Point-by-point responses for review #2

In general:

A few conferences were held for authors to discuss the comments of 3 reviewers. All co-authors converged to the point that all comments from reviewer 2 are very thoughtful and important for improving the manuscript and enhancing our understanding on the topic. Several experiments for explaining the concerns of the reviewer are performed. The paper is renewed as the reviewer's suggestions. New tables have been added in the paper. We also exchanged the location of section 2.2 and 2.3. What follows is a point-by-point reply for reviewer 2:

General comment:

This work used a simple conceptual model to provide insights on the role of atmospheric/oceanic state estimation in coupled model parameter estimation. They concluded that the accuracy of the atmospheric state is the crucial factor for such kind of parameter estimation. I regard this work is innovative and the manuscript is well structured. However, my main concern is whether the setup of the assimilation experiments and the conclusion of this work are applicable to the real world. My suggestion for this manuscript is major revision before it can be considered for formal publication. My main concerns are as follows.

RE: Thanks for the reviewer's comments. The PE without sufficient SE in a coupled system is an interesting topic. We gain a lot of benefits from this study, for example, the real analysis and prediction with the coupled data assimilation (CDA) system. While our coupled data assimilation (CDA) system was established in 2007, we have been making efforts to implement parameter estimation into CDA to improve climate analysis and prediction, but the improvement remains in a limited range or none. We have to come back to simple models to sort out the sources of noises. The simple conceptual model does have limits (added in section 4, P11L19~27), but its dynamics and transferring of the uncertainty is crystal clear. With the help of this model, we

found that since the imperfection of observing system and extra model errors have much stronger influences on coupled parameter estimation than coupled state estimation, how to enhance the signal-to-noise ratio of a parameter-state covariance is the key for successful coupled model parameter estimation. In such cases, the simple model has more visibility to demonstrate the essence of the problem.

Major comments:

{1. The setup of the SE has an assimilation interval of 5 time-steps, which is shorter than the current atmosphere analysis update interval and can be regarded as a rapid update cycle. Such setup also greatly controls the signal-to-noise of the atmospheric condition. Although the authors claim that the results are not sensitive to the choice of update interval (Page 5, line 13), the accuracy of the atmospheric state could be seriously degraded with a longer update interval (or with only x_1 observations) and shed the relationship with the parameter.

Can the authors provide PE experiments using a longer update interval (e.g. 25 TU) or assimilate x_1 only to illustrate the condition that the atmosphere state is less optimally observed? }

RE: Thanks for this important comment since the signal-to-noise ratio is different on different frequency in state variability. As suggested, we firstly tested our result with different PE update intervals. The following figures show that the major results in our studies do not depend on the PE interval settings. New lines are added in P7L20~22.
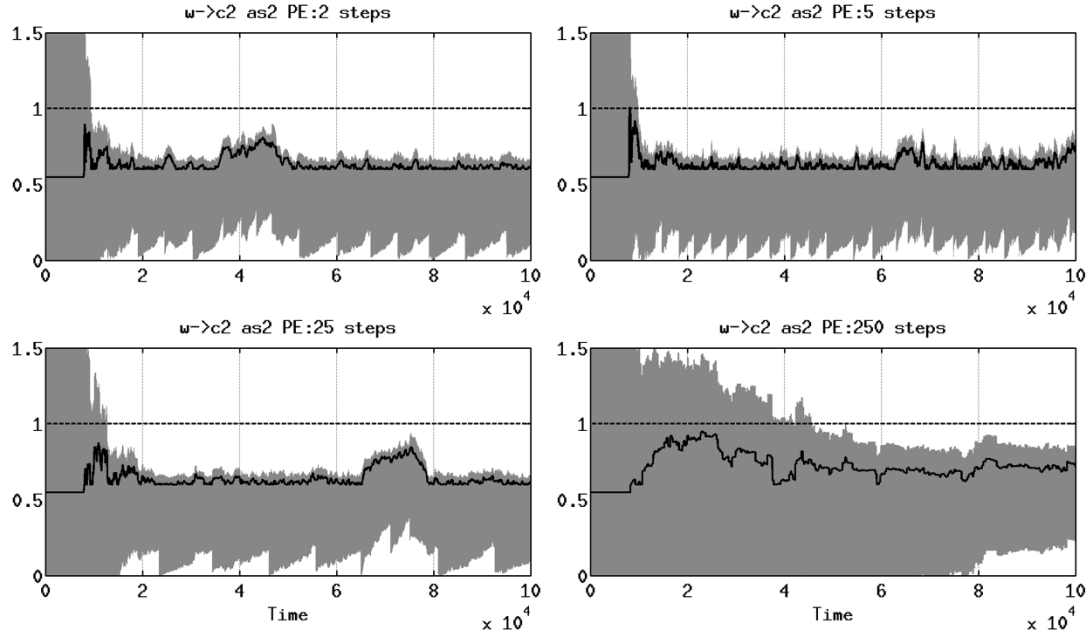
Figure caption: Time series of the estimated c_2 ensemble in the w-to-c_2 PE experiment with SE for w only, when the PE update interval is 0.02 TU (i.e. 2 time steps) (upper-left), 0.05 TU (upper-right), 0.25 TU (lower-left) and 2.5 TU (lower-right).

{2. Do the parameter spread and the amount of inflation need to be well tuned? How important are the choices for tuning the parameter spread and the amount of inflation? I suggest that the authors could link the parameter uncertainties to those appear in realistic coupled model, e.g. a_2 mimics the heat flux for atmosphere and c_2 mimics the windstress for ocean (also see the comment #3). }

RE: There are two considerations referring to the parameter spread and the amount of inflation. Firstly, as shown in the following figure, a smaller inflation level will enlarge period of the fluctuation of the black thick line. If the fluctuation period of the mean parameters is too long, then the effect is somehow similar to a slower convergence rate of the mean result during an arbitrary diagnostic window. Secondly, a too large inflation will cause the spread jump out

of the reasonable range. Within a relatively large scope, the inflation level will only change the spread of the parameter, but not change the mean of its ensemble. Because the mean value of the parameter is not as sensitive as the spread to the inflation level. It would not affect our main results in this paper.
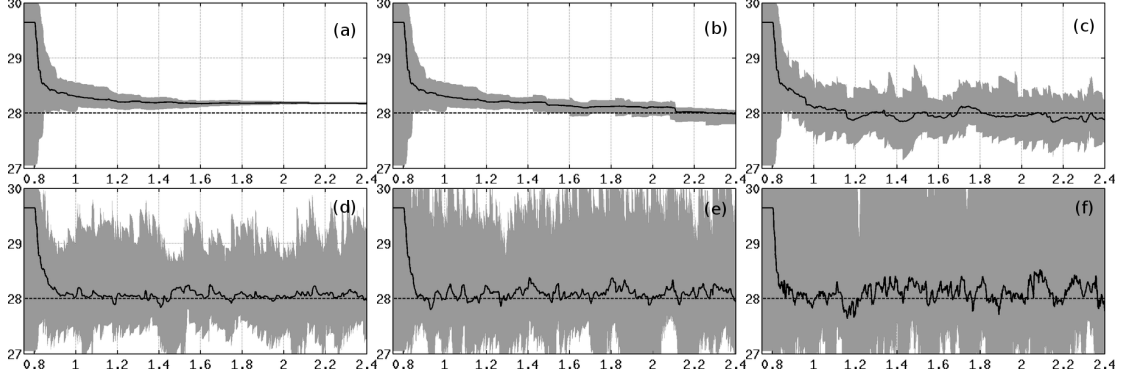

Figure caption: Time series of the estimated a_2 ensemble in the x_2-to-a_2 PE experiment with SE for x_2 only, the limited inflation value is 0.01(a), 0.05(b), 0.2 (c), 0.4(d), 0.6(e), 1.0(f). 0.2 is the value used in the paper.

{The uncertainties of parameters a_2 and c_2 (Fig. 2 and Fig. 3) are one-order different. Are they chosen on purpose? What are the averaged ensemble spreads for these two parameters? What happened if one chooses to remain a larger and same amount of uncertainty for these two parameters? }

RE: Thanks for the reviewer's thoughtful suggestion. We tried to test different initial bias combination at first, but sooner we found that in the original system, the value of a_2 is too much limited by the chaotic nature of the Lorenz equation (the a_2 can not be perturbed too much or the Lorenz equation otherwise will loose its chaotic nature and makes the experiment fail). We tried to avoid this by changing the system from two-way coupling to one-way coupling in section 3, see Appendix A. Although a_2 still cannot be changed too much, c_2 and c_6 can be changed in a wide scope. In such a circumstance, c_6 interacts with w and η, both being strongly forced by the periodic cosine function more than the Lorenz chaotic forcing. On the contrary, unlike c_6, no matter how periodic w is, c_2 is

always affected by the chaotic x_2. The experiments with varying S_s values give a lot of insights on this issue.

{If we can provide an unbiased a_2, could assimilation w-only lead to a successful parameter estimation for c_2?}

RE: For all of the experiments, only the parameter being estimated is biased from its truth. In experiment w-to-c_2, the a_2 is unbiased. And From Table 1, it clearly shows that even with an unbiased a_2, assimilation w-only will not lead to a successful parameter estimation for c_2. As in equation (1), the state variable w is calculated from c_2 and x_2. The x_2 is chaotic even with an unbiased a_2, therefore, the correlation between c_2 and w is disturbed by the chaotic x_2, and the correlation is not helpful during the estimation of the value of c_2 from the difference between w and "w observation."

{Page 4, Line 25: PE starts 40 TU later than SE. It should be clarified that the purpose is to constrain the accuracy of states (as stated at line13, Page 7). Why is it so important? }

RE: Yes, more discussions and justifications are added in the revision. Please see P6L1-4, P7L7-10.

{Compared with Fig. 2b and Fig.3b, the ensemble mean in Fig. 2c and Fig. 3c does not locate near the middle of the ensemble distribution after PE converges. Does this mean that the parameter ensemble distribution is skewed? Is there a particular reason for this result? }

RE: The thick black line indicates the ensemble mean of the parameter. As in Fig. 2b, 3b, 2c, 3c, all the thick black lines are near the referenced thin line enough to be called a successful PE. The difference between Fig. c and b is mainly about the asymmetry of the spread (shaded area). The asymmetry suggests that in the fully SE experiments (Fig. 2c, 3c) the distributions of the 20 ensemble members are not very Gaussian like. This actually exhibits advantage of the EAKF method

that it can "adjust" and keep the distribution of the ensembles. More description about EAKF are added in P4L1~10.

{3. The ensemble spread of the parameter a_2 seems to be less than 5% (and will be inflated when the spread is smaller than 0.6%). Is this realistic? In realistic setup of climate modeling, the uncertainties in the parameters associated with air-sea interaction (wind stress, heat flux) could be as large as 10%, in addition to bias in these parameters. The setup of the PE experiments may be too ideal to project the conclusion to realistic coupled modeling. In reality, there are several challenging issues in parameter estimation within atmospheric/ocean assimilation frameworks. However, such real and major obstacles cannot be explained by the results of the simple model.}

RE: The value of a_2 is too much limited by the chaotic nature of the Lorenz equation. But with the one-way coupling, we tried experiments with huge state fluctuations (different sets of S_s), the change of the forcing states is even bigger than in the real world. The main result of our paper still holds.

{In realistic parameter estimation using EnKF, how to construct a reliable error covariance between parameter and observation increments could be still challenging. In this simple model, one can easily perturb the parameter with the white noise without considering the characteristics of the horizontal structure. However, in reality, the structure of the ensemble perturbations of the parameter determines the pattern of the corrections away from the observations and how to keep a reasonable perturbation structure for parameters becomes a challenging task, especially for the parameters used in atmosphere model. }

RE: Thanks for the reviewer's comments. In this simple model, the construction of a reliable error covariance is indeed easier than in the real world. It is mainly because the observation is perfectly consistent to the model dynamics. In the real world, the structure is greatly geophysical dependent. The study is considered to be the first stepping-stone for further studies with more complex models. Therefore, we added new paragraph to fully discuss the limitation of our work in section 4, P11L19~27.

{So far, we may not have enough observation information for parameter estimation or constrain the parameter uncertainty (e.g. surface/near surface atmosphere observations that can reflect the air-sea interaction). I suggest the authors could provide some discussion about improving the accuracy of the atmosphere state for parameter estimation in real ocean modeling. What are the current limitations and what can be done? }

RE: As the parameter estimation in real ocean modeling can be very geophysical dependent, with our conceptual model, two things are suggested to be important for the further studies. The first one is our studies suggest that PE of oceanic parameter is possible to succeed with only the atmospheric observations. Considering there are also regions where the coupling effect is weak, adaptive measurements for different region seem important and necessary. Another suggestion is that the PE technique can be improved to perform separately at multiple-scales. All these require further research work to clarify. The discussion is rewritten in section 4, P11L28~P12L4.

Minor suggestions:

1. I suggest including the bias and root mean square error of the states and parameters in Table 1.

RE: Thanks for the reviewer's suggestion. A new table 2 was added in the revision showing the root mean square bias error of the state variable and the parameter during the last 100 TUs in 8 PE experiments.

2. Line 5: "tuning" procedure?

   RE: The sentence is rewritten.

3. Page 3, it will be easier for the readers if the authors can give a physical meaning for parameters a_2 and c_2.

   RE: Several lines elaborate the physical meaning of the two parameters are added in section 2.1, P3L25~28.

4. Page 6, Line 18: Shouldn't the zigzag shape mainly due to the update from assimilation of observations?

   RE: The zigzag shape is mainly due to the inflation process. The spread of the ensemble member is continuously shrinking during the PE process. After a while, when the std (spread) of the parameter ensemble is below some limit (40% of its initial spread), we inflate the ensemble by multiply a constant factor to the parameter anomalies to satisfy this STD value. More accurate description is added in P5L2.

   Sometimes the constant mag factor will be used several times for the spread to go beyond the limit. The mean value of the parameter is not sensitive to this factor. After the multiplication, the spread will generally be higher than the limited value (form the zigzag shape) to make sure the inflation would not immediately happen again.

5. I suggest that some paragraphs can be clarified or re-arranged. The first paragraph in Section 3 is somewhat confusing. I suggest starting from Table 1 and explain the differences among the experiments. Is the experiment

mentioned for Fig. 2a and Fig.3a (both atmosphere SE and ocean SE) included in Table 1?

RE: Thanks for the suggestion. The first paragraph in Section 3 (P6L6~14) has been rewritten. The experiments of Fig. 2a and Fig. 3a are not included in Table 1. The main purpose of this paper is to discuss the 8 PE experiments with partial SE. We shown the coupled SE experiment as a standard reference level for the partial SE cases. They are not very relevant to our main purpose.