Point-by-point responses for review #1

In general:

A few conferences were held for authors to discuss the comments of 3 reviewers. All comments of reviewer 1 have been fully addressed in the revision. The first authors learned a lot from such commented points, and here the authors want to express great appreciation for the advice of the reviewer. The paper is revised 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 1:

General comment:

The manuscript investigates the impact of observational constraints, through data assimilation methods, on coupled model state and parameter estimation using a conceptual 5-variable model. I found the manuscript interesting and appropriate for the journal, especially to fill the existing gap of idealized studies in coupled data assimilation experiments. It is somehow less relevant, in my opinion, for parameter estimation, given the complexity of real-world CGCM, as the authors themselves discuss in the Conclusions.

I recommend the manuscript for publication after a few issues are considered by the authors, especially to improve the readability for a general readership.

RE: Thanks for your encouragement. All issues are replied point-by-point as below. We hope the whole manuscript has been essentially improved.

1. I think the title itself "Further insight" refer to a previous paper from the author ("further" with respect to what?) and might be simplified to "Insights on" or "On the role..."

RE: The title is changed to "Insights on the Role of Accurate State Estimation in Coupled Model Parameter Estimation by a Conceptual Climate Model Study".

2. There is some literature missing that can be added: for instance

i) the parameter estimation problem (Introduction, lines 1-10) may be approached also with adjoint techniques, and I recommend the authors to mention this alternative methodology;

RE: Three different choices for the parameter estimation including the adjoint techniques and related references are added in P2L10~11.

ii) in the description of twin experiments with perturbations (P4L1-6), there are many analogies with OSSEs (Observing system simulation experiments) that can be mentioned as well.

RE: Thanks for this suggestion. Our twin experiment is a kind of OSSEs. This and references of other PE under OSSEs are added in P5L11~12.

3. The authors often refer to simple climate/coupled model. I suggest them to always use the definition of "conceptual model" as it can hardly be considered a climate model

RE: Yes, done. Thanks.

4. {The reader is too much referred to literature in the Methodology section. For instance, I had to understand only through referred papers

i) the size of the conceptual model of Eq. (1) is never discussed (is it a single-column model or a limited-size model? What are the boundary conditions of the problems, if any?)}

RE: Our model is a low-order (limited-size) conceptual model. The boundary condition is a predefined seasonally-varying solar radiation S(t)=Sm+Ss COS(2*pai*t/Spd), which is a simple and idealized approximation of the real world boundary condition. New lines of this introduction are added in P3L22~25.

{ii} little is said about the EAKF, which might be better introduced from a theoretical point of view and in terms of advantages/disadvantages w.r.t. other filters and data assimilation methods. I guess the authors choose it for its ease in the parameter estimation, but this can be better clarified}

RE: Thanks for the suggestions. More details about the EAKF method including its advantages/disadvantages was added as an independent paragraph on the beginning of section 2.2, P4L1~10.

{iii) for such a small size problem, a 20-member ensemble size appears quite small without reason. Clearly the problem size is small, but it is worth mentioning sensitivity tests performed on the ensemble size.}

RE: Thanks for the suggestion. We performed sensitivity tests on the member size. The result is shown in the following figure. Generally speaking, the RMS error of the mean parameter will increase when lowering the ensemble size. But it can also be clearly seen from the panel (d) that no matter how big the ensemble size is, the result with an oceanic SE is unacceptable. We think size 20 is enough for showing the difference between the successful and the failure cases. We added new lines to clarify this problem in P4L24~28 and P7L18~20.

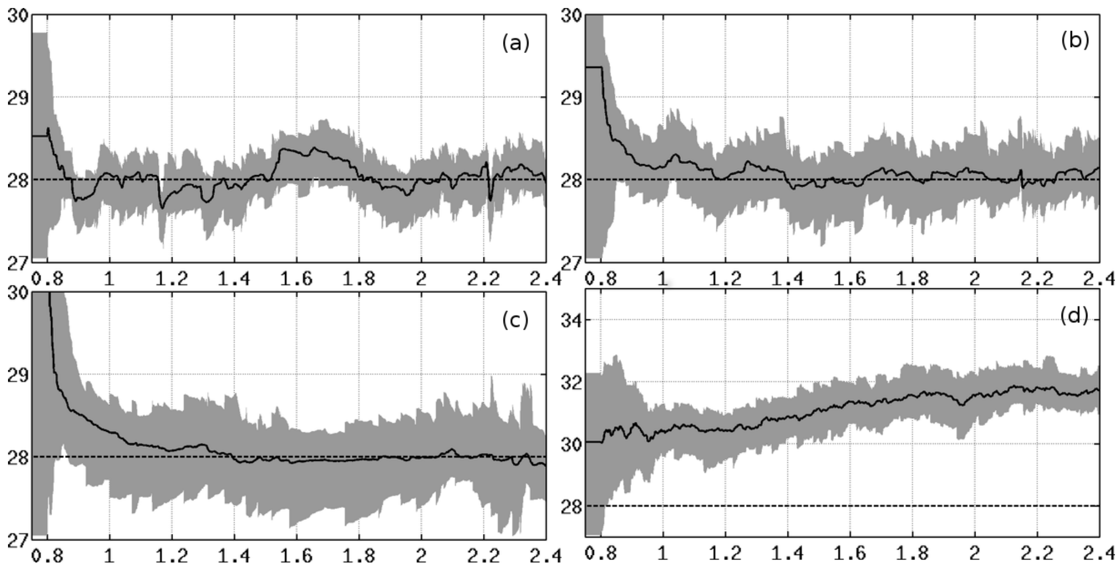

Figure caption: Time series of the parameter in 4 test cases with different

ensemble size settings. SE x_2, PE x_2 to a_2 (abc), SE w, PE w to a_2 (d). The ensemble size is 5 (a), 10 (b), 40 (c) and 40 (d) respectively.

5. {I found the conclusion in P7L3.9 on preferring atmospheric to ocean observations to determine ocean parameters very dependent on the conceptual model the authors use. First, some parameters (c2) are not ocean parameters but coupling parameters, strictly speaking;}

RE: Thanks for this comment. The c_2 is more like a coupling parameter than pure ocean parameter. Therefore we performed experiments about c_6 as an complementary (Fig. 5). The necessity of an atmospheric SE still holds.

{second, the "first guess" of the ocean parameters themselves, determining time scales and interactions, may not necessarily represent the real world; }

RE: The conceptual model cannot fully represent the real world. But it has great advantage for clarifying the PE problem without sufficient observations. The parameters are set to simulate the parameterization of CGCM. We added more description and discussion about the simple model parameters in P3L25~28.

{third, the observing network that observe ocean and atmosphere state may be not representative of the real observing networks. I would mention the limits of the conceptual model rather than emphasize this conclusion.}

RE: Thanks for the reviewer's suggestion. The real world observation generally has strong temporal and geographical dependency. The real data are always with all kinds of incomplete. All these flaws motivate us carrying out this partial SE research at the first place. Some of them cannot be represented in our model because this conceptual model does have its limits though its dynamics and transferring of the uncertainty is crystal clear. Following the suggestion, we added new lines to discuss these limits in section 4, P11L19~27.

6. Since Section 3 contains a lot of information and experiments, I suggest to add a paragraph between the 1st and 2nd paragraph of Section 4 to summarize some results from the experiments on individual/combined state and parameter estimation.

RE: A new paragraph summarizing all experiments and the direct results was added as the reviewer's suggestion in section 4, P11L4~11.

Language issues

{weak coupled → weakly coupled (P4L10 and further occurrences) }

RE: Done. Thanks.

P4L21 "And also considering...visualization" sounds very awkward

RE: The sentence was rewritten to "Therefore we set the ensemble initial values of a_2 as a Gaussian distribution N(30, 1) (30 as the mean value and 1 as the standard deviation), and the spread is enough for the model ensemble uncertainty. The ensemble initial values of c_2 are set as N(0.8, 0.5)." as in P5L30~32. Thanks.

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.

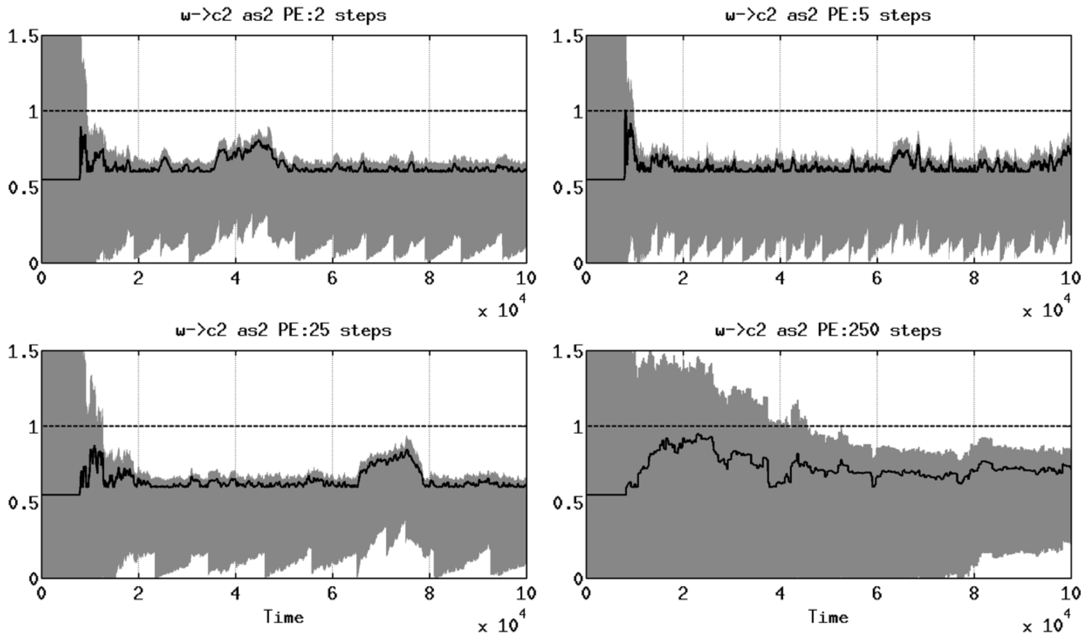

Figure caption: Time series of the estimated c_2 ensemble in the w-to-c_2 PE experiment with SE for w only, when the PE update interval is 0.02 TU (i.e. 2 time steps) (upper-left), 0.05 TU (upper-right), 0.25 TU (lower-left) and 2.5 TU (lower-right).

{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.

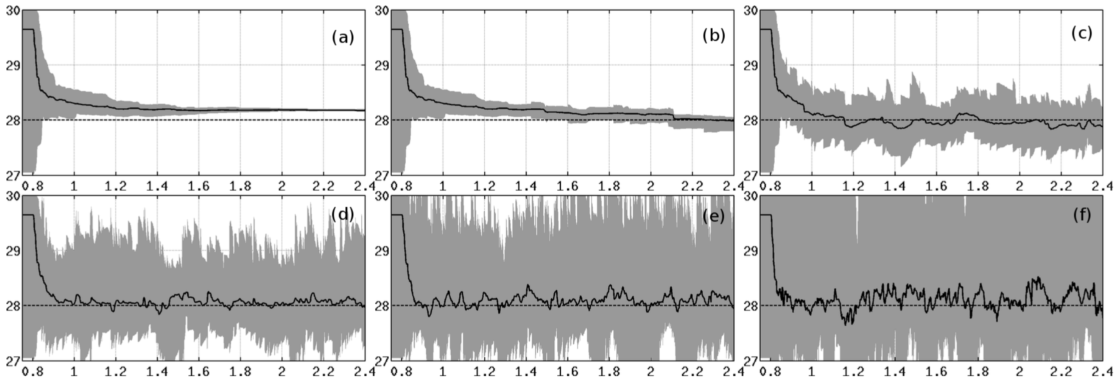

Figure caption: Time series of the estimated a_2 ensemble in the x_2-to-a_2 PE experiment with SE for x_2 only, the limited inflation value is 0.01(a), 0.05(b), 0.2 (c), 0.4(d), 0.6(e), 1.0(f). 0.2 is the value used in the paper.

{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.

Point-by-point responses for review #3

In general:

All authors appreciate greatly for the encouragements and comments from reviewer 3. Coordinating with addressing the comments of reviewers 1 & 2, all points that reviewer 3 concerns have been fully addressed in the revision. 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 3:

General comment:

This paper investigates how the model parameter estimation works in an EnKF for an atmosphere-ocean coupled system. This study performs a series of parameter estimation experiments using a low-order, toy system based on the famous Lorenz-63 three-variable model but with an extension of additional near-surface and deep ocean components. The results are somewhat interesting that the fast atmospheric component's state estimation plays a key role in the parameter estimation problem both for the ocean-atmosphere coupling coefficient c2 and the internal dynamical parameter a2 for the second atmospheric variable x2. I find the topic of parameter estimation stability jointly with state estimation stability is very interesting, and this paper is a useful contribution in the field, although could be done better. I find the value of publishing this article, but I found some issues that need to be addressed before final publication as below:

RE: Thanks for your encouragement. All issues are replied point-by-point as below. We hope the whole manuscript has been essentially improved.

1. There are a number of grammatical errors, which need to be corrected.

RE: A few rounds of reading/editing from native English speaker were conducted. The grammatical errors were fixed. Thanks.

2. {"Signal-to-noise" of the ensemble-based error covariance between the states and parameters appears repeatedly, but there is no direct investigation about it. Since this study performs idealized toy-model experiments, I would assume that the authors may find a better way of investigating and presenting the signal-to-noise more explicitly.}

RE: Thanks for the suggestion. The signal-to-noise ratio of the ensemble-based error covariance between the states and parameters is better to be diagnosed in SE only experiments. In these runs, there are no PE processes to fix the biased parameter spread so that parameter perturbations can be fully transferred to the model states. Then the state-parameter covariance can be checked without any disturbance from a PE correction. Following the previous work of Zhang et al. (2012), we defined a new index (called r_{s2n}) to measure the signal-to-noise ratio of the ensemble-based error covariance between the states and parameters. The best (worst) representation of the signal-to-noise ratio is characterized by a r_{s2n} value of 1(0). A new Table 3 is added in the paper. It shows all r_{s2n} and related values in the 8 SE only (no PE) experiments. Description of this index and related discuss is added in P9L16~27.

3. {P.7, L.7-9, "Here our results suggest that in a coupled system, to determine oceanic coefficients, it is more important to get more atmospheric measurements to constrain the atmospheric states than to get more oceanic measurements to directly apply to oceanic PE." This is an interesting hypothesis inspired by the simple toy model results, but this statement seems to be an overgeneralization. The real coupled atmosphere ocean system is much more complicated than the two-time-scale toy system with only 3 atmospheric and 2 oceanic variables. This statement should be a hypothesis or speculation at this point.}

RE: The sentences are rewritten as in P8L10~12. Thanks.

4. { 4. P.7, L.21-22, “reducing x2 uncertainty is critical”, I do not find this statement well supported or proven by the experimental results. This statement seems to be a hypothesis or speculation.}

RE: The statement was changed to “Instead, reducing x2 uncertainty (enhancing the estimation accuracy of the atmospheric states) is more relevant to the solution of the problem.” (P8L24~25)

Minor comments:

1. {Eq. (2) does not contain observation error statistics, and I am curious how to interpret this equation intuitively. I understand that this equation gives analysis increments for the ith ensemble member. The analysis increments should balance between the observation error and background error. This equation has only the background error variance in the observation space as the denominator, but does not contain the observation error variance which usually appears in the data assimilation equations as an R matrix.}

RE: The observation error variance is calculated before this projection process. The observation are firstly compared to their simulated values, the difference between them are manipulated to produce the observational increments. The production of the observational increments considers the observation error variance and its PDF. The observation error is set as a constant number in our simulation. The standard deviation of “observational” errors are 2 for the atmospheric variables $x_{1,2,3}$ and 0.2 for the oceanic variable w. New introduction of the EAKF method is added in the section 2.2, P4L1~10.

2. P.6, L.30, eta-to-c6 PE suddenly appears here, without any description about observations for eta (deep ocean state variable). Section 2.2 described only x2 and w observations, and the readers would assume the experiments use only x2 and w observations.

RE: Thanks for the reviewer's suggestion. Description about observations for eta is added in the new η-to-c_6 section (P7L31~34). It directly points out that the experiments uses η only for the PE and uses all state variables for the SE.

Insights on the Role of Accurate State Estimation in Coupled Model Parameter Estimation by a Conceptual Climate Model Study

Xiaolin Yu[1], Shaoqing Zhang[2], Xiaopei Lin[1,3], Mingkui Li[1]

[1]Physical Oceanography Laboratory of OUC, and Qingdao Collaborative Innovation Center of Marine Science and Technology, Qingdao, 266001, China
[2]GFDL/NOAA, Princeton, NJ 08542, USA
[3]Function Laboratory for Ocean Dynamics and Climate, Qingdao National Laboratory for Marine Science and Technology, Qingdao, 266001, China

Correspondence to: Shaoqing Zhang (Shaoqing.Zhang@noaa.gov)

Abstract. The uncertainties in values of coupled model parameters are an important source of model bias that causes model climate drift. The values can be calibrated by a parameter estimation procedure that projects observational information onto model parameters. The signal-to-noise ratio of error covariance between the model state and parameter being estimated directly determinates whether the parameter estimation succeeds or not. With a conceptual climate model that couples the stochastic atmosphere and slow varying ocean, this study examines the sensitivity of state-parameter covariance on the accuracy of estimated model states in different model components of a coupled system. Due to the interaction of multiple time scales, the fast varying “atmosphere” with a chaotic nature is the major source of state-parameter covariance uncertainties, and thus enhancing the estimation accuracy of atmospheric states is very important for the success of coupled model parameter estimation, especially for the parameters in the air-sea interaction processes. The impact of chaotic-to-periodic ratio in state variability on parameter estimation is also discussed. This simple model study provides a guideline when real observations are used to optimize model parameters in a coupled general circulation model for improving climate analysis and predictions.

1 Introduction

Nowadays, a coupled atmosphere-ocean general circulation model is widely used as a common tool in climate research and related applications. However, due to the approximation nature of model numeric schemes and physical parameterization, a model always has errors. In particular, one traditionally determines the values of model parameters by experience or a trial procedure which heuristically provides a reasonable estimate but usually not optimal for the coupled model. Recently, with the aid of information estimation (filtering) theory (e.g. Jazwinski, 1970), researches on optimization of coupled model parameters based on instantaneous observational information have grown quickly (e.g. Wu et al., 2013; Liu et al., 2014; Liu et al., 2014; Li et al., 2016). Traditional data assimilation that only uses observations to estimate model states (i.e. state estimation) becomes both state estimation (SE) and parameter estimation (also called optimization) (PE) with observations. Such a PE process can be implemented through a variational (adjoint) method (e.g. Stammer, 2005; Liu et al., 2012), or an ensemble Kalman filter (e.g. Zhang et al., 2012), or even a direct Bayesian approach (e.g. Jackson et al., 2004).

In the previous study with a conceptual coupled model, Zhang et al. (2012) pointed out that an important aspect of successful coupled model parameter optimization is that the coupled model states must be sufficiently constrained by observations first. This is because multiple sources of uncertainties exist in a coupled system consisting of different time scale media. If the part of uncertainties in model states, which is not correlated with parameter errors, has not been sufficiently constrained yet, the covariance between the model states and parameters being estimated is noisy (e.g. Dee & Silva, 1998; Dee, 2005; Annan et al., 2005). Without direct observational information, the noisy state-parameter covariance, the key quantity to project observed state information onto the parameter, can bring the estimated parameter toward an erroneous value (Zhang et al., 2011b). This is a general understanding about coupled model parameter estimation. However, since multiple media of the climate system have different time scale variability and different quality of observations so as to have different contributions to the uncertainty of state-parameter covariance, an outstanding question is: what is the impact of SE accuracy in different media on coupled model PE? Given the extreme importance of state-parameter covariance for PE, a clear answer for this question must further our understanding on coupled model parameter estimation.

To answer this question, here we use a simple coupled model to examine the influence of observation-constrained states in each medium on PE for different parameters in different media thoroughly. The model conceptually describes the interactions of 3 typical time scales of the climate system – chaotic (synoptic) atmosphere, seasonal-interannual upper ocean and decadal deep ocean. A twin experiment framework is used throughout the whole study.

The paper is organized as follows. After introduction, section 2 gives methodology, including brief descriptions on the simple coupled model, filtering algorithm and twin experiment framework. Section 3 first presents the results of the various PE experiments with different partial SE settings, and then analyses the conditions for successful PE with partial SE. Finally, summary and discussions are given in section 4.

2 Methodology

2.1 The model

To clearly address the issue posed in the introduction, this study employs the simple pycnocline prediction model developed by Zhang (2011ab). This conceptual coupled model is based on the Lorenz's 3-variable chaotic model (Lorenz, 1963) that is coupled a slab ocean variable (Zhang et al., 2012) interacting with a pycnocline predictive model (Gnanadesikan, 1999). For the problem that is concerned, this conceptual coupled model shares the fundamental features with a coupled general circulation model (CGCM) (see Zhang, 2011a; Han et al., 2013). The model development can be traced in Zhang (2011ab) and Zhang et al. (2012) in details. Here, we only comment on major points that are relevant to this study. The model includes 5 equations:

$$\begin{aligned}
&\dot{x}_1 = -a_1 x_1 + a_1 x_2 \\
&\dot{x}_2 = -x_1 x_3 + (1 + c_1 w) a_2 x_1 - x_2 \\
&\dot{x}_3 = x_1 x_2 - b x_3 \\
&O_m \dot{w} = c_2 x_2 + c_3 \eta + c_4 w \eta - O_d w + S_m + S_s \cos(2\pi t / S_{pd}) \\
&\Gamma \dot{\eta} = c_5 w + c_6 w \eta - O_d \eta
\end{aligned} \quad . \qquad (1)$$

The first 3 equations (Lorenz's 3-variable chaotic model) represent the dynamics of the atmosphere. The last 2 equations respectively represent the dynamics of the slab upper ocean and the pycnocline depth variation of deep ocean. There are 5 variables in the model. x_1, x_2 and x_3 are the fast-varying variables of the atmosphere with the parameters a_1, a_2, b set as 9.95, 28, and 8/3, that sustain the chaotic nature of the atmosphere. w and η are the low-frequency variables of the ocean. Equation (1) tells that the ocean in this system is driven by two kinds of forcings: the chaotic x_2 from the Lorenz equations and the periodic cosine function term serving as external forcings in the equation of w. The coupling parameter c_2, which interacts with the chaotic forcing x_2, is set as 1.0. Values of other parameters of (c_1, c_3, c_4, S_m, S_s, S_{pd}, Γ, c_5, c_6) are (10^{-1}, 10^{-2}, 10^{-2}, 10, 1, 10, 10^2, 1,10^{-3}) can be found in the literature cited before. With the parameter O_m/O_d = 10.0 (O_m = 10, O_d = 1), the time scale of the w is nearly 10 times of the time scale of the x_2. From the equation, it can be seen that the parameter a_2 has a direct influence on the variation of the state variable x_2. And the parameter c_2 has a direct influence on the variation of the state variable w. The estimation of these two parameters will be used later to interpret the relation between the accuracy of SE and successfulness of PE. Although very simple, this low-order (limited-size) conceptual model mimics very fundamental natures of interactions of three typical time scales in the real world: synoptic (chaotic) atmosphere, seasonal-interannual upper tropical oceans and decadal/multi-decadal deep ocean (Zhang 2011b). The boundary condition is a predefined seasonally-varying solar radiation $S(t)=S_m+S_s \cos(2\pi t/S_{pd})$. The state variable w mimics the surface temperature of the ocean and the x_2 mimics the surface wind of the atmosphere. Here we may mimic some parameterization of CGCM using the relation of parameters and model variables (c_2 and x_2 analogous to the drag coefficient c_d and wind for the stress on ocean, for instance).

2.2 Filtering scheme

The filtering method used in this study is the ensemble adjustment Kalman filter (EAKF, Anderson, 2001). The EAKF algorithm shares all theoretical derivation of ensemble Kalman filter (EnKF, e.g. Evensen, 1994; Houtekamer et al., 1998) that combines observational probability distribution function (PDF) with model PDF but under an adjustment idea. The algorithm can be sequentially implemented in a two-step procedure (Anderson, 2003): step 1 uses two Gaussian convolution to compute the observational increment at the observational location, and step 2 regresses the observational increment onto the related model grids or parameters by the model ensemble-evaluated covariance. While the sequential implementation provides much computational convenience for data assimilation, the EAKF maintains the nonlinearity of background flows as much as possible (Zhang and Anderson, 2003; Zhang et al., 2007). Like usual EnKFs, the EAKF has the disadvantage on dealing with model errors (also existing in variational methods without a model error compensation term).

SE and PE processes in this study are all based on the two-step EAKF implementation (see Zhang et al., 2007). In both SE and PE, the ensemble observational increment is first computed from the difference between the model simulating result and the observation. Then the ensemble observational increments are projected onto model states or/and the parameters by the following equation:

$$\Delta p_i^u = \frac{\mathrm{cov}(\Delta p, \Delta y)}{\mathrm{std}(\Delta y)^2} \Delta y_i^o, \quad \text{(i=1~N, N is the ensemble size)} \tag{2}$$

The linear regression Eq. (2) is built with the help of the 20-member ensemble, a member of the ensemble square root filter family (Tippett et al., 2003). Δy_i^o is the observational increment for the ith ensemble member. Δp_i^u is the state (parameter) increment to update the ith ensemble parameter. cov(Δp, Δy) is the error covariance computed between the model variable ensembles at the model grid and observational location (for SE) or between the ensembles of the state variable and parameter being estimated (for PE). std(Δy) is the standard deviation of the ensemble of state variable at the observational location. For example, when using x_2 to estimate c_2, on each estimating step, the ensemble of x_2 and the ensemble of c_2 are used to calculate the ratio of cov/std^2 and adjust c_2 toward a better value that minimizes the errors of model states from the observations.

Some other relevant aspects of the method are also commented here. Same as in Zhang and Anderson (2003), based on the trade-off between cost and assimilation quality, after a series of sensitivity tests on ensemble sizes of 10, 20~100, no significant difference on the quality of standard assimilation is found when the ensemble size is greater than 20. Thus a practical ensemble size of 20 (applicable for a CGCM) is chosen as a basic experiment setting. We will examine the sensitivity of major conclusions to the ensemble size in related places later. Although the intervals of the atmosphere and ocean observations are different in the real world, for convenience of comparison, we set a uniform update interval for SE (in the atmosphere and ocean) and PE as 0.05 TU (i.e. 5 time steps) in this study (we will also discuss the influence of update intervals in related places later). The inflation method must be included in the EAKF PE. Considering that the inflated parameter ensemble will influence on state variables, no inflation is applied to the model state ensemble. The PE inflation scheme follows Zhang (2011b): when the std (spread) of the parameter ensemble is below some limit (40% of the initial spread), a factor is applied to inflate the parameter ensemble spread to this value. During this process, the ensemble structure of parameters remains unchanged.

2.3 Twin experiment setup

Twin experiments are set to test the relation between coupled SE and PE. The model with the standard parameter is running 10^3 time units (TUs) after a spin-up of 10^3 TUs (2×10^3 TUs in total). Here a TU is a dimensionless time unit as defined in Lorenz (1963), roughly referring to the time scale of atmosphere going through from an attractive lob to the other (1 TU equals 100 steps of the model integrations with a Δt of 0.01). The output of last 10^3 TUs is then used as the "truth" to produce "observations" by superimposing a white noise on the "observed" variables. The standard deviation of "observational" errors are 2 for the atmospheric variables $x_{1,2,3}$ and 0.2 for the oceanic variable w. It is worth to mention what we describe here is a kind of observing system simulation experiments (OSSEs, e.g. Tong and Xue, 2008; Jung et al., 2010). The assimilation model control is an ensemble of integrations for each test case with the perturbed parameters on an erroneously-set parameter value (will be described later). The initial conditions of the ensemble for assimilation are taken from the end of 10^3-TU spin-up (the different members in the ensemble are all resulted from the parameter perturbation).

The different PE experiments can be distinguished in 3 perspectives: 1) 2 state constraint settings (i.e. SE settings) that assimilate the atmosphere "observations" (x_2) only, and the ocean "observations" (w) only; 2) 2 parameter settings – a_2 in the atmosphere equation, and c_2 in the ocean equation; 3) 2 observational settings – one atmosphere (x_2), and one ocean (w). Here the SE uses weakly coupled data assimilation as termed in the literature (e.g. Lu et al., 2015), i.e., x_2 observations impact on all x variables, and w (η if applicable) observations impact on w (η) itself (considering the different time scales of w and η, no cross-impact between them), while the PE could use different medium observations. Therefore, eventually we have a few PE cases with full SE – both x and w are constrained by their observations, and particularly 8 PE cases with partial SE – only some medium is constrained by its observations. These PE cases have different SE accuracy. Through thoroughly analysing these PE cases, we are able to detect the influence of the SE accuracy in different medium on coupled model PE.

In these PE cases, the initial value of the parameter to be estimated is deliberately set biased from the "truth" (referring to the standard parameter values described in section 2.1). To maintain the chaotic nature of the Lorenz equation, parameter values are required being within a certain range. This is a constraint for the biased amount of the initial values of a parameter. Based on some sensitivity studies, the chaotic performance is more vulnerable to the change of the atmospheric parameter a_2 than to the change of the oceanic parameters. Therefore we set the ensemble initial values of a_2 as a Gaussian distribution $N(30, 1)$ (30 as the mean value and 1 as the standard deviation), and the spread is enough for the model ensemble uncertainty. The ensemble initial values of c_2 are set as $N(0.8, 0.5)$. With the ensemble size of 20, the actually used ensemble mean values of the initial a_2 and c_2 from sampling are 29.64 and 0.56 respectively. If PE is successful, then the ensemble mean value of a_2

(c_2) should converge to 28 (1). In the 8 PE experiments, SE starts at the 40th TU while PE starts at the 80th TU in the total 10^3 TUs of assimilation period described above. The early SE constrains the model states close to the observations so as to enhance the parameter-state covariance for the coupled PE (Zhang et al., 2012). The delayed time scale of PE from SE will be discussed later .

3 Impact of SE accuracy on coupled model PE

With the method described in section 2, we designed 8 experiments to test different PE performances under different SE settings in the conceptual coupled model. Generally with a full SE (all the atmospheric $x_{1,2,3}$ and oceanic w states are estimated with the "observations" that sample the "truth"), the PE is steady and successful, no matter what observations are used to estimate which parameter. For example, the result of using observations of w (in the ocean) to estimate a_2 (parameter in the atmosphere) with all $x_{1,2,3}$ and w being estimated by their observations is shown in Fig. 1a. We can see the ensemble of a_2 successfully converges to the "truth" from the initial biased values around 30. However, if only a part of observations (only one medium observations) is used in SE, then the PE succeeds in some cases but fails in others (Fig. 1b). All the 8 cases are summarized in Table 1. Next, we will analyse and discuss these experiment results to understand the role of different medium SE on coupled model PE.

3.1 Stability, reliability and convergent rate of PE with partial SE

In Table 1, "X-to-Y" means using observations of "X" to estimate the parameter "Y" ("x_2-to-a_2" means using observations of x_2 to estimate parameter a_2, for instance). Table 1 shows that all 4 PE cases with atmospheric SE succeed while all 4 PE cases with oceanic SE fails, no matter what medium observations are used to estimate which medium parameter. The root mean square error (RMSE) of the model states and parameters during the last 100 TUs are shown in Table 2. The x_2 RMSE in the failed cases are higher than in the successful cases, while the w RMSE in two failed cases F(5) and F(6) are even smaller than the ones in successful cases. Also, in both cases of S(2) and F(6), a_2 is estimated by w observations but only when the x states are constrained by the x observations, the PE is successful, or otherwise the PE is failed although the state w is constrained by the w observations. These suggest that the uncertainty of x_2 is mainly responsible for the failure of the PE. An example of failed PE in which the observations of w are used to estimate a_2 is shown in Fig. 1b. We can see the ensemble of a_2 in Fig. 1b cannot converge to its "true" value of 28. We will thoroughly analyse such failed cases next.

The stability of PE is different among partial SE settings as shown in Figs. 2 and 3 as the time series of the ensemble mean of the estimated parameters. Figs. 2bc and 3bc show the 4 successful cases with only atmospheric SE. Compared to full SE (using observations of x_2 and w, shown in Figs. 2a and 3a), the partial SE cases show much bigger fluctuation in estimated parameter values at the beginning of spin-up period (Figs. 2bc and 3bc). From Figs. 2 and 3, it can also be seen that generally the accuracy of PE with partial SE is lower although overall the estimated parameter values converge to the truth. This can be comprehended by the lower signal-to-noise ratio of state-parameter covariance provided by the SE process, which will be discussed more details at the end of this section.

The convergence rate of PE is also obviously different with different SE settings. The case of w-to-a_2 converges much more slowly than the other cases in a_2 estimation. This phenomenon can be explained by the different time scales in different media. Figure 4 shows the variation of the state variable during SE. The observational constraint makes the mean value and the whole ensemble to follow the "truth" (see Fig. 4a for x_2 and Fig. 4e for w). It can be seen that in cases assimilating x_2, due to no direct constraint on w and η, their spread shrinks slowly. Instead they are forced by the constrained x_2 but with slower adjustment of ocean processes. As mentioned in section 2.3, the SE starts before the PE to make sure the state needed is constrained enough. Slowly shrinking of w and η spreads shall be considered in determining a longer delayed time for the PE related to w and η.

The inflation method is also important in PE (Yang & DelSole, 2009; DelSole & Yang, 2010; Zhang, 2011ab; Zhang et al., 2012). The partial and full SE cases are with the same inflation scheme (Zhang, 2011ab; Zhang et al., 2012). Shadows in Figs. 1-3 show the range of the ensemble parameter. The zigzag shapes of the shadows represent the inflation during PE. In these figures, the width of the shadows shrinks quickly once PE is activated while some of the mean parameter values move toward the "truth" slowly (for example, Fig. 2c and Fig. 3b). Also from the shapes of zigzag, we can see some inflation happens before the parameter converging to the "truth." All these imply that the designed PE is stable and its convergence rate is not much sensitive to the inflation scheme.

In addition, larger ensemble sizes are used to test the sensitivity of the above conclusion. The results show that although a large ensemble size has a positive impact on SE and PE quality but the drawn conclusion from the experiments above does not change its essence with a larger ensemble size. We also performed the experiments under different update interval settings. Test results show that for the issue we are addressing, the conclusion is not sensitive to the update interval if it is within a reasonable range.

In cases 3 and 4, we successfully estimate the oceanic parameter c_2, suggesting we can use different medium measurements to help calibrate the parameter within a coupled model. In case 3, the atmospheric observations are used for both SE and PE, while in case 4, the atmospheric observations are used for SE but the oceanic observations are used for PE. The case 3 uses only the atmospheric observations to determine an oceanic parameter and does a better job than using the oceanic observations does in case 4.

The phenomenon above in estimation of c_2 can be comprehended by the "air-sea" interaction process. What about a pure oceanic parameter (a parameter used for deep ocean, for instance)? It is interesting to see the influence of atmospheric SE accuracy on PE for a deep ocean parameter. To do that, a series of η-to-c_6 PE experiments with different SE settings is carried out. The deep ocean observation is generally sparse in the real world. But within our twin experiment framework described in section 2.2, the "observations" of η used for our PE can be produced as sufficient as other variables. All PE experiments on c_6 are conducted with η observations (observations of x_2 and w are only used in different SEs but not used in the PE). The result is shown in Fig. 5. Given the long time scale of η, the η PE experiments are extended to 10^4 TUs. The PE cases include 4 SE settings: case-1: all state variables, case-2: $x_{1,2,3}$ only, case-3: w and η, and case-4: η only. Both case-1 and case-2 succeed greatly, but the convergence rate of case-1 is faster than case-2 and the accuracy of case-1 is a little higher than case-2. In case-3, the convergence rate is fast but the estimated values remain in a bias from the truth. Case-4 apparently fails, never stably converging to any value. It is clear that the η-to-c_6 PE succeed only when the atmospheric state is constrained by observations.

It is interesting that once the atmospheric states (the Lorenz equation in this simple model) are constrained by the observations, both the atmospheric parameter (a_2) and oceanic parameters (c_2 and c_6) can be successfully estimated even in the case using the atmospheric observations (x_2) to estimate the oceanic parameter (c_2) or using the ocean observations (w) to estimate the atmospheric parameter (a_2). This seems different from our previous intuition that in-situ ocean data are always considered as the first important piece of information for determining the oceanic coefficients. Our results here strongly suggest that in the future real coupled model PE experiments, for determining the best coefficient values, no matter the atmospheric or oceanic, sufficient and accurate atmospheric measurements are crucially important. Next we will conduct more sophisticate analyses to extend our understanding on this point.

In our twin experiment setting, there are 3 types of model uncertainties: strong nonlinearity in the atmosphere (chaotic in this case), weaker nonlinearity in the ocean and biased parameter values. The SE process before PE aims to control the first and second types of the uncertainties by observational constraints on model states. Figure 6 shows the wavelet analyses for the atmospheric variable x_2 and the oceanic state variable w in the "truth" run. They represent the uncertainties of type 1 (panel a) and type 2 (panel b). With the expanded exhibition of wavelet on different periods, Fig. 6 clearly tells significantly different features of x_2 and w. The energy of x_2 is in the high frequency band and the energy of w is in the low frequency band. x_2 varies fast and represents the most uncertain mode, transferrable to low frequency w through the "air-sea" interaction. Later in section 3.2, we will show that the feedback of ocean can magnify the role of atmospheric chaotic forcings. The chaotic nature can spread out and results uncertainties in all frequency band in the system. Under such a circumstance, the method of picking a particular frequency (e.g. Barth et al., 2015) or using averaged covariance (Lu et al., 2015) to implement PE cannot essentially resolve the issue although it may relax the problem. Instead, reducing x_2 uncertainty (enhancing the estimation accuracy of the atmospheric states) is more relevant to the solution of the problem.

Without direct observations on parameter values, PE completely relies on the covariance between the parameter and model states for projecting the observational information of states onto the parameter. While the PE projection is carried out by a linear regression equation based on the state-parameter covariance (EnKF/EAKF, for instance), only a linear or quasi-linear relationship between parameters and states in ensemble is recognized. A hypothesis for all the failed cases without direct atmospheric SE could be that, under such a circumstance, the chaotic disturbances in the atmosphere (Lorenz equations in this case) continuously interacting with the parameter make difficulties for the system to build up a quasi-linear relationship between the state variable and the parameter.

To investigate the parameter-state relationship in the model background (prior PE), we conduct a series of parameter perturbation runs corresponding to 8 partial SE experiments. In these runs, there are no PE processes to fix the biased parameter spread so that parameter perturbations can be fully transferred to the model states, allowing us to check the state-parameter covariance without any disturbance from a PE correction. The results are shown in Figs. 7 and 8, where the horizontal axis is the ensemble anomaly (vs. ensemble mean) of the state variable and the vertical axis is the ensemble anomaly of the parameter, and the background black dots represent the model runs starting from different initial conditions. Since the parameter ensemble does not change with the model integration once perturbed at the initial time, the lines constructed by black dots in a perturbation run are parallel to the x-axis perfectly. However, the set of dots at the same integration time step from different initial conditions can be used to sample the relationship between the perturbed parameter and the model state. For example, 2 sets of such ensembles, which have the biggest positive and negative correlation coefficients between the parameters and the model states, are coloured (20 red dots and 20 blue dots) in each case. From Fig. 7, we can see that with SE for the atmosphere, the overall quasi-linear relationship between the model state anomalies (observational increments) and the parameter adjustments is constructed by the model. Under this circumstance, a meaningful projection from the observational increment on the parameter is gained to form a signal-dominant adjustment for the parameter ensemble. As shown in Fig. 8 (here only two examples of failed cases 5 and 8, similar in cases 6 and 7), without the atmosphere SE, the linear relationship between the parameter being estimated and the model states is not correctly built up, and thus the parameter estimation fails.

Relationship between the states and the parameters can be analysed quantitatively. Zhang et al. (2012) defined an ad hoc index to measure the signal-to-noise ratio (called r_{s2n}) of a model ensemble. Following the idea, we diagnose the signal-to-noise ratio of the ensemble-based error covariance between the states and parameters here. The new r_{s2n} is defined as R×S, where R is the averaged correlation coefficient between the parameter perturbations and the ensemble states in a selected time window, and S is the ratio of root mean square linear fitting errors of the parameter-state points in the full SE and in a partial SE (S_f/S_p). The best (worst) representation of the signal-to-noise ratio is then characterized by a r_{s2n} value of 1(0). Table 3 gives the r_{s2n} values for the SE only experiments of Fig.7 and Fig.8. Correlation coefficients of F(5) and F(8) are 0.19 and 0.24 respectively. Though the dependences of x_2 on a_2 in F(5) and w on c_2 in F(8) are fairly direct, the low R values suggest these relations can be easily interrupted by the atmospheric uncertainty. The values of r_{s2n} are much higher in the successful cases than in the failed cases. These results clearly show that reduction of the atmospheric uncertainty can greatly increase the signal-to-noise ratio of the parameter-state covariance in the system through enhancing the bonding between the state variable and the estimated parameter.

3.2 Impact of the chaotic-to-periodic ratio in forcings on oceanic PE

From the results above, we learned that the PE of c_2 or c_6 strongly relies on the SE of x. In a coupled system characterized as Eq. (1), the influence of atmosphere can thoroughly propagate to all variables of other media, although the influence may reduce for the deep ocean. However, some previous studies (e.g. Annan et al., 2005; Barth et al., 2015; Gharamti et al., 2014; Leeuwenburgh, 2008; Massonnet et al., 2014) show their successfulness in estimating parameters in ocean only using oceanic observations without constraints on atmospheric states. To understand what character of the model makes this difference, we make full use of this simple model with convenience to investigate the influence of model characteristics on coupled parameter estimation. For mimicking the real climate signals, the variability of the oceanic state variables w and η in Eq. (1) are driven by two kinds of forcings: the chaotic forcing from the atmosphere (Lorenz equations) and the periodic forcing associated with the external radiative forcing (simulated by a cosine function with the amplitude coefficient of S_s in this simple model). The oceanic states in the real world consist of both periodic and chaotic variations. The periodic characteristic of a state is naturally with high predictability and is generally easier to be detected after an averaging or filtering process. In this simple model, w (η) is directly under the influence of the parameter c_2 (c_6) - perturbations of c_2 (c_6) first directly affecting w (η) and then influencing the whole model by the interactions between w (η) and other variables. To understand the impact of periodic/chaotic variability of the ocean on oceanic parameter estimation, we modify the model in Appendix A to set a one-way coupling model. Then we define a chaotic-to-periodic ratio (CPR) in the signals of w (η) by manipulating the coefficient S_s. Eight experiments are performed here, four for w-to-c_2 PE and four for η-to-c_6 PE. Each experiment has a different S_s value of 100, 250, 500 and 1000 and thus a reducing CPR in w and η. Changes of w due to different S_s values are shown in Fig. 9. Comparing Fig. 9a to Fig. 6b, it can be seen that the chaotic signal in the one-way coupling model is much smaller than in the original two-way coupling model (with an identical S_s value of 10). The change of η is similar to w (see Fig. 10). With the increasing S_s value, the periodic part of η is magnified, and the η CPR decreases. Clearly, when the η CPR decreases, the periodic portion dominates and the η-to-c_6 PE becomes more and more successful (see Figs. 11a-d). But in the other 4 w-to-c_2 cases, for any w CPR, the w-to-c_2 PE fails (Fig. 12a). Apparently this is due to strong dependence of cov(w, c_2) (the covariance between w and c_2) on x_2 that is still chaotic without observational constraint. Though w is very periodic, the chaotic variability of x_2 sheds on w's variability (the needed variability of w for PE should come from c_2 but now comes from the chaotic x_2) and makes the PE process misjudge the difference between the simulated w and its observation, thus not producing a correct PE projection.

To further test the role of periodic signals in ocean states for oceanic PE, we conduct oceanic PE on a particular frequency band using the method described in Appendix B. Some results are shown in Fig. 12 which tells that using the covariance of η in a particular frequency and c_6 to project the corresponding η observational information can make a η-to-c_6 PE case with S_s = 250 as successful as the result of S_s=1000 with full frequencies (compare Fig. 12b to Fig. 11d). The method is designed to limit the PE process working on the 10-TU period of η information, which dramatically reduces the CPR of $w\eta$ and thus helps c_6 estimation, but given strong dependence of cov(w, c_2) on x_2, and that the CPR of x_2 is big on every frequency band, this particular frequency PE method does not help for estimation of c_2 (Fig. 12a).

4 Conclusion and discussions

The erroneous values of parameters in a coupled model are a source of model bias that can cause model climate drift. Model bias can be mitigated by parameter estimation (PE) with observational data. The signal-to-noise ratio in state-parameter covariance plays a centrally important role in the PE process. With a conceptual coupled model, we discuss the issue how to enhance the signal-to-noise ratio in coupled model PE through further understanding on various aspects of the PE process in a coupled numerical system.

We performed 3 kinds of comparisons to discuss the issue. The first kind focuses on the PE performance with a two-way coupling model. Results show that atmospheric state estimation (SE) is critically important. The second comparison is carried out by the experiments with the same parameter spread and SE settings in the first comparison but without the PE process. We use this way to examine the signal-to-noise ratio of state-parameter covariance in different SE settings. Results show that the projection of the observation increment onto the parameter can be easily interrupted under partial SE conditions. In the third kind, we changed the model structure from two-way coupling to one-way coupling, allowing the ocean state varying forced by the atmosphere without feedback to the chaotic atmosphere. The PE results are better with higher periodic and less chaotic states.

According to all these comparisons, first, we found that due to the interaction of multiple time scales in our conceptual coupled system, the fast varying components is the major source of state-parameter covariance uncertainties. Enhancing the estimation accuracy of chaotic states that interact with the parameter is the most important to maintain a signal-dominated relationship between the parameter being estimated and model states and makes successful coupled model parameter estimation. Second, the chaotic-to-periodic ratio (CPR) of the model state that closely associates with the parameter being estimated determines the requirement for the accuracy of state estimation. Given limited observational resources, the CPR shall be first investigated to increase the opportunities of successful parameter estimation.

Given the fact that observations are always imperfect, this conceptual coupled model study tries to provide some general guideline for CGCM PE application with the real observing system. However, the results have the following limitation: 1) The conceptual coupled model assumes that only the atmosphere is a chaotic uncertainty source. In the real world this is unnecessarily true (nonlinearity produced by smaller scale eddies in the ocean could be the part of chaotic uncertainty sources too, for instance). 2) The atmosphere-ocean interaction is idealized in the conceptual model. In the real world, the air-sea coupling could be complex as highly geographic dependent. 3) The twin experiment assumes that except for the parameters to be estimated the model “dynamical core” and “physics” are perfect and consistent to the “observation.” In the real world, CGCM is biased from the observations. All these aspects still need to be addressed before coupled model PE is applied to a CGCM with the real observing system.

How the accuracy of state estimation impacts on the coupled model parameter estimation is an interesting and challenging research topic. The spatial and temporal dependence of atmospheric and oceanic circulations could further complicate the issue. For example, the Kuroshio meander in the south of Japan is very different to the Kuroshio meander cross the Luzon strait. The Kuroshio cross Luzon strait is easily interrupted by the monsoon, but the meander in the south of Japan is a self-sustained dynamic system with multiple equilibria with non-periodic state changes (Taft, 1972; Yu et al., 2013); the uncertainty of the latter comes from the accumulation of the negative vorticities in the ocean. Further, we have already known that the method on a particular frequency can increase the opportunity of successfulness. When such a real problem is addressed through the PE with a CGCM, we may need to make efforts on both adaptive measurements and spectral separation. The PE method shall be improved to perform separately at different time scales. How to speed up the convergent rate in the coupled model PE process is also an important issue. All of these require further research work to clarify.

Appendices

Appendix A: One-way coupling model

A suitable scope of parameter values that maintain the model character is an important pre-condition for successful PE. For example, in Eq. (1) when a_2 is lower than 20, the variation of x_2 becomes periodic and looses the chaotic nature. When the values of the parameter of some ensemble members are numerically out of bound, different ensemble members exhibit different dynamic performance (some of them are chaotic and the rest are periodic), and the state-parameter covariance computed from the ensemble becomes unreasonable and PE must fail. In a_2 PE experiments, the values are bounded within 24 ~ 32 where nonlinearity and characteristic variability of the model maintains. For the purpose of manipulating the signal w or η, to make them become more periodic than chaotic, we changed the parameter S_s to magnify the amplitude of the cosine term that directly forces w. This causes the value of w to grow bigger according to different S_s settings. At the same time, the original two-way coupling has to be changed to one-way coupling by removing the w in the x_2 equation, which interacts with a_2 in the Lorenz equation, for keeping the ability of producing the chaotic signal. The referring x_2 equation after the modification is:

$$\dot{x}_2 = -x_1 x_3 + (1 + c_1) a_2 x_1 - x_2 \tag{A1}$$

Therefor, when using Eq. (A1), the Lorenz atmosphere cannot feel the variation of the ocean. The strength of the chaotic forcing remains the same in all cases with different S_s settings. And because the Lorenz atmosphere runs independently, there are no needs to set scope limits of the oceanic parameter S_s, c_2 and c_6 for securing the chaotic character of the system under this circumstance. The oceanic parameters can be perturbed much larger than in the two-way coupled cases.

Appendix B: The PE method on a particular frequency band

Previous studies have shown that applying the PE with an averaged covariance in particular time window can increase the signal-to-ratio ratio (Lu et al., 2015, Barth et al., 2015). Here, we propose an alternative method that has similar effect but is much easier to be implemented. This method applies PE on a particular frequency. The method succeeds to enhance the signal-to-noise ratio by using a designed filter on both the observations and the simulated ensemble results, and it can allow information focusing on a particular frequency more accurately than using the averaging method.

In this study, for the η-to-c_6 PE case with S_s=250, the periodic signal produced by the cosine function has a period of 10 TUs (1000 time steps) (defined by S_{pd} in Eq. (1), also see Fig. 10) and the chaotic signal is much slower than the periodic signal. In other words, the signal/noise ratio of η is strongest on this period. Therefore we designed a Butterworth high pass filter (BF) with a frequency pass band equal and larger than Fs/1000 (Fs is the frequency of sampling) to help the PE of η-to-c_6. The parameter update interval in the new PE method is identical to the standard full frequency PE case, but for each update step, before applied to Eq. (2), the observation and simulated ensemble results are filtered by the following BF process:

$$\begin{aligned} &\text{old}:\Delta y_i^o = \text{PE}(y^o, y_i^p) \\ &\text{new}:\Delta y_i^o = \text{PE}\left[\text{Filter}(y^o), \text{Filter}(y_i^p)\right], \quad i = 1:20 \end{aligned} \tag{B1}$$

Here y^o is the observation and y_i^p represents the simulated ensemble results. The BF is applied within a 5000 steps (or more) moving window. It means that on each PE step, the last 5000 observations and the simulated ensemble results in the same window are transformed through the same BF to produce new observations (Hobs) and new simulated results (Hens) on the particularly frequency. Then the new Δy_i^o is computed from the Hobs and Hens, and it is used with the covariance to determine the adjustment of the parameter. This new method can be used for different frequency band (low-pass, high-pass or band-pass), it succeed to improve the PE performance in our one-way coupling experiment for the η-to-c_6 PE (Fig. 12b).

Acknowledgements

This work is funded by the National Natural Science Foundation of China (41306004), the China's National Basic Research Priorities Programmer (2013CB956202) and the National Natural Science Foundation of China (41490641).

References

Anderson, J.: An ensemble adjustment Kalman filter for data assimilation, Mon. Wea. Rev., 129, 2884-2903, 2001.

Anderson, J.: A local least squares framework for ensemble filtering, Mon. Wea. Rev., 131, 634-642, 2003.

Annan, J. D., Lunt, D. J., Hargreaves, J. C., and Valdes, P. J.: Parameter estimation in an atmospheric GCM using the Ensemble Kalman Filter, Nonlinear Processes Geophys., 12(3), 363-371, 2005.

Barth, A., Canter, M., Schaeybroeck, B. V., Vannitsem, S., Massonnet, F., Zunz, V., Mathiot, P., Alvera-Azcarate, A. and Beckers, J.: Assimilation of sea surface temperature, sea ice concentration and sea ice drift in a model of the Southern Ocean, Ocean Modelling, 93, 22-39, 2015.

Dee, D. P., and Silva, A. M. D.: Data assimilation in the presence of forecast bias, Quart. J. of the Roy. Meteorological Society, 124.545, 269-296, 1998.

Dee, D. P.: Bias and data assimilation, Quart. J. of the Roy. Meteorological Society, 131.613, 3323-3344, 2005.

DelSole, T., and Yang, X.: State and parameter estimation in stochastic dynamical models, Physica D., 239, 1781-1788, 2010.

Evensen, G. : Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10143-10162, 1994.

Gharamti, M. E., Kadoura, A., Valstar, J., Sun, S., Hoteit, I.: Constraining a compositional flow model with flow-chemical data using an ensemble-based Kalman filter, Water Resources Research, 50.3, 2444-2467, 2014.

Gnanadesikan, A.: A simple predictive model for the structure of the oceanic psycnocline, Science, 283, 2077-2079, 1999.

Han, G., Wu, X., Zhang, S., Liu, Z., and Li, W.: Error covariance estimation for coupled data assimilation using a Lorenz atmosphere and a simple psycnocline ocean model, J. Climate, 26, 10218-10231, 2013.

Houtekamer, P. L. and Mitchell, H. L: Data assimilation using an ensemble Kalman filter technique, Mon. Wea. Rev., 126, 796-811, 1998.

Jazwinski, A.: Stochastic Processes and Filtering Theory, Academic Press, Cambridge, 1970.

Jackson, C., Sen, M. K. and Stoffa, P. L.: An efficient stochastic Bayesian approach to optimal parameter and uncertainty estimation for climate model predictions, J. Climate, 17, 2828-2841, 2004.

Jung, Y., Xue, M. and Zhang G.: Simultaneous estimation of microphysical parameters and the atmospheric state using simulated polarimetric radar data and an ensemble Kalman filter in the presence of an observation operator error, Mon. Wea. Rev., 138, 539-562, 2010.

Leeuwenburgh, O.: Estimation and correction of surface wind-stress bias in the Tropical Pacific with the Ensemble Kalman Filter, Tellus A, 60.4, 716-727, 2008.

Li, S., Zhang, S., Liu, Z., Yang, X., Rosati, A., Golaz, J. and Zhao M.: The Role of large-scale feedbacks in cumulus convection parameter estimation, J. Climate, 29, 4099–4119, 2016.

Liu, C., Köhl, A. and Stammer, D.: Adjoint-based estimation of eddy-induced tracer mixing parameters in the global ocean, J. Phys. Oceanogr., 42, 1186-1206, 2012.

Liu, Y., Liu, Z., Zhang, S., Rong, X., Jacob, R., Wu, S. and Lu, F.: Ensemble-based parameter estimation in a coupled GCM using the adaptive spatial average method, J. Climate, 27, 4002-4014, 2014.

Liu, Y, Liu, Z., Zhang, S., Jacob, R., Lu, F., Rong X. and Wu, S.: Ensemble-based parameter estimation in a coupled general circulation model, J. Climate, 27, 7151-7162, 2014.

Lorenz, E. N.: Deterministic non-periodic flow, J. Atmos. Sci., 20, 130-141, 1963.

Lu F., Liu, Z., Zhang, S. and Liu, Y.: Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part I: simple model study, Mon. Wea. Rev., 143, 3823-3837, 2015.

Massonnet, F., Goosse, H., Fichefet, T., Counillon, F.: Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter, J. Geo. Res.: Oceans, 119.7, 4168-4184, 2014.

Stammer, D.: Adjusting internal model errors through ocean state estimation, J. Phys. Oceanogr., 35, 1143-1153, 2005.

Taft, B.: Characteristics of the flow of the Kuroshio south of Japan, University of Tokyo Press, Tokyo, 165-216, 1972.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. : Ensemble Square Root Filters, Mon. Wea. Rev., 131, 1485-1490, 2003.

Tong, M., and Xue, M.: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments, Mon. Wea. Rev., 133, 1789-1807, 2005.

Wu, X., Zhang, S., Liu, Z., Rosati, A. and Delworth, T.: A study of impact of the geographic dependence of observing system on parameter estimation with an intermediate coupled model, Climate Dynamics, 40, 1789-1798, 2013.

Yang, X. and Delsole, T.: Using the ensemble Kalman Filter to estimate multiplicative model parameters, Tellus, 61A, 601-609, 2009.

Yu, X., Wang, F. and Wan, X.: Index of Kuroshio penetrating the Luzon Strait and its preliminary application, Acta Oceanol. Sin., 32, No. 1, 1-11, 2013.

Zhang, S.: Impact of observation-optimized model parameters on decadal predictions: simulation with a simple psycnocline prediction model, Geo. Res. Lett., 38, 1-5, 2011a.

Zhang, S.: A study of impacts of coupled model initial shocks and state-parameter optimization on climate predictions using a simple pycnocline prediction model, J. Climate, 24, 6210-6226, 2011b.

Zhang, S. and Anderson, J.: Impact of spatially and temporally varying estimates of error covariance on assimilation in a simple atmospheric model, Tellus, 55A, 126-147, 2003.

Zhang, S., Harrison, M., Rosati, A. and Wittenberg, A.: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies, Mon. Wea. Rev., 135, 3541-3564, 2007.

Zhang, S., Liu, Z., Rosati, A. and Delworth, T.: A study of enhancive parameter correction with coupled data assimilation for climate estimation and prediction using a simple coupled model, Tellus, 64A, 1-20, 2012.

Table 1: List of the successful (S) and failed (F) parameter estimation (PE) cases with partial state estimation (SE) in 8 PE experiments (in the parenthesis is the experiment serial number).

PE / SE	x_2-to-a_2	w-to-a_2	x_2-to-c_2	w-to-c_2
$x_{1,2,3}$ by x_2 obs	S (1)	S (2)	S (3)	S (4)
w by w obs	F (5)	F (6)	F (7)	F (8)

Table 2: List of root mean square error of the state variable and the parameter during the last 100 TUs in 8 PE experiments.

Exp number \ State and Param	x_2	w	a_2	c_2
S(1): x_2 obs, x_2-to-a_2	5.9224	0.0570	0.0889	N/A
S(2): x_2 obs, w-to-a_2	5.9086	0.0567	0.0895	N/A
S(3): x_2 obs, x_2-to-c_2	5.9213	0.0731	N/A	0.0250
S(4): x_2 obs, w-to-c_2	5.9174	0.0589	N/A	0.0153
F(5): w obs, x_2-to-a_2	14.6801	0.0360	1.6806	N/A
F(6): w obs, w-to-a_2	14.3177	0.0381	3.2612	N/A
F(7): w obs, x_2-to-c_2	14.4102	0.0744	N/A	0.3848
F(8): w obs, w-to-c_2	14.4004	0.0660	N/A	0.3454

Table 3: List of r_{s2n} during the last 100 TUs in 8 SE only (no PE) experiments.

	S(1): x_2 obs, x_2-to-a_2	S(2): x_2 obs, w-to-a_2	S(3): x_2 obs, x_2-to-c_2	S(4): x_2 obs, w-to-c_2	F(5): w obs, x_2-to-a_2	F(6): w obs, w-to-a_2	F(7): w obs, x_2-to-c_2	F(8): w obs, w-to-c_2
R	0.41	0.33	0.60	0.91	0.19	0.19	0.19	0.24
S_f/S_p	0.65	0.97	0.98	0.96	0.19	0.96	0.29	0.83
r_{s2n}	0.27	0.32	0.59	0.87	0.04	0.18	0.06	0.20

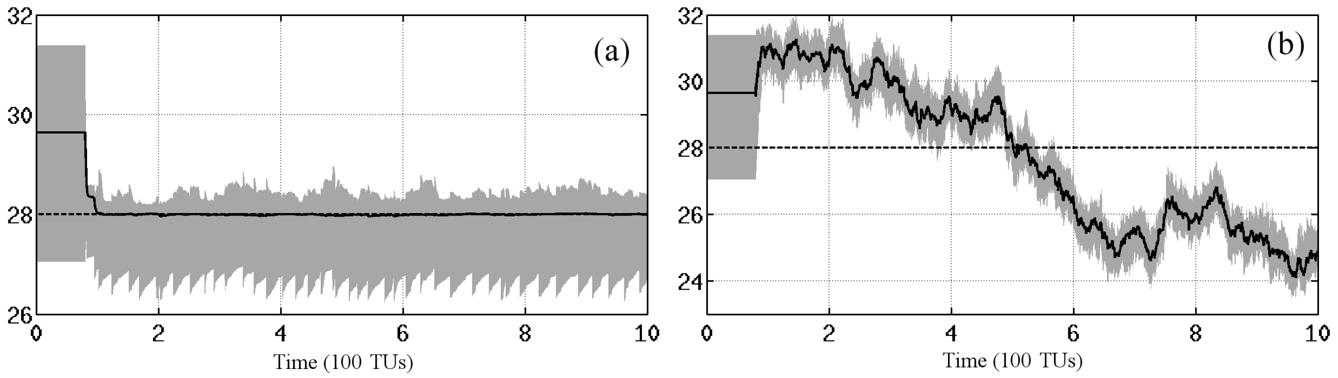

Figure 1: Time series of the ensemble mean (solid line) of the estimated parameter a_2 using observations of w (i.e. w-to-a_2) with state estimation (SE) of a) both the atmosphere ($x_{1,2,3}$) and ocean (w) from their observations ($x_{1,\,2,\,3}$ and w), and b) only w with the w observations. The dashed line marks the "true" value of the parameter a_2 and the shaded area represents the range of ensemble.

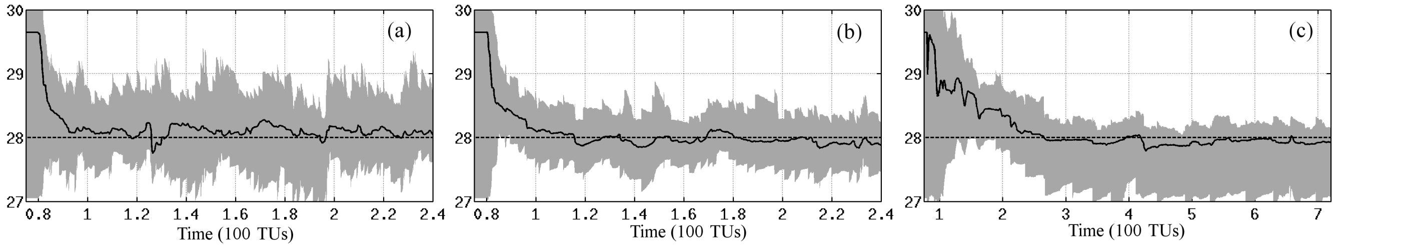

Figure 2: Time series of ensemble means (solid line) of the estimated parameter a_2 in 3 experiments, a) x_2-to-a_2 (using x_2 observations to estimate a_2) with SE for both $x_{1,2,3}$ and w, b) x_2-to-a_2 with SE for $x_{1,2,3}$ only, c) w-to-a_2 with SE for $x_{1,2,3}$ only. Any other notations are the same as in Fig. 1.

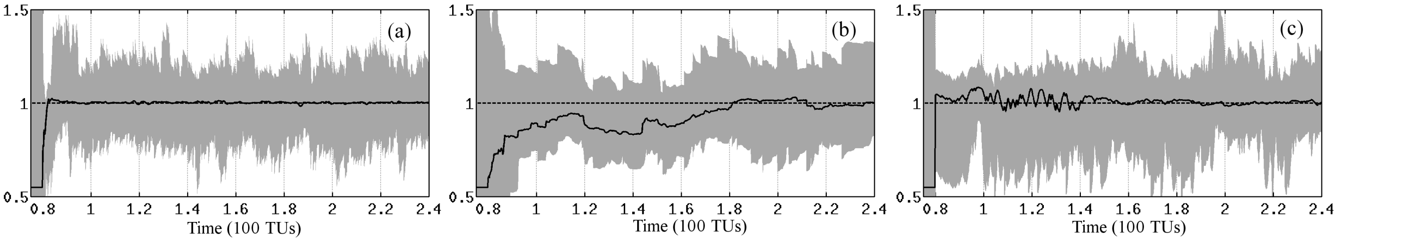

Figure 3: Time series of ensemble means of the estimated parameter c_2 in 3 experiments, a) w-to-c_2 (using w observations to estimate c_2) with SE for both $x_{1,2,3}$ and w, b) x_2-to-c_2 (using x_2 observations to estimate c_2) with SE for $x_{1,2,3}$ only, c) w-to-c_2 with SE for $x_{1,2,3}$ only. Any other notations are the same as in Fig. 1.

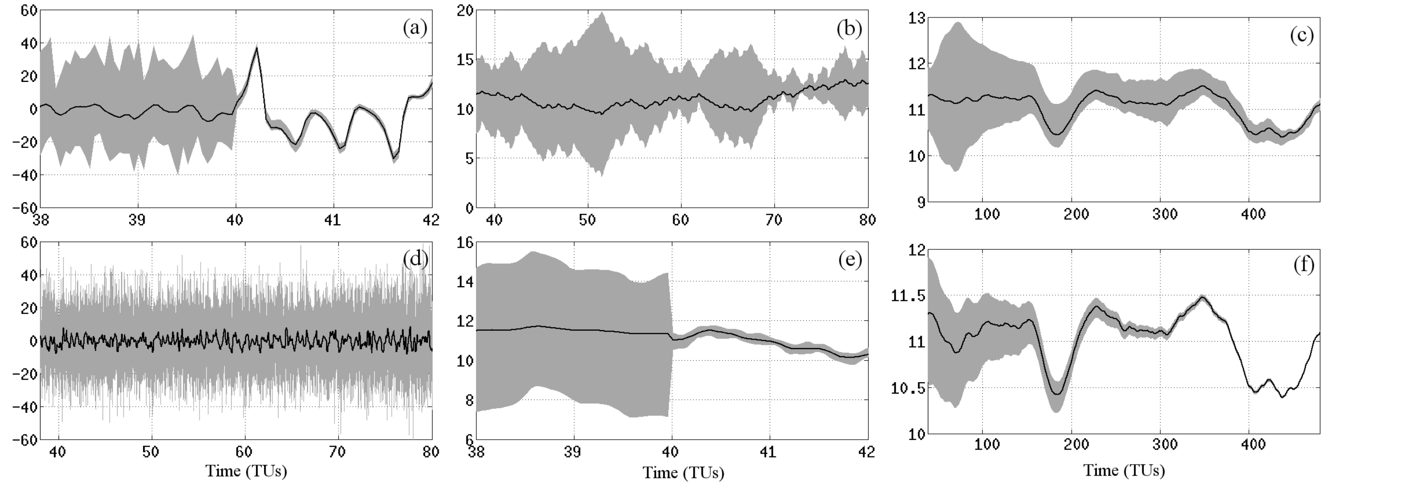

Figure 4: Time series of the state variables from the w-to-c_2 PE experiment, for ad) x_2, be) w cf) η. The upper panels abc) are from the successful case with SE for $x_{1,2,3}$, and the lower panels def) are from the failed case with SE for w. Any other notations are the same as in Fig. 1.

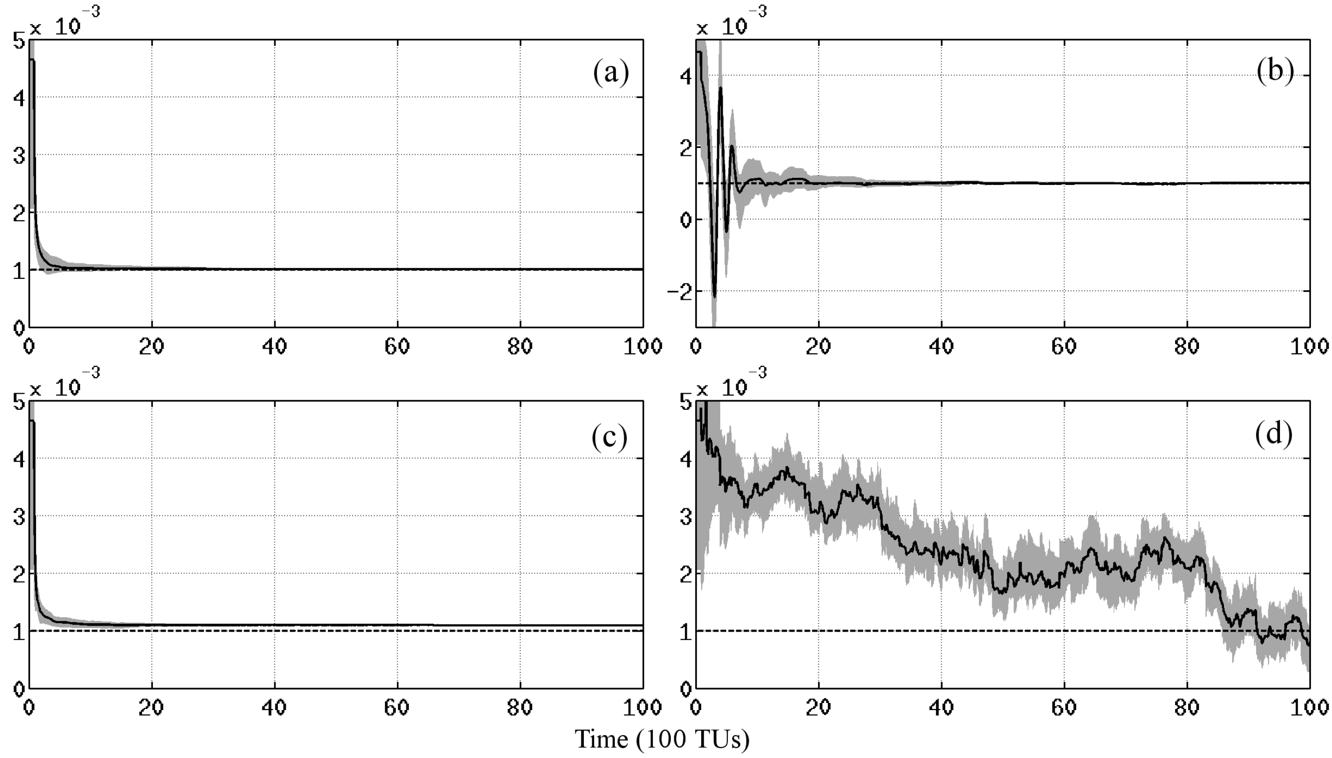

Figure 5: Time series of the ensemble of parameter c_6 from the η-to-c_6 (using η observations to estimate c_6) PE experiment in 4 different state estimation settings, a) $x_{1,2,3}$, w and η, b) x_2, only, c) w and η only and d) η only. Any other notations are the same as in Fig. 1.

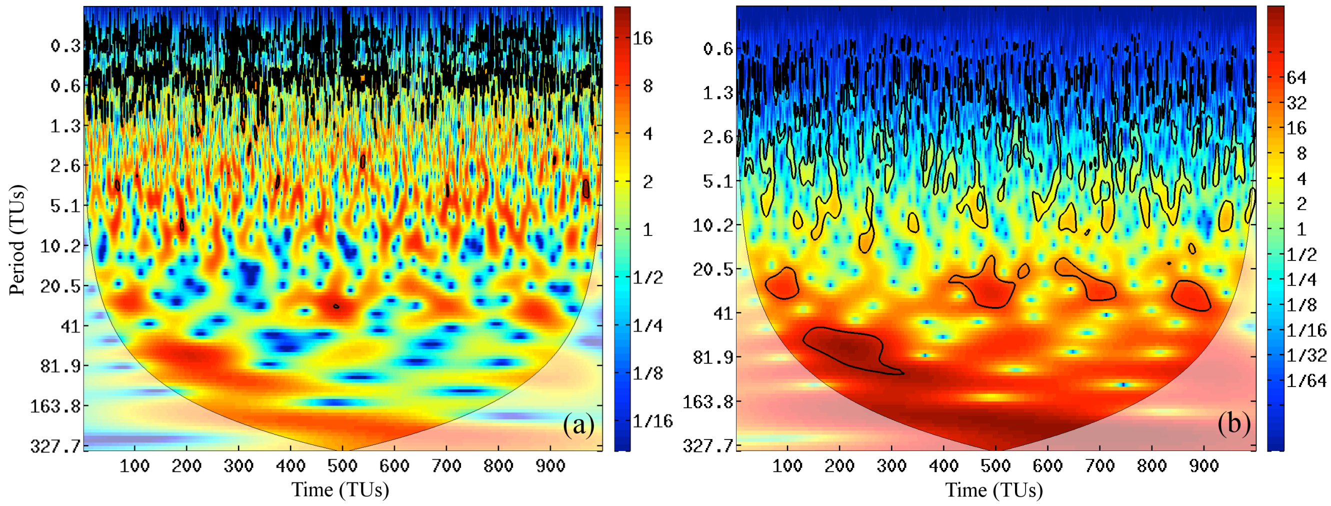

Figure 6: Wavelet analyses for a) x_2 and b) w in the "truth" model run.

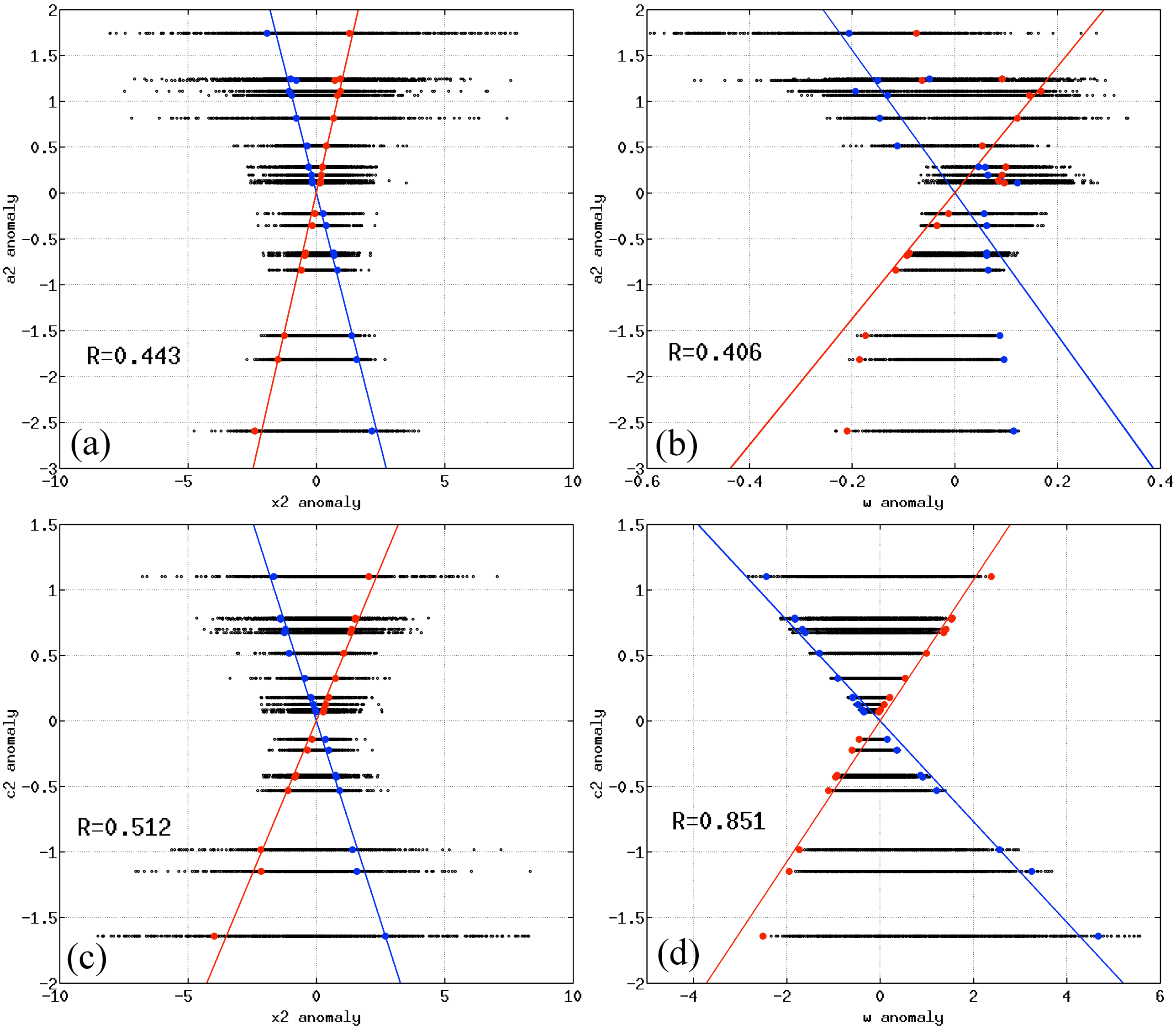

Figure 7: Sampling map of the perturbed parameter anomalies in the space of model state anomalies for a) a_2 vs. x_2 , b) a_2 vs. w, c) c_2 vs. x_2 and d) c_2 vs. w, when the atmospheric state is constrained by its observations. Dots with the same color (red or blue) represent ensembles at the same time step in the model integration. The colored line represents a linear fitting for the same color dots. Here we show two examples that have a high positive (red) and negative (blue) correlation between the parameter and model state perturbations, respectively. The R value shown in each panel is the time averaged parameter-state correlation coefficient in last 5000 time steps.

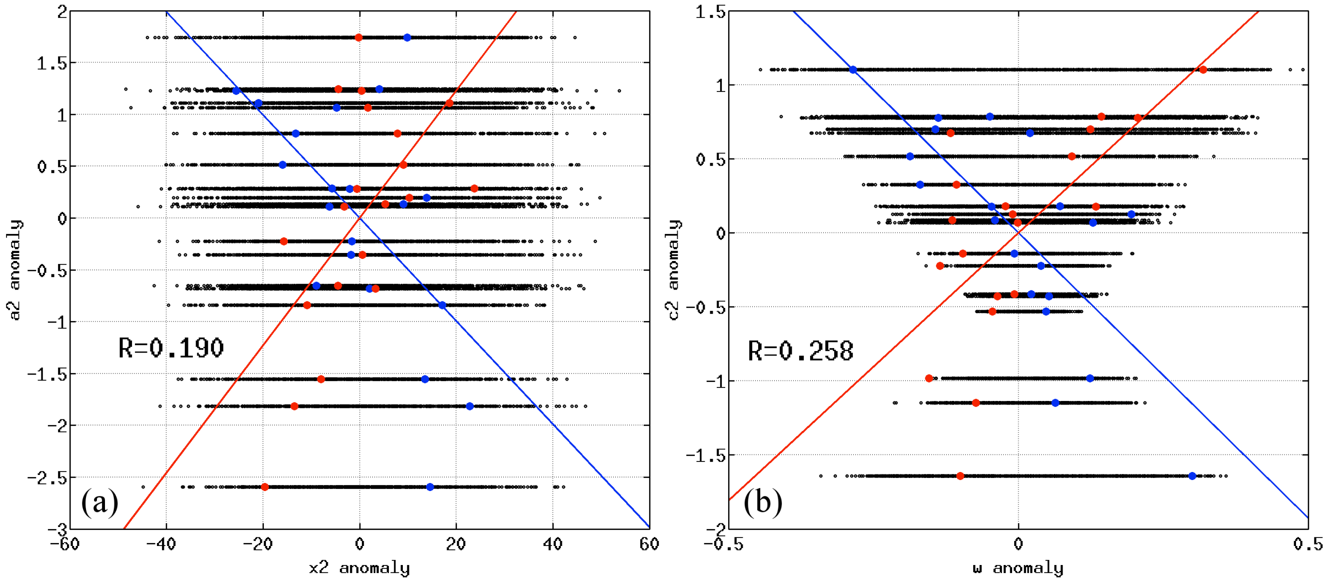

Figure 8: Same as Fig. 7 but for the case with SE of *w* only, a) a_2 vs. x_2 and b) c_2 vs. *w*. Here we show two examples that the linear fitting becomes difficult in red and blue, for which the data are taken from the same time steps as shown in Fig. 7.

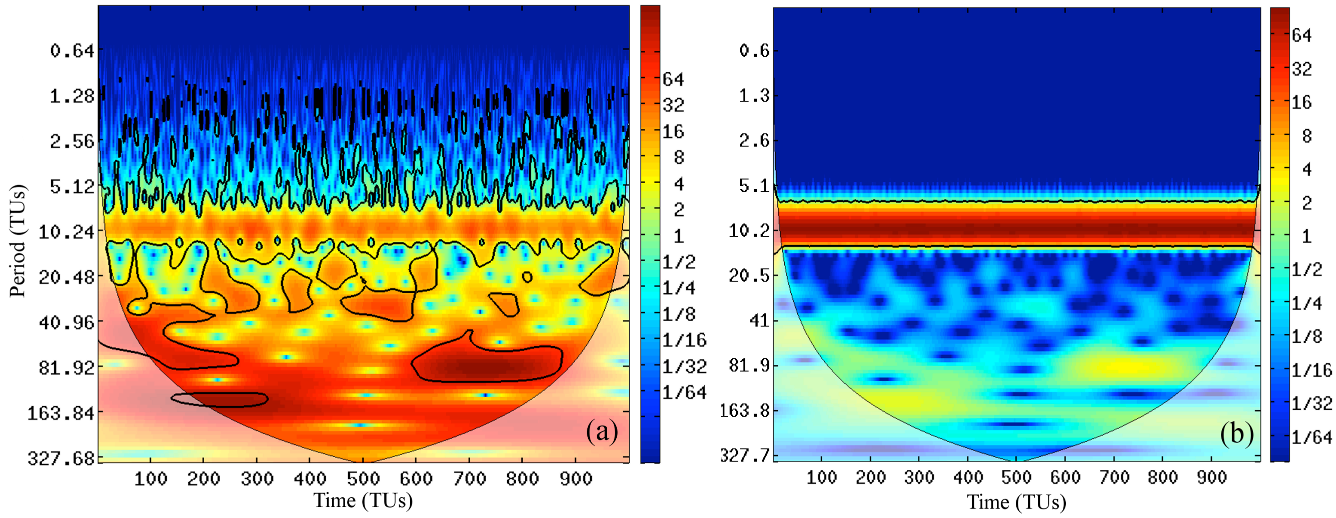

Figure 9: Wavelet analyses for *w* in the run of one-way coupling model forced by a) S_s=10 and b) S_s=250.

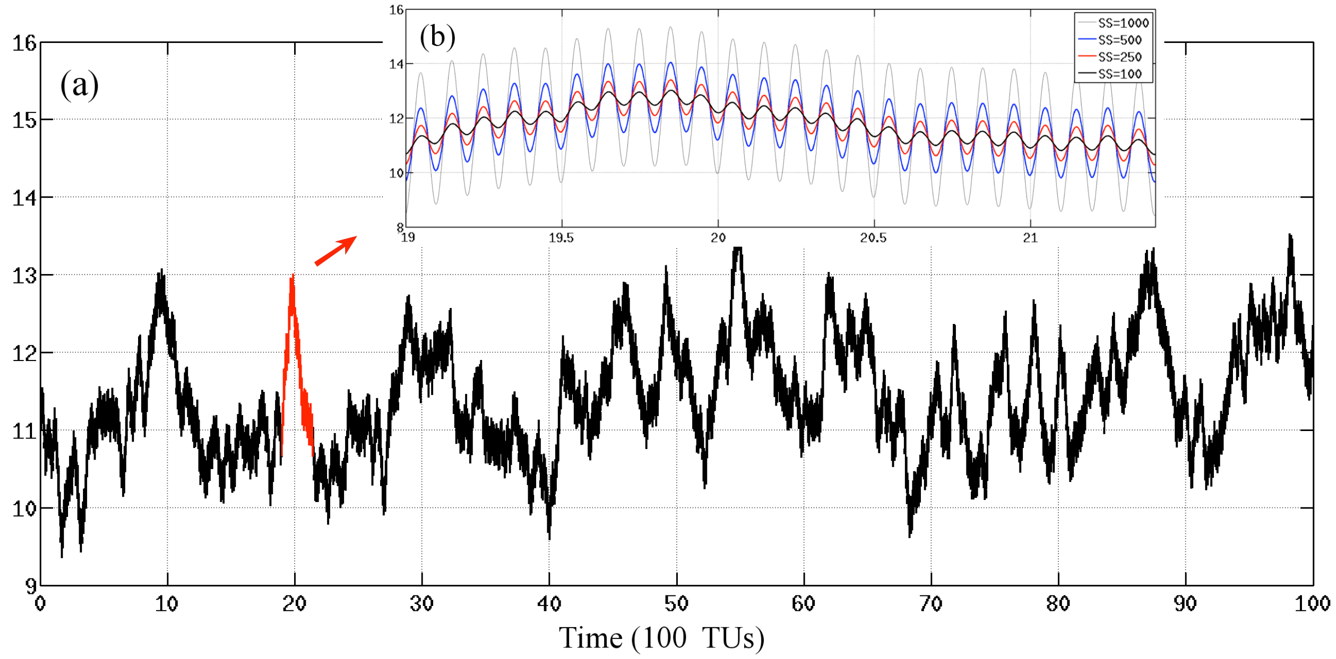

Figure 10: Time series of η with different S_s values (varying from 100 to 1000) with a one-way coupling model setting described in Appendix A. To visualize the difference induced by different S_s values, panel b) is the zoomed out version of the section marked in red in panel a).

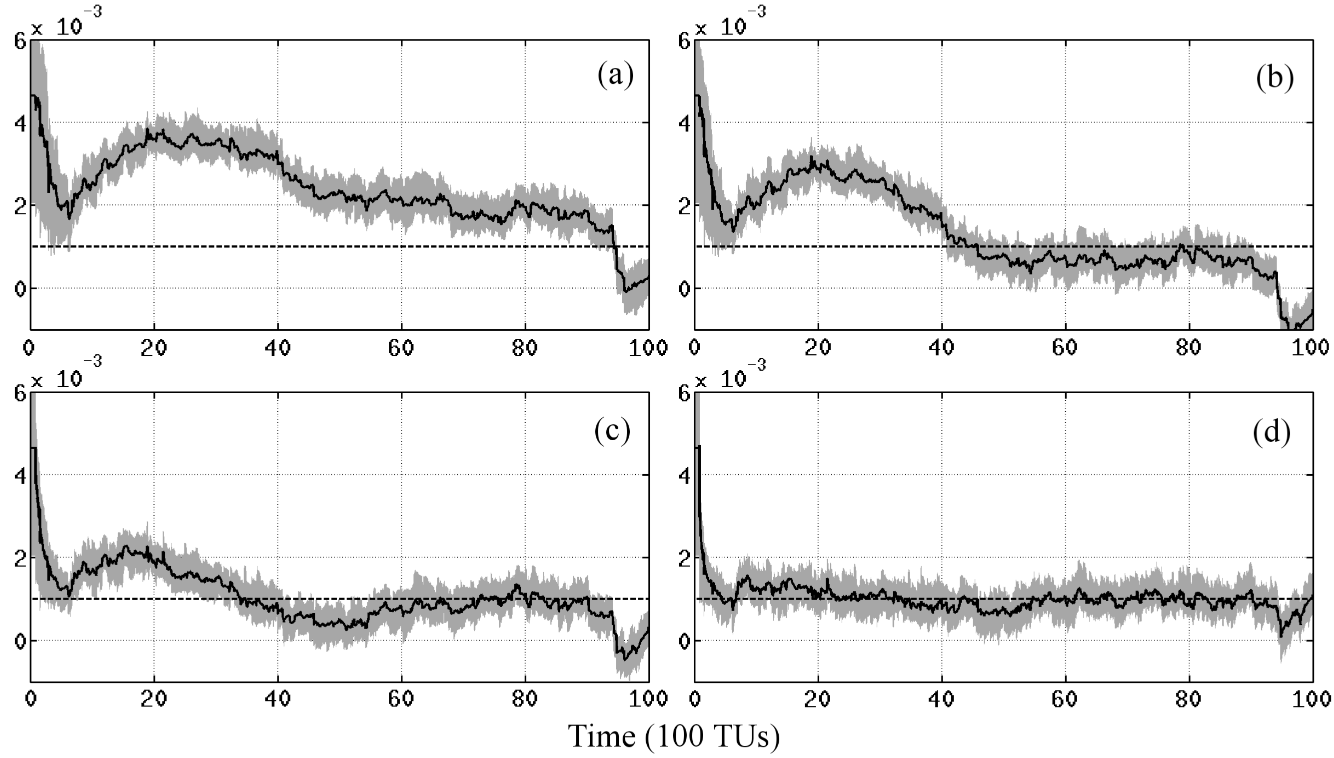

Figure 11: Time series of the ensemble of parameter c_6 in 4 *η-to-*c_6 PE experiments with different S_s values, a) 100, b) 250, c) 500 and d) 1000 with the one-way coupling model setting. In all cases, only η is constrained by its observations. Any other notations are same as Fig. 1.

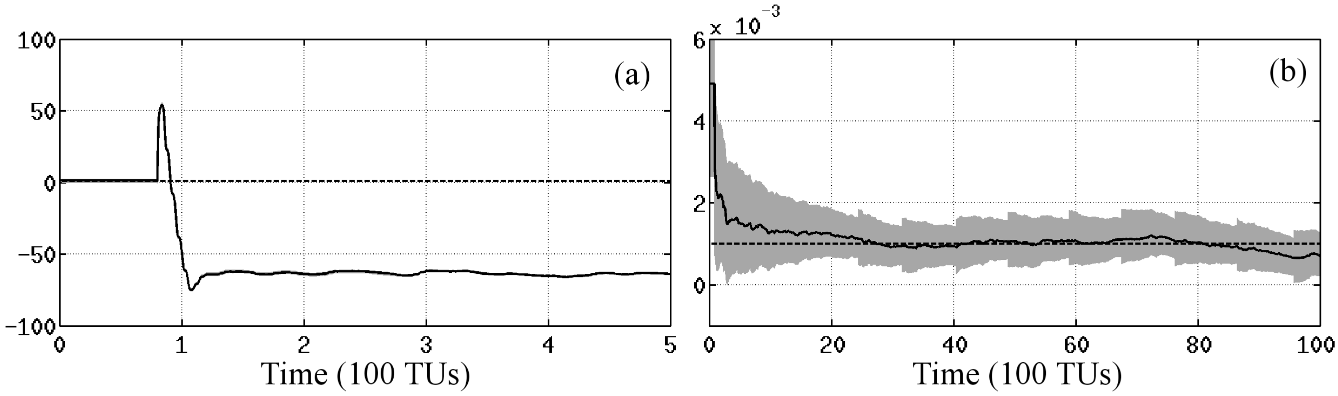

Figure 12: Time series of the ensemble of the parameter in the a) w-to-c_2 PE with SE of w only and b) η-to-c_6 PE with SE of η only using the one-way coupling model with S_s=250. Note that the initial c_2 in panel a) is approximate 0.56, and the truth is 1. Any other notations are same as Fig. 1.