MS-No.:npg-2024-24

Title: Explaining the high skill of Reservoir Computing method in El Niño prediction

Authors: Francesco Guardamagna, Claudia E. Wieners and Henk A. Dijkstra

Point-by-point reply to reviewer #3

February 9, 2025

We thank the reviewer for their careful reading and for the useful comments on the manuscript.

Overview

This manuscript investigates the prediction skill of a specific type of Recurrent Neural Network, known as Reservoir Computer (RC), in relation to ENSO forecasting. It finds that error propagation in RC is lessened compared to the Zebiak-Cane (ZC) model. While the RC demonstrates high prediction skill (e.g., an ACC greater than 0.6 at 18 months lead time), I believe this manuscript is not suitable for publication for several reasons:

Major comments:

1. *The predictions are not based on real-world data. Both the training and testing datasets are generated from the ZC model, which does not reflect actual observations. It is unclear how well RC performs when predicting realworld events, such as the ENSO events of 2014-2015.*

   Author's reply:

   First of all, the RC model's effectiveness in predicting real ENSO events has already been demonstrated in previous studies [2]. Second, we can easily clarify why we only work with data from the Zebiak and Cane (ZC) model rather than real-world observations. The objective of our study is to demonstrate that the Reservoir Computer (RC) can mitigate error propagation resulting from initial conditions perturbations more effectively than a classical dynamical numerical model. This is proposed as a potential explanation for the RC model's high performances in ENSO forecasting and its ability to overcome the Spring

Predictability Barrier problem, which was previously quantified in the ZC model in terms of sensitivity to initial conditions perturbations [3]. Such an analysis and comparison is simply impossible using real-world observations as we do not know the evolution operator of the real-world system and hence cannot determine the CNOP. By focusing on the ZC model data, we analyze the RC model's behavior and learned dynamics within a controlled environment.

Changes in manuscript:
No changes in the manuscript needed.

2. *The prediction accuracy of RC is very similar to that achieved by linear regression (LR, as shown in Fig. 2). First, error bars should be included for the LR results. Second, the performance of LR is comparable to that of RC, particularly as indicated by the proximity of the red and blue lines at lead times of 1-9 months.*

   Author's reply:
   In Fig. 2, the yellow and red lines represent the performance of the LR with and without surface wind speed anomalies (τ_c) included during training, respectively. Similarly, the blue and green lines correspond to the performance of the RC with and without τ_c included during training, respectively. To ensure a fair comparison, the yellow line should be compared with the blue line (LR vs. RC with τ_c included), and the red line with the green line (LR vs. RC without τ_c included).

   While we acknowledge that the RC does not drastically outperform the LR, our results demonstrate a clear advantage in adopting the RC, as its ability to capture nonlinear relationships between input variables, made possible by the use of a nonlinear activation function (the hyperbolic tangent in our study), leads to a consistent performance improvement, particularly in the supercritical regime, where nonlinearities play a more prominent role. This is further supported by the fact that in this regime, model performance improves when τ_c is included during training, highlighting the importance of the nonlinear effects introduced by this variable [1]. These effects are better captured by the RC, whereas the LR can only provide a linear approximation.

   It is impossible to show error bars for the LR model because, unlike the RC, the LR does not rely on random weights initialization. The LR will consistently produce the same results for a given training set, so given a specific training set, there is no variability in the LR outputs.

Changes in manuscript:
We will better describe the difference between the RC and LR performances in the "RC performances" section.

3. *The influence of wind stress in RC is inconsistent. In some instances, incorporating wind stress enhances ENSO predictions, while in others, it does not. This inconsistency undermines the conclusions drawn, as it does not provide clear insights for real-world predictions, particularly regarding whether ENSO is damped or self-exciting in actual observations.*

Author's reply:
We appreciate this critical comment of the reviewer, but our results are actually consistent and show a clear pattern.

In the supercritical regime, the RC consistently performs better across all lead times when τ_c is included during training. This highlights the importance of the nonlinear effects introduced by this variable, which the RC can efficiently capture through the use of a nonlinear activation function. In the subcritical regime, the Reservoir Computer (RC) achieves higher accuracy at shorter lead times (3–6 months) when τ_c is included, while at longer lead times (9–18 months), performance improves when τ_c is excluded. This is because τ_c plays a crucial role in capturing short-term variability, providing valuable information about the external stochastic forcing that drives the early perturbations dynamics. At longer lead times (9–18 months), improved predictive performance requires the model to rely more on the system's internal dynamics rather than the short-term influence of stochastic noise. Including τ_c during training can lead to overfitting, causing the model to focus excessively on short-term noise patterns instead of learning the internal system dynamics. As a result, model performance deteriorates at extended lead times when τ_c is included. These results clearly show how the inclusion of the variable τ_c influences the RC performances in the subcritical and supercritical regimes.

Drawing conclusions about the true nature of ENSO is not the objective (and far beyond the scope) of this study.

Changes in manuscript:
In the revised manuscript, we will clarify in the "Summary and Discussion" section that the goal of our study is not to draw conclusions about the true nature of ENSO dynamics. Rather, we aim to provide a potential explanation for the RC model's high forecasting performance. We will also better explain the influence of the variable τ_c in the subcritical and supercritical regimes in the "RC performances" section.

4. *The results from the ZC model raise concerns. For instance, in Fig. A2, the Nino3 index only fluctuates between 0.1 and -0.1.*

   Author's reply:
   Fig. A2 only illustrates the response of the deterministic ZC model to a small initial perturbation applied to the seasonal background state in the subcritical ($r_d < 0.8$) and supercritical ($r_d \geq 0.8$) regimes. In the subcritical regime, the perturbation rapidly decays, and without noise, oscillations cannot occur. In contrast, in the supercritical regime, the perturbation evolves into a stable limit cycle with a period of approximately 4 years. Fig. A2 does not show the actual long term behaviour of the ZC model in the presence of noise, which is depicted in Fig. 1.

   Changes in manuscript:
   No changes in the manuscript needed.

References

[1] Wansuo Duan, Yanshan Yu, Hui Xu, and Peng Zhao. Behaviors of nonlinearities modulating the el niño events induced by optimal precursory disturbances. *Climate Dynamics*, 40(5):1399–1413, March 2013.

[2] Forough Hassanibesheli, Jürgen Kurths, and Niklas Boers. Long-term enso prediction with echo-state networks. *Environmental Research: Climate*, 1(1):011002, jul 2022.

[3] Yanshan Yu, Wansuo Duan, Hui Xu, and Mu Mu. Dynamics of nonlinear error growth and season-dependent predictability of el niño events in the zebiak–cane model. *Quarterly Journal of the Royal Meteorological Society*, 135(645):2146–2160, 2009.