Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2019-7-RC1, 2019

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Data assimilation as a deep learning tool to infer ODE representations of dynamical models" *by* Marc Bocquet et al.

Anonymous Referee #1

Received and published: 3 April 2019

General comments:

In this paper, authors want to learn the dynamics of a system from perfect or imperfect observations. To do so, they use the observations to build a surrogate model. Compared to recent papers that try to use Machine Learning (ML) algorithms to identify the surrogate model, authors in this paper are using Data Assimilation (DA) as an optimization tool. After identifying the dynamics with their technique, they apply the surrogate model to get forecasts and compare them to true simulation runs. Note that they use local representations to deal with high-dimensional models, an important challenge almost never treated in other papers using ML approaches.

The methodology is tested on different toy models, from the Lorenz attractor to the two-scale Lorenz-96 system. Results prove the reliability of the method, even if more

details about the identification part of the dynamics would be appreciated. Finally, authors pointed out interesting perspectives in the conclusion, especially the estimation of model error statistic for their surrogate model.

The paper is overall well written with a nice review of the existing and recent literature on data-driven approaches, including analogues, diffusion maps, reduced models and neural networks. Method is described with all the details. Experiments and results are well presented and discussed. However, I have some suggestions to improve the quality of the paper and avoid some misunderstandings. It concerns the title and some technical sections: these points are detailed below.

Problem of the title:

The authors should not use "deep learning" in their title. This is a confusing point because some readers might think that this paper is dealing with ML algorithms but this is not the case. However, the discussion in Sect. 3 about the link between the presented method (based on DA) and deep learning is interesting, especially when comparing equations (17) and (18). Moreover, the authors nicely show that using a weak constraint 4D-Var is a way of controlling the backpropagation (important part of deep learning methods).

Improve the readability and understanding of the method:

I think the authors should write more context in Sect. 2. First, in Sect. 2.1, the monomial basis should be discussed, giving more explanation and illustrating for instance with the Lorenz-63 (as Brunton did). I have the same remark in Sect. 2.2.1 where the drastic reduction of columns of A due to locality could be illustrated using for instance the L96 system. Then, Sect. 2.2.2 about the homogeneity (same behavior of the model at different locations) is hard to follow: again, authors should try to illustrate. Finally, entire Sect. 5 is very technical and I suggest to keep the important results/equations in the main text and move the rest to the appendix.

Other specific points:

- p3, l4-5 –> can you give a quick explanation of the difference with the neural network approach? - p6, l5 –> what do you mean by one-dimensional state space models? The models you introduce in Sect. 4 are all multi-dimensional, please clarify. - p16, l8-9 –> this is not always the case, especially when you use recurrent neural networks, that are not deep in practice. - p21, l16-17 –> in addition to Fig. 8, for different values of observation noise, I would like to have a look at the estimated coefficients of A along the DA cycles. - p23, l20 and Fig. 9 –> what do you mean by "long space-time stripes"? This point needs clarification with maybe a zoom on Fig. 9 to make this point clear.

Typo:

- p23, l2 –> "the the"

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2019-7, 2019.