Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Discussions

Further Insights on the Role of Accurate State Estimation in Coupled Model Parameter Estimation by a Simple 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 parameters. The signal-to-noise ratio of error covariance between model states and initially perturbed parameters determinates directly the success of parameter estimation or not. With a conceptual climate model that couples the stochastic atmosphere and slow varying ocean, this study examines the sensitivity of the 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 the 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 in this study. This simple model study provides a guideline when real observations are used to optimize model parameter 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 a trial and error turning 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. Zhang, 2011ab; Zhang et al., 2012; Wu et al., 2013; Liu et al., 2014; Liu et al., 2014; Zhang et al., 2014; Li et al., 2015). Traditional data assimilation where observations are used to only estimate model states (i.e. state estimation, called SE hereafter) becomes the state-parameter optimization with observations.

In the study with a simple 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 (PE). However, given the nature of multiple media of the climate system, which have different time scale variability and different quality of observations so as different uncertainty contributions, a further question is: what is the impact of SE accuracy in different media on coupled model PE? Given the extreme importance of state-parameter covariance in PE, a clear answer for this question must advance the application of coupled model parameter optimization to climate analysis and prediction as well as climate modeling.

To answer this question, here we use a simple climate 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 typical 3 time scales of the climate system – chaotic (synoptic) atmosphere, seasonal-interannual upper ocean and decadal deep ocean. We use a twin experiment framework throughout the whole study.

The paper is organized as follows. After introduction, we describe the construction of the simple pycnocline prediction model, and the setup of twin experiment framework in section 2. Section 3 presents the details, stableness and accuracy of the various PE experiments with different partial SE settings. With more analyses, the section 3 also gives conditions for successful PE with partial SE. Finally, summary and general discussions will be given in section 4.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Discussions

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 simple climate model is based on the Lorenz's 3-variable chaotic model (Lorenz, 1963) and couples the three Lorenz chaotic atmosphere variables to a slab ocean variable (Zhang et al., 2012) interacting with a pycnocline predictive model (Gnanadesikan, 1999). For the problem that is concerned, this simple coupled model shares the fundamental features of the CGCM (Zhang, 2011a; Han et al., 2013). The model development can be traced in Zhang (2011ab) and Zhang et al. (2012). Here, we only comment on major points that are relevant to this study. It includes 5 equations:

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

The first 3 equations that represent the dynamics of the atmosphere is the Lorenz's 3-variable chaotic model (Lorenz, 1963). The last 2 equations respectively represent the dynamics of the surface ocean and the deep ocean pycnocline depth variation. 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 in this system the ocean is driven by two kinds of atmospheric forcing: 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 referred in the literature cited before. With the parameter O_m/O_d = 10.0 (O_m = 10), 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 directly influences the variation of the state variable x_2. And the parameter c_2 directly influences the variation of the state variable w. The estimation of these two parameters will be used later in this study to interpret the relation between the accuracy of coupled model SE and successfulness of PE.

2.2 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 the 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 last 10^3 TUs' output is then used as the "truth" to produce

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

“observations” by superimposing a white noise on the “observed” variables in this twin experiment framework. The standard deviation of “observational” errors are 2 for the atmospheric variables $x_{1,2,3}$ and 0.2 for the oceanic variable w. The assimilation run is an ensemble of integrations for each test case with an erroneously-set parameter value. The ensemble size in each case is 20 chosen as from previous studies (Zhang & Anderson, 2003). The initial conditions of the ensemble are taken from the end of spin-up run. Therefore the different results between the ensemble run are all from the parameter perturbations.

The PE experiment cases can be distinguished in 3 aspects: 1) 2 state constraint settings (i.e. SE settings) that assimilate the atmosphere “observations” (x_2) only and the ocean “observations” (w) only respectively; 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) – in the system are used to estimate the parameter. Here the SE uses weak coupled data assimilation as termed in the literature (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. We will analyze these PE cases throughout this study 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” (i.e. 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. And also considering convenience on visualization, we set the ensemble initial values of a_2 as a Gaussian distribution $N(30, 1)$ (with a mean value of 30 and a standard deviation of 1). The ensemble initial values of c_2 are set as $N(0.8, 0.5)$. With the ensemble size of 20, the actual used ensemble mean value of the initial a_2 and c_2 from sampling is 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 of the second 10^3 TUs described above.

2.3 Filtering scheme

The filtering method used in this study is the ensemble adjustment Kalman filter (EAKF; Anderson, 2001; 2003; Zhang & Anderson, 2003; Zhang et al., 2007). SE and PE are all based on the two-step EAKF implementation (see Zhang et al., 2007). In both SE and PE, the ensemble observational increments are 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 parameter being estimated by the following equation:

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Open Access
Discussions
EGU

$$\Delta p_i^u = \frac{\mathrm{cov}(\Delta p_k, \Delta y_k)}{\mathrm{std}(\Delta y_k)^2} \Delta y_i^o, \quad k = 1:20 \tag{2}$$

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. $\mathrm{cov}(\Delta p_k, \Delta y_k)$ is the error covariance computed among the ensemble of the state variables (needed for SE) or between the ensemble of the state variable and the ensemble of the estimated parameter (needed for PE). $\mathrm{std}(\Delta y_k)$ 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 anomaly of x_2 and the anomaly of c_2 from their ensemble mean are used to calculate the ratio of cov/std^2 and to adjust c_2 toward a better value that can minimize the errors of model states from the observations.

Some other relevant aspects of the method are also commented here. 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. Test results show, for the issue we are addressing, the conclusion is not sensitive to the update interval if it is within a reasonable range. The inflation method must be included in the EAKF for PE. The parameter will influence the state variable, so that the inflation of the atmosphere and ocean state variable is unnecessarily in this study. The PE inflation scheme follows Zhang (2011b): when the std (spread) of the parameter ensemble is below some limit (0.2 for a_2, 0.1 for c_2), we inflate the ensemble to satisfy this std value. During this process, the shape of the ensemble parameters remains unchanged.

3 Impact of SE accuracy on coupled model PE

With the method described in section 2, if all the atmospheric ($x_{1,2,3}$) and oceanic (w) states are estimated with the "observations" that sample the "truth," then 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 equation) to estimate a_2 (in the atmosphere equation) with all 4 state variables being estimated is shown in Fig. 1a, where the ensemble of a_2 successfully converges to the "truth" from the biased initial value around 30. However, if only a part of climate observations, say only observations in one medium, is used in the SE, in some cases, PE is successful while in some other cases PE fails (Fig. 1b). The successful and failed PE cases with different SE accuracy in different media are summarized in Table 1. Next, we will analyze and discuss the results of these experiments in details to discover the impact of SE accuracy 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 using the atmosphere observations succeed, while all 4 cases with oceanic SE using the ocean observations fails, no matter what medium observations are used to estimate which medium parameter. An example of failed PE in which the observations of w are used to estimate a_2 is shown in Fig. 1b, where the ensemble of a_2 cannot converge to its "true" value of 28.

The stability of PE is different among partial SE settings. It can be seen from Figs. 2 and 3 that the time series of the ensemble mean of the estimated parameters are very different. Figures 2a and 3a show the cases with both atmospheric SE and oceanic SE, while Figs. 2bc and 3bc show the 4 successful cases with only atmospheric SE. Compared to full SE (Fig. 1a), much bigger fluctuation of estimated parameter values is observed in the partial SE cases at the beginning of spin-up period. 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 may be comprehended by the lower signal-to-noise ratio of state-parameter covariance provided by the data assimilation system.

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.

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 cases 3 and 4, we successfully estimate the oceanic parameter c_2, suggesting we can use different medium measurements to calibrate the same 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 and 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 case 4.

As c_2 is a coupling parameter, similar to a parameter in air-sea interaction processes in a coupled general circulation model (CGCM), rather than a pure oceanic parameter, for example, used in the subsurface or deep ocean. 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 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: 1) all state variables, 2) $x_{1,2,3}$ only, 3) w and η, 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 c6) 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 one's intuition that in-situ ocean data are always considered as first important piece of information for determining the coefficients in ocean equations. 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. Next we will conduct more sophisticate analyses to understand this phenomenon.

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 uncertainty 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 frequencies of 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 critical for successfulness of PE.

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. In a complex climate system, the parameter and the state variable must be correlated enough for implementing PE. In the EAKF method we used, the key projection is carried out by a linear regression equation based on the state-parameter covariance, and therefore only a linear or quasi-linear relationship between parameters and states in ensemble is recognized. A critical reason for all the failure cases without direct atmospheric SE is that under such a circumstance, the chaotic disturbances in the Lorenz's atmosphere continuously interact with the parameter, which make difficulties for the system to build up a quasi-linear relationship between the state variable and the parameter for correctly projecting observational information onto the parameter.

To investigate the parameter-state relationship in the model background (prior PE), we conduct a series of parameter perturbation runs corresponding to each of the 8 partial SE experiments. 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.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Discussions
Open Access

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 colored (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 failure cases 5 and 8, similar for cases 6 and 7), without SE of the atmosphere, the linear relationship between the parameter being estimated and the model states is not correctly built, and parameter estimation fails.

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 different 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, and thus cannot produce 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 scope 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 Conclusions

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 coupled model parameter estimation. With a simple 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.

First of all, we found that due to the interaction of multiple time scales in a coupled climate system, the fast varying components is the major source of state-parameter covariance uncertainties. Enhancing the 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 so as to succeed in coupled model PE. 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.

The simple model results provide some fundamental understanding about climate model PE as a general guideline. However, when a coupled general circulation model (CGCM) is used to improve climate analysis and prediction by parameter estimation with the climate observing system, many challenges remain. In our simple model study, we assume parameter errors being the only source of model bias. One must deal with other model bias sources in application of a CGCM. For example, what is the influence of biases of dynamical core and physical schemes? And further, the uncertainties in the real world are complex. For a similar phenomenon in different regions, the dynamical mechanism may be different. For example, the Kuroshio large meander in the south of Japan is very different to the 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-states 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. The long-term meander in the south of Japan is not sensitive to the atmosphere and always has a life period in 4-7 years. 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 parameter estimation with a CGCM, we may need to make efforts on both adaptive measurements and spectral separation. 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:

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

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.

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.

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

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, 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.

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.

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.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Open Access
Discussions

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.

Zhang, S., Chang, YS., Yang, X. and Rosati, A.: Balanced and coherent climate estimation by combining data with a biased coupled model, J. Climate, 27, 1302-1314, 2014.

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).

SE \ PE	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)

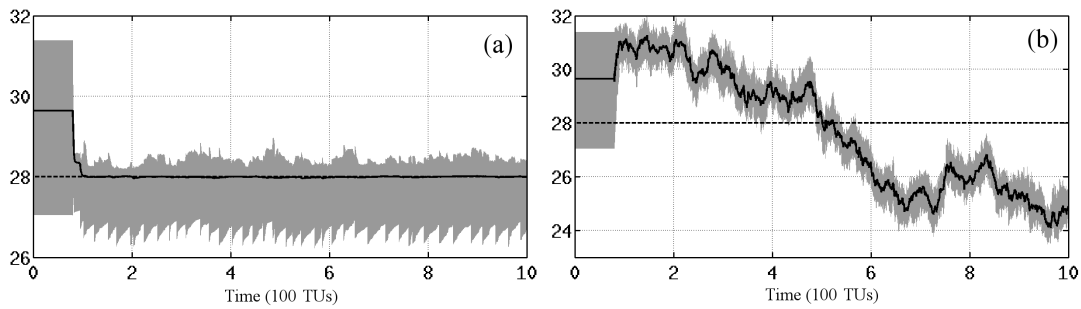


Figure 1: Time series of the ensemble mean (solid line) of the estimated parameter a_2 using observations of w (i.e. w-to-a_2) with state estimation (SE) of a) both the atmosphere ($x_{1,2,3}$) and ocean (w) from their observations ($x_{1,\,2,\,3}$ and w), and b) only w with the w observations. The dashed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

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Discussions
Open Access

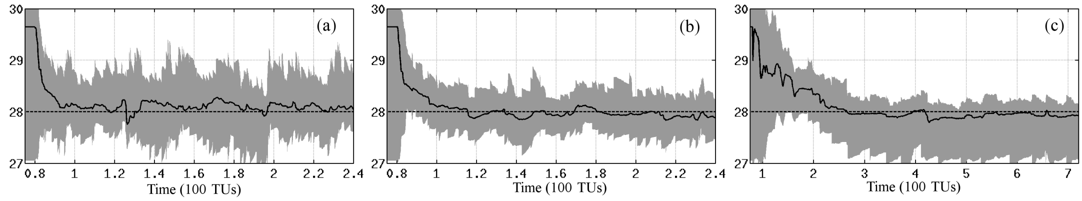


Figure 2: Time series of ensemble means (solid line) of the estimated parameter a_2 in 3 experiments, a) x_2-to-a_2 (using x_2 observations to estimate a_2) with SE for both $x_{1,2,3}$ and w, b) x_2-to-a_2 with SE for $x_{1,2,3}$ only, c) w-to-a_2 with SE for $x_{1,2,3}$ only. Any other notations are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

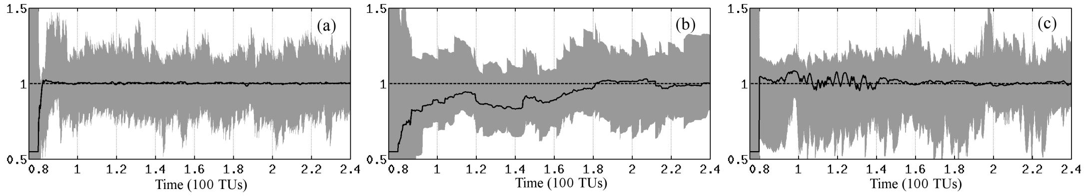


Figure 3: Time series of ensemble means of the estimated parameter c_2 in 3 experiments, a) w-to-c_2 (using w observations to estimate c_2) with SE for both $x_{1,2,3}$ and w, b) x_2-to-c_2 (using x_2 observations to estimate c_2) with SE for $x_{1,2,3}$ only, c) w-to-c_2 with SE for $x_{1,2,3}$ only.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

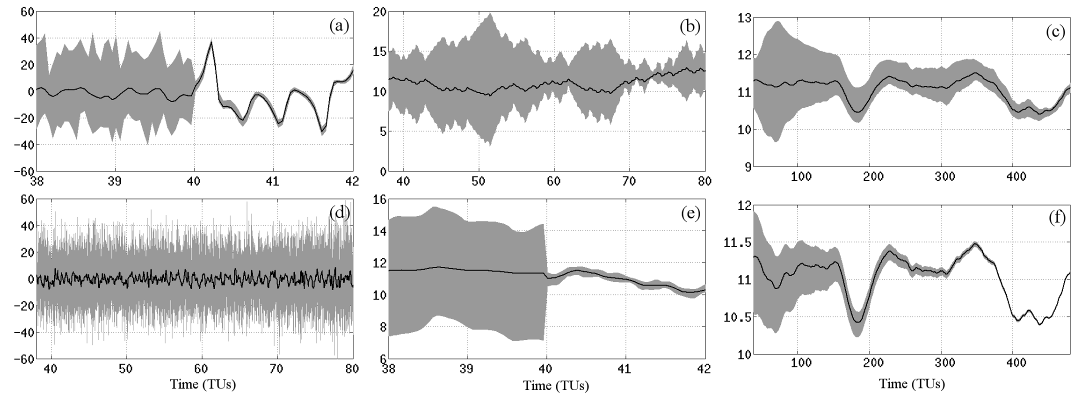


Figure 4: Time series of the state variables from the w-to-c_2 PE experiment, for ad) x_2, be) w cf) η. The upper panels abc) are from the successful case with SE for $x_{1,2,3}$, and the lower panels def) are from the failed case with SE for w. Any other notations are the same as in Fig. 1.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Open Access
Discussions

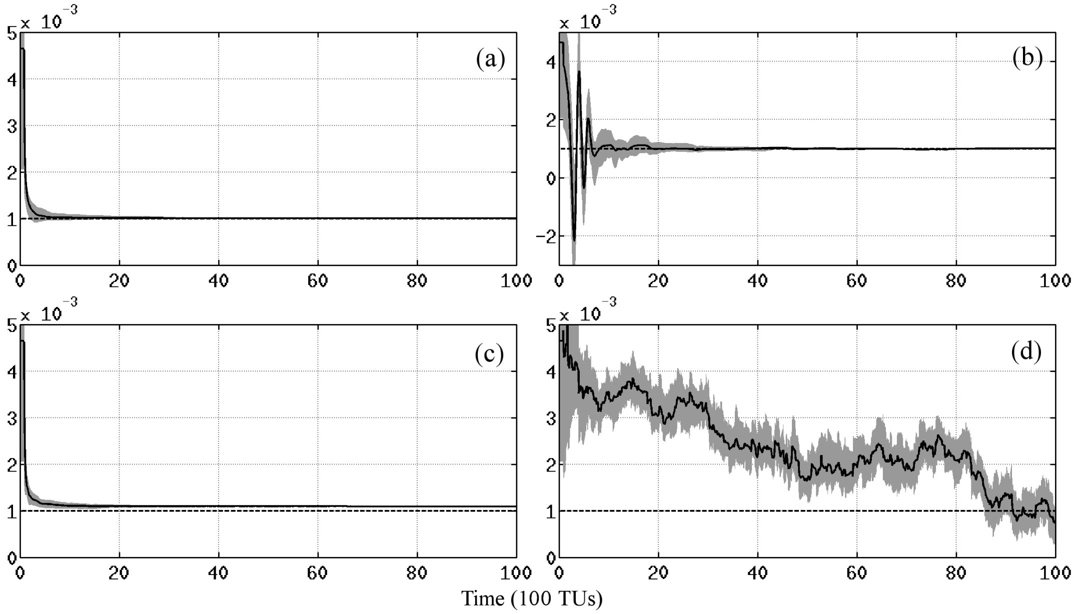


Figure 5: Time series of the ensemble of parameter c_6 from the η-to-c_6 (using η observations to estimate c_6) PE experiment in 4 different state estimation settings, a) $x_{1,2,3}$, w and η, b) x_2 only, c) w and η only and d) η only. Any other notations are the same as in Fig. 1.

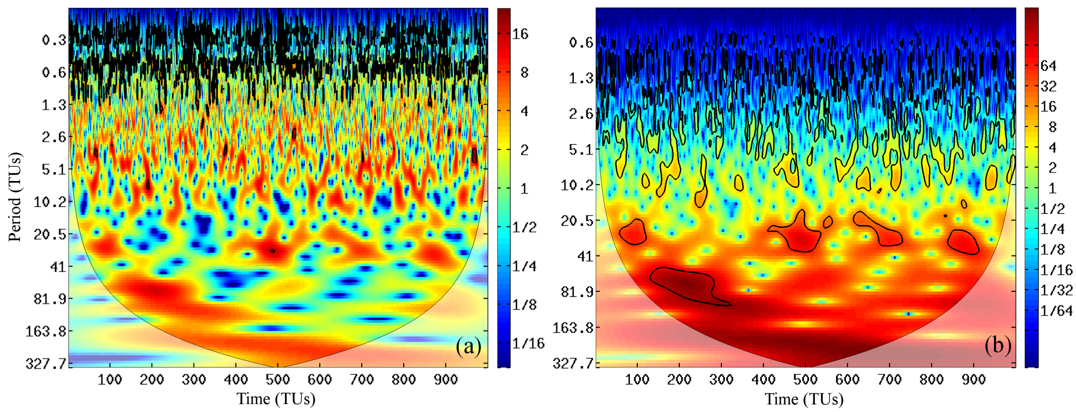


Figure 6: Wavelet analyses for a) x_2 and b) w in the "truth" model run.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Discussions
Open Access

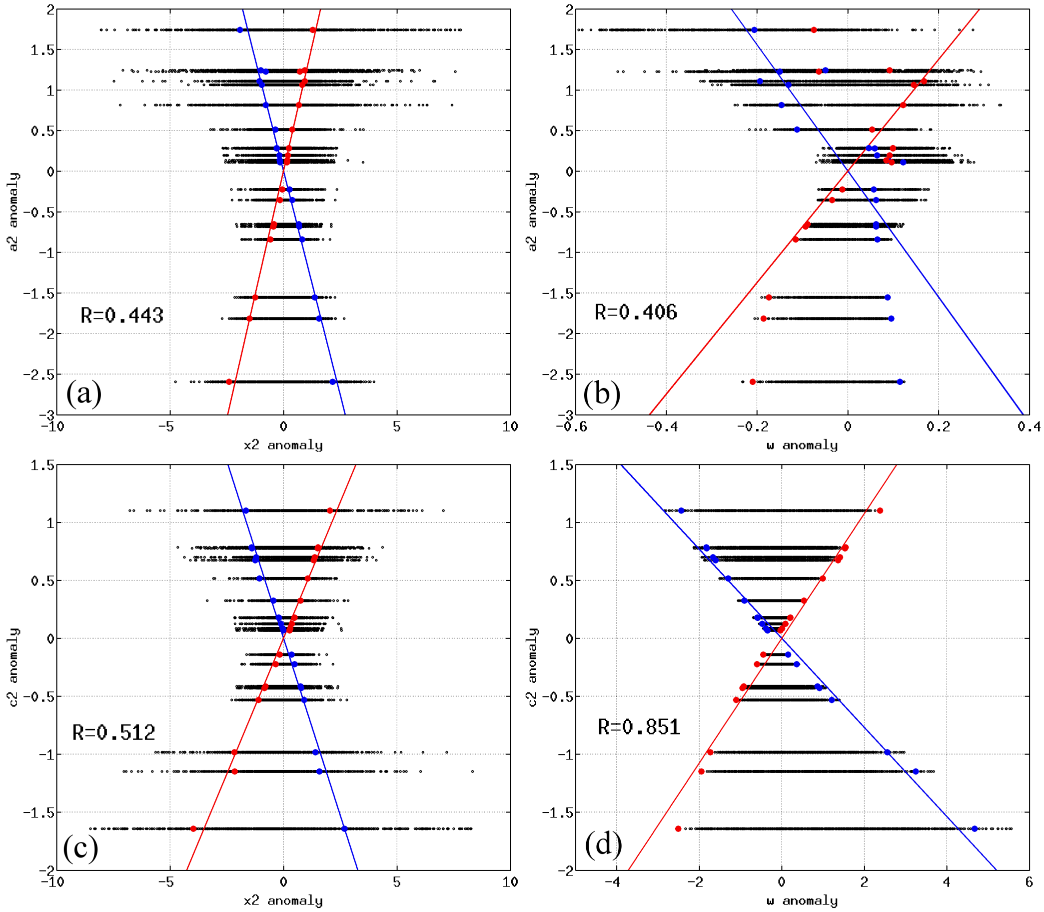


Figure 7: Sampling map of the perturbed parameter anomalies in the space of model state anomalies for a) a_2 vs. x_2 , b) a_2 vs. w, c) c_2 vs. x_2 and d) c_2 vs. w, when the atmospheric state is constrained by its observations. Dots with the same color (red or blue) represent ensembles at the same time 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.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Open Access
Discussions

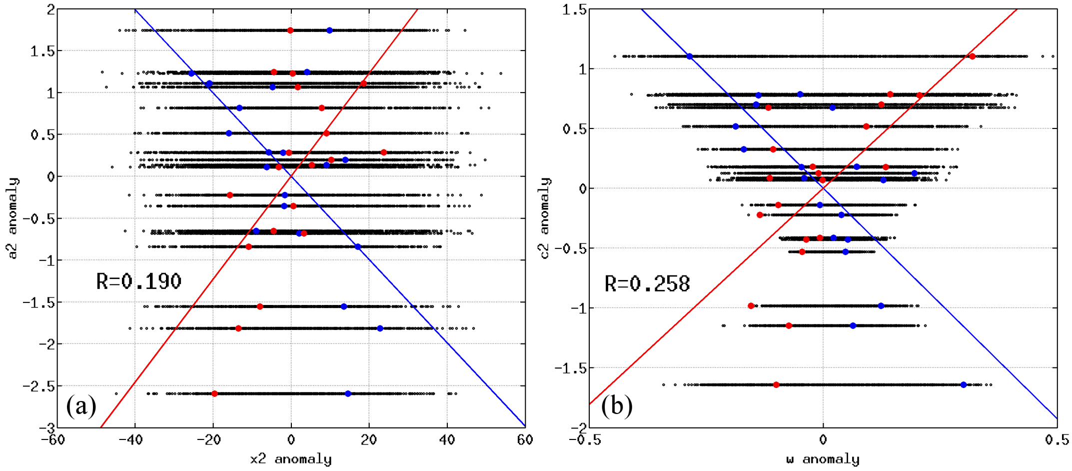


Figure 8: Same as Fig. 7 but for the case with SE of *w* only, a) a_2 vs. x_2 and b) c_2 vs. *w*. Here we show two examples that the linear fitting becomes difficult in red and blue, for which the data are taken from the same time steps as shown in Fig. 7.

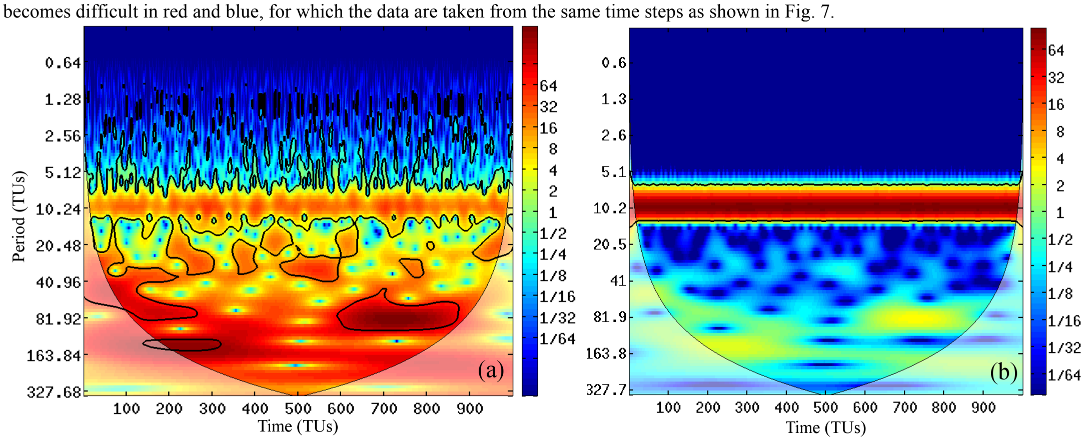


Figure 9: Wavelet analyses for *w* in the run of one-way coupling model forced by a) S_s=10 and b) S_s=250.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Open Access
Discussions

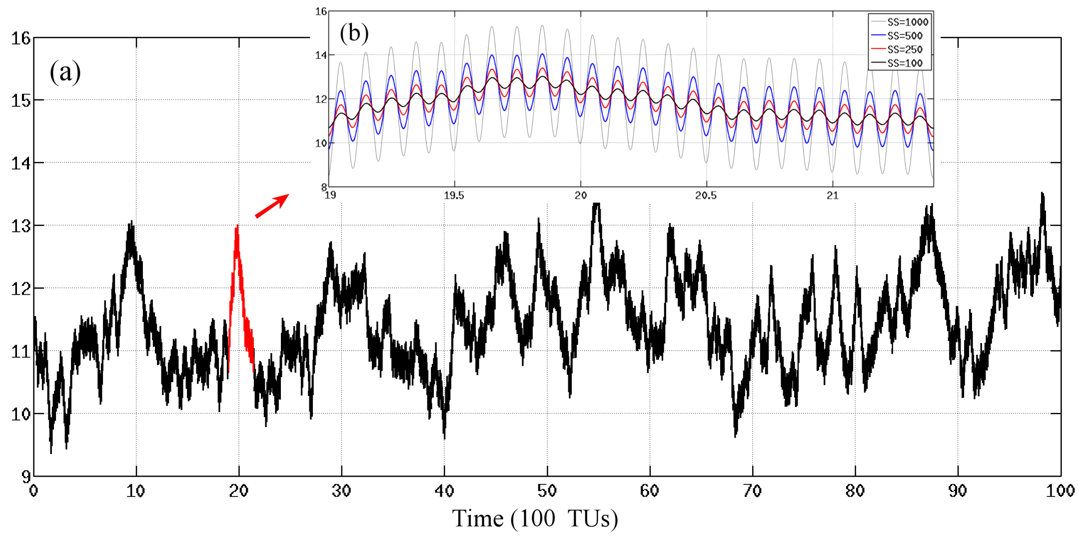
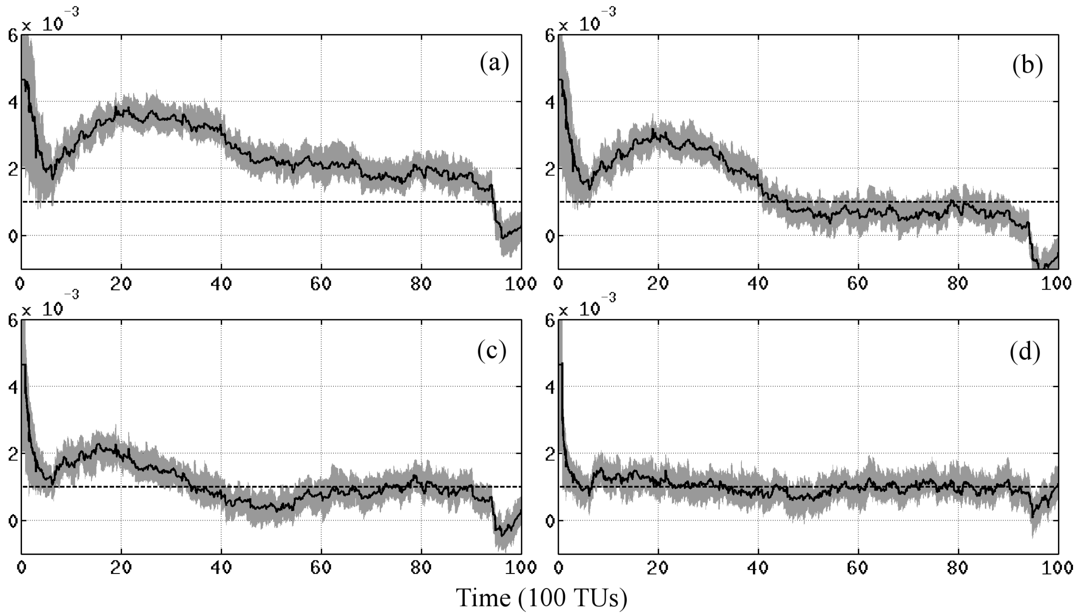


Figure 10: Time series of η with different S_s values (varying from 100 to 1000) with a one-way coupling model setting described in Appendix A. To visualize the difference induced by different S_s values, panel b) is the zoomed out version of the section marked in red in panel a).



Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-52, 2016
Manuscript under review for journal Nonlin. Processes Geophys.
Published: 27 September 2016

Nonlinear Processes in Geophysics
Open Access
Discussions

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.

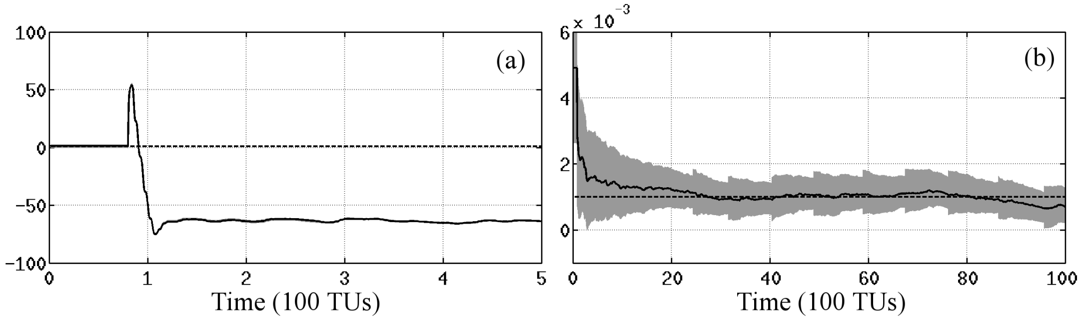


Figure 12: Time series of the ensemble of the parameter in the a) w-to-c_2 PE with SE of w only and b) η-to-c_6 PE with SE of η only using the one-way coupling model with S_s=250. Note that the initial c_2 in panel a) is approximate 0.56, and the truth is 1. Any other notations are same as Fig. 1.