Nonlin. Processes Geophys. Discuss.,
doi:10.5194/npg-2016-68-AC1, 2017

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Interactive comment on "Impact of Optimal Observational Time Window on Coupled Data Assimilation: Simulation with a Simple Climate Model" *by* Yuxin Zhao et al.

Yuxin Zhao et al.

XiongDeng407@hrbeu.edu.cn

Received and published: 21 April 2017

Reviewer 1# A few tel-conferences of all co-authors have been held to discuss the comments from reviewer #1. All authors converge to the point that all the comments are very important and useful for authors to improve the quality of this manuscript (MS). Therefore, all comments from reviewer #1 have been fully addressed in the revision.

Now we will reply to each comment point by point as following:

1 the title does not fit. You are not discussing the "Impact of Optimal Observational Time Window' but "Impact of Observational Time Window". It may be "Impact of Observational Time Window in Coupled Data Assimilation with a Simple Climate Model".

RE: Thank you very much for your generous comment and we fully agree with that. We have corrected the tile as "Impact of Observational Time Window on Coupled Data Assimilation: Simulation with a Simple Climate Model".

2 The concept of OTW in this study validates the observations in a time window to the center of the window (analysis time). It is very useful technique for data assimilation (DA). More or less it has been applied in the assimilation with real observations. Here the OTW in this study is applied in 3-dimension DA but not 4-dimension DA, which need address in the introduction. The citations of OTW are not very relevant (Page 2, line 13-15). I could not find the clear concept of OTW and how they applied in data assimilation. So you should address the technique details in the paper. I guess that you treat all the observations independently and assimilate then sequentially using their original error scales. another approach is to assimilate the average values of the observations in the OTW. It is worth to compare these two approach in the study.

RE: As Line 19-23 of Page 2, we have rewritten this part and addressed that in this study the OTW validates the observations including in a time window to the center of the window and are sequentially assimilated using their original error scales. The irrelevant citations have been deleted and we have added some ideas in a relevant one to introduce the concept of OTW using in this study. As Line 8-18 of Page 12 in the revised manuscript, we have added some discussion about the comparison of sequential assimilation and averaged-observational assimilation approaches. And the comparison of the results of the two methods are shown as Figure. 1. Here the Sqe-CDA-OTW represents the experiment that sequentially assimilates the observations including in the OTWs using their original error scales. And the Ave-CDA-OTW denotes the experiment that assimilates the average of the observations including in the OTWs. But the standard deviation of the observational error will be 1/(sqrt(N)) of the original error scales. Here all the experiments will use the biased model setting and the uni-variate adjustment scheme. From Fig. 1, we can see that the results of the Ave-CDA-OTW is almost same as those of the Sqe-CDA-OTW. And we will illustrate this phenomenon form the perspective of the calculation of the EAKF method (the corresponding interpretation are shown following Fig.1). And we can see that for a CDA system in the CGCM with a suitable inflation scheme, the assimilation approach with the time-averaged observations will have more advantages over that which sequentially assimilate the observations. It is very interesting to comparing these two approaches, but using either assimilation approach will have no influence on the essential of this study. Thanks.

3 From your result I can not connect the de-correlation time with the optimal OTW directly. The optimal OTW is much short than the de-correlation time. Like in figure7, the correlation is 0.995 for the optimal OTW. Please revise the abstract and conclusion. The time scales of the variables really can not provide the useful information to quantify the optimal OTW. I suggest you remove or shorten the discussion and analysis related to the time scales of the variables.

RE: As Line 19-27 of Page 10, We have shortened this part. On the one hand, we have deleted the discussion about the relationship between the length of the optimal OTWs and the de-correlation time scales. On the other hand, we reserved the discussion about the high correlation between the observations including in the optimal OTW and the model state at the analysis time, which makes sense to the main idea of this study. Thanks.

4 The time (s) of observations in the OTW are different from the assimilation time. The temporal offsets introduce the represent errors for the observations, which need to be consider during analysis. If not, EnKF overweights the observations The analysis will not be optimal and the analysis ensemble spread will has negative biased. Please address this part and justify your assimilation.

RE: This is a very important and necessary comment. As Line 19-33 of Page 12, we have added some discussion about the temporal offset introduced the representation errors for the observations that are different from the analysis time. We use the de-

correlation coefficients to weight the observations including in the OTWs and avoid the overweighting. The comparison of the results of the weighted and no-weighted experiment are as Fig. 2. We can see that the lengths of the optimal OTWs obtained by these two assimilation schemes are the same except that the RMSEs of the weighted observation experiment will be lower than that of the non-weighted observation experiment when using the longer OTWs (when the length of ATM-OTW is greater than 4 and that of the OCN-OTW is larger than 50). This is owing to the high correlation between the observation including in the optimal OTWs and models states at the analysis time (exceed 0.995). Thus the influence of the temporal offset can be ignored and the results obtained by these two scheme will be almost same when using the lower OTWs (no or less greater than the lengths of the optimal OTWs). When we use the longer ones, the correlation will decrease and influence of the temporal offsets will be obvious that the results of the weighted observation experiment will better. For the CGCMs, owing to the complex physics and dynamics, the influence of the time offsets will be obvious and the weights of the observations will be very necessary. But from this simple model case, we can see that whether or not using the weighted observations, the relationship between the characteristic variability time scales and the optimal OTWs will be firm and the essence of this study is established. Thanks.

5 It is hard to find the assimilation intervals for different components. Please address them together in experiment setup (2.3). I noticed that the OTW is longer than the assimilation interval, which means that you assimilate the same observations multiple times. In the perfect framework, it should degrade the filter performance from my understand and the information theory. Can you explain why you can achieve a good analysis under that circumstance. For a biased model framework, people do use observation multiple times to compensate the negative bias of forecast ensemble spread. So it make sense to see the optimal OTW in biased model cases are longer than that in perfect model case.

RE: As Line 21-22 of Page 7, in this study the assimilation interval of the atmospheric and slab oceanic states are 5 and 20 time steps, respectively. But the optimal OTWs obtained for the atmospheric and oceanic component in the perfect experiment framework are 1 and 10, which shows that no observation will be assimilated for multiple times. When the length of the OTW increases beyond that of the optimal one, the filter performance degrades, which satisfy your understanding and the information theory. Thanks.

6 The section 2.2 of "Ensemble coupled data assimilation" is incomplete. People will not understand your equations without trace back to your references. Please provide the completed two steps of EAKF.

RE: This comment is necessary. As in the section 2.2 in the revised manuscript we have provided the completely two steps of EAKF. Thanks.

7 Page 9 the last paragraph. The discussion of OTW with 4D-var is very confusion and hand-waving. The OTW in 4D-var is very different concept than yours. You should remove i

RE: Have removed it. Thanks.

8 Figure 4(a & b) show that the lower bound of RMSE reaches to 0. That does not make sense. Please check.

RE: Figures 4 and 7 are re-plotted. Thanks.

9 Figure 4 and Figure 5 show that the ensemble spread is significantly smaller than the RMSE. Please address the reasons and the effect on the assimilation.

RE: As Lines 14-18 of Page 10 of the revised manuscript, is this study, no inflation scheme was used in all the assimilation experiments. And only the data obtained in the last 5000 TUs will be used to conduct the error statistics for evaluation. Thus after first 5000 TUs assimilation, the ensemble spread has been greatly reduced and is significantly smaller the corresponding RMSE. Although in theory the reduced ensemble spread will degrade the quality of state estimation, in this study with a one-dimensional

conceptual coupled model, for simplicity and computational convenience as well as convenience for comparison, we insist not using the inflation scheme. Thanks.

10 P11 Lin13 I can not buy the statement of "since more observational information is needed to compensate the model bias and recover the characteristic variability." The reason should be "the forecast ensemble in a biased model underestimates the forecast error, which results the EnKF underweight the observations. Therefore one can improve filter performance by using observation multiple times"

RE: We fully agree with your point. In the biased model experiment framework, the length of the optimal OTWs is larger than those in the perfect experiment case. The only reason will be the influence of the bias in the biased model. As Line 21-27 of Page 11, we think that we need stronger observational constraints to reduce the mean discrepancy between the model and observation induced by the influence of the bias. Also we fully agree that it needs to assimilate the observations for multiple times to compensate the underestimation of the forecast error by the bias in the biased model. We think these two interpretations are identical that using larger OTWs to compensate the influence of the bias except from two different aspects. Thus we think these two interpretations are not contradictory. Here we reserve our original interpretation and add your point into the revised manuscript. Thanks.

11 A reading proof is required to improve the manuscript. There are too many typos and grammar errors. For example Page 2 Lin 33 The "ensemble filter" should be "ensemble Kalman filter" Page 6 Lin 13 "but they are started from different initial states." can be deleted Page 6 Lin 15 "observation" should be "truth" Page 6 Lin 33 "the coupling coefficient C1is sensitive to model stability" should be "the the model stability is sensitive the coupling coefficient C1" Page 7 Lin3-5 The whole sentence need rephrase. You can just mention that "Following Zhang and Anderson (2003) an ensemble size of 20 is applied in all experiments in this study." Page 7 Lin9-11 "absolute error" is not understandable. I guess that is "the absolution values of error". The experiments are not described clearly. Please rewrite. Page 7 Lin28 the sentence is not completed. Page

7 Lin 30 Please also give the equivalent time of atmosphere (1-2 model month?). The deep ocean should be 50 model year.

RE: Thank you very much for your indication. All of these are fixed. Here the “absolute error” just represents the “absolute value of the difference between the estimated model state value and corresponding truth”. Second in the section 2.3 of the revision manuscript, the experiments have been rewritten and described clearly. As Line 28-29 of Page 8, the equivalent time of atmosphere is 1-2 model month and that of the deep ocean should be 5 model decades.

Thank you very much again for your generous help and comments! We hope that the revised manuscript meets your requirement.

Sincerely yours, Shaoqing Zhang and Co-author

Please also note the supplement to this comment:
http://www.nonlin-processes-geophys-discuss.net/npg-2016-68/npg-2016-68-AC1-supplement.pdf

---

Interactive comment on Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-68, 2017.

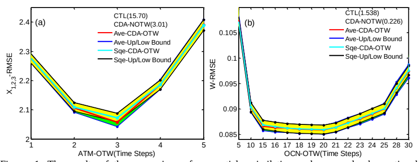

Figure. 1: The results of the comparison of sequential assimilation and averaged- observational assimilation approaches.

Here we assume the OTW including two observation a and b,

1) Sequentially assimilate these two observations using their original error scales:

$$\Delta \mathcal{Z}_{i,j} = \frac{Cov(\mathcal{Z}_j, \mathcal{Y}_k)}{(\sigma^{p}_{k,k})^2}[\left(\frac{\bar{\mathcal{Y}}^{p}_k}{1+r_k^2} + \frac{\mathcal{Y}^{o}_k}{1+r_k^{-2}} + \frac{\Delta \mathcal{Y}^{p}_{k,i}}{\sqrt{1+r_k^2}}\right) - \mathcal{Y}^{\mathcal{P}}_{k,i}]$$

in this study there is no inflation scheme, thus after many assimilation steps, the ensemble spreads of the model states will greatly decreased will be greatly smaller than the observational error. Thus we can assume $\bar{\mathcal{Y}}^{p}_k \approx \mathcal{Y}^{\mathcal{P}}_{k,i}$ and $\frac{Cov(\mathcal{Z}_j, \mathcal{Y}_k)}{(\sigma^{p}_{k,k})^2} \approx constant\ r$. For simplicity and convenience, we assume at a assimilation time the $\bar{\mathcal{Y}}^{p}_k$, $\sigma^{p}_{k,k}$ and $\Delta \mathcal{Y}^{p}_{k,i}$ will keep as constants during the sequential data assimilation process. And we can use constants C_1, C_2 and C_3 to represent the $\bar{\mathcal{Y}}^{p}_k$, r_k^2 and $\Delta \mathcal{Y}^{p}_{k,i}$, respectively.

$$\mathcal{Z}_{i,j} = \frac{1+(1-r)C_2}{1+C_2}\left[\frac{1+(1-r)C_2}{1+C_2}C_1 + \frac{rC_2}{1+C_2}a + \frac{r}{\sqrt{1+C_2}}C_3\right] + \frac{rC_2}{1+C_2}b + \frac{r}{\sqrt{1+C_2}}C_3$$

C_2 is very small and we can assume that ${C_2}^2 \approx 0$. Thus

$$\mathcal{Z}_{i,j} \approx \frac{1+2(1-r)C_2}{1+2C_2}C_1 + \frac{rC_2}{1+2C_2}(a+b) + \frac{2r+2rC_2-r^2C_2}{\sqrt{(1+C_2)^3}}C_3$$

2) assimilate the average of the observations but ${C_2}' = ({r_k}')^2 = 2C_2$; thus

$$\mathcal{Z}_{i,j}' = \frac{1+2(1-r)C_2}{1+2C_2}C_1 + \frac{2rC_2}{1+2C_2}\frac{(a+b)}{2} + \frac{r}{\sqrt{1+2C_2}}C_3$$

And we can see that the difference will be the third term $\frac{2r+2rC_2-r^2C_2}{\sqrt{(1+C_2)^3}}C_3$ and $\frac{r}{\sqrt{1+2C_2}}C_3$.

Method 1: $\frac{2r+2rC_2-r^2C_2}{\sqrt{(1+C_2)^3}}C_3 \approx \frac{2r+2rC_2-r^2C_2}{1+C_2}C_3$

Method 2: $\frac{r}{\sqrt{1+2C_2}}C_3 \approx \frac{r}{1+C_2}C_3$

Method 1-Method 2, thus:

$$\begin{aligned}\frac{2r+2rC_2-r^2C_2}{1+C_2}C_3 - \frac{r}{1+C_2}C_3 &= \frac{r+2rC_2-r^2C_2}{1+C_2}C_3 \\ &= \frac{-C_2(r-1)^2+r+C_2}{1+C_2}C_3 > 0\ and < C_3\end{aligned}$$

Here C_2 and C_3 are also very small, thus the results of these two methods are almost same. And we can extend that to longer OTWs and more observations cases in this study.

Fig. 1. The results of the comparison of sequential assimilation and averaged- observational assimilation approaches.

We use the de-correlation coefficients (As the Panel a and b of the following figure) to weight the observations including in the OTWs and avoid the overweighting. The comparison of the results of the weighted and no-weighted experiment are as following:

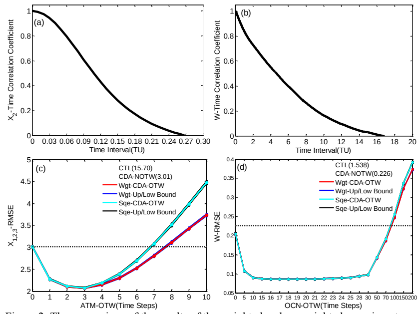

Figure 2: The comparison of the results of the weighted and no-weighted experiment.

Fig. 2. The comparison of the results of the weighted and no-weighted experiment.