Deep Residual Convolutional Neural Network Combining Dropout and Transfer Learning for ENSO Forecasting

Jie Hu, Bin Weng, Tianqiang Huang, Jianyun Gao, Feng Ye, and Lijun You

1 College of Computer and Cyber Security, Fujian Normal University, Fuzhou, China, 2 Digital Fujian Institute of Big Data Security Technology, Fuzhou, China, 3 Engineering Technology Research Center for Public Service Big Data Mining and Application of Fujian Province, Fuzhou, China, 4 Fujian Key Laboratory of Severe Weather, Fujian Institute of Meteorological Sciences, Fuzhou, China, 5 Fujian Meteorological Information Center, Fuzhou, China

Abstract

To improve El Niño-Southern Oscillation (ENSO) amplitude and type forecast, we propose a model based on a deep residual convolutional neural network with few parameters. We leverage dropout and transfer learning to overcome the challenge of insufficient data in model training process. By applying the dropout technique, the model effectively predicts the Niño3.4 Index at a lead time of 20 months during the 1984–2017 evaluation period, which is three more months than that by the existing optimal model. Moreover, with homogeneous transfer learning this model precisely predicts the Oceanic Niño Index up to 18 months in advance. Using heterogeneous transfer learning this model achieved 83.3% accuracy for forecasting the 12-month-lead El Niño type. These results suggest that our proposed model can enhance the ENSO prediction performance.

Plain Language Summary

El Niño-Southern Oscillation (ENSO) is an irregular periodic variation along with complex tropical atmosphere-ocean interaction. It impacts interannually human lives globally and locally. Hence, we contribute, the first time as we know, a deep learning model that can effectively predict El Niño strength and type. The model can transfer the knowledge learned from Niño3.4 Index prediction to Oceanic Niño Index and type prediction, respectively. We find that our proposed model has a high correlation skill and a good precision for predicting strength and type respectively in relation to an evaluation between 1984 and 2017. Moreover, our model requires smaller-sized storage against the existing deep learning model.

1. Introduction

The El Niño-Southern Oscillation (ENSO) is one of the main drivers of interannual climate variability on Earth, impacting global climate (Yang et al., 2018), agriculture (Henson et al., 2017), ecosystems (Lehodey et al., 2020), health (Heaney et al., 2019), and society (Hsiang et al., 2011). Therefore, it is valuable to predict ENSO early and accurately to minimize these effects. However, predicting the strength of ENSO remains a challenge due to its complexity (Sun et al., 2016; Timmermann et al., 2018). Also, the increasing diversity of ENSO behavior since 2000 has led to a growing interest in the type of ENSO events (Geng et al., 2020). ENSO can be mainly divided into Eastern Pacific (EP) and Central Pacific (CP) types (Yeh et al., 2009), based on the distribution of the Sea Surface Temperature Anomaly (SSTA) during its maturation phase. However, some events that the SSTA is relatively high over the central and eastern Pacific Ocean cannot be classified as CP or EP types. Zhang et al. (2019) classified ENSO into EP, CP, and a mixture of the two (MIX) types of El Niño (La Niña). To the best of our knowledge, the definition of ENSO type has not come to an agreement. Because the effects of different ENSO types vary greatly, for example, different El Niño events have a different impact on US winter temperatures (Yu et al., 2012) and the East Asian climate (Yuan & Yang, 2012). Hence, the prediction of ENSO type is important for improving the quality of climate forecasts.

Currently, both statistical (Petrova et al., 2020; Ren, Zuo, & Deng, 2018; Wang et al., 2020) and dynamical (Ren, Scaife, et al., 2018; Saha et al., 2014) methods can generate skillful predictions 6–12 months in advance. Many deep learning-based methods have emerged in recent years, for example, using Artificial Neural Networks (ANNs) (Feng et al., 2016; Petersik & Dijkstra, 2020), Recurrent Neural Networks (RNNs) (Mahesh et al., 2019), Long Short-Term Memory (LSTM) neural networks (Broni-Bedaiko et al., 2019), Convolutional Long Short-Term Memory (ConvLSTM) (Gupta et al., 2020; He et al., 2019; Mu et al., 2019), Convolutional Neural Networks (CNNs) (Ham et al., 2019; Yan et al., 2020), and Graph Neural Networks (GNNs) (Cachay et al., 2020).
The most remarkable work is the CNN-based model that can make effective forecasts 17 months in advance (Ham et al., 2019), outperforming most existing methods. This model is trained on Coupled Model Intercomparison Project phase5 (CMIP5) and reanalysis data to predict the Niño3.4 index. However, the model has few layers, only convolutional and pooling layers, does not use residual structures, and does not use some techniques to improve the predictability except for the use of transfer learning on CMIP5. Recent studies have shown that the dropout technology can improve the performance of shallow neural networks applied to temperature simulation problems (Piotrowski et al., 2020). Through extensive experiments they show that improving model performance and stability requires nodes to be discarded with much lower probability than common deep neural networks (about 1%, instead of 10–50% for deep learning). Due to a number of layers applied in our model, we consider using the dropout in more detail to further improve the prediction ability. Additionally, comparing to the existing research on ENSO prediction which only performs transfer learning on simulated data, we also consider transferring the knowledge learned from the task of predicting the Niño3.4 index to the tasks of predicting Oceanic Niño Index (ONI), so-called homogeneous transfer learning.

There are various methods of predicting ENSO types, for example, based on the random forest (Santos et al., 2020), multimodel ensemble (Ren, Scaife, et al., 2018), and CNN (Ham et al., 2019). In this work, we focus on the CNN method trained on CMIP5 data to predict El Niño types. The accuracy remains 66.7% at lead times of 12 months. However, they have only expected the types of El Niño, not yet the types of La Niña and the normal events. Besides, using transfer learning in the index prediction leads to a slight performance improvement, while in the type prediction, no transfer learning is used. Nevertheless, we can transfer the knowledge learned from the task of index prediction to type prediction. This method is called heterogeneous transfer learning, thereby further improving the prediction ability.

In this work, the main contributions are summarized as follows:

1. We propose a deep Residual Convolutional Neural Network (Res-CNN) model for ENSO predictions, including the Niño3.4 index, ONI, and types. It is worth noting that our model requires only a few changes for different tasks. We find that the Res-CNN model can effectively predict the Niño3.4 index for up to 20 months in advance, 3 months more than the previous CNN-based model.
2. Keeping the network structure intact, we show the ONI can be skillfully predicted 18 (12) months in advance with (without) homogeneous transfer learning, which provides us a new strategy for further enhancing the predictive ability of ENSO.
3. We apply heterogeneous transfer learning to enhance the type prediction. We show that the knowledge learned from the index prediction task can be transferred to the type prediction task by changing only the output layer of the model trained for the index prediction task and retraining on the type prediction task. The accuracy of El Niño type prediction can reach 83.3% 12 months in advance, while the current best is 66.7%.

2. Data and Methods

2.1. Data

The predictors are three consecutive months SSTA and Heat Content Anomaly (vertically averaged oceanic temperature anomaly in the upper 300 m) over 0°–360°E, 55°S–60°N at a resolution of 5° × 5°. The simulated data set is CMIP5 3.2 (core) (https://esgf-node.llnl.gov/projects/cmip5/) (Bellenger et al., 2014). The reanalysis data set is the simple ocean data assimilation version 2.2.4 (SODA) (https://climatedataguide.ucar.edu/climate-data/soda-simple-ocean-data-assimilation) (Giese & Ray, 2011). Global Ocean Data Assimilation System (GODAS) (Behringer & Xue, 2004) can be found in https://www.esrl.noaa.gov/psd/data/gridded/data.godas.html.

2.2. Res-CNN Model

The input data for three consecutive months are recorded as $x_{t-2}$, $x_{t-1}$, $x_t$, the output data of the Niño3.4 index, ONI or type all referred to as $y$, and the forecast result can be described by

$$y_{t+l} = F(x_{t-2}, x_{t-1}, x_t)$$

(1)

where $F$ denotes the Res-CNN model, and $l$ is the forecast lead months from 1 to 23.
Res-CNN shown in Figure 1 uses a seven-layer convolutional neural network, a three-layer max-pooling to extract features, two-layer skip connections, and one-layer fully connected layer to generate the final result. In index predicting task, the output is a single value; while in type predicting task, the output is the probability of various categories. The convolution process of Res-CNN is the most efficient computational tool for extracting features as follows:

\[ v_{l,x,x} = \sum_{m=1}^{M_{l}} \sum_{p=1}^{P_{l}} \sum_{q=1}^{Q_{l}} w_{p,q} u_{l-1,m}^{(x+p-1/2,y+q-1/2)} (l-1) x_m + b_{l} \]  

(2)

where \((x, y)\) is the dimensions of the feature map, \(l\) denotes the \(l\)th convolution layer, and \(f\) is for the \(f\)th feature map. \(M\) means the number of feature maps, and \((P, Q)\) is the dimensions of the \(l\)th filter. \(b\) is the bias units, \(w\) is the weight at grid point \((p, q)\) in the convolution kernel and \(v_{l,f}\) denotes one value of the \(l\)th filter and the \(f\)th feature map.

The parameters of our model are learned through multiple iterations of the minimization loss function of Mean Square Error in predicting index or Cross Entropy in predicting type. In our model, the residual structure can be defined as follows:

\[ y = R(x, [W]) + x \]  

(3)

where \(x\) and \(y\) are the input and output vectors of the considered layer, and the function \(R\) denotes the residual mapping to be learned. The operation \(R + x\) is performed by a shortcut connection and element-wise addition. The other details are same as those in K. He et al. (2016), except that we use the Tanh activation function instead of the rectified linear unit and do not use the batch normalization (Ioffe & Szegedy, 2015). Because our network is shallow compared to a standard residual network, small changes in network parameters have little effect when the network is not deep. Also, because our data is insufficient and complex, if the input of each layer of the network is kept the same distribution, the model cannot be trained well. Setting the number of residual connections to 0, 1, 2, and 3, our model has various structures. In order to further improve performance, 11 different dropout rates are token, namely 0, 0.01, 0.03, 0.05, 0.07, 0.1, 0.3, 0.5, 0.7, 0.9, 0.99. Thus, for each advance month, there are 44
models (Figure S1 in Supporting Information S1). The final model would be the best result from the model that determines the number of residual connections. See Text S1 in Supporting Information S1 for details on dropout and transfer learning techniques.

2.3. Indexes Forecast

In predicting index includes the Niño3.4 index and ONI. The number of the unit in fully connected layer is one, and the Adam (Kingma & Ba, 2014) optimization algorithm is used. The specific parameter settings can be found in the Text S2 in Supporting Information S1. We use the correlation coefficient function $I$ as a measure of the ENSO index prediction:

$$I_l = \frac{1}{12} \sum_{t=1}^{12} \frac{\sum_{j=1}^{e} (O_{y,t} - \bar{O}_t)(F_{y,t,l} - \bar{F}_t)}{\sqrt{\sum_{j=1}^{e} (O_{y,t} - \bar{O}_t)^2 \sum_{j=1}^{e} (F_{y,t,l} - \bar{F}_t)^2}}$$

(4)

where $O$ and $F$ denote the observed and the predicted values, respectively. $\bar{O}_t$ and $\bar{F}_t$ denote the temporal climatology concerning the calendar month $m$ (from 1 to 12) and the forecast lead months $l$ (from 1 to 23). The label $y$ means the forecast target year. Finally, $s$ and $e$ denote the earliest and latest validation or test year.

2.4. Types Forecast

We conduct two kinds of experiments: one is to predict three types, that is, EP, CP, and MIX of El Niño; and the other is with seven types, that is, EP, CP, MIX of El Niño (La Niña), and Normal Year (NY). See the Text S3 in Supporting Information S1 for more details.

3. Results

The All-season Correlation Skill (ACS) is shown in Figure 2 for the CNN (b) and the Res-CNN (c). The ACS of the 3-month-moving-averaged Niño3.4 index between 1984 and 2017 in the Res-CNN model is higher than almost all state-of-the-art dynamic models and the CNN model (Figure 2a). It is worth noting that, except for the Res-CNN model, the CNN model fails to perform optimally when the lead time is less than 6 months. The correlation coefficient of the CNN model exceeds 0.5 only 17 months in advance, and worse than the Scale Interaction Experiment-Frontier (SINTEX-F) dynamic prediction model for a lead of 23 months, while the Res-CNN model reaches 20 months in advance and outperforms the SINTEX-F dynamic prediction model in all advance months. Thus, we conclude that the Res-CNN model can skillfully predict ENSO 20 months in advance, which is better than all the compared models. The Res-CNN model exhibits a higher correlation coefficient than the CNN model in almost all target seasons, especially in spring and autumn seasons. For example, when the target season is MJJ (May-June-July), the SINTEX-F model predicts a correlation coefficient above 0.5 for only four months (Table S3 in Supporting Information S1), the CNN model for 11 months (Figure 2b), and the Res-CNN model for 17 months (Figure 2c), suggesting that our model is less affected by the spring prediction barrier (SPB) than the CNN and SINTEX-F model. Typically, the SPB phenomenon is more severe in statistical models than in dynamic models (Jan van Oldenborgh et al., 2005). The Res-CNN model is less affected than other statistical methods because it is likely to make fuller use of the heat content information than other statistical methods, and accurate initialization of heat content can improve spring forecasting (McPhaden, 2003). Nevertheless, the skills are much lower for summer time than winter, which may be related to the predictability. And the “spring barrier” (western Pacific Ocean) may be a main factor to impact the summer predictability.

The ACS of the 3-month-moving-averaged Niño3.4 index from 1982 to 2001 and from 2002 to 2017 is shown in Figure S2 in Supporting Information S1. Whether it is 82-01 or 02-17, the prediction skill in the Res-CNN model is higher than the CNN model in all months ahead. However, compared with the skill between 1982 and 2001, 2002–2017 declines sharply, with its effective prediction only 12 months. Similarly, the CNN model drops to only 10 months, which is inseparable because the behavior of ENSO becomes more diverse after 2000 (Barnston et al., 2012). To assess the effect of the actual prediction more clearly in the Res-CNN model, we generated curves of the Niño3.4 index predicting the DJF season 18 months in advance for the years 1982–2017 (Figure S3 in Supporting Information S1). Compared to the SINTEX-F model, the Res-CNN correlation coefficient is over
0.2 higher. Besides, the Res-CNN also has correlation coefficient of 0.2 higher than the CNN model 20 months in advance (Figure S4 in Supporting Information S1), better predicting years with higher Niño3.4 index, such as 1982/1983, 1997/1998.

The ACS of the ONI is shown in Figure 3 for the Gaussian Density Neural Network (GDNN) (a), the Quantile Regression Neural Network (QRNN) (b) (Petersik & Dijkstra, 2020), and the Res-CNN (c). Instead of our method, which is to predict a single value, the GDNN and the QRNN are used to estimate the prediction uncertainty of ENSO forecast. Compared to the GDNN and QRNN methods, the ACS of the ONI between 1984 and 2017 in the Res-CNN model using homogeneous transfer learning is highest (Figure 3d). Notably, the correlation coefficients of GDNN and QRNN in predicting ONI from 2002 to 2011 drop below 0.5 at 7 months ahead, while our method still has 0.6 at 10 months ahead, which is almost consistent with Niño 3.4 in predicting 2002 to 2017. Besides, comparing the results of predicting Niño3.4 index and ONI from 1982 to 2017, ONI gives better results until the advance month is 12, while Niño3.4 index gives better results after that, suggesting that for index prediction with greater than 1 year, the amount of data has an impact on the model.

Figure 2. Correlation skill for various lead times and methods. All-season correlation skill of the 3-month-moving-averaged Niño3.4 index for multiple lead times from 1984 to 2017 in the Residual Convolutional Neural Network (Res-CNN) model (red), Convolutional Neural Network (CNN) model (Ham et al., 2019) (dodger blue), Scale Interaction Experiment-Frontier (SINTEX-F, Luo et al., 2008) dynamical forecast system (sky blue), other dynamical forecast systems (Kirtman et al., 2014) included in the North American MultiModel Ensemble (NMME) project (the other colors), and the multimodel mean (Mean). The correlation skill of the Niño3.4 index for each season in the CNN model (b) and the Res-CNN model (c). The black dashed line indicates that the correlation coefficient is equal to 0.5.
The results of 3, 6, 9, 12, 18, 23 months ahead forecasts of El Niño types from 1982 to 2017 are shown in Table 1. We found that compared to using one-step, using two-step achieves better results in all five scenarios of A–E; comparing A-two with B-two and C-two with E-two, the results of A are not as good as those of B and C's are not as good as E's. This indicates that the pretraining and soda training methods are not as good as just using the pretraining method in type prediction, in the meantime, the distribution of SODA and GODAS data is very inconsistent, possibly due to the significant difference in the frequency of occurrence of various types in the SODA data set (Yeh et al., 2009) and the diversity after 2000 (Barnston et al., 2012). Compared with A-two and C-two, our model can still achieve 67% accuracy under 12 months ahead using heterogeneous transfer learning. Also, it can predict all the type of super ENSO 12 months in advance, especially the 2015/2016 El Niño, the strongest events in history, which can still be predicted at a lead time of 18 months. At present, almost all models cannot predict the event one year in advance (Tang et al., 2018). By comparing the results of A-two and D-two, the accuracy of A-two is lower than that of D-two, indicating that transfer training on SODA instead reduces the accuracy of most of the models initially trained on CMIP5. This suggests that fine-tuning on SODA does not yield better
### Table 1: Prediction of ENSO Types

| Methods  | Lead months |
|----------|-------------|
| Methods  | 3 6 9 12 18 |
| A-one    | 27 22 21 19 14 |
| A-two    | 32 26 22 22 17 |
| B-one    | 25 20 17 15 13 |
| B-two    | 30 24 18 18 16 |
| C-one    | 29 22 19 19 15 |
| C-two    | 32 27 24 24 20 |
| D-one    | 28 23 18 18 16 |
| D-two    | 29 26 24 21 17 |
| E-one    | 26 24 16 14 13 |
| E-two    | 28 26 19 17 17 |
| Super ENSO | 3/3 2/3 2/3 2/3 2/2 |
| Accuracy (%) | 89/89 72/75 61/67 61/67 47/56 |

*Note. Results of forecasting the types of ENSO 3, 6, 9, 12, 18 months in advance from 1982 to 2017. There are 36 events in total, A-E in the table represents the number of correct predictions. H and N denote the use and nonuse of heterogeneous transfer learning, respectively. One means one-step seven classes prediction, two means two-step seven classes prediction, that first predicts El Niño, La Niña, and normal year events, and then predicts whether El Niño or La Niña will be EP, CP, or MIX. Super ENSO means that A-two/C-two correctly predicted the number of 1982/1983, 1997/1998, 2015/2016 El Niño. Accuracy indicates that the accuracy of A-two/C-two. A: Train in CMIP5. B: Train in CMIP5 and then train in SODA. C: Heterogeneous transfer learning the index model to CMIP5. D: Homogeneous transfer learning the C to SODA. E: Heterogeneous transfer learning the index model to CMIP5 and then training in SODA. The model of using heterogeneous transfer is the optimal model for predicting the respective lead and target of the Niño3.4 index. The values in bold mean the best results among these options for different advance months.*

Results, probably because heterogeneous transfer learning has been able to resolve, to some extent, the problem of unbalanced data distribution between CMIP5 and SODA.

Finally, to evaluate the performance of our model, we compare it with the CNN model. Figure S5 in Supporting information S1 shows our model achieves 83.3% accuracy 12 months earlier on the period from 1984 to 2017 compared to the CNN model (66.7% accuracy). These results indicate that the Res-CNN model predicts the ENSO index and type better than the CNN model.

### 4. Discussions

Through various ablation experiments, we found that we got better and more stable results at a lower dropout rate (0–0.3) than those at a higher dropout rate (0.5–0.9) (Figure S6 in Supporting information S1). The finding differs from the conventional deep learning approach usually set with 0.4–0.6 of the dropout rate. The achievement is because high dropout rate is not suitable for this task. Therefore, too large dropout rate will lead to serious loss of regional features, and only some of the regional features will reduce the accuracy of ENSO forecast. To find the appropriate number of residual connections, we conducted ablation experiments. The obtained results (Figure S7 in Supporting information S1) show that the effective prediction months were about 17, 18, 20, and 16 when the number of residual connections was 0, 1, 2, and 3, respectively. It was selected as our optimal model since the model with residual connections of two predicted best. Furthermore, Figure S8 in Supporting information S1 shows that using unnormalized two residual connections achieved better prediction results in comparison to using normalization, indicating that data normalization does not improve the model performance in deep learning-based ENSO prediction. Additionally, the model with a residual number of 3 can only predict the effective forecast for 17 months, indicating no further improvement of a higher residual connection number.

### 5. Conclusions

Although this study showed remarkable results, there are still some limitations. In predicting the Niño3.4 index, the predictive ability of Res-CNN is notably improved in all seasons. However, by comparing the correlation coefficient for lead months from 1 to 23 months, we found that they were nearly the lowest from late spring to fall (Table S1 in Supporting Information S1), the same as CNN (Table S2 in Supporting Information S1) and SINTEX-F (Table S3 in Supporting Information S1). This suggested that the SPB is still prevalent (Levine & McPhaden, 2015) and requires further study. Moreover, there is a large negative anomaly of the predicted SST for the first 10 years for both CNN and our model, whether this implies a change in climate or for other reasons we also need to investigate further. Holding the model structure constant to predict ONI, surprisingly, Res-CNN can effectively predict for 12 months (Figure S9 in Supporting information S1) despite using only a small amount of data. However, the correlation coefficients were unstable at times high and low under different months of advance, and did not show a stable downward trend. To alleviate the problem, we predicted ONI using homogeneous transfer learning, and the skill was significantly enhanced. Since our model initially predicted the Niño3.4 index well, we assume that our model could learn to predict it. The ONI definition is closer to the Niño3.4 index, so the model is able to learn lots of knowledge and only needs less training to predict the ONI well. By varying only the number of the unit in the output layer of our model to predict the El Niño type, the result in Table 1 is still almost 20% points higher than the CNN. Moreover, two-step and heterogeneous transfer learning were used in this work to predict ENSO types, with some predictive performance improvement.
In summary, this study showed that the Res-CNN-based model can improve the long-term prediction of ENSO. Also, we found that the predictive ability can be better improved by using transfer learning and dropout techniques. The future extensions would be using different numbers of predictors and input months under different prediction months, for example, intuitively, trying fewer predictors and input months under shorter advance months.

Data Availability Statement

The research data can be found in the website (https://doi.org/10.5281/zenodo.4646653).

Acknowledgments

This study was supported by National Key R&D Program Special Fund Grant (No. 2018YFC1505805), National Natural Science Foundation of China (Nos. 62072106, 61070062), General Project of Natural Science Foundation in Fujian Province (No. 2020J1168), Open project of Fujian Key Laboratory of Severe Weather (No. 2020KFKT04).

References

Barnston, A. G., Tippett, M. K., L’Heureux, M. L., Li, S., & DeWitt, D. G. (2012). Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability improving? Bulletin of the American Meteorological Society, 93(5), 631–651. https://doi.org/10.1175/bams-d-11-00111.1
Behringer, D., & Xue, Y. (2004). Evaluation of the global ocean data assimilation system at ncep: The Pacific Ocean. In Proceedings of the eighth symposium on integrated observing and assimilation systems for atmosphere, oceans, and land surface. American Meteorological Society.
Bellenger, H., Guiyouri, É., Leloup, J., Lengaigne, M., & Vialard, J. (2014). ENSO representation in climate models: From cmip3 to cmip5. Climate Dynamics, 42(7), 1999–2018. https://doi.org/10.1007/s00382-013-1783-z
Brom-Bedak, C., Katsumi, F. A., Unemi, T., Asumi, M., Abdulai, J.-D., Shinomiya, N., & Owusu, E. (2019). El Niño-southern oscillation forecasting using complex networks analysis of LSTM neural networks. Artificial Life and Robotics, 24(4), 445–451. https://doi.org/10.1007/s10015-019-00540-2
Cachy, S. R., Erickson, E., Bucker, A. F. C., Pokropek, E., Potosnak, W., Osei, S., & Lütjens, B. (2020). Graph neural networks for improved El Niño forecasting. arXiv preprint. arXiv:2012.01598.
Feng, Q. Y., Vasile, R., Segond, M., Gorgolchi, A., Wang, Y., Havlin, S., et al. (2016). Climatelearn: A machine-learning approach for climate prediction using network measures. Geoscientific Model Development Discussions, 18. https://doi.org/10.5194/gmd-2015-273
Geng, T., Cai, W., & Wu, L. (2020). Two types of ENSO varying in tandem facilitated by nonlinear atmospheric convection. Geophysical Research Letters, 47, e2020GL088784. https://doi.org/10.1029/2020GL088784
Giese, B. S., & Ray, S. (2011). El Niño variability in simple ocean data assimilation (soda), 1871–2008. Journal of Geophysical Research, 116, C02024. https://doi.org/10.1029/2010JC006695
Gupta, M., Kodalama, H., & Sandeep, S. (2020). Prediction of ENSO beyond spring predictability barrier using deep convolutional LSTM networks. IEEE Geoscience and Remote Sensing Letters, 1–5. https://doi.org/10.1109/lgrs.2020.3032353
Ham, Y. G., Kim, J. H., & Luo, J. J. (2019). Deep learning for multi-year ENSO forecasts. Nature, 573(7755), 568–572. https://doi.org/10.1038/s41586-019-1559-7
He, D., Lin, P., Liu, H., Ding, L., & Jiang, J. (2019). Dlenso: A deep learning ENSO forecasting model. In 32nd International conference on machine learning (pp. 448–456). Springer. https://doi.org/10.1007/978-3-030-39911-8_2
He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In 2016 IEEE conference on computer vision and pattern recognition (CVPR) (pp. 770–778). IEEE. https://doi.org/10.1109/CVPR.2016.90
Heaney, A. K., Shaman, J., & Alexander, K. A. (2019). El Niño-southern oscillation and under-5 diarrhea in Botswana. Proceedings of the National Academy of Sciences, 10(1), 1–9. https://doi.org/10.1038/s41467-019-13584-6
Henson, C., Market, P., Lupo, A., & Guinan, P. (2017). ENSO and PDO-related climate variability impacts on midwestern united states crop yields. International Journal of Biometeorology, 61(5), 857–867. https://doi.org/10.1007/s00484-016-1263-3
Hsiang, S. M., Meng, K. C., & Cane, M. A. (2011). Civil conflicts are associated with the global climate. Nature, 476(7361), 438–441. https://doi.org/10.1038/nature10311
Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International conference on machine learning (pp. 448–456).
Jan van Oldenborgh, G., Balmaseda, M. A., Ferranti, L., Stockdale, T. N., & Anderson, D. L. T. (2005). Did the ECMWF seasonal forecast model outperform statistical ENSO forecast models over the last 15 years? Journal of Climate, 18(16), 3240–3249. https://doi.org/10.1175/jcli3420.1
Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint. arXiv:1412.6980.
Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., Paolino, D. A., Zhang, Q., et al. (2014). The north American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. Bulletin of the American Meteorological Society, 95(4), 585–601. https://doi.org/10.1175/bams-d-12-00050.1
Lehohley, P., Bertrand, A., Hobday, A. J., Kiyofuji, H., McClatchie, S., Menkès, C. E., et al. (2020). ENSO impact on marine fisheries and ecosystems. In El Niño southern oscillation in a changing climate (pp. 429–451). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119548164.ch19
Levine, A. F., & McPhaden, M. J. (2015). The annual cycle in ENSO growth rate as a cause of the spring predictability barrier. Geophysical Research Letters, 42, 5304–5311. https://doi.org/10.1002/2015GL064309
Luo, J.-J., Masson, S., Behera, S. K., & Yamagata, T. (2008). Extended ENSO predictions using a fully coupled ocean–atmosphere model [Journal Article]. Journal of Climate, 21(1), 84–93. https://doi.org/10.1175/jcli4121.1
Mahesh, A., Evans, M., Jain, G., Castillo, M., Lima, A., Lunghino, B., et al. (2019). Forecasting El Niño with convolutional and recurrent neural networks. In 3rd conference on neural information processing systems (NEURIPS 2019) (pp. 8–14).
McPhaden, M. J. (2003). Tropical Pacific Ocean heat content variations and ENSO persistence barriers. Geophysical Research Letters, 30(9), 1480. https://doi.org/10.1029/2003GL016873
Mu, B., Peng, C., Yuan, S., & Chen, L. (2019). ENSO forecasting over multiple time horizons using convLSTM network and rolling mechanism. In 2019 international joint conference on neural networks (IJCNN) (pp. 1–8). IEEE. https://doi.org/10.1109/IJCNN.2019.8851967
Petersik, P. J., & Dijkstra, H. A. (2020). Probabilistic forecasting of El Niño using neural network models. Geophysical Research Letters, 47, e2019GL086423. https://doi.org/10.1029/2019GL086423
Petrova, D., Ballester, J., Koopman, S. J., & Rodó, X. (2020). Multiyear statistical prediction of ENSO enhanced by the tropical pacific observing system. Journal of Climate, 33(1), 163–174. https://doi.org/10.1175/JCLI-D-18-0877.1

Piotrowski, A. P., Napierkowski, J. J., & Piotrowska, A. E. (2020). Impact of deep learning-based dropout on shallow neural networks applied to stream temperature modelling. Earth-Science Reviews, 201, 103076. https://doi.org/10.1016/j.earscirev.2019.103076

Ren, H.-L., Scaife, A. A., Dunstone, N., Tian, B., Liu, Y., Ineson, S., & et al. (2018). Seasonal predictability of winter ENSO types in operational dynamical model predictions. Climate Dynamics, 52(7–8), 3869–3890. https://doi.org/10.1007/s00382-018-4366-1

Ren, H.-L., Zuo, J., & Deng, Y. (2018). Statistical predictability of Niño indices for two types of ENSO. Climate Dynamics, 52(9–10), 5361–5382. https://doi.org/10.1007/s00382-018-4453-3

Saha, S., Moorhi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., & et al. (2014). The NCEP climate forecast system version 2. Journal of Climate, 27(6), 2185–2208. https://doi.org/10.1175/JCLI-D-12-00082.1

Santos, M. A. D. C., Vega-Oliveros, D. A., Zhao, L., & Bertoni, L. (2020). Classifying El Niño-southern oscillation combining network science and machine learning. IEEE Access, 8, 55711–55723. https://doi.org/10.1109/access.2020.2982035

Sun, Y., Wang, F., & Sun, D.-Z. (2016). Weak ENSO asymmetry due to weak nonlinear air-sea interaction in cmip5 climate models. Advances in Atmospheric Sciences, 33(3), 352–364. https://doi.org/10.1007/s00376-015-5018-6

Tang, Y., Zhang, R.-H., Liu, T., Duan, W., Yang, D., Zheng, F., & et al. (2018). Progress in ENSO prediction and predictability study. National Science Review, 5(6), 826–839. https://doi.org/10.1093/nsr/nwy105

Timmermann, A., An, S. I., Kug, J. S., Jin, F. F., Cai, W., Capotondi, A., & et al. (2018). El Niño-southern oscillation complexity. Nature, 559(7715), 535–545. https://doi.org/10.1038/s41586-018-0252-6

Wang, X., Slawinska, J., & Gianikas, D. (2020). Extended-range statistical ENSO prediction through operator-theoretic techniques for nonlinear dynamics. Scientific Reports, 10(1), 1–15. https://doi.org/10.1038/s41598-020-59128-7

Yan, J., Mu, L., Wang, L., Ranjan, R., & Zomaya, A. Y. (2020). Temporal convolutional networks for the advance prediction of ENSO. Scientific Reports, 10(1), 1–15. https://doi.org/10.1038/s41598-020-65070-5

Yang, S., Li, Z., Yu, J.-Y., Hu, X., Dong, W., & He, S. (2018). El Niño-southern oscillation and its impact in the changing climate. National Science Review, 5(6), 840–857. https://doi.org/10.1093/nsr/nwy046

Yeh, S. W., Kug, J. S., Dewitte, B., Kwon, M. H., Kirtman, B. P., & Jin, F. F. (2009). El Nino in a changing climate. Nature, 461(7263), 511–514. https://doi.org/10.1038/nature08316

Yu, J., Zou, Y., Kim, S. T., & Lee, T. (2012). The changing impact of El Niño on us winter temperatures. Geophysical Research Letters, 39, L15702. https://doi.org/10.1029/2012GL052483

Yuan, Y., & Yang, S. (2012). Impacts of different types of El Niño on the east Asian climate: Focus on ENSO cycles. Journal of Climate, 25(21), 7702–7722. https://doi.org/10.1175/JCLI-D-11-00576.1

Zhang, Z., Ren, B., & Zheng, J. (2019). A unified complex index to characterize two types of ENSO simultaneously. Scientific Reports, 9(1), 1–8. https://doi.org/10.1038/s41598-019-44617-1