Manuscript ID: GMD-2018-48

Title*: Evaluation of iterative Kalman smoother schemes for multi-decadal climate analysis with comprehensive Earth system models

Authors: J García-Pintado, André Paul

Date: 05/09/2018

*Modified from the discussion paper

Responses to Reviewer #1 (anonymous)

We thank the reviewer for their positive comments, which will help improve the manuscript. Below we give each comment and describe how we are altering the manuscript to address the reviewer's concerns.

Let us note that indirectly prompted by Reviewer # 2, we have considered a new title for the manuscript as indicated above. In general, we are finishing a substantial rewriting of the manuscript, due to (a) specific request from the reviewers to shorten introductory parts and expand the results and discussions, and (b) reviewers' comments also have indirectly suggested us that some parts of the manuscript were in need of further explanation.

Thus the Introduction is now longer and the description of the paleoclimate context has been slightly expanded, but former section 2 (Problem definition) has been now dropped and compacted within the Introduction. The description of the nonlinear relation between the control variables and the observation space in the experiment with the Community Earth System Model (CESM) now receives more attention. The analysis with ETKF-GA (the Gaussian anamorphosis) is now more detailed, and the described scheme of the iterated Kalman smoother is now also included in the CESM experiment. Also, the first experiment, with the 1D energy balance model, has been updated with a new more adequate 4D-Var benchmarking.

Responses

Rev#1:

The authors present innovative low-cost strategies for online model parameter estimation, which can potentially be applied for climate field reconstruction with coupled GCMs. The manuscript is written in a transparent way and authors explain very well all the assumptions used in their study. The paper moves the time-averaged data assimilation efforts a small but very important step forward. I recommend publishing the paper after minor revision.

General Comments:

Rev#1:

1) Scientific significant:

The manuscript represent a substantial contribution to modeling science within the scope of this journal. The ideas and methods are original.

2) Scientific quality:

The scientific approach and applied methods are valid and the assumptions are well introduced. However, the results are not discussed in a balanced way and could be improved (see below). The models, technical advances and/or experiments described have the potential to perform calculations leading to significant scientific results.

Auth: According to both reviewers, we have done a thorough revision of the manuscript, shortening the introductory sections and expanding the result sections.

Rev#1:

3) Scientific reproducibility

The modeling science seems not to be reproducible. I think if the authors share their code, this problem will be solved. However, their methodology is well described and traceable.

Auth: We are sharing the code.

Rev#1:

4) Presentation quality

The presentation quality is fair and could be improved in a new version. Number of figures presenting the results can be revised.

Auth: See answer to general comment 2 above. The updated version involves additional figures showing an example of the nonlinear relation between a model parameter and the dual of the observation space for experiment 2 (CESM experiment) as well as the effect of Gaussian transformation, and an example of sensitivity estimates of T2m to a model parameter (as example of atmospheric variable, according with comment 23 below).

Rev#1:

5) Overall, the manuscript is understandable for experts working in the field but not easy to follow for general readers. There are too many acronyms in the manuscript. Please spell out when possible.

Auth: We hope the new version is more clear.

Rev#1:

6) Given that this paper belongs to the category of "Development and technical papers"

(see:https://www.geoscientific-model-development.net/about/manuscript_types.html#item2),

I encourage the authors to make their code available online (at least for the 1D experiment). This might help the community very much and improve the code itself. According to the journal policies, the authors have to include the model's version in the title (e.g., Model XXX (version Y)).

Auth: We now are sharing the code online for the 1D experiment. We also share the DA code for CESM code under request to the authors. Let us note there is a number of more efficient available DA software based on compiled language (C, F90), as , for example, SANGOMA, EMPIRE or PDAF. Our code is instead for prototyping and research purposes, mostly based on R scripting with netCDF4 as interchange format, and resorting to calls to CDO, and occasionally to bash and perl. Note the computational demand of the assimilation is minor in our case by comparison with the computational demand for the model integrations, with CESM runs in ensemble mode for mutidecadal time spans. We are currently arranging our CESM+DA software as an R package with included documentation, in a way it can be directly usable by the paleoclimate community.

Rev#1: 7) Page1Line3 (P1L3): Explain how model's parameters have relationship with proxies!

Auth: This sentence has been removed in the new abstract.

Rev#1: 8) P1L6-7 is too complicated!

Auth: Removed in the new abstract.

Rev#1: 9) P1L6 : Authors mix two approaches: model's parameter estimation (atmosphere and ocean) and fresh water melting parameters estimation. Have they done separate experiments? They describe, the latter might be essential for North Atlantic circulations.

However, there might be nonlinear relationships between these two and they could contaminate each other. Which one impacts the error reduction larger?

Auth: The sensitivity experiment for the parameter-space schemes (now renamed as Numerical Local Sensitivity (NLS) schemes and “fractional” substituted by “multistep”(i.e. NLS-IKS, and NLS-MKS), are based on individual sensitivity analysis. We have now described which parameter has shown a higher sensitivity in the analysis (specifically, the atmosphere parameter cldlfrc_rhminl: minimum relative humidity fro low stable cloud formation.

Rev#1: 10) P2L12: Could you provide references for that?

Auth: We have modified the order of the paragraphs. The reference now given for this is Annan et al. (2005b), which was indicated later in the former manuscript.

Rev#1: 11) P2L32-33: How should one do this? Reference?

Auth: We give now a few references to sensitivity analysis. Please see updated text.

Rev#1: 12) P2L5: Explain the “generally positive results”!

Auth: Explained.

Rev#1: 13) P3L20: Aren’t the parameters not being updated at each DAW? Aren’t they time varying when the new observation is available? How could one do that in future projections without observations? What are the challenges? How could tuning the model for the past improve projections? Please clarify!

Auth: Yes. Their estimates vary, but this does not mean that they are time-varying. This has been clarified. We also clarify that the goal is to conduct past climate field reconstruction at long time scales, and the the model is assumed to be previously tuned. The control variables in the assimilation (model deterministic parameters, and other inputs) are used to carry on the responsibility to generate an uncertain background. Thus the estimated control variables at each DAW serve as a mechanism to minimise the cost functions and to obtain a climate filed reconstruction fusing model and data. Differences (or increments) between corresponding tuned parameters (for present day dense datasets) and those estimated by the assimilation based on proxy databases of past climates may serve to diagnose model differences, and be a very useful tool (as opposed to direct updating of the climatic full-field as in standard EnKFs, here there is a physical mechanism in the model explaining the climatic increments resulting from the assimilation). But estimated parameters as part of the paleoclimate assimilation based on proxy data would not in principle be meant to replace the originally tuned model for future projections. We have clarified this in the manuscript.

Rev#1: 14) P6L10: Explain briefly the gradient descent algorithm, learning rate, number of iterations, etc…

Auth: There are several options for this (conjugate gradient methods in general). Then, tipically 4D-Var uses about 3 inner loops and one (e.g.; UKMO) or two outer loops (e.g.; ECMWF). But we think it is better not to expand too much on this here as it is not central to the manuscript. Currently operational centres have mostly moved to hybrid methods ('En4DVar') and either use ensemble of 4D-Vars (as ECMWF) or use EnKF to get hybrid B matrices, and the scenario of options is rather wide. Alternatively, for the interested reader, we have given the new sentence:

“The current implementation of variational assimilation (with atmospheric models) is now different in each operational NWP center, who have mostly moved to hybrid methods. A recent review of operational methods of variational and ensemble-variational daa assimilation is given by Bannister (2017)”

Rev#1: 15) P8L6-7: However, aren’t the parameters updated based on time-averaged obs?

Auth: Yes. The estimates of the parameters. Please see response to comment 13.

Rev#1: 16) P11L18-19: How do you define the learning rate in gradient descent?

Auth: We have removed these lines. The strategy for the (now called) NLS-MKS scheme stepping is described later in the section.

Rev#1: 17) P15L19: Why analyzing only 10 years? 90 years for spin up of 1d model?

Auth: This follows the protocol in Paul and Losch (2012) to make the experiment more comparable with theirs.

Rev#1: 18) P18L27: “not shown”, but is interesting to put in supplementary.

Auth: Included in supplementary material

Rev#1: 19) P19L2 : is it a typical set-up of CESM?

Auth: Yes, it is the scientifically validated compset with short name B1850CN, as found in http://www.cesm.ucar.edu/models/cesm1.2/cesm/doc/modelnl/compsets.html

Rev#1: 20) P20L9: Could you explain and discuss the problems of multi-component DA in your set-up? Differences in time-scales of proxies, etc.

Auth: Please note we have indicated this now by the more standard (“strongly coupled” data assimilation). A brief discussion has been included.

Rev#1: 21) P20L13: Have you done other experiments with other sets of parameters? For example more or fewer numbers of parameters?

Auth: For other model configuration yes, but not specifically for this study. We understand that additional experiments based on a wider number of error sources (e.g. from biases in other parameters not included in the control vector) would be very illustrative. We comment now on this in the introduction of the new manuscript in reference to error compensation in real applications. We also indicate the study should be expanded in further research.

Rev#1: 22) P24L10: In figure 3 it is really hard to see the differences. Maybe centering the colorbar with zero might help. How do you explain that the error reduction is due to DA and not the lack of sensitivity of the model to perturbation of the parameters. For example Figure 4 upper left panel shows that the model is not sensitive to changes of cldfrc_rhminl in Arctic and Antarctic regions.

Auth: We have tried to add observation locations to Figures 4 and 5, and believe this render plots too noisy in these cases. Also, making colour scales centred around 0 made a bit more difficult to see some patterns in the shown cases (not in other cases). The neatest solution we have found is to include isolines at level 0 for both Figure 5 and Figure6, and to center Figure 3 at longitude 180 to match Figures 4 and 5, as seen in the revised version.

Rev#1: 23) You focus on the ocean where the observations where assimilated, how about the atmospheric variables? Is there any error reduction happening there? Could you show for example global T2m quantities?24) Figures similar to Fig.4 for other perturbed parameters could be shown in the supplementary. This will clarify the sensitivity of CESM.

Auth: We have added sensitivity plots for T2m in the manuscript.

Rev#1: Specific Comments:

1) P3L6: “in section?”

Auth: Corrected.

Rev#1: 2) P3L7: “problem of CFR” which problem?

Auth: This sentence has been removed.

Rev#1: 3) P22L12: How about the uneven time resolution of observations?

Auth: We obtain model equivalent of the observations at the resolution of the observation time. This includes, for example, seasonal means or annual means during specific sub-spans of the DAW, which are specific for each proxy type. As the paleoclimate proxies represent an integrated effect longer that the model timestep (e.g. commonly get monthly output from CESM) forward models (proxy system models) can include this integration effect. In this study we do not touch the problem of paleoclimate proxy modelling and the forward model only does the time integrations. Now we mention in the introduction that this is the other stumbling stone (along with the high computational demand) hampering assimilation with comprehensive ESMs for past climates.

Rev#1: 4) figure 2, 3, 4: for the results one has to switch between figure 2, 2 and 4 to follow theline of discussions. You could at least put the observation locations on figure 3 and 4. Note that the land-sea mask is also shifted between the figure 2 and 3-4.

Auth: We have reorganized the discussion. Please, see also the answer to comment 22 above.

References (not included in the paper)

Alll references here are included in the manuscript