 5, the AE model shows promising fidelity in predicting ENSO monthly variability with at least a 6-month lead time, and notably most accurate in a 12-month lead time, with excellent results out to 18 months. In previous studies, the accuracy of the predictions decreased as lead time increased (e.g., Ham et al. 2019; Qiao et al. 2023).

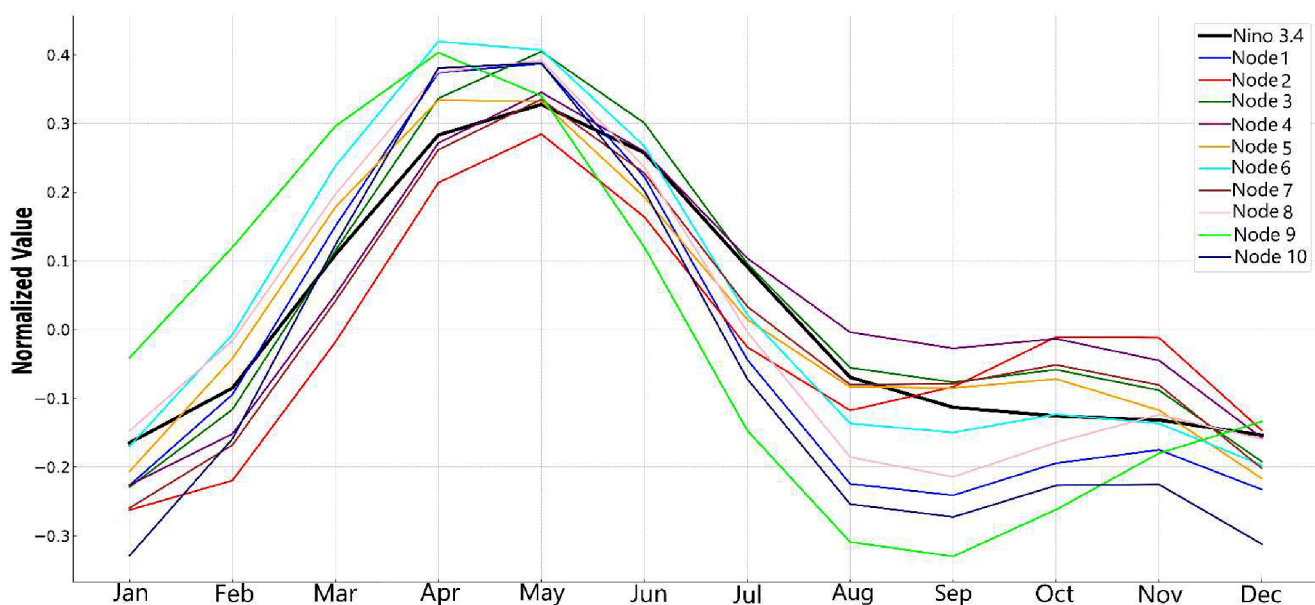
From Fig. 5, the correlation between the standardized actual and predicted monthly Niño 3.4 index during the testing period, is 0.91 at lag 6, 0.94 at lag 12 and 0.91 at lag 18. The larger predictive accuracy at lag 12 can be traced to the findings that the encoded patterns are more correlated to the Niño 3.4 index at lag 12 (Figures A3 and A3), better capturing the annual cycles of the ENSO phenomenon as shown in Fig. 6, beyond the SPB. In some months outside the DJFM (when the predictive accuracy peaked), such as July 2016 for lag 6, September 2016 for lag 18, and May 2020 for lag 12, the difference between the actual and predicted monthly index exceeded 1.0 (Figure A4). This is mostly because of the SPB and the limitation of the current model in capturing the full variability of the reference Niño 3.4 index. Generally, for Lag 6, 72.94% of biases have an absolute magnitude less than 0.5, and 27.06% of biases have an absolute magnitude greater than or equal to 0.5. For Lag 12, 83.53% of biases have an absolute magnitude less than 0.5 whereas 16.47% of biases have an absolute magnitude greater than or equal to 0.5. For Lag 18, 75.29% of biases have an absolute magnitude less than 0.5 whereas 24.71% of biases have an absolute magnitude greater than or equal to 0.5 (Figure A4). Figure A4 also shows a negative slope in the biases during the test period. Hence for all lags, the Niño 3.4 Index, showed a trend towards increased predictability over time, with the effect being most pronounced for Lag 6. This could be due to various factors such as changes in the underlying data patterns.

From Fig. 6, for the East Pacific ENSO patterns (nodes 5, 6, 8, 9, 10), during the analysis period, the nodes had

**Fig. 5** Evaluation metrics of the time series predictions from our AE-LSTM model







**Fig. 6** Annual cycle of the Niño 3.4 index and the ten encoded patterns used as predictors

their maximum amplitudes in May 2015, April 2016, April 2016, March 2016, and May 2015, respectively. These East Pacific ENSO patterns recorded their minimum amplitudes during October 1955 for nodes 5, 6, 8, and 9; and January 1974 for node 10. For the Central Pacific ENSO patterns (nodes 2, 4), the nodes recorded their minimum amplitude in January 1974 and their maximum amplitudes in November 2015. Node 3 recorded its maximum amplitude in November 2015 and its minimum amplitude in January 1974. For the remaining more irregular patterns (nodes 1 and 7), the maximum amplitudes were in April 2016; node 1 had its minimum amplitude in October 1955, whereas node 7 had its minimum amplitude in January 1974. Node 10 was dominant during the period when many nodes reached their minimum values (e.g., in 1955 and 1974). Conversely, node 4, which is the most related to the Niño 3.4 index (Figure A2 and Table A1), was dominant in the recent period (2015–2016) when several nodes peaked.

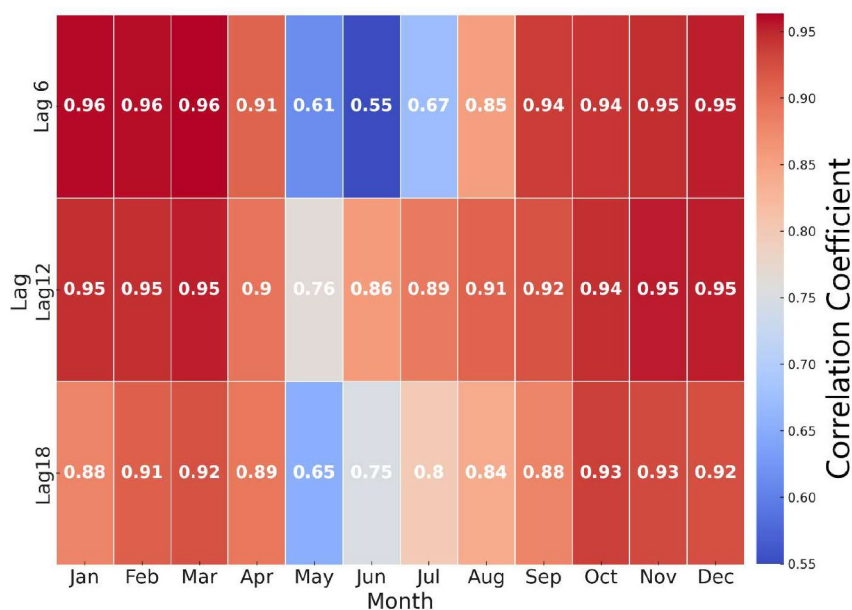
Different studies that applied other deep learning methods have reported success in forecasting ENSO in several lead times (Liu et al. 2022, 2023; Kim et al. 2022; Wang et al. 2023; Mu et al. 2021; Chen et al. 2023b; Chen et al. 2023; Jonnalagadda and Hashemi 2023). Ham et al. (2019) applied a convolutional LSTM to achieve skillful ENSO forecasts for lead times of up to one and a half years. In their study, the correlation skill of the Niño 3.4 index in the CNN model was above 0.5 for up to 17-months lead. The deep learning forecasts were also reported to be better than the dynamical forecast system. Wang and Huang (2023) also reached the same conclusion that CNN outperforms the dynamical forecast system. Thus, the promising predictions reached in this study (Figs. 3 and 5) using the AE and

LSTM deep learning model are supportive of previous convolutional LSTM results for ENSO prediction.

Different studies that applied other deep learning methods have reported success in forecasting ENSO in several lead times (Liu et al. 2022, 2023; Kim et al. 2022; Wang et al. 2023; Mu et al. 2021; Chen et al. 2023b; Chen et al. 2023; Jonnalagadda and Hashemi 2023). Ham et al. (2019) applied a convolutional LSTM to achieve skillful ENSO forecasts for lead times of up to one and a half years. In their study, the correlation skill of the Niño 3.4 index in the CNN model was above 0.5 for up to 17-months lead. The deep learning forecasts were also reported to be better than the dynamical forecast system. Wang and Huang (2023) also reached the same conclusion that CNN outperforms the dynamical forecast system. Thus, the promising predictions reached in this study (Figs. 3 and 5) using the AE and LSTM deep learning model are supportive of previous convolutional LSTM results for ENSO prediction.

Figure 7 shows the correlation for each month and, to address uncertainty in the correlations, Table A2 shows the bootstrapped observed mean correlation coefficients and the interquartile range (IQR) confidence intervals of the actual bootstrapped correlations. The most accurate result at lag 12, when considering all months in the evaluation period (Fig. 5), is shown to be associated with predictions from the months of May to August when they were relatively superior to lags 6 and 18 predictions (Fig. 7). In other months, lag 6 correlations were the same or larger than the correlations of lags 12 and 18 (Fig. 7). The same argument holds for lag 18 from May to July. Overall, the correlations are relatively smallest at lag 18, yet still remarkably large given the long lead time. The most accurate predictions were

**Fig. 7** Monthly correlations between actual and predicted Niño 3.4 index at 6-, 12-, and 18-month lead times. All the correlations are statistically significant at a 95% confidence level



achieved from December to March and the least accurate were achieved from May to July. Ham et al. (2019) also found better predictions in the early months of the year, consistent with those of Fig. 7. From Fig. 7, the decrease of the correlation coefficient from 0.91 in April (IQR confidence interval: 0.90–0.94) to 0.61 in May (IQR confidence interval: 0.54–0.71) at lag 6, is indicative of the SPB. Also, for lag 12 and lag 18, similar decreases in the correlation values from April are also evident but were less pronounced compared to those for lag 6. Nonetheless, a correlation of 0.61 in May at lag 6 and 0.76 (confidence interval: 0.66–0.86) at lag 12, is relatively large compared to traditional expectations for ENSO forecasting during the SPB (Barnston et al. 2012; Larson and Kirtman 2017; Mukhin et al. 2021). The modest spring barrier decrease for our approach suggests that the AE model improved predictions during those months. A similar predictability pattern, but with somewhat smaller correlations exceeding 0.5, during spring was also achieved by Ham et al. (2019). However, the result in Fig. 7 from April to May exceeds the correlations achieved in other studies and reinforces the capability of deep learning predictive models to improve ENSO predictions during spring (Gupta et al. 2020; Wang et al. 2022). Key processes contributing to the SPB include the seasonal shift in the position of the Intertropical Convergence Zone and the associated changes in oceanic and atmospheric circulation patterns (Duan and Wei 2013). These transitions create a period of uncertainty, significantly impacting the accuracy and reliability of ENSO predictions and making the SPB a focal point in the field of climate modeling. Our results show that the combination of AE and LSTM networks aids in identifying essential physical processes and their interactions when applied to SST in the tropical Pacific, and this advancement

could be critical for understanding and overcoming the SPB. The AE component is particularly effective in data compression and feature extraction, a crucial aspect in identifying the essential physical variables and interactions underlying the SPB. For example, Fig. 3 demonstrates the autoencoder model's proficiency in detecting diverse physical processes that influence ENSO dynamics.

Furthermore, given our model training period (January 1950 to June 2007) and the test period starting from November 2008, the SPB has different degrees of impact on forecasts with 6-month, 12-month, and 18-month lead times. Forecasts with a 6-month lead time initiated just before or during the early part of the SPB (for example, forecasts initiated in late fall or early winter) will have their target prediction period fall directly within the spring months when the SPB is most influential. In contrast, 18-month forecasts, despite also encountering the SPB, have the advantage of a longer temporal window to adjust to these changes, potentially leading to somewhat lessened SPB impacts (Fig. 7). A 12-month lead time forecast is less directly affected by the SPB since the prediction period spans a full annual cycle, including periods both before and after the SPB. This allows the model to integrate information from a complete seasonal cycle (Fig. 6), potentially offering a more comprehensive understanding of the ENSO dynamics. The ability of the model to integrate information from a complete seasonal cycle of ENSO might account for the reason why the forecast accuracy, measured by the anomaly correlations, during this spring months is relatively highest in a 12-month lead time and least accurate in a 6-month month-lead time (Figs. 6 and 7). This also has an impact on the overall accuracy of the all-season predictions during different lead times (Fig. 5).

Other evaluation metrics based on ENSO events, i.e., hit rate, false alarm ratio, and the CSI, in addition to the count of events, are shown in Figs. 8 and 9. From Fig. 8, forecasting very strong ENSO events is more accurate (i.e., large hit rate, small false alarm ratio, and large CSI) at lag 6 compared to other lags with an 85% hit rate. The forecast of very strong ENSO events was the least accurate at lag 18 with a 55% hit rate and a 70% hit rate at lag 12. Similarly, the false alarm ratio for very strong ENSO events was also the lowest at lag 6, which is the lag associated with the largest CSI. However, strong events and moderate events were forecast with improved accuracy at extended lead times of at least 1 year. Strong events generated a 74% hit rate at lag 12, compared to 65% at lag 6 and 71% at lag 18. Moderate events have a larger accuracy of 86% at lag 12, compared to 85% at lag 18 and 78% at lag 6. Also, the CSI for strong and moderate events was highest at lag 12. However, lag 6 has the smallest false alarm ratio for strong events, which was larger for lags 12 and 18. This suggests that as the lead time increases the model accuracy in forecasting very strong ENSO events decreases. Considering moderate events, lag 12 had the smallest false alarm ratio.

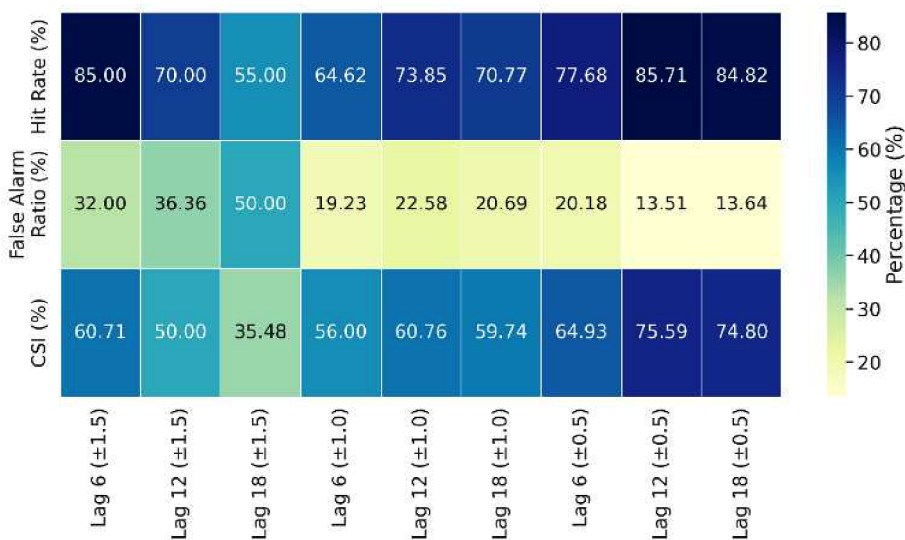
From Fig. 9, the capability of the AE—LSTM forecast model to accrue true positives more than false positives or false negatives can be seen. Again, for very strong ENSO events, the count of true positives is relatively higher at lag 6. Nonetheless, for strong and moderate events there are relatively more counts of true positives under lag 12. The evaluations in Figs. 8 and 9 indicate these forecasts are trustworthy, as the improved success achieved in forecasting ENSO at greater lead times is crucial for stakeholders’ applications of the forecasts (e.g., Pagano et al. 2001). Other studies applying ANN have recorded promising percentages of hit rates in forecasting the Niño 3.4 index. For example, in the 1984–2017 validation period, Ham et al. (2019)

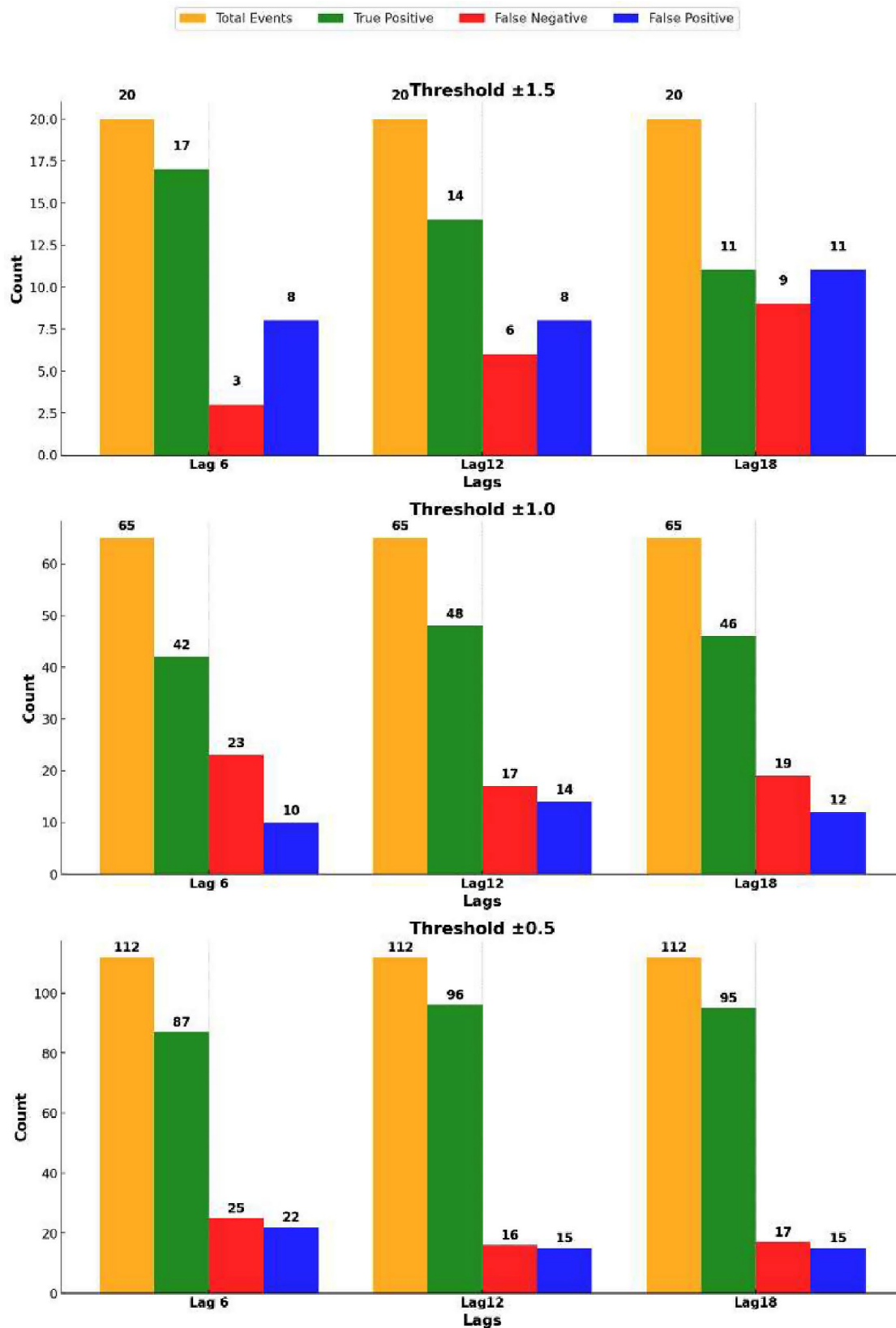
applied CNN and reported a hit rate of 66.7% at 12 months lead. Our model also reasonably captured the annual variability of ENSO decay rates (with a correlation of ~0.8) calculated as the difference between the DJF Niño 3.4 index in the preceding winter and the MAM Niño3.4 index (Li et al., 2023), possibly because the input predictors captured the annual cycle and monthly variability of ENSO (Figure A5).

A drawback of the current model is in forecasting event duration, particularly for strong ENSO events. Figures A6 to A8 indicate that a phase error discrepancy of at least 30 days can be expected in forecasting strong ENSO events with the current model. A typical example was that at lag 6, an actual event was from April 2015 to April 2016, but the predicted event was from May 2015 to June 2016. Therefore, if such a bias is found to be consistent, a bias adjustment of the forecast from our model by subtracting one month might improve the accuracy of the forecast event durations.

In summary, the novel deep learning model presented here adds to the increasing body of literature suggesting that ANN has the capacity to improve ENSO forecast at several lead times (Zhou et al. 2021; Chen et al. 2023b; Chen et al. 2023). Recently, Zhou and Zhang (2023) introduced a transformer-based model for ENSO prediction, signifying further advancements in the application of neural networks for ENSO forecasting. The major difference between the novel method introduced here and the similar, but more popular, convolutional LSTM that has been successfully applied by Guputa et al. (2020) is that, whereas AE’s strength lies in data compression and capturing nonlinearities, CNNs excel at spatial pattern recognition directly from raw data. In actuality, AE is explicitly designed for data compression and dimensionality reduction. Forecasting large-scale space-time problems, such as the one presented herein, involves a high dimensionality for the input data. The AE nodes can capture and represent essential patterns in data

**Fig. 8** Evaluation metric for predicting very strong (columns 1–3), strong (columns 4–6), and moderate (columns 7–9) ENSO events. An event can either be El Niño or La Niña. Hence the prediction is considered accurate only when it aligns with the sign and magnitude of the actual event. The values in the brackets are the thresholds for calculating the ENSO events





**Fig. 9** Further evaluation metrics of the Niño 3.4 predictions based on counts of events at different lags and thresholds. Total events are the sum of the true positives and false negatives



in a lower-dimensional space, making the model training more tractable and hence efficient. Conversely, CNNs focus primarily on spatial feature extraction but are not inherently designed for compression. Both AE and CNN can capture non-linear patterns, given their use of non-linear activation functions. However, the nature of the non-linearities they capture might differ, owing to their architecture. AEs, through their encoding and decoding process, capture non-linear relationships to reconstruct the data. CNNs capture spatial non-linearities through their convolutional layers. CNNs may have a potential advantage when it comes to spatial pattern recognition owing to their convolutional nature. They can automatically identify local spatial patterns across a grid of data. Similarly, AEs identify spatial patterns, but the patterns recognized are more directly related to data reconstruction and compression. Further, AEs allow for a space-time description of the data for comparing their non-linear patterns to traditional linear decompositions, such as PCA.

## 4 Conclusions

In this study, a new deep learning model combining AE neural networks and the LSTM deep learning model was introduced to forecast the Niño 3.4 index in 6- to 18-month lead times. AE was used to encode the non-linear SST patterns in the tropical Pacific that act as predictors of ENSO. The encoded patterns were used as predictors in the LSTM model to forecast ENSO. AE could detect SST patterns in the tropical Pacific that serve as effective predictors of the Niño 3.4 index using the LSTM deep learning predictive model. The predictors capture the annual cycle of the Niño 3.4 index. The Niño 3.4 index is used over all months during the test period and the model forecasts were most accurate at lag 12, compared to those at lags 6 and 18. The all-month correlation between the predicted and actual index is 0.94 at lag 12 and 0.91 at lags 6 and 18. Although the predictions were slightly more accurate from December to March., the correlations exceeded a value of +0.9 from January through April and from September through December for 6-month lead times, from January through April and from July through December for 12-month lead times, and from February to March and from October through December for 18-month leads. Typically, as ENSO events emerge in the boreal autumn months, these very large correlations suggest the AE with LSTM technique should have good practical performance for their prediction. Forecasting extreme ENSO events recorded the highest hit rate of 85% at lag 6, 70% at lag 12, and 55% at lag 18. However, moderate ENSO events were forecast with larger hit rates at lag 12. A drawback of the model, at this stage in its development, is in

capturing event durations. Bias correction may be useful in lowering the phase errors found. Moreover, considering the different number of cases of very strong, strong, and moderate El Niños and La Niñas, we have an imbalanced design. Thus, our AE-LSTM model will likely fit most accurately in the largest class. Nonetheless, there are there are methods to accommodate imbalanced data sets (e.g., Jafarigol and Trafalis 2023) which we will consider in future work, in addition to adjusting the biases in the predictions.

Furthermore, our results showed that, although the SPB was still present as evidenced by the drop in the fidelity of the predictions from May, the overall high correlations (always greater than 0.6 and as large as 0.76 for 12-month lead time forecasts) in the spring months suggest that the impact of the SPB was lowered, compared to previous studies. This improvement represents an advance in modeling and predicting ENSO. Finally, given that ENSO prediction is important for climate modelers, the techniques introduced here provided results that had improved accuracy to various measures of the ENSO phenomenon and hence may be useful for such scientists to help lessen the cold pool bias in the current group of GCMs and, further, helping to diagnose the double Inter-tropical Convergence Zone problem (Tian and Dong 2020).

Finally, this study shows that, in terms of temporal patterns, LSTM, when combined with AE can capture sequential or temporal patterns in time series data, making it suitable for time series forecasting problems, like predicting ENSO. When there is a need to capture sequential patterns in time series data in a computationally efficient manner, this research suggests AE and LSTM may well be a preferred machine learning approach.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00382-024-07180-8>.

**Author contributions** All authors worked equally on all aspects of this manuscript.

**Funding** Dr Ibebuchi is funded as postdoctoral researcher at Kent State University through NOAA Award Number NA22OAR4310142 (PI: Dr Cameron C Lee).

**Data availability** SST data is available at <https://psl.noaa.gov/data/gridded/tables/sst.html>. The Niño 3.4 index is available at <https://psl.noaa.gov/data/climateindices/list/>.

## Declarations

**Conflict of interest** There are no conflicts of interest in this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the

source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- An SI, Jin FF (2004) Nonlinearity and asymmetry of ENSO. *J Clim* 17:2399–2412
- Ashok K, Behera SK, Rao SA, Weng H, Yamagata T (2007) El Niño Modoki and its possible teleconnection. *Geophys Res (Oceans)*, 112, C11007
- Barnston AG, Tippett MK, L'Heureux ML, Li S, DeWitt DG (2012) Skill of real-time seasonal ENSO model predictions during 2002–11: is our capability increasing? *Bull Am Meteorol Soc* 93:631–651
- Cai W, McPhaden MJ, Grimm AM, Rodrigues RR, Taschetto AS, Garreaud RD et al (2020) Climate impacts of the El Niño–southern oscillation on South America. *Nat Reviews Earth Environ* 1:215–231
- Cai W, Santoso A, Collins M, Dewitte B, Karamperidou C, Kug JS et al (2021) Changing El Niño–Southern oscillation in a warming climate. *Nat Reviews Earth Environ* 2:628–644
- Chen Y, Huang X, Luo J, Lin Y, Wright J et al (2023a) Prediction of ENSO using multivariable deep learning. *Atmospheric Ocean Sci Lett*. 100350
- Chen Y, Huang X, Luo JJ, Lin Y, Wright JS, Lu Y et al (2023b) Prediction of ENSO using multivariable deep learning. *Atmospheric and Oceanic Science Letters*, p 100350
- Chen H, Jin Y, Shen X, Lin X, Hu R (2023) El Niño and La Niña asymmetry in short-term predictability on springtime initial condition. *Npj Clim Atmospheric Sci* 6(1):121
- Chen HC, Tseng YH, Hu ZZ, Ding R (2020) Enhancing the ENSO predictability beyond the spring barrier. *Sci Rep* 10:984
- Cho D, Yoo C, Son B, Im J, Yoon D, Cha D-H (2022) A novel ensemble learning for post-processing of NWP model's next-day maximum air temperature forecast in summer using deep learning and statistical approaches. *Weather Clim Extremes* 35:100410
- Choi D, Shallue CJ, Nado Z, Lee J, Maddison CJ, Dahl GE (2019) On empirical comparisons of optimizers for deep learning. *arXiv preprint arXiv:1910.05446*
- Dasgupta P, Roxy MK, Chattopadhyay R, Naidu CV, Metya A (2021) Interannual variability of the frequency of MJO phases and its association with two types of ENSO. *Sci Rep* 11:11541
- Dawson A, O'Hare G (2000) Ocean-atmosphere circulation and global climate: the El-Niño-Southern oscillation. *Geography: J Geographical Association* 85(3):193
- Duan W, Wei C (2013) The 'spring predictability barrier' for ENSO predictions and its possible mechanism: results from a fully coupled model. *Int J Climatol* 33:1280–1292
- Dufrénot G, Ginn W, Pourroy M (2023) ENSO Climate Patterns on Global Economic Conditions
- Forouzes M, Thiran P (2021) Disparity between batches as a signal for early stopping. In *Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part II* 21 (pp. 217–232). Springer International Publishing
- Geng T, Cai W, Wu L (2020) Two types of ENSO varying in tandem facilitated by nonlinear atmospheric convection. *Geophys Res Lett*, 47, e2020GL088784.
- Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. *AISTATS*
- Gupta M, Kodamana H, Sandeep SJ (2020) Prediction of ENSO beyond spring predictability barrier using deep convolutional LSTM networks. *IEEE Geosci Remote Sens Lett* 19:1–5
- Ham YG, Kim JH, Luo JJ (2019) Deep learning for multi-year ENSO forecasts. *Nature* 573:568–572
- Hidalgo HG, Alfaro EJ (2015) Skill of CMIP5 climate models in reproducing 20th century basic climate features in Central America. *Int J Climatology* 35:3397–3421
- Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. *Science* 313:504–550
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
- Hrudya PH, Varikoden H, Vishnu R (2021) A review on the Indian summer monsoon rainfall, variability and its association with ENSO and IOD. *Meteorol Atmos Phys* 133:1–14
- Hsiang SM, Meng KC, Cane MA (2011) Civil conflicts are associated with the global climate. *Nature* 476:438–441
- Hu ZZ, Huang B, Zhu J, Kumar A, McPhaden MJ (2019) On the Variety of Coastal El Niño Events. *Clim Dyn* 52:7537–7552
- Hu K, Huang G, Huang P, Kosaka Y, Xie SP (2021) Intensification of El Niño-induced atmospheric anomalies under greenhouse warming. *Nat Geosci* 14:377–382
- Huang B, Thorne PW, Banzon VF, Boyer T et al (2017) Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *J Clim* 30:8179–8205
- Ibebuchi CC (2024) Redefining the North Atlantic Oscillation index generation using autoencoder neural network. *Science and Technology, Machine Learning*
- Ibebuchi CC, Richman MB (2024) Non-linear modes of global sea surface temperature variability and their relationships with global precipitation and temperature. *Environmental Research Letters*
- Jafarigol E, Trafalis T (2023) A review of machine learning techniques in Imbalanced Data and Future trends. *arXiv Preprint. arXiv:2310.07917*
- Jin EK, Kinter JL, Wang B, Park C-K et al (2008) Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Clim Dynam* 31:647–664
- Jonnalagadda J, Hashemi M (2023) Long lead ENSO Forecast using an adaptive graph Convolutional recurrent neural network. *Eng Proc* 39(1):5
- Kao H, Yu J (2009) Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *J Clim* 22:615–632
- Kim J, Kwon M, Kim SD, Kug JS, Ryu JG, Kim J (2022) Spatio-temporal neural network with attention mechanism for El Niño forecasts. *Sci Rep* 12(1):7204
- Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. *arXiv Preprint* 415 arXiv:1412.6980.
- L'Heureux ML, Levine AF, Newman M, Ganter C, Luo J, Tippett MK, Stockdale TN (2020) ENSO prediction. *El Niño South Oscillation Chang Clim*, 227–246
- Larson S, Lee SK, Wang C, Chung, ES, Enfield D (2012) Impacts of non-canonical El Niño patterns on Atlantic hurricane activity. *Geophys Res Lett* 39(14).
- Larson SM, Kirtman BP (2017) Drivers of coupled model ENSO error dynamics and the spring predictability barrier. *Clim Dyn* 48:3631–3644
- Lee T, McPhaden MJ (2010) Increasing intensity of El Niño in the central-equatorial Pacific. *Geophys Res Lett* 37:14
- Levine AF, McPhaden MJ (2015) The annual cycle in ENSO growth rate as a cause of the spring predictability barrier. *Geophys Res Lett* 42:5034–5041

- Li G, Chen L, Lu B (2023) A physics-based empirical model for the seasonal prediction of the central China July precipitation. *Geophys Res Lett* 50(3):e2022GL101463.
- Liu X, Li N, Guo J, Fan Z, Lu X, Liu W, Liu B (2022) Multistep-ahead prediction of Ocean SSTA based on hybrid empirical Mode Decomposition and Gated Recurrent Unit Model. *IEEE J Sel Top Appl Earth Observations Remote Sens* 15:7525–7538
- Liu Y, Duffy K, Dy JG, Ganguly AR (2023) Explainable deep learning for insights in El Niño and river flows. *Nat Commun* 14(1):339
- Lopez H, Kirtman BP (2014) WWBs, ENSO predictability, the spring barrier and extreme events. *J Geophys Research: Atmos* 119:10–114
- Mohan VS, Vinayakumar R, Soman KP, Poornachandran P (2018) S.P.O.O.F Net: Syntactic Patterns for identification of Ominous Online Factors, 2018 *IEEE Security and Privacy Workshops (SPW)*, San Francisco, CA, USA, pp. 258–263
- Mu B, Ma S, Yuan S, Xu H (2020), July Applying convolutional LSTM network to predict El Niño events: Transfer learning from the data of dynamical model and observation. In 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC) (pp. 215–219). IEEE
- Mu B, Qin B, Yuan S (2021) ENSO-ASC 1.0. 0: ENSO deep learning forecast model with a multivariate air–sea coupler. *Geosci Model Dev* 14:6977–6999
- Mu B, Qin B, Yuan S (2022) ENSO-GTC: ENSO Deep Learning Forecast Model with a global spatial-temporal Teleconnection Coupler. *J Adv Model Earth Syst*, 14, e2022MS003132.
- Mukhin D, Gavrilov A, Seleznev A, Buyanova M (2021) An atmospheric signal lowering the spring predictability barrier in statistical ENSO forecasts. *Geophys Res Lett*, 48(6), e2020GL0912
- Odériz I, Silva R, Mortlock TR, Mori N (2020) El Niño–Southern oscillation impacts on global wave climate and potential coastal hazards. *J Geophys Research: Oceans*, 125, e2020JC016464.
- Ortega G, Arias PA, Villegas JC, Marquet PA, Nobre P (2021) Present-day and future climate over central and South America according to CMIP5/CMIP6 models. *Int J Climatol* 41:6713–6735
- Pagano TC, Hartmann HC, Sorooshian S (2001) Using climate forecasts for water management: Arizona and the 1997–1998 El Niño. *JAWRA J Am Water Resour Association* 37:1139–1153
- Pal KK, Sudeep KS (2016) Preprocessing for image classification by convolutional neural networks, 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 1778–1781
- Patil KR, Jayanthi VR, Behera S (2023) Deep learning for skillful long-lead ENSO forecasts. *Front Clim* 4:1058677
- Patro SG, Sahu KK (2015) Normalization: a Preprocessing Stage. *Int Adv Res J Sci Eng Technol*. 2<https://doi.org/10.17148/IARJSET.2015.2305>
- Phan QT, Wu YK, Phan QD (2021) A hybrid wind power forecasting model with XGBoost, data preprocessing considering different NWP. *Appl Sci* 11:1100
- Qiao S, Zhang C, Zhang X et al (2023) Tendency-and-attention-informed deep learning for ENSO forecasts. *Clim Dyn*. <https://doi.org/10.1007/s00382-023-06854-z>
- Reddy PJ, Perkins-Kirkpatrick SE, Ridder NN, Sharples JJ (2022) Combined role of ENSO and IOD on compound drought and heatwaves in Australia using two CMIP6 large ensembles. *Weather Clim Extremes* 37:100469
- Saha M, Nanjundiah RS (2020) Prediction of the ENSO and EQUINO indices during June–September using a deep learning method. *Meteorol Appl* 27:e1826
- Schaefer JT (1990) The critical success index as an indicator of warning skill. *Wea Forecast* 5:570–575
- Tian B, Dong X (2020) The double-ITCZ bias in CMIP3, CMIP5, and CMIP6 models based on annual mean precipitation. *Geophys Res Lett*, 47, e2020GL087232.
- Timmermann A, An SI, Kug JS, Jin FF, Cai W, Capotondi A et al (2018) El Niño–southern oscillation complexity. *Nature* 559:535–545
- Trenberth KE, Hoar TJ (1996) The 1990–1995 El Niño–Southern Oscillation event: Longest on record. *Geophys Res Lett* 23:57–60
- Van Houdt G, Mosquera C, Nápoles G (2020) A review on the long short-term memory model. *Artif Intell Rev* 53:5929–5955
- Wang T, Huang P (2023) Superiority of a convolutional neural network Model over Dynamical models in Predicting Central Pacific ENSO. *Advances in Atmospheric Sciences*
- Wang GG, Cheng H, Zhang Y, Yu H (2022) ENSO analysis and prediction using deep learning: a review. *Neurocomputing*
- Wang H, Hu S, Li X (2023) An interpretable deep learning ENSO forecasting model. *Ocean-Land-Atmosphere Res* 2:0012
- Wu X, Okumura YM, DiNezio PN (2021) Predictability of El Niño duration based on the onset timing. *J Clim* 34:1351–1366
- Yeh SW, Kug JS, Dewitte B, Kwon MH, Kirtman BP, Jin FF (2009) El Niño in a changing climate. *Nature* 461:511–514
- Ying J, Huang P, Lian T, Tan H (2019) Understanding the effect of an excessive cold tongue bias on projecting the tropical Pacific SST warming pattern in CMIP5 models. *Clim Dyn* 52:1805–1818
- Zhang RH, Gao C, Feng L (2022) Recent ENSO evolution and its real-time prediction challenges. *Natl Sci Rev* 9(4):nwac052
- Zhao Y, Sun D (2022) ENSO asymmetry in CMIP6 models. *J Clim* 35:5555–5572
- Zhou L, Zhang R-H (2023) A self-attention-based neural network for three-dimensional multivariate modeling and its skillful ENSO predictions. *Sci Adv* 9(10):adf2827
- Zhou P, Huang Y, Bingyi HU, Wei J (2021) Spring Predictability Barrier Phenomenon in ENSO Prediction Model based on LSTM Deep Learning Algorithm. *Acta Scientiarum Naturalium Universitatis Pekinensis* 57(6):1071–1078

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.