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

Jie Hu[1], Bin Weng[2], Tianqiang Huang[1], Jianyun Gao[3], Feng Ye[1], and Lijun You[4]

[1]College of Mathematics and Informatics, Fujian Normal University, Digital Fujian Institute of Big Data Security Technology, Engineering Technology Research Center for Public Service Big Data Mining and Application of Fujian Province
[2]College of Mathematics and Informatics
[3]Fujian Key Laboratory of Severe Weather, Fujian Institute of Meteorological Sciences
[4]Fujian Key Laboratory of Severe Weather, Fujian Meteorological Bureau

November 22, 2022

Abstract

To improve EI Niño-Southern Oscillation (ENSO) amplitude and type forecast, we pro-pose a model based on a deep residual convolutional neural network with few parame-ters. 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 EI Niño type. These results suggest that our proposed model can enhance the ENSO prediction performance.

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

Jie Hu[1,2,3*], Bin Weng[1,2,3†], Tianqiang Huang[1,2,3], Jianyun Gao[4], Feng Ye[1,2,3], Lijun You[5]

[1]College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350007, China
[2]Digital Fujian Institute of Big Data Security Technology, Fuzhou 350007, China
[3]Engineering Technology Research Center for Public Service Big Data Mining and Application of Fujian Province, Fuzhou 350007, China
[4]Fujian Key Laboratory of Severe Weather, Fujian Institute of Meteorological Sciences, Fuzhou 350007, China
[5]Fujian Meteorological Information Center, Fuzhou 350007, China

Key Points:

- Deep residual convolutional neural network is designed to forecast the amplitude and type of ENSO
- The prediction skill is improved by applying dropout and transfer learning
- Our method can successfully predict 20 months in advance for the period between 1984 and 2017

[*]Co-first authors
[†]Co-first authors

Corresponding author: Tianqiang Huang, `fjhtq@fjnu.edu.cn`

Corresponding author: Jianyun Gao, `fzgaojyun@163.com`

Abstract

To improve EI 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 EI 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 EI 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 1984-2017. Moreover, our model requires smaller-sized storage against the existing deep learning model.

1 Introduction

The EI Niño-Southern Oscillation (ENSO) is one of the main drivers of inter-annual 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 (Timmermann et al., 2018; Sun et al., 2016). 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 EI 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, e.g., different EI 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 (Saha et al., 2014; Ren, Scaife, et al., 2018) methods can generate skillful predictions 6-12 months in advance. Many deep learning-based methods have emerged in recent years, e.g. using Artificial Neural Networks (ANNs) (Petersik & Dijkstra, 2020; Feng et al., 2016), 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) (Mu et al., 2019; Gupta et al., 2020; D. He 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, e.g. based on the random forest (Santos et al., 2020), multi-model 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 EI Niño types. The accuracy remains 66.7% at lead times of 12 months. However, they have only expected the types of EI 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, three 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 EI 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 dataset is CMIP5 3.2 (core) (1861-2004) (Bellenger et al., 2014). The reanalysis dataset is simple ocean data assimilation version 2.2.4 (SODA) (1871-1973) (Giese & Ray, 2011) and Global Ocean Data Assimilation System (GODAS) (1982-2017) (Behringer & Xue, 2004).

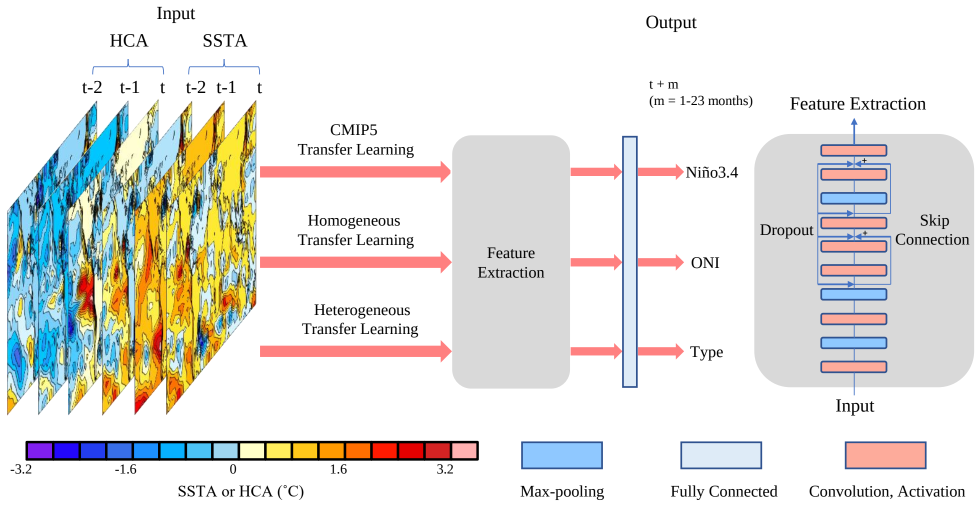

Figure 1. The architecture of the Res-CNN model. The variables of the input layer correspond to the sea surface temperature (in units of °C) anomaly and the oceanic heat content (in units of °C) anomaly from time t - 2 months to t months, between 0°–360°E and 55°S–60°N. The three-month-averaged Niño3.4 index, ONI and ENSO type from time t + 1 months to t + 23 months is used as a variable for the output layer.

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

$$y_{t+l} = F_l\left(x_{t-2}, x_{t-1}, x_t\right) \tag{1}$$

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 7-layer convolutional neural network, a 3-layer max-pooling to extract features, 2-layer skip connections, and 1-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:

$$v_{l,f}^{x,y} = \sum_{m=1}^{M_{l-1}} \sum_{p=1}^{P_l} \sum_{q=1}^{Q_l} w_{l,f,m}^{p,q} v_{(l-1),m}^{(x+p-P_l/2, y+q-Q_l/2)} + b_{l,f} \tag{2}$$

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_l, Q_l) 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

$$y = R(x, \{W\}) + x \tag{3}$$

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). The final model would be the best result from the model that determines the number of residual connections. See Text 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. We use the correlation coefficient function I as a measure of the ENSO index prediction:

$$I_l = \sum_{t=1}^{12} \frac{\sum_{y=s}^{e} \left(O_{y,t} - \bar{O}_t\right) \left(F_{y,t,l} - \bar{F}_{t,l}\right)}{\sqrt{\sum_{y=s}^{e} \left(O_{y,t} - \bar{O}_m\right)^2 \sum_{y=s}^{e} \left(F_{y,t,l} - \bar{F}_{t,l}\right)^2}} \tag{4}$$

where, O and F denote the observed and the predicted values, respectively. $\bar{O}_{t,l}$ and $\bar{F}_{t,l}$ 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, i.e. EP, CP, and MIX of EI Niño; and the other is with seven types, i.e. EP, CP, MIX of EI Niño (La Niña), and Normal Year (NY). See the Text S3 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 three-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 [40] 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), 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

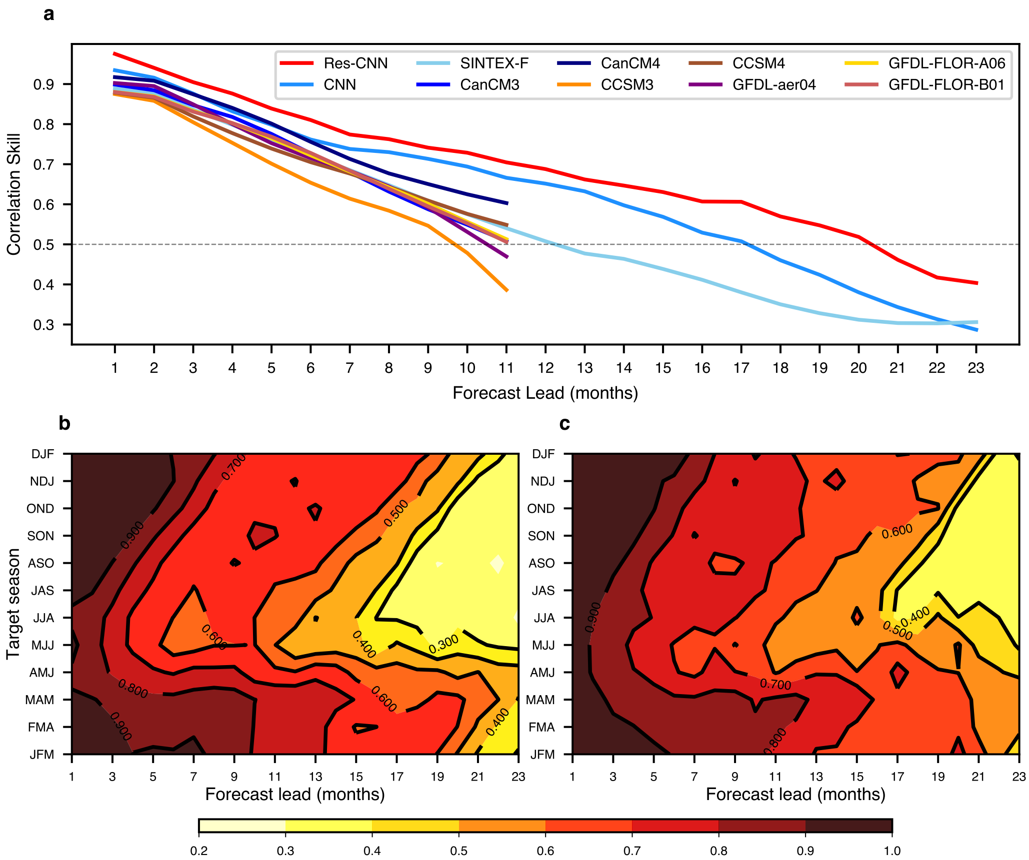

Figure 2. Correlation skill for various lead times and methods. All season correlation skill of the three-month-moving-averaged Niño3.4 index for multiple lead times from 1984 to 2017 in the Res-CNN model (red), CNN model (Ham et al., 2019) (dodger blue), SINTEX-F (Luo et al., 2008) dynamical forecast system (sky blue), and other dynamical forecast systems (Kirtman et al., 2014) included in the North American Multi-Model Ensemble (NMME) project (the other colors). 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.

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 three-month-moving-averaged Niño3.4 index from 1982 to 2001 and from 2002 to 2017 is shown in Figure S2. 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 ten 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 to 2017 (Figure S3). 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), 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), and the Res-CNN (c). Compared to the GDNN and QRNN methods (Petersik & Dijkstra, 2020), 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 one year, the amount of data has an impact on the model.

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 pre-training and soda training methods are not as good as just using the pre-training 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 dataset (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 super ENSO 12 months in advance, especially the 2015/2016 EI 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 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 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

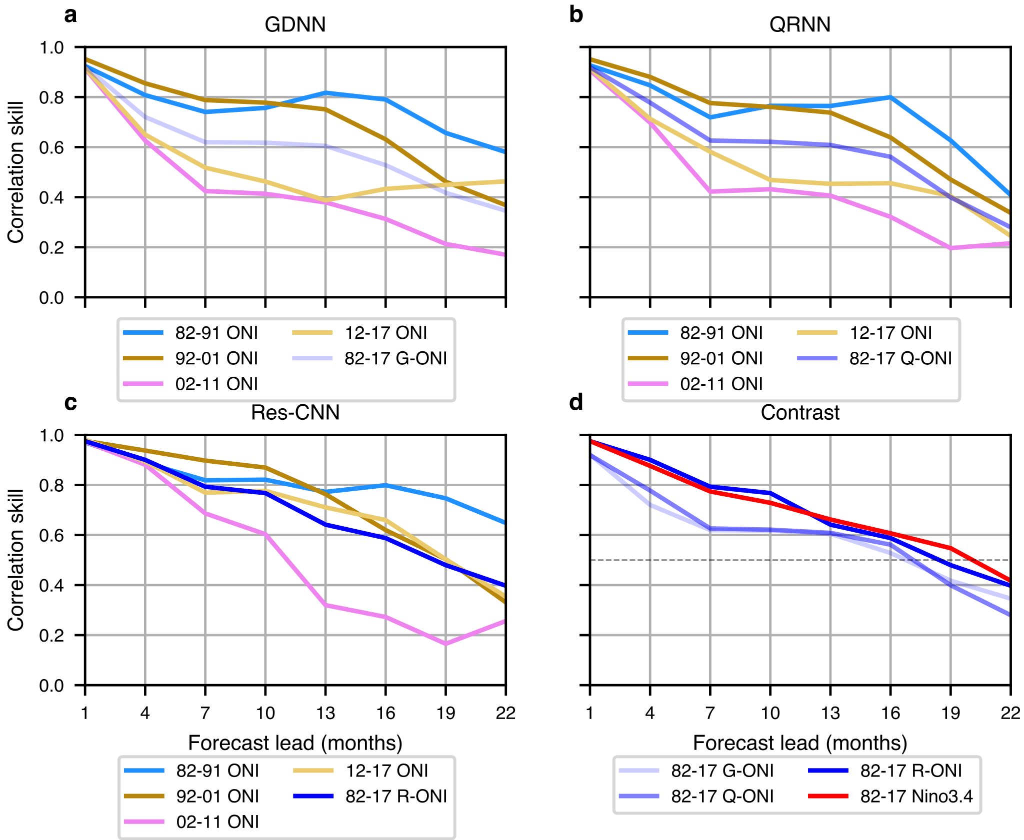

Figure 3. Correlation skill of the ONI for various lead times, decades, and models. The all-season correlation skill of the ONI from 1982 to 1991, from 1992 to 2001, from 2002 to 2011, from 2012 to 2017, from 1982 to 2017 using GDNN (a), QRNN (b) and Res-CNN (c) with transfer learning for various lead times. (d) The all-season correlation skill of the ONI between 1982 and 2017 using GDNN, QRNN, and Res-CNN at various lead times, Niño3.4 index using Res-CNN (red). The black dashed line indicates that the correlation coefficient is equal to 0.5.

Table 1. Prediction of ENSO types

Methods	Lead months 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

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 non-use 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 EI 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.

indicate that the Res-CNN model predicts the ENSO index and type better than the CNN model.

4 Discussions

Through various dropout 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). The finding differs from the conventional deep learning approach usually set with 0.4-0.6 of the dropout rate. The achievement is due to fewer parameters of the network. Therefore, too large dropout rate will lead to significant reduction of the learnable parameters of the network during each round of iterations. Hence such training process would not produce suitable results. To find the appropriate number of residual connections, we conducted ablation experiments. The obtained results (Figure S7) 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 2 predicted best. Furthermore, Figure S8 shows that using un-normalized 2 residual connections achieved significantly 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), the same as CNN (Table S2) and SINTEX-F (Table S3). 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) 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 EI Niño type, the result in Table 1 is still almost 20 percentage 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, e.g., intuitively, trying fewer predictors and input months under shorter advance months.

Acknowledgments

This study was supported by National Key R&D Program Special Fund Grant (No. 2018YFC1505805), National Natural Science Foundation of China (No. 62072106, No. 61070062), General Project of Natural Science Foundation in Fujian Province (No. 2020J01168), Open project of Fujian Key Laboratory of Severe Weather (No. 2020KFKT04). The research data can be found in (`https://doi.org/10.5281/zenodo.4646653`), and the code used in this paper are available at this site (`https://github.com/icodeworld/Deep-learning-ENSO`).

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 increasing? [Journal Article]. *Bulletin of the American Meteorological Society*, *93*(5), 631-651. doi: 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 *Proc. eighth symp. on integrated observing and assimilation systems for atmosphere, oceans, and land surface.*

Bellenger, H., Guilyardi, É., Leloup, J., Lengaigne, M., & Vialard, J. (2014). Enso representation in climate models: From cmip3 to cmip5. *Climate Dynamics*, *42*(7), 1999–2018. doi: 10.1007/s00382-013-1783-z

Broni-Bedaiko, C., Katsriku, F. A., Unemi, T., Atsumi, 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. doi: 10.1007/s10015-019-00540-2

Cachay, S. R., Erickson, E., Bucker, A. F. C., Pokropek, E., Potosnak, W., Osei, S., & Lütjens, B. (2020). Graph neural networks for improved el ni\˜ no forecasting. *arXiv preprint arXiv:2012.01598*.

Feng, Q. Y., Vasile, R., Segond, M., Gozolchiani, A., Wang, Y., Havlin, S., ... Dijkstra, H. A. (2016). Climatelearn: A machine-learning approach for climate prediction using network measures [Journal Article]. *. CC*, 18. doi: 10.5194/gmd-2015-273

Geng, T., Cai, W., & Wu, L. (2020). Two types of enso varying in tandem facilitated by nonlinear atmospheric convection [Journal Article]. *Geophysical Research Letters*, e2020GL088784. doi: 10.1029/2020gl088784

Giese, B. S., & Ray, S. (2011). El niño variability in simple ocean data assimilation (soda), 1871–2008. *Journal of Geophysical Research: Oceans*, *116*(C2). doi: 10.1029/2010jc006695

Gupta, M., Kodamana, H., & Sandeep, S. (2020). Prediction of enso beyond spring predictability barrier using deep convolutional lstm networks. *IEEE Geoscience and Remote Sensing Letters*. doi: LGRS.2020.3032353

Ham, Y. G., Kim, J. H., & Luo, J. J. (2019). Deep learning for multi-year enso forecasts [Journal Article]. *Nature*, *573*(7775), 568-572. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/31534218` doi: 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 *Pacific rim international conference on artificial intelligence* (pp. 12–23). doi: 10.1007/978-3-030-29911-8_2

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition [Conference Proceedings]. In *2016 ieee conference on computer vision and pattern recognition (cvpr)* (p. 770-778). Las Vegas, NV, USA: IEEE. doi: 10.1109/CVPR.2016.90

Heaney, A. K., Shaman, J., & Alexander, K. A. (2019). El niño-southern oscillation and under-5 diarrhea in botswana. *Nature communications*, *10*(1), 1–9. doi: 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 [Journal Article]. *International journal of biometeorology*, *61*(5), 857-867. doi: 10.1007/s00484-016-1263-3

Hsiang, S. M., Meng, K. C., & Cane, M. A. (2011). Civil conflicts are associated with the global climate [Journal Article]. *Nature*, *476*(7361), 438-441. doi: 10.1038/nature10311

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *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 Article]. *Journal of Climate*, *18*(16), 3240-3249. Retrieved from `<GotoISI>://WOS:000232051600011` doi: 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., ... others (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. doi: 10.1175/bams-d-12-00050.1

Lehodey, P., Bertrand, A., Hobday, A. J., Kiyofuji, H., McClatchie, S., Menkès, C. E., ... Tommasi, D. (2020). Enso impact on marine fisheries and ecosystems [Journal Article]. *El Niño Southern Oscillation in a Changing Climate*, 429-451. doi: 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*(12), 5034–5041. doi: 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. doi: 10.1175/2007jcli1412.1

Mahesh, A., Evans, M., Jain, G., Castillo, M., Lima, A., Lunghino, B., ... others (2019). Forecasting el niño with convolutional and recurrent neural networks. In *33rd conference on neural information processing systems (neurips 2019), vancouver, canada* (pp. 8–14).

McPhaden, M. J. (2003). Tropical pacific ocean heat content variations and enso persistence barriers [Journal Article]. *Geophysical Research Letters*, *30*(9). Retrieved from `<GotoISI>://WOS:000182873800003` doi: 10.1029/2003gl016872

Mu, B., Peng, C., Yuan, S., & Chen, L. (2019). Enso forecasting over multiple time horizons using convlstm network and rolling mechanism [Conference Proceedings]. In *2019 international joint conference on neural networks (ijcnn)* (p. 1-8). Budapest, Hungary: IEEE. doi: 10.1109/IJCNN.2019.8851967

Petersik, P. J., & Dijkstra, H. A. (2020). Probabilistic forecasting of el niño using neural network models [Journal Article]. *Geophysical Research Letters*, *47*(6), e2019GL086423. doi: 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 Article]. *Journal of Climate*, *33*(1), 163-174. doi: 10.1175/jcli-d-18-0877.1

Piotrowski, A. P., Napiorkowski, 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. doi: 10.1016/j.earscirev.2019.103076

Ren, H.-L., Scaife, A. A., Dunstone, N., Tian, B., Liu, Y., Ineson, S., ... MacLachlan, C. (2018). Seasonal predictability of winter enso types in operational dynamical model predictions [Journal Article]. *Climate Dynamics*, *52*(7-8), 3869–3890. doi: 10.1007/s00382-018-4366-1

Ren, H.-L., Zuo, J., & Deng, Y. (2018). Statistical predictability of niño indices for two types of enso [Journal Article]. *Climate Dynamics*, *52*(9-10), 5361–5382. doi: 10.1007/s00382-018-4453-3

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., ... others (2014). The ncep climate forecast system version 2. *Journal of climate*, *27*(6), 2185–2208.

Santos, M. A. D. C., Vega-Oliveros, D. A., Zhao, L., & Berton, L. (2020). Classifying el niño-southern oscillation combining network science and machine learning [Journal Article]. *IEEE Access*, *8*, 55711–55723. doi: 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 [Journal Article]. *Advances in Atmospheric Sciences*, *33*(3), 352-364. Retrieved from `<GotoISI>://WOS:000370030300008` doi: 10.1007/s00376-015-5018-6

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

Timmermann, A., An, S. I., Kug, J. S., Jin, F. F., Cai, W., Capotondi, A., ... Zhang, X. (2018). El niño–southern oscillation complexity [Journal Article]. *Nature*, *559*(7715), 535-545. Retrieved from https://www.nature.com/articles/s41586-018-0252-6 doi: 10.1038/s41586-018-0252-6

Wang, X., Slawinska, J., & Giannakis, D. (2020). Extended-range statistical enso prediction through operator-theoretic techniques for nonlinear dynamics. *Scientific reports*, *10*(1), 1–15. doi: 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. doi: 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 [Journal Article]. *National Science Review*, *5*(6), 840-857. doi: 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 [Journal Article]. *Nature*, *461*(7263), 511-4. doi: 10.1038/nature08316

Yu, J., Zou, Y., Kim, S. T., & Lee, T. (2012). The changing impact of el niño on us winter temperatures [Journal Article]. *Geophysical Research Letters*, *39*(15). doi: 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 Article]. *Journal of Climate*, *25*(21), 7702-7722. doi: 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 [Journal Article]. *Scientific reports*, *9*(1), 1-8. doi: 10.1038/s41598-019-44617-1

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

Jie Hu[1,2,3] *, Bin Weng[1,2,3] †, Tianqiang Huang[1,2,3], Jianyun Gao[4], Feng Ye[1,2,3], Lijun You[5]

[1]College of Mathematics and Informatics, Fujian Normal University, Fuzhou 350007, China

[2]Digital Fujian Institute of Big Data Security Technology, Fuzhou 350007, China

[3]Engineering Technology Research Center for Public Service Big Data Mining and Application of Fujian Province, Fuzhou 350007, China

[4]Fujian Key Laboratory of Severe Weather, Fujian Institute of Meteorological Sciences, Fuzhou 350007, China

[5]Fujian Meteorological Information Center, Fuzhou 350007, China

1. Text S1
2. Text S2
3. Text S3

Corresponding author: Tianqiang Huang, fjhtq@fjnu.edu.cn

Corresponding author: Jianyun Gao, fzgaojyun@163.com

*Co-first authors

†Co-first authors

4. Figures S1 to S9

5. Table S1 to S3

Text S1. Dropout and Transfer Learning

The dropout technique is used to alleviate the overfitting issue. In this work we use the base version (Srivastava et al., 2014), i.e., in each iteration, each node may be dropped out with probability p according to the Bernoulli distribution. The p is the probability that the node is set to zero, it's then scaled by a factor of $1/(1-p)$. This method can facilitate the interaction between feature detectors (hidden layer nodes). Detector interaction means that some detectors rely on other detectors to function. In other words, it lets the activation value of a neuron stop working with a certain probability as forwarding propagation. This can make the model more general because it does not rely too much on some local characteristics.

For transfer learning (Zhuang et al., 2020), the feature space of the data is represented by attributes ($X_s = X_t$) and labels (Y_s, Y_t), with s representing the source object and t representing the target object. Y_s and Y_t are equal for homogeneous transfer learning and unequal for heterogeneous transfer learning. The most straightforward method is applying the source domain model directly to the target domain before training. For example, the ONI (type) is predicted by using homogeneous (heterogeneous) transfer learning, which uses the model parameters of predicting the Niño3.4 index as the initialization parameters before training.

Text S2. Indexes forecast

The batch size is 400 (20) and epoch is 100 (50) during transfer learning (training). The learning rate is set to 0.001 and a learning rate scheduler is used. To compare the prediction performance of Res-CNN, the North American Multi-model Ensemble phase 1 (Kirtman et al., 2014), Scale Interaction Experiment-Frontier (SINTEX-F) (Luo et al., 2008), and CNN are used during the valid period from 1984 to 2017. To predict ONI, we divide the valid data into four-time slices, 1982-1991, 1992-2001, 2002-2011, and 2012-2017. First, one slice data is randomly taken as a test set, then the first 20 years of the remaining slice data are considered as a training set, while the remaining time slice data as a validation set. Since the test sets are independent, there is no overlap between training data, validation data, and test data. The batch size is 100, the epoch is 100, and the learning rate is 0.0005.

Text S3. Types forecast

We measure the performance using accuracy, which is the number of correct predictions divided by the total number of predictions. All types are defined by the Niño3 index (SSTA averaged over 150°- 90°W, 5°S-5°N), denoted as N_3, and the Niño4 index (SSTA averaged over 160°E-150°W, 5°S-5°N), denoted as N_4. The calculation of classification is as follows:

$$
\begin{aligned}
r &= \sqrt{(N_3 + N_4)^2 + (N_3 - N_4)^2} = \sqrt{2\left(N_3^2 + N_4^2\right)} \\
\theta &= \begin{cases} \arctan\frac{(N_3 - N_4)}{(N_3 + N_4)} & N_3 + N_4 > 0 \\ \arctan\frac{(N_3 - N_4)}{(N_3 + N_4)} - 180 & N_3 + N_4 < 0 \end{cases} \\
\theta &\in \begin{cases} (15^\circ, 90^\circ) & \text{EP EI Niño} \\ (-15^\circ, 15^\circ) & \text{MIX EI Niño} \\ (-90^\circ, -15^\circ) & \text{CP EI Niño} \\ (-165^\circ, -90^\circ) & \text{EP La Niña} \\ (-195^\circ, -165^\circ) & \text{MIX La Niña} \\ (-270^\circ, -195^\circ) & \text{CP La Niña} \\ \text{other} & \text{NY} \end{cases}
\end{aligned} \tag{1}
$$

Finally, the ENSO events occur when r in the December-January-February (DJF) season is greater than their standard deviation (Ham et al., 2019).

References

Ham, Y. G., Kim, J. H., & Luo, J. J. (2019). Deep learning for multi-year enso forecasts [Journal Article]. *Nature*, *573*(7775), 568-572. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/31534218` doi: 10.1038/s41586-019-1559-7

Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., Paolino, D. A., Zhang, Q., ... others (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. doi: 10.1175/bams-d-12-00050.1

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. doi: 10.1175/2007jcli1412.1

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, *15*(1), 1929–1958.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, *109*(1), 43–76. doi: 10.1109/jproc.2020.3004555

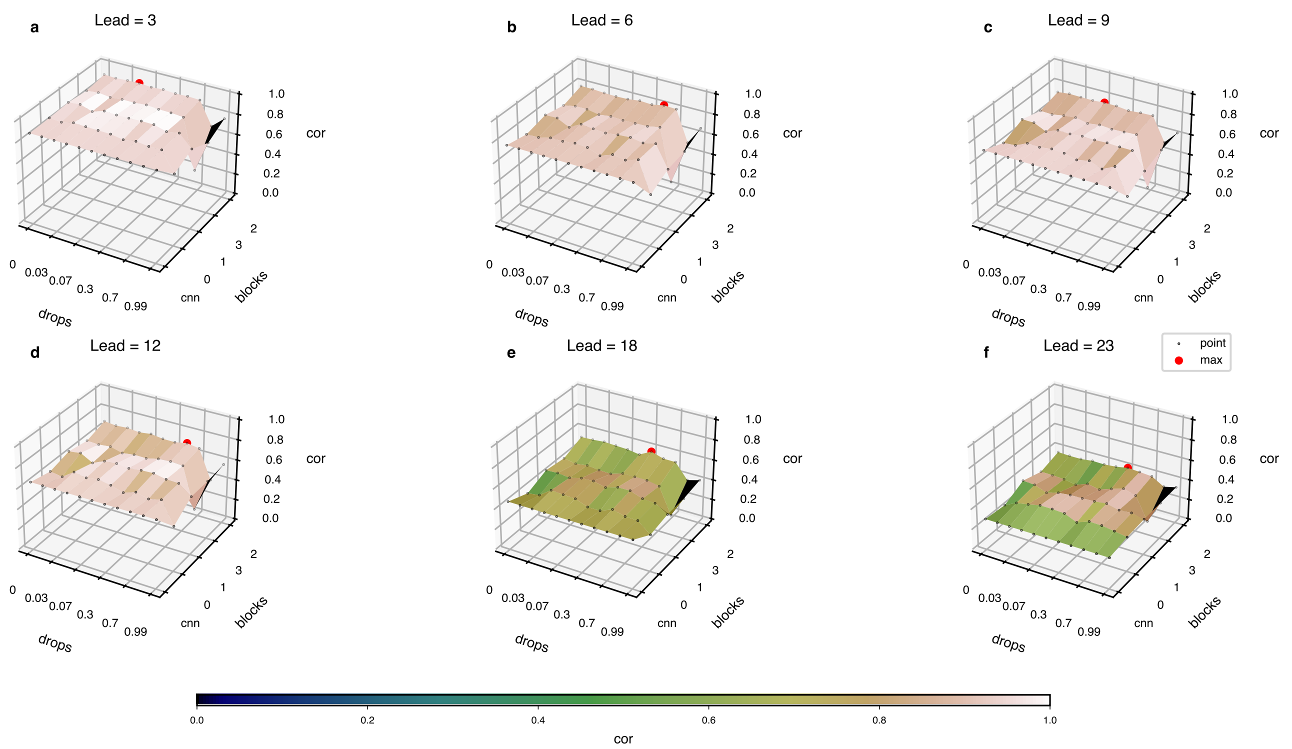

Figure S1. Model selection. The optimal model for a given lead month is selected from many different dropout rates and the number of residuals connected. Here only the advance months are shown as 3, 6, 9, 12, 18, 23, and the corresponding is a-f. The best number of residual blocks is 2, and the best dropout rate is between 0.05 and 0.5.

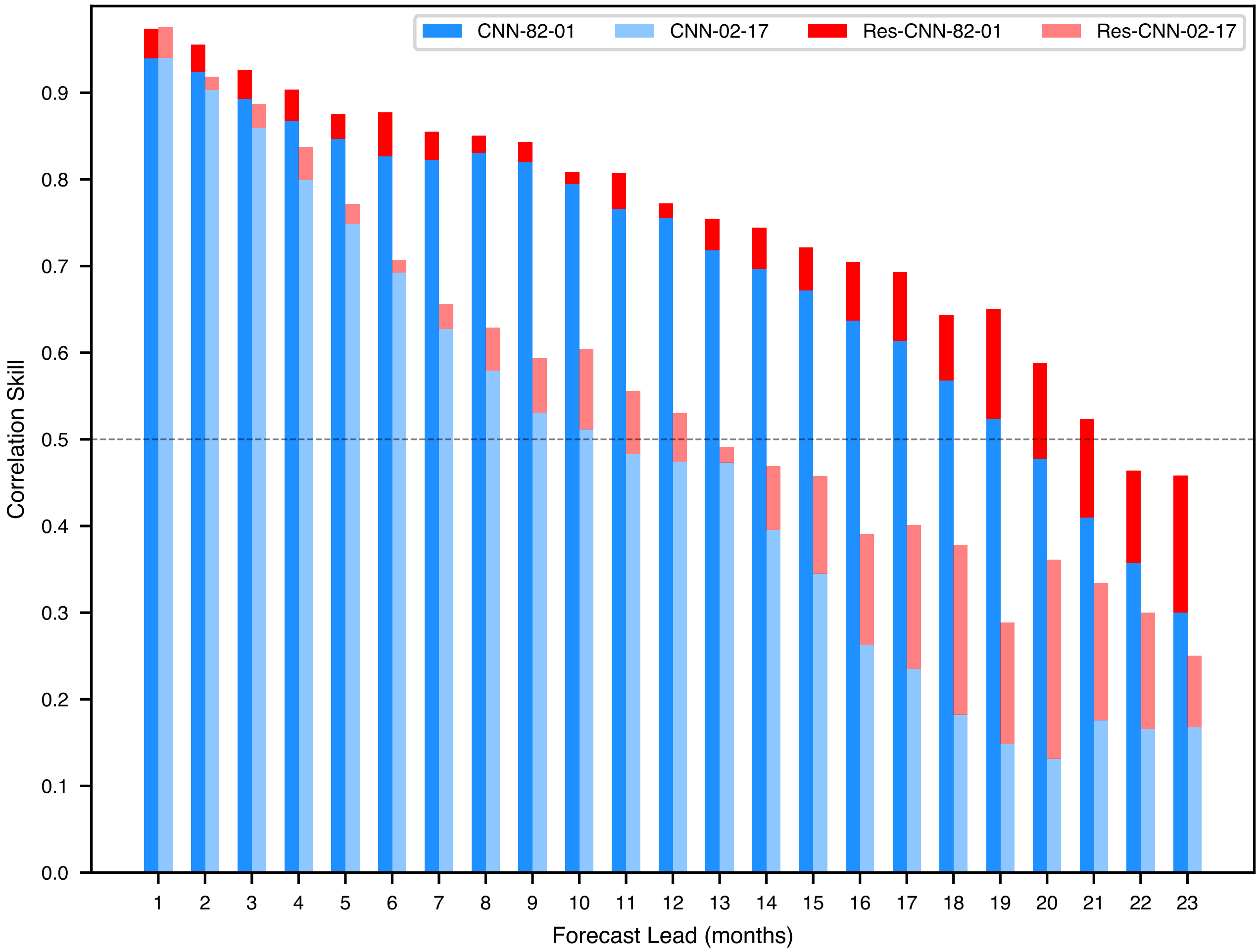

Figure S2. Correlation skill of the CNN and the Res-CNN for various lead times and decades. The all-season correlation skill of the three-month-moving-averaged Niño3.4 index from 1982 to 2001 (deep) and from 2002 to 2017 (shallow) as a function of the prediction lead month using the CNN model (dodger blue) and the Res-CNN model (red). The black dashed line indicates that the correlation coefficient is equal to 0.5.

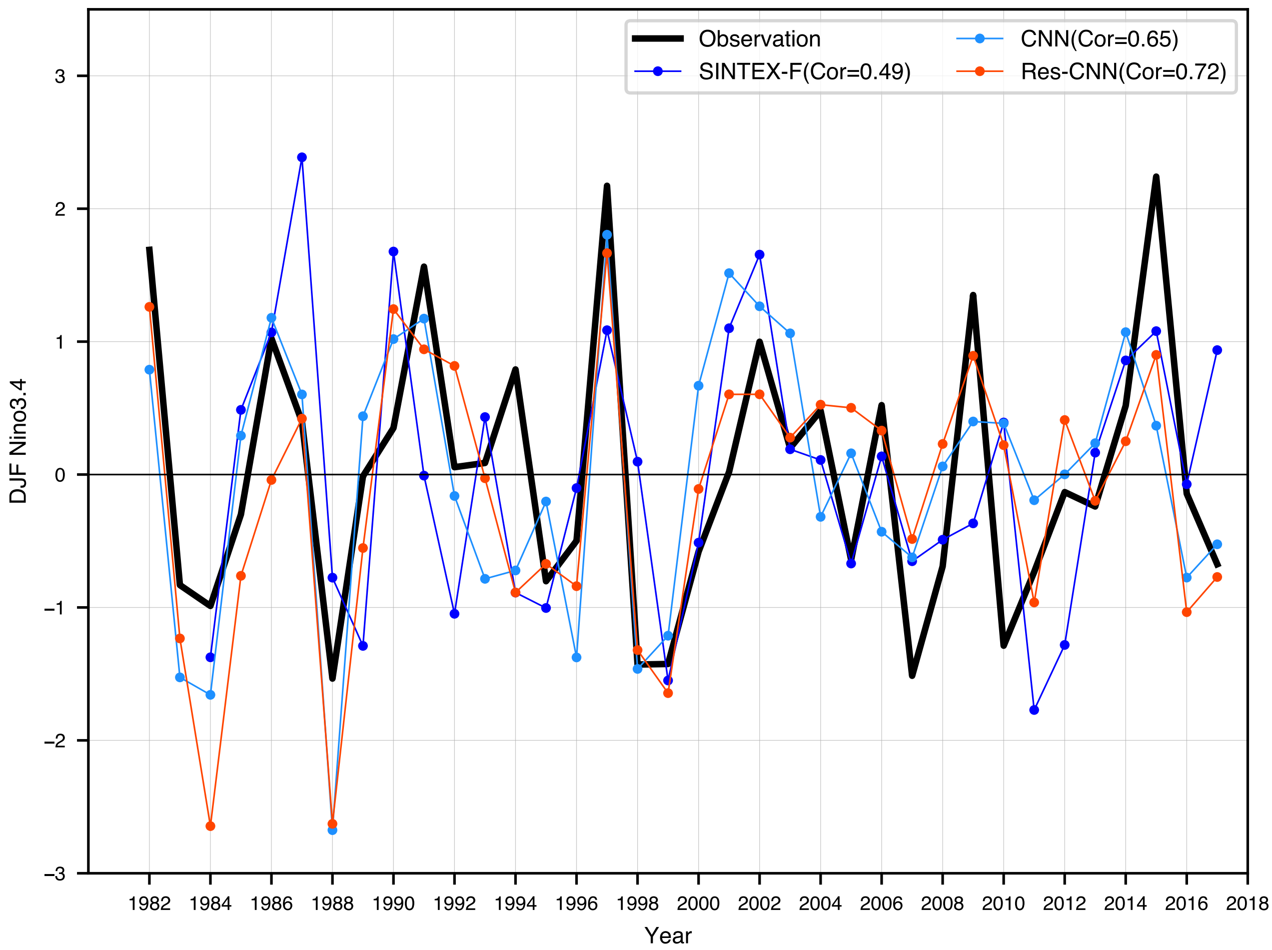

Figure S3. Time series 18 months in advance. Time series of DJF season Niño3.4 indexes for 18-month-lead prediction using the CNN model (dodger blue) and the Res-CNN model(red). Cor means correlation skill for DJF season from 1982 to 2017.

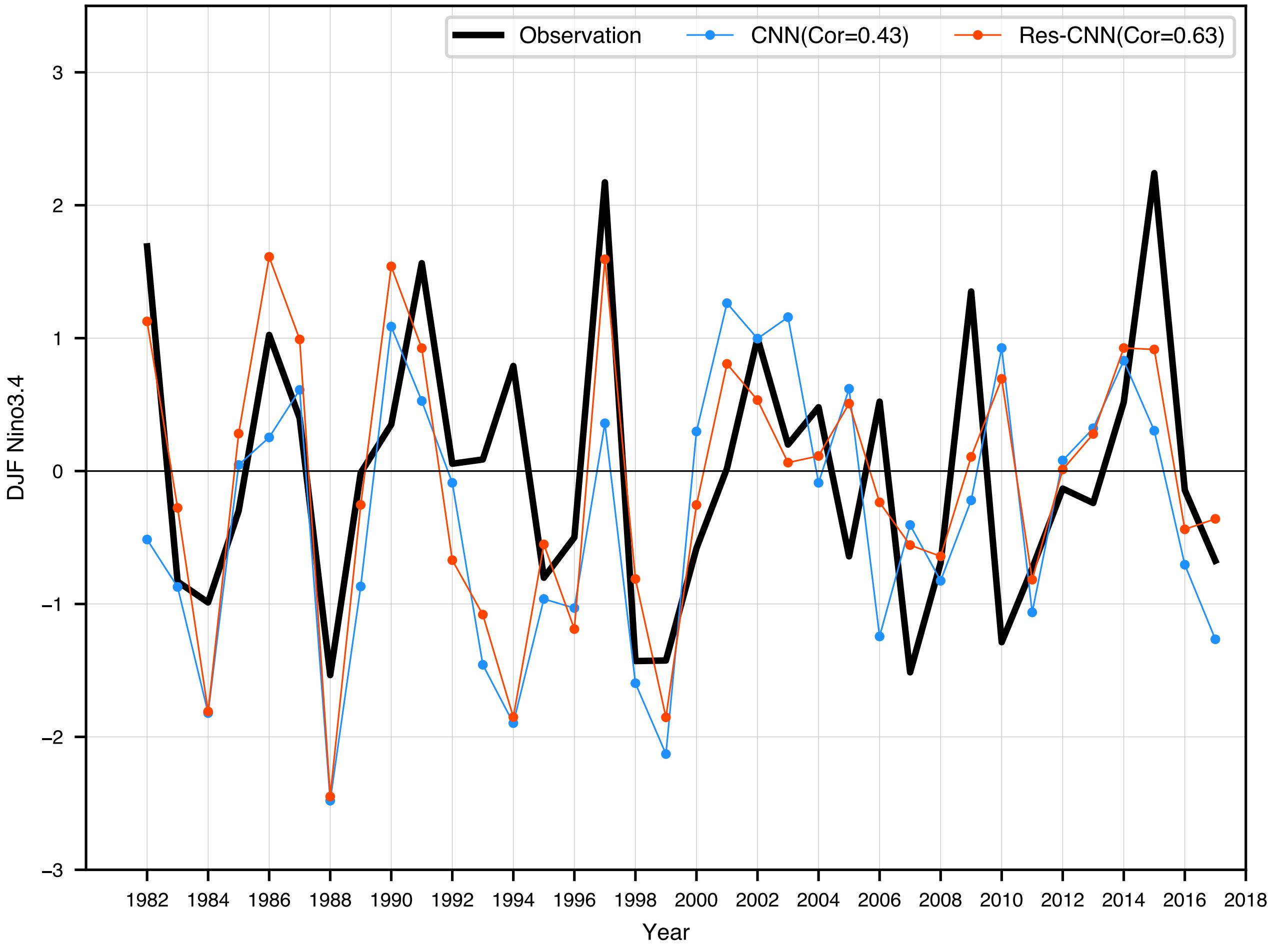

Figure S4. Niño3.4 prediction 20-month in advance. Time-series of DJF season Niño3.4 indexes for 20-month-lead prediction using the CNN model (dodger blue), and the Res-CNN model(red). Cor means correlation skill for DJF season from 1982 to 2017.

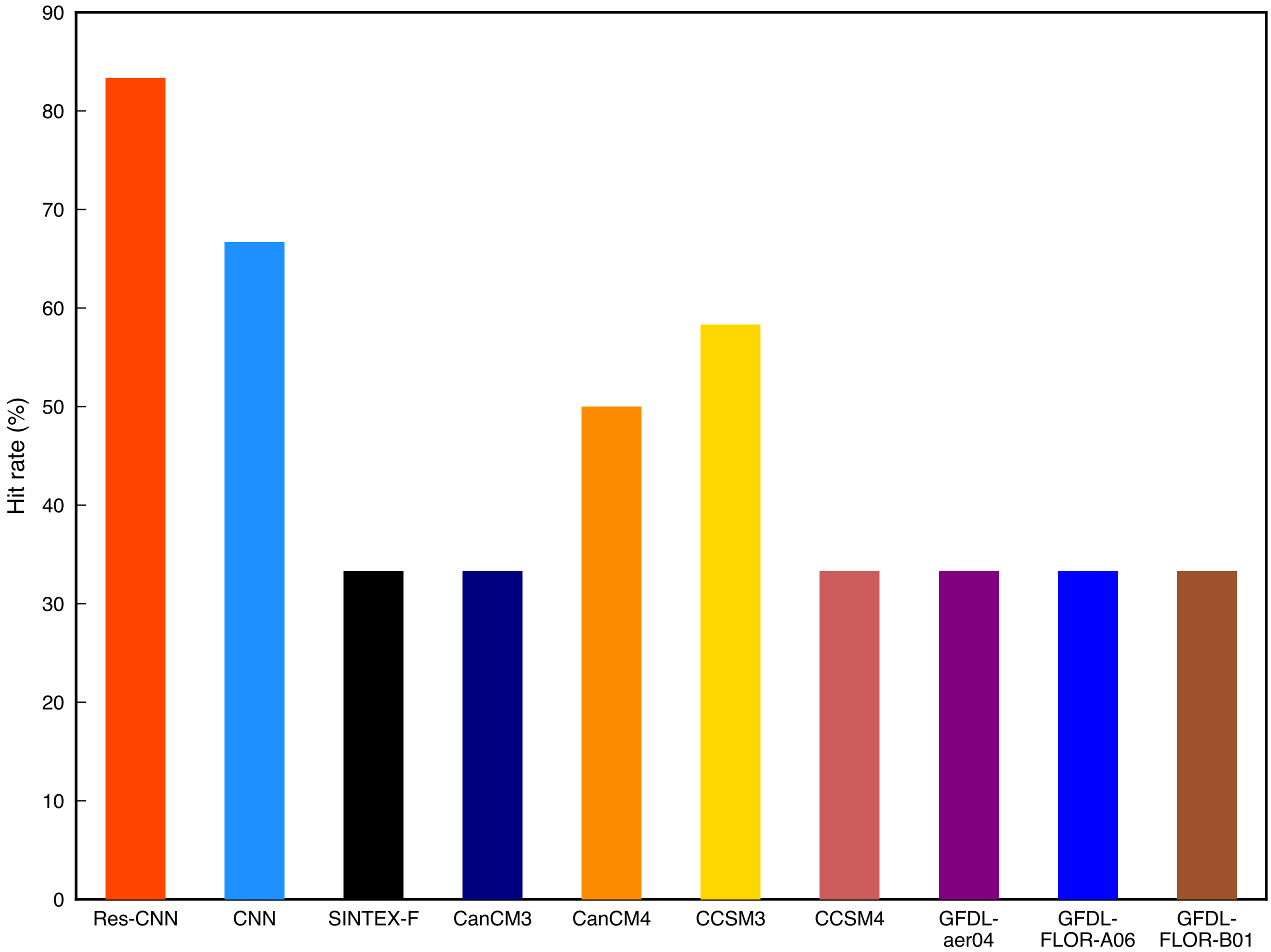

Figure S5. Prediction of EI Niño type. The prediction accuracy of EI Niño types (EP, CP, MIX) 12 months in advance using Res-CNN (red), CNN (dodger blue), and other models from 1982 to 2017.

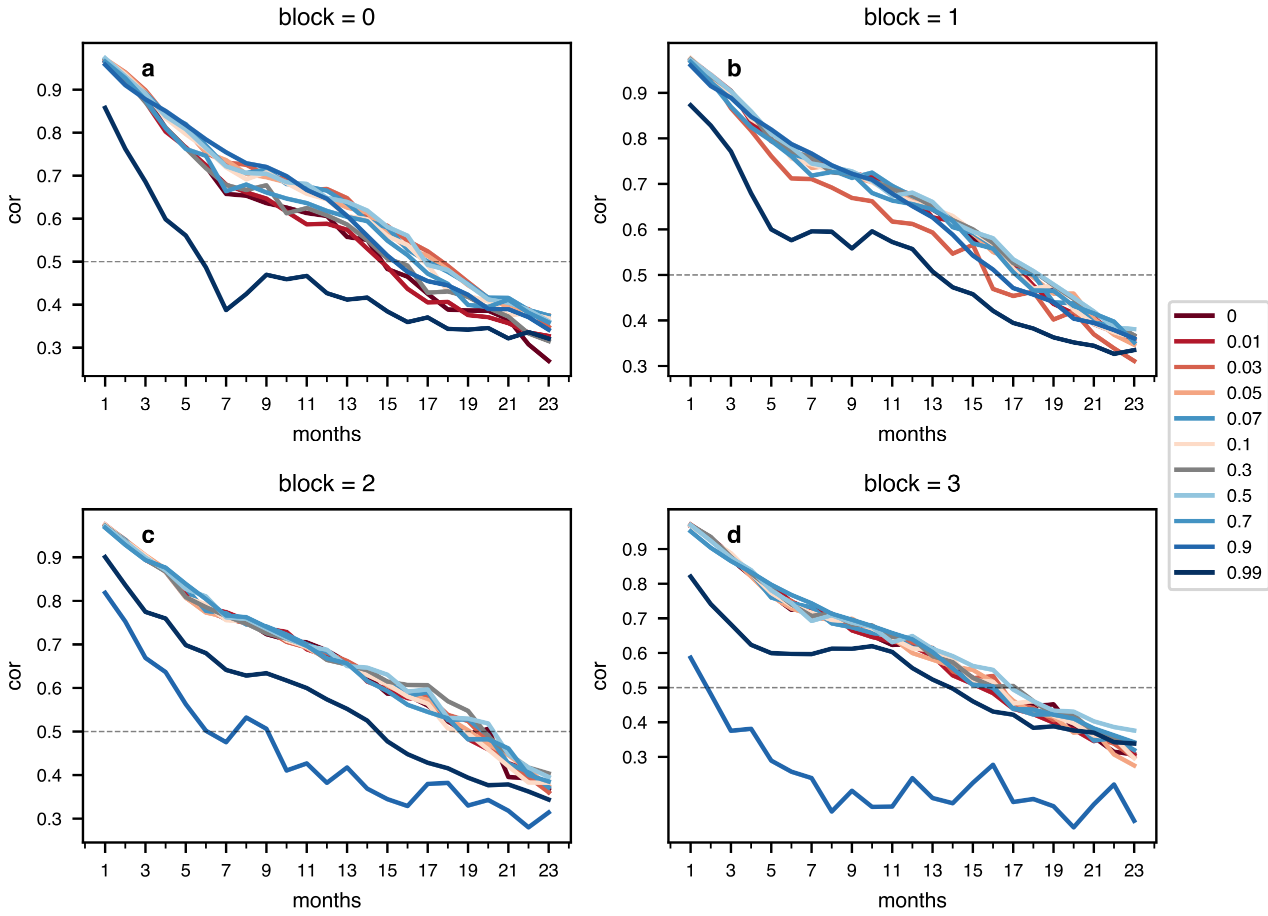

Figure S6. The impact of the dropout. The effect of the dropout rate under different skip connection conditions. Here the skip connections of 0-3 are shown as a-d, respectively. The black dashed line indicates that the correlation coefficient is equal to 0.5. Notice that the dropout rate of 0.99 is better than the dropout rate of 0.9 for c and d, probably because the number of skip connections is higher, and its effect is greater than the dropout.

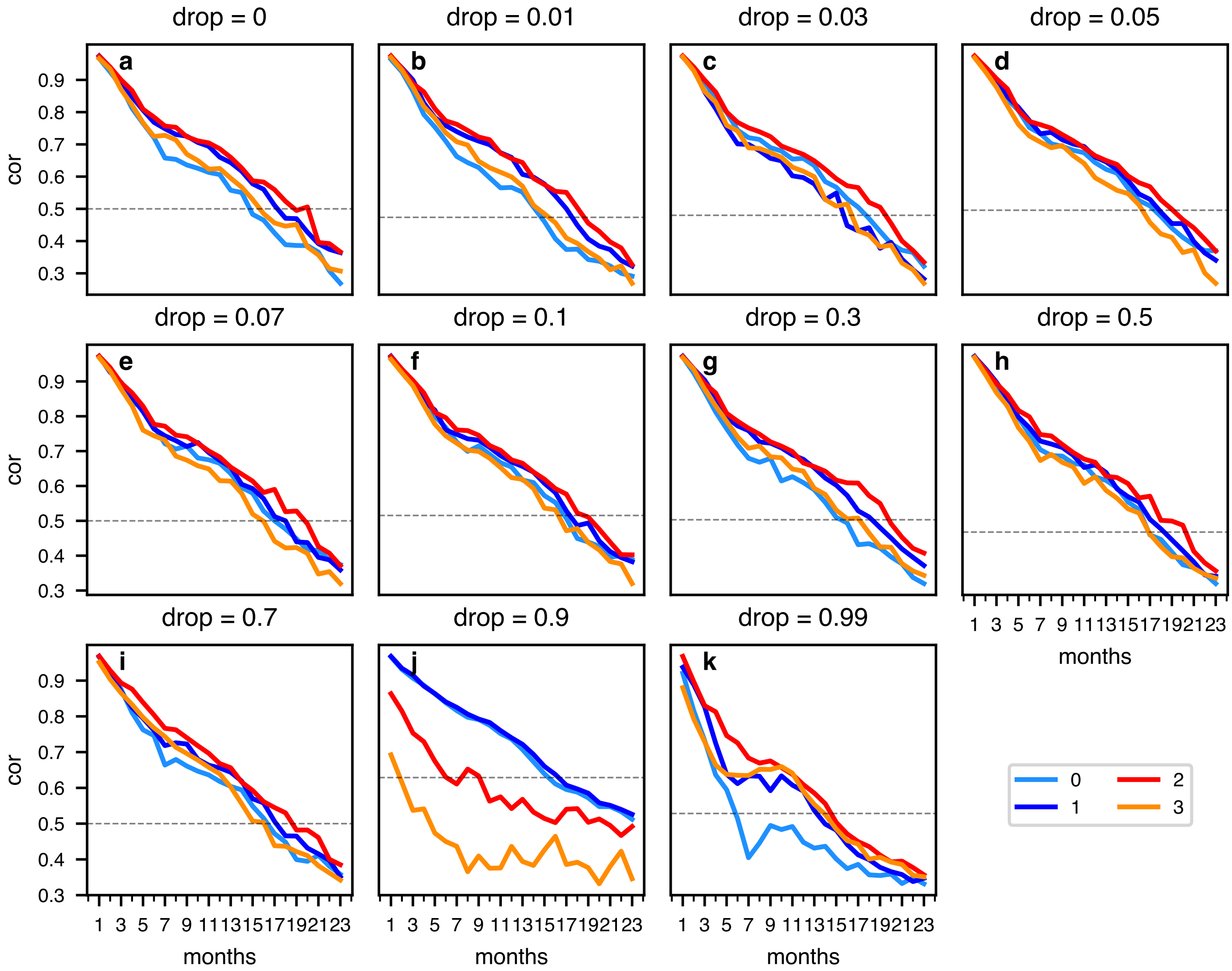

Figure S7. The impact of skip connection. The effect of the skip connection under different dropout rate conditions. Here the dropout rates are 0, 0.01, 0.03, 0.05, 0.07, 0.1, 0.3, 0.5, 0.7, 0.9, 0.99, shown in a-k, respectively. The black dashed line indicates that the correlation coefficient is equal to 0.5.

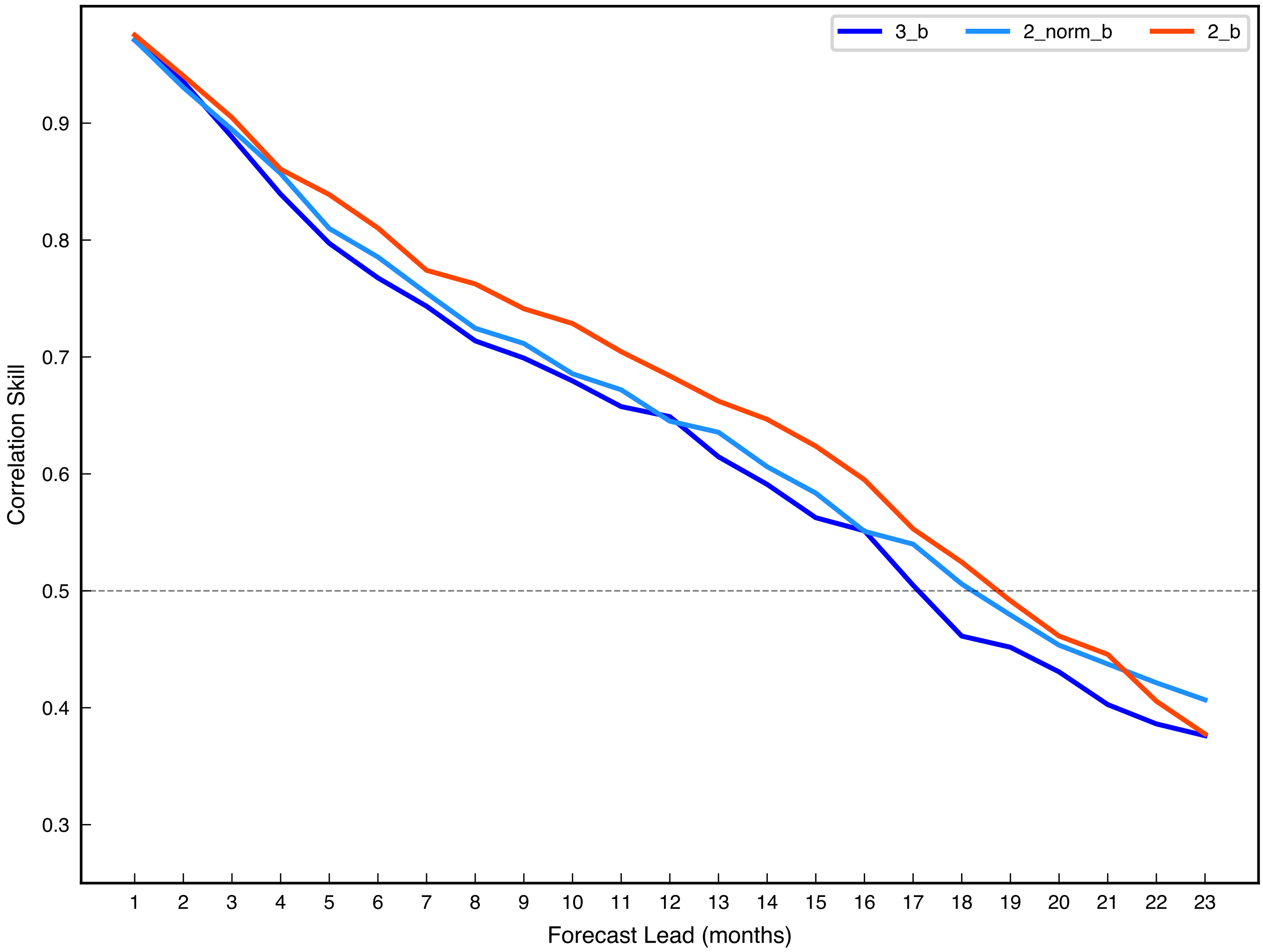

Figure S8. Compares the skill between un-normalized 3 residual connections (blue), un-normalized 2 residual connections (red), and normalized 2 residual connections (dodger blue). The black dashed line indicates that the correlation coefficient is equal to 0.5.

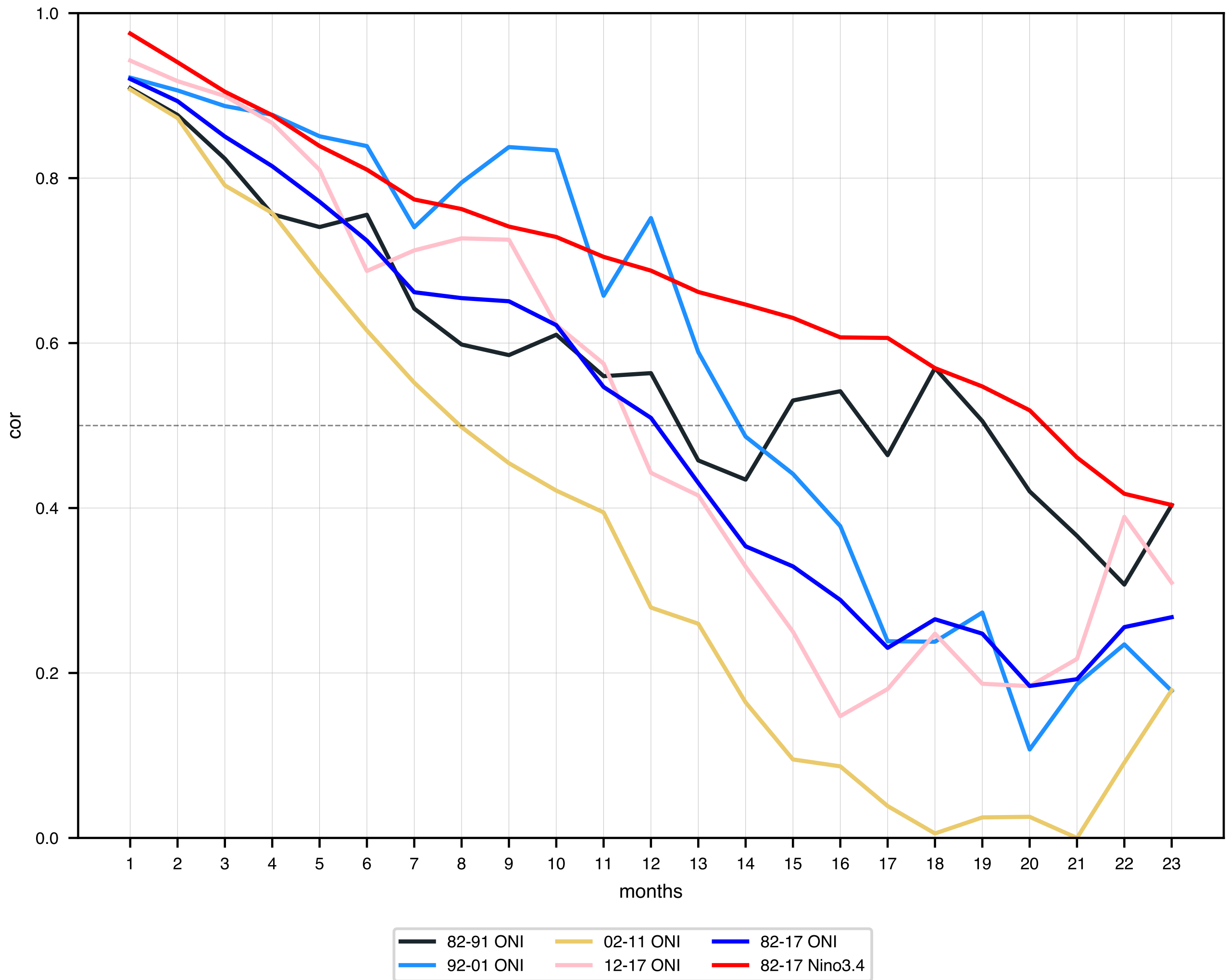

Figure S9. All-season correlation skill of the ONI and Niño3.4 index. The all-season correlation skill of the three-month-moving-averaged ONI from 1982 to 1991(black), ONI from 1992 to 2001(dodger blue), ONI from 2002 to 2011(khaki), ONI from 2012 to 2017(pink), ONI from 1982 to 2017(blue), Niño3.4 from 1982 to 2017(red) using the Res-CNN without transfer learning for various lead times. The black dashed line indicates that the correlation coefficient is equal to 0.5.

Table S1. Correlation skill - Res-CNN

	Target season											
Forecast lead	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF
1	0.98	0.97	0.99	0.96	0.95	0.97	0.97	0.98	0.97	0.97	0.99	0.99
2	0.95	0.96	0.93	0.89	0.89	0.90	0.95	0.94	0.97	0.97	0.97	0.97
3	0.94	0.92	0.88	0.84	0.85	0.83	0.88	0.92	0.93	0.97	0.97	0.95
4	0.93	0.90	0.86	0.80	0.78	0.82	0.83	0.87	0.91	0.92	0.95	0.93
5	0.92	0.88	0.82	0.79	0.76	0.80	0.75	0.79	0.86	0.90	0.91	0.90
6	0.89	0.87	0.82	0.71	0.69	0.79	0.77	0.74	0.79	0.83	0.89	0.93
7	0.89	0.87	0.81	0.68	0.64	0.73	0.72	0.77	0.69	0.75	0.83	0.89
8	0.88	0.86	0.83	0.74	0.69	0.71	0.74	0.68	0.74	0.70	0.77	0.80
9	0.85	0.85	0.84	0.70	0.59	0.70	0.72	0.69	0.76	0.72	0.70	0.78
10	0.81	0.81	0.79	0.68	0.66	0.66	0.71	0.72	0.74	0.71	0.74	0.69
11	0.79	0.81	0.82	0.59	0.52	0.62	0.69	0.70	0.74	0.72	0.71	0.74
12	0.80	0.79	0.84	0.64	0.55	0.56	0.63	0.62	0.71	0.71	0.73	0.69
13	0.68	0.76	0.77	0.69	0.55	0.50	0.60	0.67	0.68	0.67	0.68	0.68
14	0.68	0.74	0.80	0.69	0.58	0.50	0.53	0.59	0.64	0.63	0.73	0.65
15	0.66	0.67	0.80	0.60	0.53	0.63	0.55	0.57	0.63	0.63	0.64	0.64
16	0.66	0.67	0.67	0.62	0.60	0.52	0.53	0.53	0.57	0.65	0.65	0.61
17	0.64	0.69	0.67	0.74	0.63	0.37	0.39	0.54	0.59	0.65	0.66	0.70
18	0.68	0.66	0.66	0.66	0.56	0.40	0.35	0.41	0.56	0.61	0.58	0.70
19	0.65	0.67	0.63	0.61	0.53	0.52	0.39	0.36	0.42	0.60	0.58	0.61
20	0.73	0.68	0.67	0.60	0.61	0.40	0.34	0.30	0.34	0.36	0.58	0.61
21	0.64	0.57	0.56	0.56	0.46	0.49	0.33	0.34	0.34	0.33	0.35	0.57
22	0.53	0.58	0.54	0.51	0.46	0.40	0.36	0.30	0.32	0.30	0.34	0.36
23	0.43	0.53	0.58	0.48	0.38	0.38	0.39	0.32	0.37	0.36	0.35	0.29

Table S2. Correlation skill - CNN

	Target season											
Forecast lead	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF
1	0.96	0.95	0.91	0.89	0.92	0.88	0.93	0.94	0.97	0.96	0.96	0.96
2	0.94	0.94	0.90	0.85	0.84	0.84	0.91	0.93	0.97	0.96	0.95	0.95
3	0.92	0.91	0.89	0.79	0.75	0.77	0.83	0.90	0.93	0.95	0.95	0.94
4	0.90	0.89	0.83	0.75	0.68	0.68	0.75	0.81	0.90	0.92	0.95	0.93
5	0.91	0.86	0.82	0.71	0.63	0.66	0.68	0.74	0.81	0.91	0.92	0.92
6	0.90	0.88	0.83	0.69	0.55	0.60	0.68	0.65	0.74	0.83	0.90	0.90
7	0.93	0.88	0.84	0.73	0.56	0.58	0.59	0.65	0.65	0.76	0.82	0.87
8	0.88	0.90	0.86	0.72	0.60	0.61	0.65	0.66	0.69	0.66	0.73	0.80
9	0.82	0.85	0.84	0.73	0.60	0.66	0.63	0.71	0.65	0.68	0.66	0.74
10	0.80	0.80	0.80	0.65	0.60	0.60	0.66	0.68	0.72	0.67	0.67	0.68
11	0.73	0.76	0.79	0.62	0.50	0.59	0.59	0.64	0.70	0.69	0.68	0.70
12	0.77	0.79	0.79	0.61	0.40	0.56	0.58	0.62	0.66	0.68	0.70	0.65
13	0.73	0.75	0.76	0.68	0.41	0.40	0.56	0.62	0.69	0.71	0.64	0.65
14	0.66	0.65	0.73	0.65	0.44	0.43	0.48	0.55	0.64	0.67	0.66	0.61
15	0.63	0.59	0.70	0.56	0.44	0.40	0.46	0.47	0.58	0.66	0.67	0.66
16	0.65	0.60	0.63	0.53	0.33	0.30	0.40	0.48	0.51	0.62	0.68	0.63
17	0.66	0.64	0.55	0.55	0.41	0.29	0.27	0.41	0.48	0.50	0.65	0.67
18	0.64	0.69	0.56	0.52	0.37	0.26	0.22	0.24	0.37	0.48	0.54	0.63
19	0.60	0.66	0.63	0.52	0.27	0.29	0.22	0.20	0.27	0.41	0.45	0.57
20	0.49	0.60	0.63	0.47	0.37	0.26	0.22	0.21	0.22	0.23	0.42	0.45
21	0.39	0.48	0.55	0.50	0.31	0.24	0.23	0.24	0.25	0.23	0.27	0.43
22	0.38	0.38	0.50	0.45	0.29	0.28	0.24	0.17	0.26	0.26	0.23	0.33
23	0.28	0.38	0.40	0.44	0.27	0.17	0.25	0.27	0.28	0.22	0.25	0.24

Table S3. Correlation skill - SINTEX-F

	Target season											
Forecast lead	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF
1	0.95	0.95	0.93	0.80	0.71	0.73	0.87	0.91	0.93	0.96	0.96	0.97
2	0.94	0.93	0.88	0.71	0.71	0.78	0.87	0.92	0.93	0.95	0.96	0.96
3	0.93	0.90	0.82	0.70	0.61	0.73	0.79	0.87	0.91	0.92	0.94	0.93
4	0.91	0.88	0.80	0.59	0.56	0.65	0.75	0.80	0.87	0.92	0.92	0.93
5	0.92	0.87	0.81	0.58	0.40	0.62	0.69	0.74	0.78	0.88	0.93	0.91
6	0.91	0.89	0.80	0.60	0.36	0.42	0.71	0.66	0.73	0.78	0.89	0.92
7	0.92	0.88	0.81	0.61	0.40	0.34	0.54	0.71	0.61	0.72	0.79	0.89
8	0.90	0.88	0.79	0.57	0.41	0.38	0.44	0.60	0.69	0.59	0.72	0.79
9	0.83	0.86	0.78	0.54	0.34	0.39	0.47	0.50	0.61	0.67	0.60	0.71
10	0.81	0.80	0.76	0.54	0.30	0.29	0.50	0.53	0.51	0.59	0.67	0.61
11	0.68	0.78	0.71	0.53	0.30	0.25	0.36	0.58	0.54	0.50	0.58	0.67
12	0.72	0.63	0.72	0.49	0.30	0.24	0.35	0.43	0.61	0.53	0.52	0.55
13	0.55	0.67	0.55	0.55	0.27	0.23	0.31	0.43	0.45	0.62	0.54	0.55
14	0.63	0.52	0.58	0.44	0.40	0.19	0.32	0.40	0.47	0.44	0.64	0.55
15	0.54	0.62	0.45	0.42	0.30	0.31	0.17	0.42	0.44	0.46	0.47	0.65
16	0.65	0.49	0.57	0.34	0.24	0.24	0.31	0.18	0.47	0.45	0.49	0.51
17	0.54	0.58	0.41	0.40	0.24	0.18	0.23	0.35	0.16	0.47	0.47	0.52
18	0.54	0.55	0.46	0.29	0.20	0.23	0.18	0.25	0.37	0.15	0.48	0.49
19	0.54	0.52	0.51	0.30	0.20	0.11	0.26	0.21	0.24	0.38	0.17	0.48
20	0.47	0.52	0.42	0.43	0.19	0.24	0.11	0.28	0.23	0.22	0.40	0.21
21	0.25	0.43	0.46	0.25	0.33	0.21	0.35	0.14	0.30	0.26	0.24	0.44
22	0.50	0.25	0.38	0.36	0.13	0.30	0.28	0.42	0.16	0.29	0.31	0.26
23	0.24	0.50	0.26	0.34	0.25	0.15	0.32	0.34	0.44	0.18	0.30	0.35