Nonlinear Processes in Geophysics Discussion. Authors: Zhe An, Daniel Rey, JingXin Ye and Henry D.I. Abarbanel

24 May, 2016

Title: *Estimating the State of a Geophysical System with Sparse Observations : Time Delay Methods to Achieve Accurate Initial States for Prediction*

Recommendation: **Minor Revision**

General Comment

This is a very interesting and relevant study, timely and central on one of the most appealing research topic in data assimilation nowadays.

The methods describes a new formulation of nudging using time delay observations. While the approach has been already presented in previous papers from some of the same authors, the present manuscript adds a relevant application to a dynamical system of larger dimension so making this study appealing in the perspective of applications to more realistic model-observation scenarios.

I suggest the manuscript to be accepted subject to a minor revision directed mainly to improve the readability and the presentation of the results. In fact in a few places the manuscript lacks of the necessary details to make it clearer to a slightly broader audience than just experts in data assimilation.

Find below specific comments that I would ask the authors to address during their revision.

Specific Comment

1. Abstract. The authors should not introduce in the abstract alone important quantities, such as the critical threshold, L_s. This must be re-introduced in the Introduction or at least somewhere else in the main body of the article. Note for instance that it is only in the abstract that it is stated clearly that the work deals with chaotic dynamics. While this is patent in the rest of the paper it must be said explicitly.

2. Introduction. Lines 15-17, page 1. The statement is strictly true only for chaotic systems (this relates to my previous comment).

3. Page 2, lines 7-11. When talking about the necessity to control the unstable modes, it is relevant to mention methods, like the Assimilation in the Unstable Subspace, in which the analysis update is explicitly designed for this purposes; (see *e.g.* Palatella, L., A., Carrassi and A. Trevisan, 2013. Lyapunov vectors and Assimilation in the Unstable Subspace: theory and applications. J. Phys. A: Math. Theor. 46, 254020)

4. Page 2, line 12. The reference is Pazo et al., 2016 (Pazo, D., A. Carrassi and JM. Lopez, 2016. Data Assimilation by delay-coordinate nudging. Q. J. Roy. Meteor. Soc. 142, 12901299)

5. Page 2, line 16. L_s is not properly defined. The authors should explain a bit more precisely what it is meant by "critical threshold" ?

6. Page 3, line 5. Notation or text must be improved. Is $\mathbf{y}$ a L-dimension vector? If so, you better state that you take L observations that are then collected in the observation vector $\mathbf{y}$.

7. Page 3, line 13. Again notation or text must be improved. The equation $y_l(t) = x_l(t) + noise$ suggests that $L = D$ which is not the case in your experiments, and the fact that $L \ll D$ is indeed one of your key point. You should say that you use an operator H that only observes a portion of the state-vector (mainly the heights in the experiments that follow).

8. Page 3, Eq.(2) and lines 24-30. Eq.(2) requires observations at each time-step of integration, something which is usually obtained by interpolating in between successive observations in real applications. For the sake of completeness, it must be also added that the condition of a negative conditional Lyapunov exponent implies the setting for the strength of the nudging forcing, $\mathbf{g}$ and not just on L. In classic literature in fact this is usually the case, and the observational network is given.

9. Page 4, Eq.(7). I think $\mathbf{g}$ and $\mathbf{G}$ should not be bold.

10. Page 7, line 15. How is L_s obtained ? Does it come from the simple nudging case Eq.(2) ?

11. Page 7, line 18. "Cardinali" is written twice. Line 19: "observed" should be "observe".

12. Page 8, line 20. The sentence about parameter estimation relates to the chosen "perfect model" scenario. It would be better if the author states this clearly.

13. Page 8, line 22-24. You might want to say something more on this regard. How is it optimized? What does it mean average mutual information? Is it the time decorrelation scale?

14. Page 9, line 4-6. Are you computing the error using only the observed components of the state vector also for $t > T$?

15. Page 10, line 1. Do you mean Fig.5?

16. Page 10, line 5-6. Do you have suggestions on how to select the coupling?

17. Page 10, line 11-13. In fact the values chosen for L are very close to each other and results highlight a strong sensitivity of your method to this (error diverge when $L = 248$). Can you comment more on this? Another aspect regards how those observations are placed. One can always achieve a better control of the error by a proper deployment of the observations (possibly with the use of target observations). What would it happen if observations were denser in the proximity of the most dynamically active areas?

18. Page 10, line 18-22. The final paragraph of Section 5 is important but it necessitates some improvements: (1) You might want to say that, as known from classical synchronization results, the optimal forcing strength (g and G in the present context) depends on the number and distribution of the observations; (2) the conclusion relating the model resolution and observations network is unclear. Even with a high resolution model one may still necessitate a growing number of observations to keep under control the unstable modes. Please clarify this point.

19. Section 5.1. A number of details are unclear in the present version of the Section. In particular one gets easily confused by the mix of information on how observations are made noisy given in the caption of Fig.7 and the fields "data" in lines 29-30. (1) What is ϕ in the first equation? Did you mean ψ?; (2) From the 2^{nd} equation one sees that the observations have zero mean, which seems inconsistent. (3) In the experiments so far you have only observed h, why are you then showing how observational noise is simulated for the velocities if the latter are not assimilated? (4) What are the values of C_{height} and C_{data} and how are they chosen? We do not know how these values scales with respect,

for instance, to the system climate variance in the same variables. It is consequently impossible to judge to which extent this observational error is small or big. (5) Please make consistent the text in Section 5.1 with the Fig. 7 caption. From the latter one learns how observations are being perturbed with normal distribution with variance σ.

20. Page 11, line 6. The reference should be Mariano et al (2002). Isn't?

21. Page 11, line 18-30. This part is very interesting but key details are missing on the state-augmentation formulation that the authors seem to have used to incorporate drifts. Please improve presentation on this aspect and provide more details.

22. Page 13, line 23-25. Although its meaning may be clear to a reader from the data assimilation community, R_m is undefined? It may be convenient to define it in relation with Eq. (12), *i.e.* with R_f.

23. Page 12 - 13. The long discussion on the equivalence with smoother (4DVar) is very interesting but, in my opinion, very badly placed in the middle of the conclusion. That is not a conclusion, but rather a discussion. I think the authors should move it in the main body of the manuscript, perhaps when the method is presented and before the numerical results. In any case an independent section would be ideal.

24. Page 14, line 27-34. This comes too late and it would be better seen at the beginning of the Conclusion, when you recall the motivation behind the research effort.

25. Page 15, line 1. You have repeated already many times the model you have adopted.

26. Page 15, line 5-6 and line 13. Check the reference to Cardinali (2013).

27. References. Some entries in the list have typos that require corrections, particularly in the page information which has "?" instead of "-". This is the case, for instance, of Kuznetsov et al, 2003, but the problem is present elsewhere as well. Make consistent use of journal names abbreviations throughout all entries.

28. Figure 6. Correct labels and box in the top panels.