Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2020-39-AC2, 2020

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Boosting performance in machine learning of geophysical flows via scale separation" *by* Davide Faranda et al.

Davide Faranda et al.

davide.faranda@lsce.ipsl.fr

Received and published: 18 December 2020

The authors are utilizing Echo State Networks to predict filtered dynamics in the perturbed Lorenz 1963 equations, the Pomeau-Manneville 89 intermittent map, and the Lorenz 1996 equations. A moving average filter is utilized for scale separation in time. The filtered dynamics are smoother and easier to predict. A residual term is added, either sampled from the training data, or based on an analytic formula derived from the moving average filter. Assuming that the filter width is smaller than the associated large timescales of the processes involved, the large scale processes can be successfully predicted. The authors claim that modeling only the spatially coarse grained and time averaged state can boost performance of ESN. However, the generalization of this argument to more realistic systems is not sufficiently supported by the results, as

elaborated in the comment section below. The idea of utilizing a moving average filter for noise reduction and scale separation, or spatial coarse graining is known. I am not sure that the novelty of the paper to apply ESNs to (spatially/time) filtered dynamics, is enough to guarantee publication in the journal. The effect of the unmodeled dynamics (the information lost during filtering) is not taken into account in the model. In most interesting applications, the effect of the unmodeled modes is the problem, and a field of study by itself (closure models in turbulence, small scale models in weather etc.)

We respect the opinion of the reviewer on our work, but we feel that the motivations provided here and in the following comments for rejection are not supported and sometimes do not contain any element that could help us to improve the manuscript. For example, the reviewer states many times that our results are "known" or "existing in the literature" or "hardly surprising" but not a single reference to previous works which should contain our results is provided. We are therefore unable to assess which part of our work could be not original or to provide an adequate rebuttal to state why our work is instead original. Furthermore, the reviewer says that our results are "not sufficiently supported by the results", but it is not said in which way this is the case. We stress that, for all cases presented, we have used at least three statistical metrics to assess performances, and scanned a large range of coarse graining and noise intensities, as well as performed several realisations of our systems. If this is not enough to warrant publication, then we would like to know why our suggested metrics are not sufficient to support our conclusion. Besides these remarks on the structure of the proposed comments, we will do our best to answer here the reviewer's comments and thus to improve the manuscript.

1 Comments

1. In the three-dimensional Lorenz system, it is logical that the moving average filter produces better results. By construction, noise is added to the system. It does not come as a surprise that the ESN predicting the filtered dynamics (which are smoother) and augmented with the random residual terms, shows superior performance. However, there is no complex multiscale effect taking place, as the whole state information is given to the system (no hidden state, at least nothing is mentioned in the text about it). Moreover, as a reference time-scale, the Lyapunov time of the deterministic system is used, although the system is augmented with noise, which means that the effective Lyapunov time is in essence much shorter, as stochasticity accelerates the divergence of nearby trajectories. In any case, it is important to be critical about the conclusions drawn from this case.

We agree with the reviewer that "it is logical that the moving average filter produces better results" but what we would like to show is that there is a dependence on the noise intensity on the quality of the results obtained and particularly that the moving average filters is very useful for intermediate noise intensity, namely when the stochastic component starts to affect the deterministic dynamics, but not at the same order of magnitude. We disagree with the statement that " the effective Lyapunov time is in essence much shorter, as stochasticity accelerates the divergence " as this is also dependent on the noise level. The most interesting performances for the filtered ESN are obtained when the noise is yet 2-3 order of magnitude smaller than the typical scales of the deterministic component, and we expect the perturbation on the lyapunov exponents to be of these orders as well.

2. In the Pomeau-Manneville intermittent map, it is not a surprise that the ESN cannot capture the dynamics, as they are changing very rapidly, even visually they look completely stochastic. A deterministic ESN with tanh (smooth, continuous) activation

function cannot be expected to produce trajectories that look spiking/stochastic/rapidly changing. Most previous studies on ESNs were handling relatively smooth signals, and not such rapidly changing signals. At least the nature of the signal has to be taken into account in the selection of the activation function of the reservoir. Thus, it does not come as a surprise that utilizing the ESN on the time averaged dynamics and then adding a stochastic residual improves performance. As expected, the plain ESN diverges, as demonstrated also in previous studies with such non-smooth signals.

We thank the reviewer for the comment. Indeed this can be a good explanation of our results. The reviewer says again that "it is not a surprise" or "demonstrated also in previous studies". However, no references are provided for us to improve the quality of the manuscript or to give credits to those who have already analysed this problem on another angle. We would be more than happy to include and discuss those references in the manuscript. Furthermore, the reviewer makes a confusion between a visual analysis and what signals truly are. The PM system is piecewise continuous differentiable, and is hence "relatively" smooth from the mathematical point of view.

3. In the Lorenz 96 system, as demonstrated in Figure 8, the method fails to capture the long-term climate, as the dynamics predicted by the ESN are clearly different from the groundtruth.

We only partially agree with the reviewer about the results obtained for Lorenz96. Although the detailed dynamics does look different from that of the original system, there are a few things correctly captured by the ESN, namely the quasiperiodic spatio-temporal oscillations and the fact that ESN produces non-divergent dynamics.

4. In the sea-level pressure, the moving average filter ESN does not achieve any significant improvement based on the results in Figure 9.

Here, we would like to gently disagree with the referee comment. The ESN with filter does produce significant improvements, in terms of all the metrics considered, and as noted by the other reviewer.

5. In the abstract, the authors claim that "multiscale dynamics and intermittency introduce severe limitations on the applicability of recurrent neural networks, both for short-term forecasts, as well as for the reconstruction of the underlying attractor". This is shown for Echo State Networks in the document, but not in general for Recurrent Neural Networks. The argument has to be relaxed to take into account only ESNs, or a relevant reference for other RNN architectures should be given.

We agree with this comment of the referee, in the new version of the manuscript we will clearly restrict our attention to ESNs.

6. There is a contradiction in the text, in page 3, the authors state that "We aim at understanding this sensitivity in a deeper way, while assessing the possibility to reduce its impact on prediction through simple noise reduction methods", although one sentence before, they claim that they choose the ESN framework for "...its ability to forecast chaotic time series and its stability to noise". These sentences are contradicting each other. Later in the text, the authors state "Since Echo State Networks are known to be sensitive to noise (see e.g. [34]), ...".

What we mean here is that ESN are less sensitive to noise than other techniques, but also that our goal is precisely to evaluate such sensitivity and the improvement coming from noise reduction techniques. We will make this clearer in the

new version of the manuscript

7. The analysis of the performance of the proposed method based on different parameters e.g. intermittency of dynamics/degree of coarse graining, etc. is interesting. However, this is not adequate to warrant publication.

We are delighted to see that the reviewer admits that our results are interesting. This encourages us to pursue their publication.

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2020-39, 2020.