Atmospheric and Oceanic Science Letters 16 (2023) 100350

Contents lists available at ScienceDirect

Atmospheric and Oceanic Science Letters

journal homepage:
http://www.keaipublishing.com/en/journals/atmospheric-and-oceanic-science-letters/

Prediction of ENSO using multivariable deep learning

Yue Chen [a], Xiaomeng Huang [a,*], Jing-Jia Luo [b], Yanluan Lin [a], Jonathon S. Wright [a], Youyu Lu [c], Xingrong Chen [d], Hua Jiang [d], Pengfei Lin [e]

[a] Ministry of Education Key Laboratory for Earth System Modeling and Department for Earth System Science, Tsinghua University, Beijing, China
[b] Institute for Climate and Application Research (ICAR)/CICFEM/KLME/ILCEC, Nanjing University of Information Science and Technology, Nanjing, China
[c] Ocean and Ecosystem Sciences Division, Fisheries and Oceans Canada, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada
[d] National Marine Environmental Forecasting Center, State Oceanic Administration, Beijing, China
[e] State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Keywords:
ENSO forecast
Deep learning
Spring predictability barrier
High-dimensional spatiotemporal predictors
关键词:
ENSO预报
深度学习
春季预报障碍
多维时空预报因子

ABSTRACT

A novel multivariable prediction system based on a deep learning (DL) algorithm, i.e., the residual neural network and pure observations, was developed to improve the prediction of the El Niño–Southern Oscillation (ENSO). Optimal predictors are automatically determined using the maximal information for spatial filtering and the Taylor diagram criteria, enabling the best prediction skills at lead times of eight months compared with most operational prediction models. The hindcast skill for the most challenging decade (2011–18) outperforms the multi-model ensemble operational forecasts. At the six-month lead, the correlation (COEF) skill of the DL model reaches 0.82 with a normalized root-mean-square error (RMSE) of 0.58 °C, which is significantly better than the average multi-model performance (COEF = 0.70 and RMSE = 0.73°C). DL prediction can effectively alleviate the long-standing spring predictability barrier problem. The automatically selected optimal precursors can explain well the typical ENSO evolution driven by both tropical dynamics and extratropical impacts.

摘要

本文基于残差神经网络和观测数据构建了一套深度学习多因子预报测模型, 以改进厄尔尼诺-南方涛动(ENSO)的预报. 该模型基于最大信息系数进行因子时空特征提取, 并根据泰勒图的评估标准可自动确定关键预报因子进行预报. 该模型在超前8个月以内的预报性能要优于当前传统的业务预报模式. 2011–2018年间, 该模型的预报性能优于多模式集成预报的结果. 在超前6个月预报时效上, 模型预报相关性可达0.82, 标准化后的均方根误差仅为0.58°C, 多模式集成预报的相关性和标准化后的均方根误差分别为0.70和0.73°C. 该模型春季预报障碍问题有所缓解, 并且自动选取的关键预报因子可用于解释热带和副热带热动力过程对于ENSO变化的影响.

1. Introduction

The El Niño–Southern Oscillation (ENSO) is the dominant ocean–atmosphere coupled phenomenon on Earth, typically with 2–7-year periodicity (McPhaden et al., 2006; Wang et al., 2017; Capotondi and Sardeshmukh, 2015). Since ENSO variability can induce global environmental changes through local atmospheric responses, ocean circulation, and teleconnections, it may cause significant losses of human lives and property (Kiladis and Diaz, 1989; McPhaden, 2015; Cai et al., 2018; Jiménez-Muñoz et al., 2016). Thus, ENSO forecasting draws significant attention from both scientific and public communities (Latif et al., 1994, 1998; Huang et al., 2017).

Continuous efforts have been made to develop theories to better understand ENSO dynamics, as well as methods and models to forecast ENSO, since the late 1980s (Barnston et al., 2012). Bjerknes (1969) first proposed the positive feedback of an ocean–atmosphere coupled system as the major process for the development of ENSO. Since then, precursors related to tropical dynamics, such as the sea surface temperature (SST), warm water volume (WWV), and surface winds, have long been considered to be useful predictors for the generation of ENSO (Tang et al., 2018; Clarke, 2014; Tseng et al., 2017; Chen et al., 2004; Capotondi and Sardeshmukh, 2015). The recharge-discharge process emphasizes the key role of the upper Pacific Ocean heat content in the phase transition of ENSO (Jin, 1997a, 1997b; Jin and An, 1999). Since 2000, the level of prediction skill using the WWV has reduced, which is possibly related to changes in the coupled system in the tropical Pacific (Horii et al., 2012; McPhaden, 2012; Bunge and Clarke, 2013; Ren et al., 2019). The subsurface forcing on SST in the central-eastern equatorial Pacific is mentioned in Gao et al. (2022). Signals outside the tropical Pacific, such as extratropical atmospheric forcing and the In-

* Corresponding author.
E-mail address: hxm@mail.tsinghua.edu.cn (X. Huang).

https://doi.org/10.1016/j.aosl.2023.100350
Received 2 December 2022; Revised 8 February 2023; Accepted 20 February 2023
Available online 28 February 2023

dian Ocean–Atlantic remote influence, draw more attention (Lim and Hendon, 2017; Luo et al., 2010, 2017; Chen et al., 2020; Chikamoto et al., 2015; Ham et al., 2013). Many different important precursors have been shown to affect the development of ENSO over different time scales, leading to well-known ENSO complexity (Timmermann et al., 2018; Zhang et al., 2022; Chen et al., 2022). Geng et al. (2020) emphasized the role of nonlinear atmospheric convection in the increasing occurrence of strong EP (eastern Pacific) El Niño and strong CP (central Pacific) La Niña. Chen et al. (2022) built a multiscale stochastic model considering the variables over different time scales and succeeded in reproducing the spatiotemporal dynamical evolution of different types of ENSO events. Unfortunately, the individual roles and relative importance of these predictors in the predictability of ENSO remain unclear.

Statistical and dynamic models are two common alternatives used for ENSO forecasting (Clarke, 2014; Goddard et al., 2001). Statistical models use observations to construct a simplified relationship between the predictor and ENSO dynamics based on both linear (such as discriminant analysis, canonical variate analysis, and multiple linear regression) (Cécile, 1996; Mason and Mimmack, 2002; Clarke and Gorder, 2003; Petrova et al., 2017) and nonlinear (such as neurological networks (Hsieh, 2009)) methods. A recent statistical model for ENSO prediction was constructed and successfully forecasted the three-year La Niña event in March 2022 (Fang and Zheng, 2021; Fang et al., 2022). Conventional statistical models evolve slowly and are relatively low in computational cost (Barnston et al., 2012). However, their forecast skills are limited by short historical observations, the complexity of statistical methods, and the current understanding of ENSO dynamics (Tang et al., 2018; McPhaden, 1998). Dynamic models consider the fully nonlinear physical equations governing complicated ocean–atmosphere coupled dynamics (Barnston et al., 2012). These models are computationally more expensive and have been significantly upgraded (Tang et al., 2018; Lai et al., 2018) since the first coupled model was designed (Cane et al., 1986; Zebiak and Cane, 1987). However, the complexity of the ENSO system, systematic model error, missing physical processes, initialization errors, and model drift still exist, thus reducing the predictive skill (Tang et al., 2018; Rashid and Hirst, 2016; Hu and Duan, 2016; Zhu et al., 2017).

More than 20 models support the real-time ENSO forecast on the International Research Institute for Climate and Society (IRI) website (http://iri.columbia.edu/climate/ENSO/currentinfo/update.html). Most of these models can reasonably forecast ENSO at least six months ahead, with the correlation skill reaching approximately 0.65 on average (Barnston et al., 2012). For the hindcast period of 2011–18, the statistical models (not including deep learning (DL) models) have comparable skill to dynamic models, and even show better performances at longer lead times. At a lead of six months, the Pearson correlation coefficient of 0.72 for the ensemble of statistical models is higher than 0.67 for the ensemble of dynamic models (Fig. 1). Both dynamic and statistical models have improved in the last decade. However, accurate prediction of the onset, duration, and magnitude of ENSO events, especially extreme ENSO events, continues to be a challenging task (Jin et al., 2008). Considering that ENSO is a damped mode triggered by stochastic forcing or external noisy signals (Chen et al., 2004; Fedorov and Philander, 2000; Thompson and Battisti, 2010), the limited predictability of the ENSO system at lead times longer than six months may come from the inherent uncertainty of the complex dynamic system, including its spring predictability barrier (SPB) (Larson and Kirtman, 2017; Yu et al., 2009) and the signal-to-noise ratio problem (Rowell, 1998).

2. ENSO prediction based on deep learning

ENSO prediction can be considered an inversion problem by assuming that the associated dynamics are constrained by a large amount of data within the coupled ocean–atmosphere system, both temporally and spatially (Wang et al., 2020; Dominiak and Terray, 2005). Thus, new methods of digging through these data (data mining) may further reveal the hidden dynamics to improve ENSO prediction (Chen et al., 1996; Lee and Siau, 2001; Ham et al., 2019, 2021). Machine learning has unique advantages in data mining owing to its autonomic and powerful self-learning ability in digging useful hidden information out of the raw data (Lecun et al., 2015; Lary et al., 2016; Sejnowski, 2020). DL is a branch of machine learning with a deep neural network as the basic structure (Lecun et al., 2015). Compared with shallow networks, DL tends to better handle complex function approximations and avoid the underfitting problem with lower sample complexity (Mhaskar et al., 2017; Liang and Srikant, 2016), including the challenging ENSO prediction issue (Wang et al., 2020; Yan et al., 2020).

To date, ENSO prediction has normally involved specific predictors (e.g., SST and/or ocean heat content) based on prior knowledge/experience and well-known theory. Such prediction methods may be limited by simplified human choice and judgment without taking full advantage of DL. A recently developed Taylor-ResNet model (Fig. S1, Tables S2–S3) can automatically select the optimal predictors at different lead times (Table S4) and hence is promising for improving ENSO prediction. The model uses the first 80% of samples for training and the rest for testing for all the experiments implemented in this study. It is noteworthy that the testing samples used for prediction are separated from the training samples. Thus, the model could not learn any information of the testing samples before the prediction, as the real forecast. The precursors are superimposed successively when the performance improves according to the evaluation criterion. Thus, the model can determine the optimal combination of precursors being used for different lead times.

The prediction performance based on optimal precursors is shown in Fig. 1. The correlation skill (COEF) of the Taylor-ResNet model remains greater than 0.9 up to a five-month lead with small forecast errors measured by the RMSE (red dotted curves in Fig. 1), which is superior to all the operational forecast models gathered by IRI and the Convolutional Neural Networks (CNN) model (Ham et al., 2019). The hindcast skills of different models are widely spread at leads beyond six months (0.4–0.82 for COEF and 0.6–1 for RMSE). The hindcast skills of all the models decrease quickly at longer lead times (Barnston et al., 2012; Clarke, 2014; Chen et al., 2004), but the Taylor-ResNet model still preserves the highest COEF and the lowest RMSE.

At a lead of six months, the correlation skill of the DL Taylor-ResNet model is 0.82, with an nRMSE (normalized RMSE) of 0.59. These values are superior to the ensemble of dynamic models (COEF = 0.67 and nRMSE = 0.76 for an average of eight models) and the ensemble of statistical models (COEF = 0.72 and nRMSE = 0.69 for an average of six models). The nRMSE is reduced by 22% and 14% compared with the dynamic and statistical ensembles, respectively. Most of the operational models have low skills at leads longer than nine months. For the Taylor-ResNet model, at lead times greater than eight months, its hindcast performance shows a marked decline; however, its COEF is still greater than 0.5. During 2011–18, the ensemble of statistical models does not show a much lower prediction skill than the ensemble of dynamic models. The dynamic models outperform the statistical models at lead times shorter than four months, but the reverse occurs at lead times longer than five months. By averaging the model uncertainties and unpredicted noises, ensemble forecasting may improve the skill compared with using a single model (Tang et al., 2018; Kirtman and Min, 2009; Palmer et al., 2000). Here, the skill of the ensemble based on all models (including eight dynamic and six statistical models from IRI, CNN, and Taylor-ResNet) is mostly determined by the Taylor-ResNet model for lead times up to eight months. For longer lead times, the ensemble based on the Taylor-ResNet model and dynamical models outperforms both the single models and the ensemble of all models.

3. Observational data used

The data/reanalysis that support this study are publicly available online: OISST (Reynolds et al., 2007); NCEP/NCAR Reanalysis 1 project (Kalnay et al., 1996); GODAS (Behringer and Xue, 2004); outputs

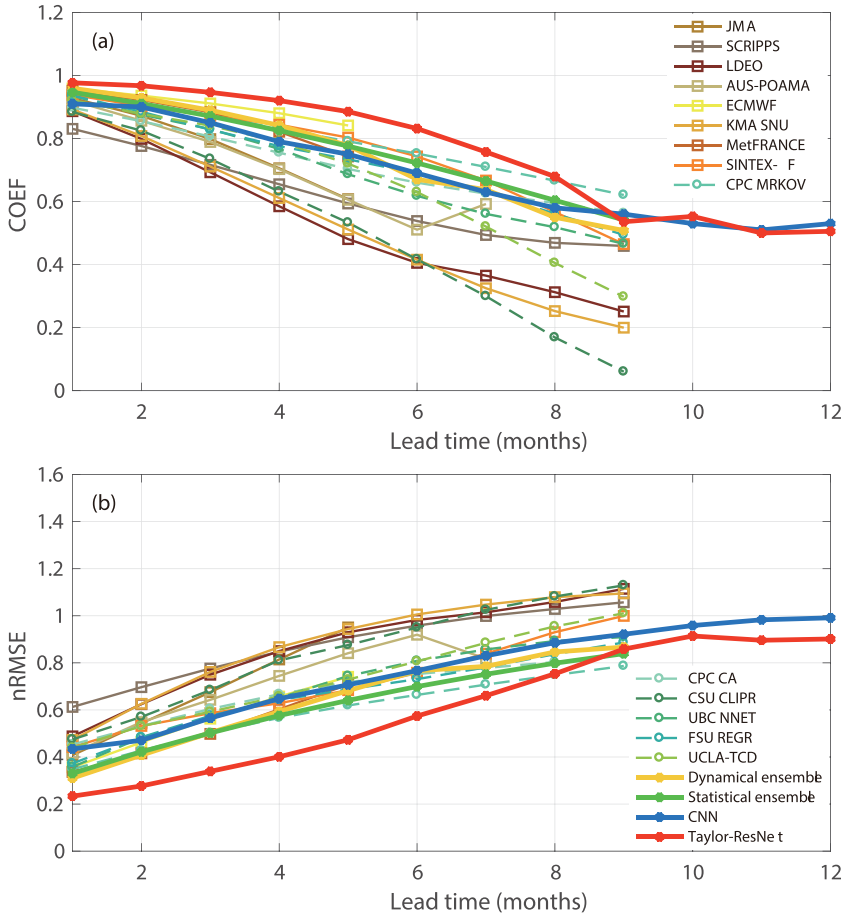

Fig. 1. The (a) correlation skill and (b) nRMSE (RMSE divided by the standard deviations of the observations) of the Niño3.4 index for various model hindcasts relative to observations as a function of lead time (in months) for all target seasons during 2011–18. The yellow and green solid dotted curves denote the ensembles of dynamical models (solid lines with square symbols) and statistical models (dashed lines with circle symbols), respectively. The red solid dotted line is the Taylor-ResNet model, while the blue line comes from the CNN model of Ham et al. (2019).

of operational forecast models, https://iri.columbia.edu/~forecast/ensofcst/Data/; and Oceanic Niño Index from ERSST.v5 (Huang et al., 2017).

4. Optimal precursors for ENSO prediction

The key precursors automatically selected at different lead months are listed in Table S4. The SST is confirmed to be very useful for ENSO prediction throughout all lead months, as expected (Clarke, 2014; Lai et al., 2018). The 10 m zonal wind (UWND) also works effectively at many lead months, followed by latent heat (LH) flux. The sea surface zonal current (UVEL) is more effective with a 4–7-month lead. The sea surface height, as an approximate measure of ocean heat content, is relatively important at lead times within five months, which is somewhat beyond expectation.

The spatial pattern of the maximal information coefficient (MIC, Fig. S2, Reshef et al., 2011) at different lead months provides further information about the individual roles of the key precursors in the evolution of ENSO. The values for MIC range from 0 to 1. Here, a larger value denotes a closer dependence between the precursor and the ENSO event. At short lead times (1–3 months, shown by the first column in Fig. S2), the SST over the equatorial eastern-central Pacific, accompanied by the UWND confined to the equatorial western-central Pacific, shows a close relationship with ENSO (high MIC values in Fig. S2), confirming the dominant role of the tropical SST and UWND in Table S4. This finding is consistent with classical ENSO theories, as basin-scale ocean–atmosphere coupled Bjerknes feedback can predominantly control the ENSO evolution over a short period (Bjerknes, 1969; Jin, 1997a,b; Chen et al., 2015). This equatorial Pacific feedback can spread throughout the tropical Pacific Ocean and affect the global climate, which is the key source for forecasting the global climate (Chen et al., 2020). These processes can be captured well by both dynamic models and other machine learning models, leading to skillful predictions at short lead times (Ham et al., 2019; Yan et al., 2020; Pal et al., 2020).

The prediction skill is gradually decreased at lead times beyond four months. The coupled relationship between SST and UWND during ENSO development weakens (confirmed by the MIC map, shown by the second column in Fig. S2), indicating that the coupled feedback is only accurately captured up to a three-month lead time. The local initialization of Bjerknes feedback starts from the central equatorial Pacific, consistent with the large MIC value for local SST at a 4–6-month lead in combination with the UWND signal to the west (Clarke, 2014; Bunge and Clarke, 2013; Yu et al., 2010). Wind anomalies in the western and central equatorial Pacific tend to generate eastward-propagating Kelvin waves, providing an important source of WWV accumulation in the tropical Pacific (Schopf and Suarez, 1988; Suarez and Schopf, 1988). Correspondingly, UVEL along the equatorial Pacific plays an important role in the development of ENSO through the zonal advection process after SST anomalies have occurred (Yu et al., 2010; Kug et al., 2009), which confirms the key role of the zonal transport of WWV in ENSO development (Tseng et al., 2017). The LH flux in the northwestern Pacific Ocean also plays a role and shows better correlation skill than other predictors when the prediction starts in March and April (Fig. S3).

As expected, the impacts of extratropical precursors tend to play more important roles when the lead time is longer than six months (Chen et al., 2020; Yu et al., 2010). Simultaneously, the tropical contribution decreases at a 7–9-month lead (shown by the third column in Fig. S2), consistent with the seasonal phase change for the WWV and SST (Chen et al., 2020). The evolution of extratropical precursors is consistent with the large MIC in the subtropical eastern Pacific for LH flux, SST, and UWND from lead times of 7–9 months. The North Pacific Oscillation–like wind fields induce surface heat flux changes, contributing to warm subtropical SST anomaly footprinting (i.e., the Pacific Meridional Mode (Chang et al., 2007)). The warm SST anomaly can be effectively extended to the central equatorial Pacific (Yu et al., 2010; Di Lorenzo et al., 2015) through positive wind–evaporation–SST feedback, i.e., commonly leading ENSO events by three seasons (Yu et al., 2010; Chiang and Vimont, 2004; Yu and Kim, 2011; Ding et al., 2015). These features confirm the involvement of SST, LH, and UWND in improving ENSO prediction after 2000.

At 10–12-month leads, the MIC suggests that the key SST precursor is likely confined to the central equatorial Pacific and western North Pacific (shown in the last column in Fig. S2). The memory of the equatorial upper ocean may affect the evolution of ENSO approximately one year ahead due to the frequent phase change in ENSO in successive years from 1983 to 2018. UWND over the equatorial Pacific and the eastern Indian Ocean (not shown) also plays some role, suggesting the possibility of teleconnection between the two basins. This finding has been addressed in previous studies, as Indo-Pacific equatorial winds can modulate the formation of ENSO events in the following year through the Walker Circulation, potentially overcoming the SPB problem (Clarke and Gorder, 2003; Izumo et al., 2010; Gutzler and Harrison, 1987). In the mid–high latitudes of the Pacific Ocean, the dominant MIC areas for UWND and LH are confined to north of approximately 20°N. The subtropical atmospheric impact on the tropical ocean through the Victoria Mode may be denoted here (Ding et al., 2015). Correspondingly, the superior correlation skill of UWND and LH for the boreal spring at lead times longer than 10 months in Fig. S3 may also confirm the potential role in alleviating the SPB problem. For the latter half of the year, UVEL along the equator and the northeastern Pacific Ocean perform better at long lead times, as revealed by the larger correlation skill shown in Fig. S3.

Since ENSO mainly matures during boreal winter, the models' correlation skills as a function of the target season can provide more seasonal skill information (Fig. S4). All models show degraded forecast skill for boreal summer even at very short lead times, which extends to later seasons for longer lead times, as Barnston et al. (2012) mentioned. For the prediction period of 2011–18, the Taylor-ResNet model shows better correlation skill than all the other models when the prediction starts between March and August. Notably, both the dynamic and statistical model ensembles are generally better than any single model (Tang et al., 2018; Kirtman and Min, 2009). However, the dynamic model SINTEX-F tends to perform better when the hindcast starts in March, and the statistical model CPC MRKOV shows better skill when the hindcast starts between April and May for lead times longer than six months. The correlation skill pattern for the CNN model is similar to that of the statistical model ensemble. The Taylor-ResNet model is comparable with the dynamic and statistical ensembles and even shows better performance at longer lead times. This result indicates the potential of our model to capture the physical processes of the ENSO system. Moreover, the Taylor-ResNet model surpasses the other models in alleviating the SPB problem. In particular, the predictions starting from boreal spring and summer at medium to long lead times (three to eight months) show better performances.

5. Conclusions

DL can effectively extract hidden information from a vast volume of input data. In this study, a new prediction system was proposed for ENSO prediction, which blends a DL technique with multidimensional variables from multisource observations. The system can automatically determine the optimal precursors without human intervention by introducing the evaluation criterion of the Taylor diagram. In view of the combination of optimal precursors, the prediction performance of our system surpasses those from the operational prediction models archived by IRI. For the prediction period of 2011–18, the correlation skill of ENSO prediction with DL reaches 0.94, 0.82, 0.58, and 0.54 at the 3-, 6-, 9-, and 12-month leads, respectively. The standard RMSEs at the four lead times are 0.35, 0.58, 0.85, and 0.96. Moreover, the SPB problem is greatly alleviated in our system.

However, only a few essential optimal precursors can improve the prediction performance. Additionally, long-range prediction may be a major bottleneck of machine learning (Yan et al., 2020; Mu et al., 2019). With an optimal combination of precursors using DL, despite short to medium lead times, the hindcast skill can be relatively good (Fig. 1), at long leads, the prediction performance is still not high enough. This finding applies for both the CNN and the Taylor-ResNet model, in particular for the latest decade (2011–18), when ENSO prediction becomes much more difficult. Of course, another reason may be related to the short duration for the precursors to be validated and to represent ENSO variability.

This study demonstrates that the combination of DL and geoscience big data helps deepen our understanding of the Earth system. ENSO prediction is a multidimensional integration of time, space, and physical variables while the present study carried out qualitative analyses of different parameters, we are not yet able to give a full dynamic interpretation of the constructed model. This data-driven discovery process may also be limited by the length of available observational data. Furthermore, a better understanding of ENSO dynamics will lead to improved designs of model algorithms. Although the method used in this study mainly focused on hindcasting the Niño3.4 index, in the future, different types of Niño index calculated over different regions can also be learnt, which could verify the predictability limits for different types of ENSO events, as discussed in Fang et al. (2022). Thus, ENSO complexity can be further explored.

Acknowledgments

This work is supported by the National Natural Science Foundation of China [grant Nos. 42125503 and 42075137] and the National Key Research and Development Program of China [grant Nos. 2020YFA0608000 and 2020YFA0607900].

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.aosl.2023.100350.

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? Bull. Am. Meteorol. Soc. 93 (5), 631–651.

Behringer, D. W., Xue, Y., 2004. Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. Eighth Symposium on Integrated Observing and Assimilation System for Atmosphere, Ocean, and Land Surface. AMS 84th Annual Meeting, Washington State Convention and Trade Center, Seattle, Washington, 11-15 January 2004.

Bjerknes, J., 1969. Atmospheric teleconnections from the equatorial Pacific. Mon. Wea. Rev. 97, 163–172.

Bunge, L., Clarke, A.J., 2013. On the warm water volume and its changing relationship with ENSO. J. Phys. Oceanogr. 44 (5), 1372–1385.

Cai, W.J., Wang, G.J., Dewitte, B., Wu, L.X., Santoso, A., Takahashi, K., Yang, Y., Carréric, A., McPhaden, M.J., 2018. Increased variability of eastern Pacific El Niño under greenhouse warming. Nature 564 (7735), 201–206.

Cane, M.A., Zebiak, S.E., Dolan, S.C., 1986. Experimental forecasts of El Niño. Nature 321, 827–832.

Capotondi, A., Sardeshmukh, P.D., 2015. Optimal precursors of different types of ENSO events. Geophys. Res. Lett. 42, 9952–9960.

Cécile, P., 1996. A stochastic model of Indo-Pacific sea surface temperature anomalies. Phys. D Nonlinear Phenom. 98, 534–558.
Chang, P., Zhang, L., Saravanan, R., Vimont, D.J., Chiang, J.C.H., Ji, L., Seidel, H., Tippett, M.K., 2007. Pacific meridional mode and El Niño-Southern Oscillation. Geophys. Res. Lett. 34, L16608.
Chen, D., Cane, M.A., Kaplan, A., Zebiak, S.E., Huang, D., 2004. Predictability of El Niño over the past 148 years. Nature 428 (6984), 733–736.
Chen, D., Lian, T., Fu, C., Cane, M.A., Tang, Y., Murtugudde, R., Song, X., Wu, Q., Zhou, L., 2015. Strong influence of westerly wind bursts on El Niño diversity. Nat. Geosci. 8 (5), 339–345.
Chen, H.C., Tseng, Y.H., Hu, Z.Z., Ding, R.Q., 2020. Enhancing the ENSO predictability beyond the spring barrier. Sci. Rep. 10, 984.
Chen, N., Fang, X., Yu, J.Y., 2022. A multiscale model for El Niño complexity. npj Clim. Atmos. Sci. 5, 1–13.
Chen, M.-S., Han, J.W., Yu, P. S., 1996. Data mining: An overview from a database perspective. IEEE Trans. Knowl. Data Eng. 8 (6), 866–883.
Chiang, J.C.H., Vimont, D.J., 2004. Analogous pacific and Atlantic meridional modes of tropical atmosphere–ocean variability. J. Clim. 17 (21), 4143–4158.
Chikamoto, Y., Timmermann, A., Luo, J.-J., Mochizuki, T., Kimoto, M., Watanabe, M., Ishii, M., Xie, S.-P., Jin, F.-F., 2015. Skillful multi-year predictions of tropical trans-basin climate variability. Nat. Commun. 6, 6869.
Clarke, A.J., 2014. El Niño physics and El Niño predictability. Annu. Rev. Mar. Sci. 6 (1), 79–99.
Clarke, A.J., Gorder, S.V., 2003. Improving El Niño prediction using a space-time integration of Indo-pacific winds and equatorial pacific upper ocean heat content. Geophys. Res. Lett. 30 (30), 325–348.
Di Lorenzo, E., Liguori, G., Schneider, N., Furtado, J.C., Anderson, B.T., Alexander, M.A., 2015. ENSO and meridional modes: a null hypothesis for Pacific climate variability. Geophys. Res. Lett. 42, 9440–9448.
Ding, R., Li, J., Tseng, Y.-H., Sun, C., Guo, Y., 2015. The Victoria mode in the North Pacific linking extratropical sea level pressure variations to ENSO. J. Geophys. Res. Atmos. 120, 27–45.
Dominiak, S., Terray, P., 2005. Improvement of ENSO prediction using a linear regression model with a southern Indian Ocean sea surface temperature predictor. Geophys. Res. Lett. 32, L18702.
Fang, X., Zheng, F., 2021. Effect of the air–sea coupled system change on the ENSO evolution from boreal spring. Clim. Dyn. 57, 109–120.
Fang, X., Zheng, F., Li, K., Hu, Z.-Z., Ren, H., Wu, J., Chen, X., Lan, W., Yuan, Y., Feng, L., 2022. Will the historic southeasterly wind over the Equatorial Pacific in March 2022 trigger a Third-year La Niña Event? Adv. Atmos. Sci. 40 (1), 6–13.
Fedorov, A.V., Philander, S.G., 2000. Is El Niño changing? Science 288 (5473), 1997–2002.
Gao, C., Chen, M.N., Zhou, L., Feng, L.C., Zhang, R.-H., 2022. The 2020-21 prolonged La Niña evolution in the tropical Pacific. Sci. China Earth Sci. 65 (12), 2248–2266.
Geng, T., Cai, W., Wu, L., 2020. Two types of ENSO varying in tandem facilitated by nonlinear atmospheric convection. Geophys. Res. Lett. 47, e2020GL088784.
Goddard, L., Mason, S., Zebiak, S., Ropelewski, C., Basher, R., Cane, M., 2001. Current approaches to seasonal-to-interannual climate predictions. Int. J. Climatol. 21, 1111–1152.
Gutzler, D.S., Harrison, D.E., 1987. The structure and evolution of seasonal wind anomalies over the near-equatorial eastern Indian and western Pacific oceans. Mon. Wea. Rev. 115, 169–192.
Ham, Y.G., Kim, J.H., Kim, E.S., On, K.W., 2021. Unified deep learning model for El Niño/Southern Oscillation forecasts by incorporating seasonality in climate data. Sci. Bull. 66 (13), 1358–1366.
Ham, Y.G., Kim, J.H., Luo, J.J., 2019. Deep learning for multi-year ENSO forecasts. Nature 573, 568–572.
Ham, Y.G., Kug, J.S., Park, J.Y., Jin, F.F., 2013. Sea surface temperature in the north tropical Atlantic as a trigger for El Niño/Southern Oscillation events. Nat. Geosci. 6, 112–116.
Horii, T., Ueki, I., Hanawa, K., 2012. Breakdown of ENSO predictors in the 2000s: decadal changes of recharge/discharge-SST phase relation and atmospheric intraseasonal forcing. Geophys. Res. Lett. 39, L10707.
Hsieh, W.W., 2009. Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels. Cambridge University Press, Cambridge, pp. 157–169.
Hu, J., Duan, W., 2016. Relationship between optimal precursory disturbances and optimally growing initial errors associated with ENSO events: Implications to target observations for ENSO prediction. J. Geophys. Res. Oceans 121, 2901–2917.
Huang, B., Shin, C.-S., Shukla, J., Marx, L., Balmaseda, M.A., Halder, S., Kinter, J.L., 2017. Reforecasting the ENSO events in the past fifty-seven years (1958-2014). J. Clim. 30 (19), 7669–7693.
Izumo, T., Vialard, J., Lengaigne, M., Boyer Montegut, C., Behera, S.K., Luo, J.-J., Cravatte, S., Masson, S., Yamagata, T., 2010. Influence of the state of the Indian Ocean Dipole on the following year's El Niño. Nat. Geosci. 3, 168–172.
Jiménez-Muñoz, J.C., Mattar, C., Barichivich, J., Santamaría-Artigas, A., Takahashi, K., Malhi, Y., Sobrino, J.A., van der Schrier, G., 2016. Record-breaking warming and extreme drought in the Amazon rainforest during the course of El Niño 2015–2016. Sci. Rep. 6, 33130.
Jin, E.K., Kinter, J.L., Wang, B., Park, C.-K., Kang, I.-S., Kirtman, B.P., Kug, J.-S., 2008. Current status of ENSO prediction skill in coupled ocean–atmosphere models. Clim. Dyn. 31, 647–664.
Jin, F.F., 1997a. An equatorial ocean recharge paradigm for ENSO. Part I: conceptual model. J. Atmos. Sci. 54, 811–829.
Jin, F.F., 1997b. An equatorial recharge paradigm for ENSO. II: A stripped-down coupled model. J. Atmos. Sci. 54, 830–8847.
Jin, F.F., An, S.I., 1999. Thermocline and zonal advective feedbacks within the equatorial ocean recharge oscillator model for ENSO. Geophys. Res. Lett. 26, 2989–2992.
Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., et al., 1996. The NCEP/NCAR 40-year reanalysis project. Bull. Am. Meteorol. Soc. 77 (3), 437–472.
Kiladis, G.N., Diaz, H.F., 1989. Global climatic anomalies associated with extremes in the southern oscillation. J. Clim. 2 (9), 1069–1090.
Kirtman, B.P., Min, D., 2009. Multimodel ensemble ENSO prediction with CCSM and CFS. Mon. Wea. Rev. 137 (9), 2908–2930.
Kug, J.S., Jin, F.F., An, S.I., 2009. Two types of El Niño events: cold tongue El Niño and warm pool El Niño. J. Clim. 22, 1499–1515.
Lai, A.W.C., Herzog, M., Grad, H.F., 2018. ENSO forecasts near the spring predictability barrier and possible reasons for the recently reduced predictability. Bull. Am. Meteorl. Soc. 815–838.
Larson, S.M., Kirtman, B.P., 2017. Drivers of coupled model ENSO error dynamics and the spring predictability barrier. Clim. Dyn. 48 (11-12), 1–14.
Lary, D.J., Alavi, A.H., Gandomi, A.H., Walker, A.L., 2016. Machine learning in geosciences and remote sensing. Geosci. Front. 7 (1), 3–10.
Latif, M., Anderson, D., Barnett, T.P., Cane, M.A., Kleeman, R., Leetmaa, A., 1998. A review of the predictability and prediction of ENSO. J. Geophys. Res. 103 (C7), 14375.
Latif, M., Barnett, T.P., Cane, M.A., Flügel, M., Zebiak, S.E., 1994. A review of ENSO prediction studies. Clim. Dyn. 9 (4), 167–179.
Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436.
Lee, S.J., Siau, K., 2001. A review of data mining techniques. Ind. Manag. Data Syst. 101 (1), 41–46.
Liang, S.Y., and Srikant, R., 2016. Why deep neural networks? arXiv preprint. arXiv:1610.04161.
Lim, E.-P., Hendon, H.H., 2017. Causes and predictability of the negative Indian Ocean Dipole and its impact on La Niña during 2016. Sci. Rep. 7, 12619.
Luo, J.J., Liu, G., Hendon, H., Alves, O., Yamagata, T., 2017. Inter-basin sources for two-year predictability of the multi-year La Niña event in 2010–2012. Sci. Rep. 7, 2276.
Luo, J.-J., Zhang, R., Behera, S., Masumoto, Y., Jin, F.-F., Lukas, R., Yamagata, T., 2010. Interaction between El Niño and extreme Indian Ocean Dipole. J. Clim. 23 (3), 726–742.
Mason, S.J., Mimmack, G.M., 2002. Comparison of some statistical methods of probabilistic forecasting of ENSO. J. Clim. 15 (15), 8–29.
McPhaden, M.J., 1998. The Tropical Ocean–Global Atmosphere observing system: A decade of progress. J. Geophys. Res. 103(C7), 142169–142240.
McPhaden, M.J., 2012. A 21st century shift in the relationship between ENSO SST and warm water volume anomalies. Geophys. Res. Lett. 39, L09706.
McPhaden, M.J., 2015. Commentary: Playing hide and seek with El Niño. Nat. Clim. Change 5 (9), 791–795.
McPhaden, M.J., Zebiak, S.E., Glantz, M.H., 2006. ENSO as an integrating concept in earth science. Science 314 (5806), 1740–1745.
Mhaskar, H., Liao, Q., Poggio, T.A., 2017. When and why are deep networks better than shallow ones? In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence 31 (1), 2343–2349.
Mu, B., Peng, C., Yuan, S., Chen, L., 2019. ENSO forecasting over multiple time horizons using ConvLSTM network and rolling mechanism. In: Proceedings of the International Joint Conference on Neural Networks (IJCNN), pp. 1-8.
Pal, M., Maity, R., Ratnam, J.V., Nonaka, M., Behera, S.K., 2020. Long-lead prediction of ENSO Modoki index using machine learning algorithms. Sci. Rep. 10 (1), 365.
Palmer, T.N., Brankovic, C., Richardson, D.S., 2000. A probability and decision-model analysis of PROVOST seasonal multimodel ensemble integrations. Q. J. R. Meteorol. Soc. 126, 2013–2034.
Petrova, D., Koopman, S.J., Ballester, J., Rodó, X., 2017. Improving the long-lead predictability of El Niño using a novel forecasting scheme based on a dynamic components model. Clim. Dyn. 48 (3–4), 1249–1276.
Rashid, H.A., Hirst, A.C., 2016. Investigating the mechanisms of seasonal ENSO phase locking bias in the ACCESS Coupled Model. Clim. Dyn. 46, 1075–1090.
Ren, H.L., Zuo, J., Deng, Y., 2019. Statistical predictability of Niño indices for two types of ENSO. Clim. Dyn. 52, 5361–5382.
Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., Mcvean, G., Turnbaugh, P. J., Lander, E. S., Mitzenmacher, M., Sabeti, P. C., 2011. Detecting Novel Associations in Large Data Sets. Science 334, 1518–1524.
Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K., Schlax, M.G., 2007. Daily high-resolution-blended analyses for sea surface temperature. J. Clim. 20, 5473–5496.
Rowell, D., 1998. Assessing potential seasonal predictability with an ensemble of multi--decadal GCM simulations. J. Clim. 11, 109–120.
Schopf, P.S., Suarez, M.J., 1988. Vacillations in a coupled ocean–atmosphere model. J. Atmos. Sci. 45, 549–566.
Sejnowski, T.J., 2020. The unreasonable effectiveness of deep learning in artificial intelligence. Proc. Natl. Acad. Sci. U. S. A. 117 (48), 30033–30038.
Suarez, M.J., Schopf, P.S., 1988. A delayed action oscillator for ENSO. J. Atmos. Sci. 45, 3283–3287.
Tang, Y.M., Zhang, R.-H., Liu, T., Duan, W.S., Yang, D.J., Zheng, F., Ren, H.L., et al., 2018. Progress in ENSO prediction and predictability study. Natl. Sci. Rev. 5 (06), 48–61.
Timmermann, A., An, S.I., Kug, J.S., Jin, F.F., Zhang, X., 2018. El Niño–Southern oscillation complexity. Nature 559 (7715), 535–545.
Thompson, C.J., Battisti, D.S., 2010. A linear stochastic dynamical model of ENSO. part ii: analysis. J. Clim. 14 (4), 445–466.
Tseng, Y.H., Hu, Z.H., Ding, R.Q., Chen, H., 2017. An ENSO prediction approach based on ocean conditions and ocean–atmosphere coupling. Clim. Dyn. 48, 2025–2044.
Wang, C.Z., Clara, D., Yu, J.Y., DiNezio, P., Clement, A., 2017. El Niño and Southern Oscillation (ENSO): A review. Coral Reefs of the Eastern Pacific. Springer, Dordrecht, pp. 85–106.

Wang, X., Slawinska, J., Giannakis, D., 2020. Extended-range statistical ENSO prediction through operator-theoretic techniques for nonlinear dynamics. Sci. Rep. 10, 2636.
Yan, J., Mu, L., Wang, L., Ranjan, R., Zomaya, A.Y., 2020. Temporal convolutional networks for the advance prediction of ENSO. Sci. Rep. 10 (1), 8055.
Yu, J.Y., Kao, H.Y., Lee, T., 2010. Subtropics-related interannual sea surface temperature variability in the Central Equatorial Pacific. J. Clim. 23, 2869–2884.
Yu, J.Y., Kim, S.T., 2011. Relationships between extratropical sea level pressure variations and the central pacific and eastern pacific types of ENSO. J. Clim. 24 (3), 708–720.
Yu, Y.S., Duan, W.S., Xu, H., Mu, M., 2009. Dynamics of nonlinear error growth and season-dependent predictability of El Nino events in the Zebiak–Cane model. Q. J. R. Meteorol. Soc. 135, 2146–2160.
Zebiak, S.E., Cane, M.A., 1987. A model El Niño-southern oscillation. Mon. Wea. Rev. 115, 2262–2278.
Zhang, R.H., Gao, C., Feng, L.C., 2022. Recent ENSO evolution and its real-time prediction challenges. Natl. Sci. Rev. 9 (4), nwac052.
Zhu, J., Kumar, A., Wang, W., Hu, Z., Huang, B., Balmaseda, M., 2017. Importance of convective parameterization in ENSO predictions. Geophys. Res. Lett. 44, 6334–6342.