Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2017-52-RC2, 2017

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Interactive comment on "Feature-based data assimilation in geophysics" *by* Matthias Morzfeld et al.

Anonymous Referee #2

Received and published: 20 December 2017

Manuscript ID: npg-2017-52

General Comments:

This study evaluates the concept of "feature-based data assimilation (DA)" in four simple geophysical models by employing three numerical techniques: Markov Chan Monte Carlo, direct sampling, and Global Bayesian Optimization. The true test would come when evaluating "feature based DA" in more-complex geophysical modeling systems. Nonetheless, the results presented offer encouragement, especially because the basic ideas presented in this article can be applicable to some of the machine learning algorithms used in data assimilation. The topic of the manuscript is within the scope of this journal. However, some of the concepts presented require more detailed explanations, which need to be addressed before final publication. Specific details are provided in

Printer-friendly version

Discussion paper

the following sections.

Specific Comments:

1. Writing style needs to be improved in some parts of the manuscript. The phrase "model output and data" or the words "model" and "data" are repeatedly used throughout the text. This makes sense, since they are the central themes of this paper. However, some sentences could be rephrased to avoid repetition and thus continue to engage the reader. Only some of the typos and grammatical deficiencies encountered were included in the "Technical Corrections" section.

2. P1, L5: Clarify this statement "assimilate data of a complex-system into a lower-dimensional model." In data assimilation problems, the most common scenario is to have a model state with dimensions much larger than those of an observation vector. Can you provide a few examples in the text rather than pointing the reader directly to some references.

3. Please, elaborate on the applicability of "feature-based DA" when assimilating data into higher-dimensional models (e.g. coupled atmosphere-chemistry or coupled atmosphere-ocean numerical models)? See the work by Yablonski and Ginis, 2007 for reference.

4. P9, L10: Explain what is meant by "different scales" are you referring to time or space?

5. P13, L34: Provide an explanation to the strong correlation between the parameters, but not so much to the initial conditions. Also, did you evaluate correlations between pairs of parameters?

6. P14, L4-9: Further clarify the differences on the trajectories for the Here and Lynx populations. Is the model evaluated at each cycle-step, whereas the data are cumulative? Whys is this a limitation? In general data assimilation problems, better fits to the observations are found by assimilating continuous data and not just data at a

given cycle. What limitations in LV model would lead to such different behavior on the trajectories of those critters.

7. P16, L16-19: What are the advantages of using an “averaging window”? Is it possible that trying to find an optimal multiplicative parameter ïĄś(t) through minimization of a variational cost function would produce better results? Did you consider this approach as an alternative to particle filters?

Technical Corrections:

1. Labeling figures using letters e.g. (a), (b), etc. will ease finding their reference in the text.

2. P1, L6: Delete the comma before “and.” Two nouns that appear There are many instances of this throughout the text Don’t separate two nouns that appear together as a compound subject or compound object.

3. P3, L23: Choose “and” or “but”

4. P4, L10: Replace “all over” with “in”

5. P4, L19: Choose either “as in the case” or “for example”

6. P5, L24: Replace “one thus wants” with “one thus want”

7. P5, L25: Same as in 2

8. P9, L23: The abbreviation for a “million years” is myr or Myr, correct throughout the text

9. P9, L30: Same as in 2

10. P10, L1: Insert a comma before “but”

11. P10, L26: Replace “form” with “from”

12. P11, L5: “posterior distributions” is repeated twice

13. P13, L22: Insert “of” in between “mean” and “zero”

14. P13, L26: Replace the period next to “as follow” with a colon

15. P13, L30: Move the comma inside the quotation marks

16. P15, L13: What does R stands for?

17. P16, L13: Use the past participle form of to vary

18. P21, L1: Replace the period after “follows” with a comma

19. P21, L22: Delete the comma after “process”

20. P21, L25: Same as in 18

21. P22, L5: replace the colon after “terms” with a comma

22. P23, L7: missing period

23. P23, L11: Use plural form of “issue” and revise comma use

24. P24, l12: Use “leads” or “led”

25. P25, L8: Revise the use of the word “base”

Figures:

1. Figure 1: label figures (a) – (d) and insert those into the body of the manuscript for easier reading, repeat for all Figures with multiple plots

References:

Yablonsky, R. M., and I. Ginis 2008: Improving the Ocean Initialization of Coupled Hurricane–Ocean Models Using Feature-Based Data Assimilation, Monthly Weather Review, 136, 2592–2607.

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2017-52, 2017.