**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 #1

February 9, 2025

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

### Overview

*This study focuses on explaining the origins of high forecasting skills of recently emerging deep learning models for ENSO predictions. More specifically, the authors firstly build a skillful Reservoir Computers (RC) model for simulating Zebiak-Cane numerical model, and then investigate and compare the sensitivities on initial perturbations of RC and ZC models (also including an LC model). The authors find RC models are less susceptible to initial perturbations no matter for short and long lead months, which is also the possible reason of weakening the impact of spring predictability barrier (SPB) and extending the effective lead time. In general, this study crafts some novel experiments, and I have the following major questions for the author to answer:*

### Major comments:

1. *$W_{in}$ in Equation (4a) and $W^{in}$ in the following illustration is not consistent.*

   **Author's reply:**
   We thank the reviewer for pointing out this imprecision.

   **Changes in manuscript:**
   We will correct this imprecision in lines 99, 100, and 101 on page 4 of the revised manuscript.

2. *What is the value of $N_x$ in your study, which is quite an important configuration for forecasting skill of RC model.*

   **Author's reply:**

   For both the supercritical and subcritical regimes, as well as for each set of training variables, we determined the optimal hyperparameter sets using a Bayesian search. The search consistently converged on large reservoir dimensions $N_x$, with notable differences between the supercritical and subcritical regimes. In the supercritical regime, the optimal reservoir dimension is approximately 400, regardless of whether zonal surface wind speed anomalies ($\tau_C$) are included in the training variable set. In the subcritical regime, the optimal reservoir dimension is larger, with $N_x = 476$ when $\tau_C$ is included and $N_x = 534$ when $\tau_C$ is excluded.

   **Changes in manuscript:**
   We will specify the $N_x$ values used in our study in the "Training and validation of the RC" section on page 8. Moreover, a summary table reporting the optimal hyperparameter sets for each regime and training variable set will be added to the Appendix.

3. *Around line 169, the authors mention that "the initial reservoir state for each prediction was determined using 5 years of data before the time t ". Can you explain this expression more to those unfamiliar with RC models?*

   **Author's reply:**
   We agree with the reviewer that this expression may not be clear to those unfamiliar with the Reservoir Computer (RC) model. Discarding the initial $x(n)$ reservoir states for $0 \leq n \leq n_{transient}$ before training and evaluation is a standard practice in Reservoir Computing. This step is necessary to mitigate the impact of initial transients caused by the arbitrary initialization of the reservoir state, which is typically set to $x(0) = 0$ or initialized randomly. In our case, the reservoir state was initialized as $x(0) = 0$. This initialization creates an artificial starting state that is unlikely to recur once the reservoir dynamics stabilize. A warmup period is therefore introduced to allow the reservoir to reach a

stable dynamical regime before training or inference. The length of the warmup period depends on the reservoir's memory capacity and the specific learning task. Based on our experiments, a warmup period of 5 years is sufficient to stabilize the reservoir dynamics and eliminate the effects of initial transients. As a result, before inference, the reservoir states corresponding to the 5 years preceding the initial time step $t(n)$ (the starting point of our forecast) are discarded. In our notation, this means discarding $x(n)$ reservoir states for $t(n) - n_{transient} \leq n \leq t(n)$, where $n_{transient} = 180$, given our 10-days time step.

**Changes in manuscript:**
We will better explain why, when working with a RC model, a warm-up period is necessary to properly initialize the Reservoir state in the "Training and validation of the RC" section on page 8.

4. *Around line 224, the authors mention that "This result aligns with expectations, as non-linearities play a more important role in the supercritical regime". I am wondering is it true that the supercritical regime exhibits more nonlinear than subcritical regime, which favors the performance of the RC model? What's the relationship between nonlinearity of regime and RC model performance?*

**Author's reply:**
ENSO can be described by two different theoretical frameworks. According to one perspective, it is a stable (damped) mode sustained primarily by random atmospheric noise (subcritical regime). Alternatively, it can be viewed as a self-sustained oscillatory mode (supercritical regime). In the latter scenario, nonlinearity is essential in modulating ENSO behavior. In the Zebiak and Cane (ZC) model, nonlinearities come from three main sources: heat advection, wind stress anomalies, and subsurface water temperature variations [1].

The Reservoir Computing model can capture complex nonlinear relationships between input variables by employing a nonlinear activation function (in our study, the hyperbolic tangent). In contrast, a Linear Regressor can only estimate linear relationships between input variables. Consequently, the performance gap between the Reservoir

Computing model and the Linear Regressor is expected to be more pronounced in the supercritical regime, where nonlinear effects are more important.

**Changes in manuscript:**
We will provide a clearer explanation of why nonlinearities play a more significant role in the supercritical regime compared to the subcritical regime in the "Zebiak and Cane" model section on page 2. Additionally, in the "Reservoir Computer" section on page 4, we will explain why the Reservoir Computing model is better suited for solving problems involving nonlinear relationships between input variables, compared to a simple Linear Regressor.

5. *Around line 248, do the authors use values of the CNOP objective function to assess whether the model is susceptible?*

   **Author's reply:**

   For both the RC and the ZC model, we used the Cobyla optimization algorithm to maximize the CNOP objective function and estimate the optimal initial perturbations. The resulting maximal error growth, calculated for a specific lead time, was then taken as our measure of the model's sensitivity to initial perturbations.

   **Changes in manuscript:**
   We will better clarify this point in the "CNOPs Analysis" section on page 12.

6. *For figure 5, why the CNOP results for ZC model is quite symmetrical while the CNOP results for RC models are usually biased?*

   **Author's reply:**
   The ZC model's optimal initial perturbations exhibit a notably symmetrical distribution, evident in both SST perturbations at shorter lead times (3 months) and thermocline depth perturbations at longer lead times (6 and 9 months). This symmetry suggests that the model is

sensitive to both negative and positive initial perturbations, depending on the specific event.

In contrast, the RC generally demonstrates greater sensitivity to initial SST perturbations across both shorter and longer lead times, and the distribution of these optimal initial SST perturbations consistently shows a clear preference for either positive or negative values.

To better understand these differences in behavior, we plan to identify the specific initial conditions and times of year when the ZC model shows a preference for negative or positive perturbations. We will then compare the behavior of the RC for the same types of events, focusing on how the optimal perturbations evolve for both models. This additional analysis will provide valuable insights into why the ZC model's optimal initial perturbations exhibit a more symmetrical distribution.

**Changes in manuscript:**
We will add and discuss the results of this additional analysis in the "CNOP analysis" section of the revised manuscript on page 11.

7. *Around line 297, the authors mention that the inclusion or exclusion of wind speed anomalies has a large effect on the different variables of CNOP in the RC models for the subcritical regime. I am wondering why this is not obvious in the RC models for the supercritical regime?*

**Author's reply:**

In the subcritical regime, ENSO variability is primarily sustained by atmospheric noise, introduced as stochastic wind stress forcing. This noise influences the thermocline slope, activating mechanisms that lead to the development of perturbations. When the variable $\tau_c$ is included during training, the RC explicitly learns the relationship between wind anomalies and thermocline adjustments, and the state of the surface winds is provided as an initial condition. Consequently, smaller thermocline perturbations can be amplified by wind anomalies, leading to larger deviations from the reference trajectory. In contrast, when $\tau_c$ is not included, the RC only learns the direct relationship between SST and thermocline depth anomalies without explicit knowledge of how wind anomalies influence thermocline slope adjustments. As a result, in the absence of wind-forcing information, a larger initial thermocline perturbation is required to generate significant error propagation over time.

In the supercritical regime, thermocline depth perturbations are not purely activated by atmospheric noise but also emerge due to the internal instability of the system. The influence of stochastic atmospheric noise is weaker than in the subcritical case. This means that even if the model does not explicitly account for the effect of wind forcing on the thermocline slope, strong initial thermocline depth anomalies are not needed to maximize error propagation.

However, even in the supercritical regime, when $\tau_c$ is not included, the RC remains more sensitive to initial thermocline perturbations at longer lead times than when $\tau_c$ is included, but the difference is much less pronounced than in the subcritical regime.

In every case, the Reservoir Computer is consistently less sensitive to thermocline depth perturbations at longer lead times compared to the Zebiak and Cane model. This suggests that the RC effectively mitigates error propagation from thermocline perturbations.

**Changes in manuscript:**

We will provide a clearer explanation in the "CNOPs Analysis" section on page 11 of why the inclusion of $\tau_c$ has a greater impact on the RC's optimal initial perturbations in the subcritical regime than in the supercritical regime.

8. *I have noticed there is another similar study that revealing the initial perturbations of SST for ENSO predictions in AI model (https://doi.org/10.1002/qj.4882). Maybe this is a more comprehensive way to detecting the detailed patterns and physical variables of initial perturbations related to SPB from index models (such as RC models used in this study) to spatial models.*

**Author's reply:**

We thank the reviewer for bringing to our attention a study similar to ours that was not cited in our manuscript. Although the analysis by Qin et al. [2] bears similarities to our experiments, there are substantial differences that make both studies valuable:

(a) In [2], the GFDL CM2p1 dynamical numerical model is used solely to validate the optimal initial perturbations computed for the Deep Learning model employed in their study. However, they do not compute the optimal initial perturbations for the GFDL CM2p1 model itself. As a result, their analysis does not explore whether the GFDL CM2p1 and Deep Learning models exhibit similar optimal initial perturbations or how optimal initial errors propagate in both models. This comparison is a central aspect of our study.

(b) In [2], the CNOP objective function for the deep-learning model is optimized using automatic differentiation, a feature available in most modern deep-learning frameworks, such as PyTorch and TensorFlow. However, this approach does not apply to dynamical numerical models since these models are not implemented using modern deep-learning frameworks. Moreover, for these models, computing or approximating gradients with respect to their outputs is often highly complex or practically infeasible, making it challenging to apply gradient-based optimization techniques. To enable a meaningful comparison of how optimal initial perturbations evolve in both machine learning and numerical models, a gradient-free optimization method, like the Cobyla algorithm employed in our study, is essential.

These differences in the analysis performed and the methods adopted underscore the complementary contributions of our study and that in [2].

**Changes in manuscript:**

In the "Summary and Discussion" section, we will reference [2], discuss their interesting results and underline the key differences between their study and ours.

9. *The AI model appears to be less sensitive to the initial perturbations, or the initial perturbations do not grow as fast as those in numerical models, which is the reason for the higher skill of the AI model. Similar conclusions are also obtained in another similar study (https://doi.org/10.1029/2023GL105747). Can the authors discuss the pros and cons of this characteristics? I think this will be significantly valuable for the future modelling, as well as further understanding, of earth system.*

   **Author's reply:**
   We agree with the referee that discussing the pros and cons of this characteristic of AI models compared to dynamical models will be highly valuable.

   **Changes in manuscript:**
   In the "Summary and Discussion" section, we will discuss the pros and cons of this characteristic of AI models, referring to the interesting results presented by Selz et al. in [3].

# 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] Bo Qin, Zeyun Yang, Mu Mu, Yuntao Wei, Yuehan Cui, Xianghui Fang, Guokun Dai, and Shijin Yuan. The first kind of predictability problem of el niño predictions in a multivariate coupled data-driven model. *Quarterly Journal of the Royal Meteorological Society*, 150(765):5452–5471, 2024.

[3] T. Selz and G. C. Craig. Can artificial intelligence-based weather prediction models simulate the butterfly effect? *Geophysical Research Letters*, 50(20):e2023GL105747, 2023. e2023GL105747 2023GL105747.