Ensemble Data Assimilation in a Simple Coupled Climate Model: The Role of Ocean-Atmosphere Interaction

Zhengyu Liu [1,2,*], Shu Wu[2], Shaoqing Zhang[3], Yun Liu[2], Xinyao Rong[4]

1. Laboratory for Climate, Ocean and Atmosphere Studies, Peking University, China
2. Nelson Institute Center for Climatic Research & Dept. Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, USA
3. Geophysical Fluid Dynamics Lab./NOAA, USA
4. Meteorological Research Institute, Chinese Meteorological Administration, China

Adv. Atmos. Sci.

Submitted, 24, Oct, 2012

Revision, 11, Jan., 2013

Final, 20, Jan., 2013

[*] Corresponding author address: Zhengyu Liu, 1225 W. Dayton St., Madison, WI 53706, USA, zliu3@wisc.edu

Abstract

A conceptual coupled ocean-atmosphere model is used to study coupled ensemble data assimilation schemes with the focus on the role of ocean-atmosphere interaction in the assimilation. The optimal scheme is the fully coupled data assimilation scheme that employs the coupled covariance matrix and assimilates observations in both the atmosphere and ocean. The assimilation of synoptic atmospheric variability that captures the temporal fluctuation of the weather noise is found critical for the estimation of not only the atmospheric, but also oceanic states. The synoptic atmosphere observation is especially important in the mid-latitude system, where oceanic variability is driven by weather noise. The assimilation of synoptic atmospheric variability in the coupled model improves the atmospheric variability in the analysis and the subsequent forecasts, reducing error in the surface forcing and, in turn, in the ocean state. Atmospheric observation can further improve the oceanic state estimation directly through the coupled covariance between the atmosphere and ocean states. Relative to the mid-latitude system, the tropical system is influenced more by ocean-atmosphere interaction and, thus, the assimilation of oceanic observation becomes more important for the estimation of the ocean and atmosphere.

1. Introduction

As a flow-dependent data assimilation scheme, Ensemble Kalman Filter (EnKF) (Evensen, 1994; Tippett et al., 2003) in principle is equivalent to the 4-dimensional Variational Assimilation (4D-Var) scheme. Yet, EnKF is much more promising for the application to complex models such as coupled ocean-atmosphere general circulation models (OAGCMs), because it does not require an adjoint model. In an OAGCM, EnKF is critical in the model initialization for climate predictions (e.g. Zhang et al., 2009, 2010). Since the memory of the climate system lies in the ocean, most prediction studies have focused on the improvement of the initial state of the ocean. Previous works for the initialization in OAGCMs either used crude nudging schemes (e.g. Latif et al., 1993; Rosati et al., 1997; Luo et al., 2005; Smith et al., 2007; Keenlyside et al., 2008), or applied data assimilation in the component model separately (e.g. Ji et al., 1995; Rosati et al., 1997; Fuji et al., 2009). Recently, an EnKF scheme is implemented in an OAGCM for the assimilation of both atmospheric and oceanic data (Zhang et al., 2007). This scheme is found to improve the initial coupled state, and in turn, the seasonal climate prediction, significantly over that from a traditional 3-dimensional Variational Assimilation (3D-Var) ocean initialization (Zhang et al., 2008). However, except for a few studies in simplified coupled climate models (e.g. Sun et al., 2002; Zhang et al., 2011, Zhang, 2011a,b), EnKF has not been explored extensively in coupled climate models. This is due partly to the relatively new development of the EnKF method itself and partly to the more complex nature of the coupled climate system, especially the different time scales between the atmosphere and ocean. Therefore, important issues on EnKF assimilation in OAGCMs remain to be explored. Here, we are concerned with two questions. First, how important is the assimilation of synoptic atmospheric variability for coupled climate prediction? Second, what is the role of ocean-atmosphere coupling in coupled data assimilation and for the initialization and climate prediction?

There have been studies that suggest the importance of the assimilation of atmospheric observations in climate prediction, notably El Nino Southern Oscillation (ENSO) prediction. Using a simple nudging scheme, forecast is improved using the initial ocean state that is forced by the observed surface wind (Cane et al., 1986; Latif et al., 1993), and furthermore, the initialization is obtained by assimilating the observed surface wind in the coupled mode, instead of forcing the ocean in the ocean-alone mode (Chen et al., 2002). Using an EnKF, ENSO forecast is improved by including the assimilation of atmospheric observations in the coupled model, relative to that initialized using the ocean-alone 3D-VAR assimilation (Zhang et al., 2008). Yet, there have been no studies that systematically explored the roles of coupled assimilation and atmospheric observation in the coupled system.

Here, we will explore the role of coupled assimilation and the role of atmospheric observation in coupled EnKF data assimilation systematically. As a pilot study here, we will apply EAKF (a type of EnKF, Anderson, 2001, 2003) to a simple conceptual coupled ocean-atmosphere model. We will compare various coupled assimilation schemes with the focus on the role of ocean-atmosphere coupling in the coupled system. Special attention is also paid to the role of synoptic atmospheric observations in the coupled assimilation. The coupled climate will be studied in two settings, a mid-latitude-like system and a tropical-like system, the former being driven completely by weather noises. Our study shows that the fully coupled assimilation scheme, which assimilates both oceanic and atmospheric observation through the coupled covariance matrix, gives the best analysis. This optimal analysis is achieved because the assimilation of synoptic atmospheric variability improves the surface atmospheric forcing to the ocean. In particular, high frequency atmospheric data captures the temporal behavior of the weather noise and therefore improves the surface "stochastic" atmospheric forcing to the ocean. The weather noise forcing is particularly important in the mid-latitude system. In addition, the coupled covariance between the atmospheric and oceanic states further improves the oceanic state directly in the analysis through the background covariance between the atmosphere and ocean.

The paper is arranged as follows. We will describe our conceptual coupled climate model in section 2. We will then compare different coupled assimilation schemes in the mid-latitude and tropical systems in section 3 and 4, respectively. A summary and discussion will be given in section 5.

2. The Model

The simple climate model consists of a fast and chaotic "atmosphere" and a slowly oscillating "ocean". The atmospheric "wind", or "weather noise", is governed by the Lorenz63 model (Lorenz, 1963)

$$\begin{aligned} m_l \frac{dx_1}{dt} &= a_l(x_2 - x_1) \\ m_l \frac{dx_2}{dt} &= b_l x_1 - x_2 - x_1 x_3 \\ m_l \frac{dx_3}{dt} &= x_1 x_2 - c_l x_3 \end{aligned} \quad , \qquad (1)$$

where the factor $m_l = 1/6$ is used to match the time steps of the Lorenz model with the rest of model equations. The "surface air temperature" T_a is determined by an idealized thermodynamic model

$$m_a \frac{dT_a}{dt} = c(T - T_a) - \mu_a T_a + c_4 x_2 . \tag{2}$$

The slow ocean consists of the "sea surface temperature" (SST) T and "thermocline depth" h, which are described by an oscillator model (Jin, 1997) :

$$\begin{aligned} \frac{dT}{dt} &= RT + \gamma h + c(T_a - T) + c_2 x_2 - e_n (h + bT)^3 \\ \frac{dh}{dt} &= -rh - \alpha bT \end{aligned} \tag{3}$$

The default model parameters are

$$a_l = 10,\ b_l = 28,\ c_l = 8/3,\ m_a = 1/20,\ \mu_a = 1/3, \tag{4}$$

for the atmosphere,

$$\alpha = 0.125,\ \ \gamma = 0.75,\ \ r = 0.25,\ b_0 = 2.5,\ \mu = 0.5,\ b = b_0 \mu,\ R = \gamma b - 1 = 0.3125,\ \ e_n = 1, \tag{5}$$

for the ocean,

$$c = 1, \tag{6a}$$

for thermal coupling, and

$$c_2 = 0.05,\ c_4 = 0.1. \tag{6b}$$

for the forcing of weather noise. All variables are in the nondimensional form, with a nondimensional time *t~1* corresponding to a dimensional time ~2 months. The model is solved using a 4-th order Runge-Kutta method, with a time step of *dt=0.002* (~2.88 hrs, or 250 steps ~ 1 month).

In this conceptual coupled model, the Lorenz63 model can be thought to represent internal atmospheric variability of, say, “wind”; this wind component is induced by the chaotic instability of the atmosphere itself and is independent of oceanic feedback. The wind variability acts as a weather noise that drives the air temperature (via the term $c_4 x_2$) and SST (via the term $c_2 x_2$) variability.[1] The air temperature is coupled with SST through a negative ocean-atmosphere feedback $c(T\text{-}T_a)$ and thus represents the part of atmospheric variability that is strongly coupled with the ocean. The ocean model was originally derived for the tropical coupled ocean-atmosphere system (as the recharge oscillator model, Jin, 1997) with an internal oscillation mode of ~ 2-3 years. This oscillator is used here symbolically to represent an ocean-alone system. To avoid confusion, this model will be called the ocean oscillator model hereafter.

In spite of its simplicity, the conceptual model captures the essential feature of a coupled system, with a fast atmosphere (days) coupled with a slowly varying ocean (months to years). The model parameters for the atmosphere wind model (1) and the oceanic model (3) are the the standard parameters of Lorenz (1963) and Jin (1997), respectively, except for the tunable relative coupling strength μ. Other model parameters are tuned such that the coupled model captures some important statistical features of the coupled variability in a much more realistic system (see later discussion on Figs.2 and 4) such that this model may be of relevance to more complex climate systems. We constructed two model settings, a mid-latitude-like and a tropics-like coupled systems. In the mid-latitude system, parameters take the default values in eqns. (4)-(6). In particular,

[1] The internal variability “wind” can also be thought as “precipitation”, which forces salinity variability in the ocean but with little feedback from the salinity.

the oceanic instability parameter is small ($\mu = 0.5$ in (5)) such that the oceanic mode is a damped oscillating mode. As such, the mid-latitude system is driven completely by the atmospheric noise using large forcing parameters $c_2=0.05$, $c_4=0.1$ in (6b). In the tropical system, the atmospheric forcing effect is reduced by 10 times to $c_2=0.005$, $c_4=0.01$. Furthermore, the instability is enhanced with $\mu = 1.5$ such that the oceanic mode becomes self-exciting. Mathematically, the mid-latitude system is a damped system forced by strong stochastic noise, while the tropical system is a self-exciting system modified by weak stochastic noise. [2]

In the mid-latitude system, the atmospheric wind exhibits fast and chaotic variability (Fig.1b). The ocean exhibits slow irregular oscillation punctuated by rapid events associated with the atmospheric forcing (Fig.1a); the air temperature consists of fast variability due to the wind and slow variability due to SST feedback (Fig.1b). The mid-latitude system captures some major features in a state-of-art OAGCM, the National Center for Atmospheric Research Community Climate System Model version 3.5 (NCAR CCSM3.5), as seen by comparing the lagged correlation in the mid-latitude North Atlantic in the OAGCM CCSM3.5 (Fig.2a) and in the simple model (Fig.2b). In the CCSM3.5 (Fig.2a) and the simple model (Fig.2b), both auto-correlations imply a short decorrelation time less than a month for the surface wind and a long decorrelation time of several months for the SST. Both autocorrelations of the air temperature decline rapidly in the first month and then slowly for several months, both attributed by the fast atmospheric wind and slow SST feedback. Both cross-correlations between wind and

[2] The intensity of noise forcing plays the critical role here. The result remains robust for the mid-latitude system when the instability parameter is increased to $\mu = 1.5$, and remains robust for the tropical system when the instability parameter is reduced to $\mu = 0.5$.

SST are higher for wind leading SST than for SST leading wind, suggesting that the wind is a major driving agent for SST variability with little feedback from SST. In comparison, both cross-correlations between air temperature and SST are more symmetric with lead-lags, although the correlations are still stronger for air temperature leading SST. This reflects the nature of the negative ocean-atmosphere feedback in the mid-latitude, with the air-sea heat flux playing a dual role of first driving and later damping the SST (Frankignoul et al., 1998). Therefore, the simple model captures some statistical features of ocean-atmosphere feedback in more realistic systems.

In the tropics, the ocean exhibits a self-exciting oscillation without any perturbation. Fig.3a and b show a self-exciting solution perturbed weakly by the chaotic atmosphere. In comparison with the mid-latitude in Figs.1a, b, the tropical solution exhibits a much more regular cycle perturbed by weak noise. Due to the weak impact of weather noise, the lagged correlation shows that, in both the OAGCM (meridional wind, Fig.4a) and the simple model (Fig.4b), the air temperature almost co-vary with SST, while the wind is almost uncorrelated with SST.

In short, in spite of its idealized nature, the simple model captures important features of the coupled ocean-atmosphere system and therefore provides a useful tool for exploring the role of ocean-atmosphere interaction in coupled assimilation.

3. Coupled Assimilation in the Mid-latitude System

We now study different schemes of data assimilation in the coupled mid-latitude model in the perfect model scenario, with the focus on the ocean state, whose long memory is critical for climate predictability. First, a control simulation is performed with the initial condition $h=0$, $T=0$, $T_a=0.15$, $x_1=x_2=x_3=0.0001$ (Figs.1, 3). The model is spun off and then integrated for 200 yrs to represent the "truth". A synthetic observation is constructed by adding an observational noise onto the truth. The observational error for each variable is an independent Gaussian noise with a standard deviation 10% that of the control simulation. Unless otherwise specified, the coupled model assimilates the observation every 10 steps (~1.2 days) for the atmosphere and 40 steps (~5 days) for the ocean. Each ensemble has 20 members and each assimilation is integrated for 200 years with no inflation on the background covariance. The initial condition for the ensemble member is constructed from the observation at the time with a small random perturbation. Here, we discuss the results with all observational variables assimilated. When a subset of the observational states are assimilated, the results remain qualitatively consistent. Further sensitivity experiments show that our major conclusion remains qualitatively valid for other settings, including assimilation time steps, ensemble members， the magnitude of the observational error and the inflation factors.

We first compare three coupled assimilation schemes in the mid-latitude system, all using the coupled background covariance matrix in the filter analysis: CP-A assimilates the atmospheric observation only, CP-O assimilates the oceanic observation only, and CP-AO assimilates both atmospheric and oceanic observations (Table 1). We will compare the results of these schemes in terms of the normalized RMSEs (root mean square error normalized by the standard deviation of the control)[3]. The most comprehensive scheme is the fully coupled assimilation scheme CP-AO, which assimilates observations of both the atmosphere and ocean. The RMSE is reduced to 30%

[3] To reduce the impact of the outlier problem in EAKF (Lawson and Hansen, 2004; Anderson, 2010; Liu et al., 2012), a simple approach is used here: for each scheme, the RMSE is calculated with the top 5% of the RMSEs excluded (the result is similar if the top 1% is excluded). This way, our major conclusions become robust for different assimilation settings and model parameters.

(~0.03) and 3% (~0.003) of the observational errors for the atmosphere and ocean, respectively (Fig.5). (note, in Fig.5, the uncoupled scheme As-O will be discussed later in section 3b). If only the ocean observation is assimilated (CP-O), the RMSE is reduced to 20% (~0.02) and 85% (~0.085) of the observational errors for *h* and SST, respectively (Fig.5), but remains comparable with the control for the atmospheric variables, with the RMSEs of 0.55 and 0.9[4] (both off scale in Fig.5) for air temperature and winds, respectively. The modest oceanic errors, especially for SST, are much larger than those in CP-AO, suggesting the importance of the atmospheric observation for the ocean state in the coupled assimilation. The poor constrain of the ocean observation on the atmosphere is expected because the wind does not respond to SST (as in eqn. (1), and the poor correlation < 0.2, Fig.A1b), and the air temperature is driven primarily by the stochastic wind forcing with only a weak response to SST (correlation < 0.4, Fig.A1b).

In contrast, when the atmospheric observation is assimilated into the coupled model (CP-A), the analysis is improved dramatically. The RMSE of CP-A is reduced to almost the same level as in CP-AO (Fig.5). This suggests that, for the mid-latitude system, atmospheric observation can play a much more important role than the oceanic observation for the coupled state. It is interesting that the atmospheric observation is even more important than the oceanic observation itself for the ocean state. The critical importance of the atmospheric observation here can be understood, partly, from the dynamic nature of the mid-latitude coupled system. The SST variability is forced by synoptic atmospheric variability, which is often considered as stochastic noise at the slow

[4] Even though the atmospheric wind is forcing the air temperature and SST dynamically, with no dynamic feedback at all as shown in eqn. (1), the wind is still improved slightly by oceanic observations (normalized RMSE below 1 in CP-O). Our further experiments show that this improvement is due to the background covariance between the wind and air temperature used in the analysis. Therefore, SST observation improves the air temperature, and in turn, wind. The instantaneous covariance allows the “response” variable to improve the “forcing” variable.

ocean (and climate) time scale (Frankignoul and Hasselmann, 1977). This dominant role of atmospheric forcing on SST is shown clearly in the lagged correlation between SST and air temperature (Fig.A1f), where the maximum correlation (~0.6) occurs when air temperature leads SST (by ~80 steps). Therefore, as synoptic atmospheric forcing is improved, the ocean state is also improved.

a) The role of synoptic atmospheric forcing

We now further explore the role of synoptic atmospheric observation on the coupled assimilation. As atmospheric observation becomes less frequent, we speculate that the effect of the atmospheric observation on the coupled, in particularly the oceanic, state, will be reduced. Less frequent atmospheric observation should increase the analysis error in both CP-A and CP-AO and, furthermore, the error will increase faster in CP-A than in CP-AO because the latter is constrained by the ocean observation. This speculation is confirmed by two sets of assimilation experiments in CP-A and CP-AO, in which the atmospheric observational steps are increased from 10 to 640 steps systematically (while the ocean observation remains fixed at 40 steps). Fig.6 shows the RMSE ratio between the CP-A and CP-AO experiments as a function of the atmospheric assimilation steps. Since ocean variability is forced by the entire history of the atmospheric forcing, as a measure of the error of the atmospheric forcing, the RMSEs here are accumulated over both analysis and forecast steps[5]. Overall, as the steps of the atmospheric observation increase, the RMSE ratio tend to increase for the ocean (Fig.6a) and air temperature (Fig.6b), indicating a faster increase of RMSE in CP-A than in CP-AO. Therefore, ocean observations become more important for the ocean and air

[5] The variation of the RMSE ratio also remains similar for the analysis RMSEs (not shown).

temperature as atmospheric observations become less frequent. (The ratio of RMSEs for wind remains ~1 (not shown) because of the lack of oceanic impact on wind). In Fig.6b, the RMSE ratio for air temperature increases from 1 (at step 10) to 1.15 (at step 640) (the slight decreases at steps 20 and 80 are likely caused by sampling error). Therefore, the RMSE of air temperature increases slightly faster in CP-A than in CP-AO, reflecting the weak impact of SST on air temperature (Fig.6b). The faster error growth in the atmospheric forcing then leads to a faster error growth in the ocean in CP-A than in CP-AO, and the RMSE ratios for oceanic variables increase eventually much beyond 1 for large atmospheric observational steps (Fig.6a). Indeed, in the limit of very large atmospheric observational steps, the RMSE of oceanic variability in CP-A saturates towards the control (~60% of control at step 640, not shown) because the CP-A scheme now uses virtually no observations in the atmosphere and ocean; the RMSE of oceanic variability in CP-AO, however, saturates towards that of CP-O (about 5%-10% of the control at step 640, not shown), because CP-AO now still uses full oceanic observations (every 40 steps). Since the RMSE of the ocean is much larger in CP-A than in CP-O, the RMSE ratio between CP-A and CP-AO grows very large in the ocean, especially for *h*.

In spite of this overall increase trend of the RMSE ratio, it is important to note that the RMSE ratio remains close to ~1 for air temperature (Fig.6b) and ocean (Fig.6a) for sufficiently high frequency of atmospheric observations, notably at steps 10, 20 and even 40. This occurs because the atmospheric observation is so frequent that the forecast error has not grown significantly in the atmosphere, therefore, the error of the atmospheric forcing is not much larger in CP-A than in CP-AO (as seen in the RMSE ratio of air temperature in Fig.6b). The atmospheric forcing is therefore sufficiently accurate in CP-A such that the addition of oceanic observations in CP-AO does not improve the ocean state significantly (Fig.6a). This argument also implies that the critical frequency of atmospheric observation should be significantly shorter than the saturation time of forecast error, or crudely the persistence time. The atmospheric decorrelation time is less than ~40 steps for wind (Figs.A1c-e), and less than ~150 steps for air temperature (using a cut off correlation of ~0.2). Therefore, the critical frequency beyond which the RMSE ratio increases above 1 should be shorter than ~40 – 150 steps, consistent with the ~40 steps in Fig.6a. In short, if the atmospheric observation is sufficiently shorter than its persistence time, the atmospheric observation is able to improve the atmospheric forcing and, in turn, the oceanic variability, significantly, in the coupled system.

b) Coupled vs. uncoupled assimilation schemes

We now compare the fully coupled scheme against an uncoupled assimilation scheme As-O (Table 1). The As-O scheme assimilates both atmospheric and oceanic observations, but separately in a two-tier approach: first, the atmospheric observation is assimilated in the atmosphere model forced by the SST observation (Specifically, the SST forcing at each step is derived from the SSTs at the observational steps using a linear interpolation). Second, the atmospheric forcing (at analysis and forecast steps) is used to force the ocean model in its assimilation of oceanic observations. The atmospheric analysis here is equivalent to the standard atmospheric reanalysis product. For the oceanic state, the As-O scheme is equivalent to an ocean data assimilation forced by an atmospheric reanalysis product. In a sense, As-O is similar to many previous works for the initialization of the ocean state for climate predictions in coupled climate models (e.g.

Cane et al., 1986; Latif et al., 1993; Rosati et al., 1997) (although the assimilation schemes there are not ensemble filters). A comparison of the RMSEs in As-O and CP-AO (Fig.5) shows that, even with the same atmospheric and oceanic observations, the RMSE is significantly higher in As-O than in CP-AO, especially for the ocean. The improved analysis in CP-AO over As-O is due, partly, to the improvement of the SST forcing (to the atmosphere) through the coupled dynamics. Indeed, the RMSE of the SST analysis in CP-AO is reduced from the observational error (~0.1, Fig.5) (which is the error for the SST forcing in As-O) to less than 5% of the observational error (< 0.005, Fig.5). Relative to As-O, the improved SST forcing CP-AO improves the atmosphere dynamically, which then improves the ocean dynamically. Indeed, even with the additional assimilation of oceanic observations, the analysis of As-O is significantly poorer than that in the coupled scheme CP-A for the ocean state and air temperature (Fig.5), even though the latter only assimilates the atmospheric observation. This is consistent with the critical importance of synoptic atmospheric observations as discussed in Fig.6.

To further evaluate the role of atmospheric surface forcing, we performed another uncoupled oceanic assimilation (not shown) that is the same as As-O except that the atmospheric forcing is replaced by that in CP-AO at every time step. The RMSE in the ocean is now reduced by about half of that in As-O (due to the improved atmospheric forcing), but the RMSE still remains significantly higher than in CP-AO, even though both ocean assimilations used the same atmospheric forcing. This implies that the improved surface atmospheric forcing through the coupled dynamics is not the only cause for the improved assimilation in the coupled scheme CP-AO over the uncoupled scheme As-O.

c) The role of coupled background covariance

In principle, ocean-atmosphere coupling affects the coupled data assimilation not only through the coupled dynamics, but also through the coupled covariance in the filter analysis. To further explore the difference between the coupled and uncoupled schemes, especially the role of the ocean-atmosphere interaction through the coupled covariance, we further compare the fully coupled scheme CP-AO with another coupled scheme: the dynamically coupled scheme CP-ABOB (Table 1). In CP-ABOB, atmospheric and oceanic observations are assimilated as in CP-AO except that the background covariance matrices for the atmosphere and ocean only use the sub-matrices for each component separately. Specifically, denote the transposes for atmospheric and oceanic variables as $\mathbf{A}=[x_1, x_2, x_3, T_a]^T$ and $\mathbf{O}=[T, h]^T$, respectively, the background covariance matrix is

$$\mathbf{B}=\begin{bmatrix} \mathbf{B_{AA}} & \mathbf{B_{AO}} \\ \mathbf{B_{AO}} & \mathbf{B_{OO}} \end{bmatrix}, \tag{7}$$

in CP-AO, but

$$\mathbf{B_{ABOB}}=\begin{bmatrix} \mathbf{B_{AA}} & 0 \\ 0 & \mathbf{B_{OO}} \end{bmatrix}. \tag{8}$$

in CP-ABOB. Here $\mathbf{B_{AA}}=<\mathbf{A},\mathbf{A}>$, $\mathbf{B_{OO}}=<\mathbf{O},\mathbf{O}>$, $\mathbf{B_{AO}}=<\mathbf{A},\mathbf{O}>$.

A comparison of CP-ABOB and CP-AO (Fig.7) shows that the RMSEs are comparable for the atmosphere, but is significantly greater in CP-ABOB than CP-AO for the ocean. Therefore, atmospheric observations can improve the ocean significantly in the fully coupled scheme CP-AO directly through the coupled covariance. This improvement is further shown to be caused completely by the impact of the atmospheric observation on ocean. This is shown in two additional partially coupled experiments CP-A2OB and CP-O2AB, which respectively use the coupling covariance $\mathbf{B_{AO}}$ on the ocean and atmosphere, respectively, with the corresponding background covariance matrices

$$\mathbf{B_{A2OB}} = \begin{bmatrix} \mathbf{B_{AA}} & 0 \\ \mathbf{B_{AO}} & \mathbf{B_{OO}} \end{bmatrix}, \qquad \mathbf{B_{O2AB}} = \begin{bmatrix} \mathbf{B_{AA}} & \mathbf{B_{AO}} \\ 0 & \mathbf{B_{OO}} \end{bmatrix}. \tag{9}$$

Fig.7 shows almost the same RMSEs in CP-A2OB and the fully coupled CP-AO, but almost the same RMSEs in CP-O2AB and the dynamically coupled CP-ABOB. Therefore, for the mid-latitude system here, the impact of the coupled covariance on the coupled analysis is due to the atmospheric impact on ocean, with little oceanic impact on the atmosphere.

It is also interesting to compare CP-ABOB with the uncoupled scheme As-O. Fig.7 shows that the RMSE is smaller in CP-ABOB than in As-O for air temperature, thermocline and SST. The error reduction in air temperature confirms that atmospheric observations improve the atmosphere state more in the coupled model than in the uncoupled atmospheric model, because the SST forcing is improved over the observation (used in As-O) by the coupled dynamics in the coupled model. For the ocean state, we may attribute the reduced RMSE from CP-ABOB to CP-AO to the coupled covariance, and from As-O to CP-ABOB to the improvement of atmospheric forcing in the coupled model.

In short, high frequency synoptic atmospheric observation improves the coupled state significantly because of its improvement on the atmospheric analysis and, in turn, the surface forcing to the ocean. The fully coupled assimilation CP-AO improves the ocean significantly over the uncoupled scheme As-O for two reasons: the coupled dynamics improves the atmospheric forcing by improving the SST forcing to the atmosphere (from As-O to CP-ABOB), and, the coupled background covariance allows the atmospheric observation to improve the ocean state through the analysis directly.

4. Coupled Assimilation in the Tropical System

We now discuss the tropical system briefly, in comparison with the mid-latitude system. We will show that the major conclusions in the mid-latitude system still hold qualitatively in the tropical system: the fully coupled scheme gives the optimal coupled state and high frequency synoptic atmospheric observations can improve the ocean state significantly. Quantitatively, however, the stronger ocean-atmosphere coupling in the tropics renders synoptic atmospheric observation less important than in the mid-latitude, while oceanic observations become more important.

As in the mid-latitude system (Fig.5), the normalized RMSEs in CP-AO, CP-A and CP-O (Table 1) a minimum in CP-AO, and almost the same in CP-A and CP-AO. Therefore, CP-AO is the optimal scheme and synoptic atmospheric observation plays a dominant role. Meanwhile, the assimilation of the ocean observation in CP-O reduces the RMSEs by half compared with the mid-latitude system for ocean (*h* and *T*, ~0.01, ~0.045 in Fig.8, vs. ~0.02 and ~0.09 in Fig.5) and air temperature (~0.35 vs. ~0.65, off scale in Fig.8 and Fig.5), due to the stronger ocean-atmosphere coupling and the weaker weather noise forcing in the tropical system. Indeed, the stronger ocean-atmosphere coupling can be seen in the much larger correlation between SST and air temperature in the tropical (~0.9, Fig.A2b,f) than in the mid-latitude (~0.4, Fig.A1b,f) systems. The weaker weather noise forcing can also be seen in the lagged cross-correlation, which peaks almost simultaneously in the tropical system (Fig.A2f), rather than when the air temperature leads SST in the mid-latitude system (Fig.A1f). The increased role of oceanic observations in the tropical system can also be seen the RMSE ratio between CP-A and CP-AO in Fig.9. Although qualitatively similar to the mid-latitude system (Fig.7), an increase of atmospheric observational steps increases the RMSE more in CP-A than in CP-AO, quantitatively, the RMSE ratio increases significantly beyond 1 for ocean (Fig.9a) and air temperature (Fig.9b) at 20 steps, while it remains close to 1 even till ~40 steps in the mid-latitude system.

Coupling also improves the estimation, as in the mid-latitude. The RMSE is reduced from the uncoupled As-O to the coupled CP-AO (Fig.8), similar to the mid-latitude system (Fig.5). Quantitatively, the RMSE is reduced by 10 times in the tropical system (0.06 in As-O to 0.007 in CP-AO), but only by a half in the mid-latitude system (from 0.022 to 0.013), because of a greater role of ocean-atmosphere coupling in the tropical system. The coupled covariance also improves the estimation (Fig.10) as in the mid-latitude (Fig.7), in comparison of the fully coupled CP-AO with the dynamically coupled CP-ABOB. Quantitatively, however, the improvement is much less than in the tropics, as the RMSE in CP-ABOB is not much greater than in CP-AO for air temperature and ocean (Fig.10). Therefore, unlike the mid-latitude, where the coupled covariance is the major mechanism that improves the coupled over the uncoupled schemes, the improvement of the atmospheric forcing is the major mechanism that improves the coupled assimilation in the tropics. This is consistent with a stronger ocean-atmosphere coupling and, in turn, a stronger feedback of SST on air temperature in the tropical system.

5. Summary and Discussions

We studied several coupled schemes of EAKF in a simple coupled ocean-atmosphere model in the perfect model scenario, with the focus on the role of ocean-atmosphere interaction in the assimilation. Our study confirms that the optimal assimilation scheme is the fully coupled data assimilation scheme that assimilates observations in both the atmosphere and ocean and that employs the coupled covariance matrix. It is further found that the assimilation of synoptic atmospheric variability is critical for the improvement of not only the atmospheric state, but also the oceanic state, especially in the mid-latitude system, where oceanic variability is driven predominantly by weather noise. Furthermore, atmospheric observation can also improve the oceanic state through the coupled covariance, especially in the mid-latitude system. Relative to the mid-latitude system, the tropical system is influenced more by oceanic dynamics and ocean-atmosphere interaction. Therefore, the assimilation of oceanic observation becomes more important. This study suggests that the analysis of the coupled climate state variables are best derived in the fully coupled model using both the atmospheric and oceanic observations. Furthermore, synoptic atmospheric observations are critical for the improvement of the coupled analysis. Finally, coupled covariance between the ocean and atmosphere should also be employed to achieve the best analysis.

The importance of synoptic atmospheric observation for improving the ocean state has important implication for climate predictions: although the memory of the climate system lies in the ocean, synoptic atmospheric observations can significantly improve the ocean initial state and, in turn, climate prediction of slow oceanic variables. Therefore, the synoptic atmospheric observation alone is able to improve the coupled initial state in a balanced way (in both atmosphere and ocean), which will help improving climate prediction. We performed ensemble climate prediction experiments initialized by the coupled state of different assimilation schemes. Since each of our schemes (Table 1) improves the coupled state in both the atmosphere and ocean in a balanced way, it also improves the climate prediction of slow ocean state. For example, the RMSE is smaller in CP-AO than As-O in both the ocean and air temperature (Figs.5, 8), which in turn is smaller than those in CP-O; accordingly, the climate prediction of T and h deteriorate from CP-AO to As-O and finally to CP-O (not shown). One extreme example of unbalanced initial condition is the perfect ocean experiment (PO), as being used in some early studies of experimental decadal climate predictions (Collins et al., 2002). In PO, the ocean initial condition is the truth, while the atmosphere initial state is selected randomly from the control. A comparison of the climate prediction (Fig.11) shows that the prediction of the ocean state eventually becomes much worse in PO than in CP-AO after a very short lead time when ocean is almost perfect in PO. This occurs because the very large initial error in the atmosphere in PO quickly drives the ocean away from the truth.

It is interesting that the major conclusions of our conceptual model study seem to be consistent with previous studies in more realistic models. The importance of the atmospheric observations has been recognized even in the early stage of ENSO prediction, where less advanced assimilation schemes such as nudging are used for initialization (e.g. Cane et al., 1986; Latif et al., 1993). These studies found that a better forecast is achieved using the initial ocean state that is forced by the observed surface wind and the addition of further oceanic observation may not improve climate prediction significantly. Our conclusion that the assimilation in the coupled scheme (e.g. CP-A) improves the coupled state than the uncoupled assimilation (e.g. As-O) also appears to be consistent with Chen et al. (2002). They found that their ENSO prediction is improved if the initialization is obtained by assimilating the observed surface wind in the coupled mode, instead of forcing the ocean in the ocean-alone mode. The importance of synoptic wind for improving climate prediction is consistent with the EAKF study in an OAGCM (Zhang et al., 2008). This study shows that ENSO forecast is improved using the EAKF in the coupled model compared with the ocean-alone 3D-VAR assimilation.

Much further studies are needed, especially in more realistic models. One surprising result in our model is the overwhelming importance of synoptic atmospheric observation, such that the assimilation of synoptic atmospheric observation alone (CP-A) improves the coupled state almost the same as assimilating additionally oceanic observations (CP-AO). Equivalently, the assimilation of oceanic observation has little impact on the atmosphere, even the air temperature, as shown in CP-O. Previous studies with more realistic models, including OAGCMs show that the assimilation of oceanic observations in the coupled model can indeed improve the atmospheric state, especially in the tropics (Ji et al., 1995; Rosati et al., 1997; Luo et al., 2005; Fuji et al., 2009). The overwhelming role of synoptic atmospheric observation in our study could be related to the lack of dynamic ocean-atmosphere feedbacks in our idealized model, especially in the tropics. In a more realistic tropical system, the (zonal) wind anomaly is significantly correlated with SST, because of the strong dynamic response of the atmosphere to tropical SST anomaly (Gill, 1980; Lindzen and Nigam, 1987). This zonal wind effect is absent in our tropical system, which only simulates the meridional wind (Fig.4a) and therefore lacks the dynamic ocean-atmosphere feedback.

Acknowledgement

This work is supported by NSF, Chinese 2012CB955201 and NSFC 41130105. SZ is supported by NOAA.

Appendix: Lagged cross-correlations among model variables

To help us understand the nature of the covariance among different model variables, and in turn the ensemble filter analysis, the lagged cross-correlations among different model variables are shown for the mid-latitude system in Fig.A1 and for the tropical system in Fig.A2. See the text for discussions.

References

Anderson, J., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev*., **129**, 2894-2903

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

Anderson, J., 2010: A non-Gaussian ensemble filter update for data assimilation. *Mon. Wea. Rev*., **138**, 4186-4198

Cane, M., S. Zebiak and S. Dolan, 1986: Experimental forecasts of El Nino. *Nature*, **321**, 827-832

Chen, D., S. Zebiak, A. Busalacchi and M. Cane, 1995: An improved procedure for El Nino forecasting: implications for predictability. *Science*, **269**, 1699-1702

Collins, M., 2002: Climate Predictability on Interannual to Decadal Time Scales: The Initial Value Problem. *Clim. Dyn*. **19**, 671-692.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99,** 1043–1062.

Frankignoul C. and K. Hasselmann, 1977: Stochastic climate models. Part II: Application to sea-surface temperature variability and thermocline variability. *Tellus*, **29,** 284–305.

Frankignoul, C, A. Czaja, and B. L’Heveder, 1998: Air–sea feedback in the North Atlantic and surface boundary conditions for ocean models. J. Climate, 11, 2310–2324.

Fuji, Y., T. Nakaegawa, S. Matsumoto, T. Yasuda, G. Yamanaka, and M. Kamachi, 2009: Coupled Climate Simulation by Constraining Ocean Fields in a Coupled Model with Ocean Data, *J. Climate*, **22**, 5541-5557.

Gill, A. E., 1980: Some simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc.*, **106,** 447–462.

Ji, M., A. Leetmaa and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies. *Mon. Wea. Rev*., **123**, 460-481

Jin, F., F., 1997: An equatorial ocean recharge paradigm for ENSO. Part I: A conceptual model. *J. Atmos. Sci*., **54**, 811-829

Keenlyside, N., M. Latif, J. Jungclaus, L. Kornblueh and E. Roeckner, 2008: Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, **453**, 84-887.

Lawson, G. W. and J. A. Hansen, 2004: Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Monthly Weather Review*, **132**, 1966–1981.

Latif M., A. Sterl, E. Maier-Reimer and M. Junge, 1993: Structure and predictability of the El Nino/Southern Oscillation phenomenon in a coupled ocean-atmosphere general circulation model. *J. Climate*, **6**, 700-708

Lindzen R. S. and S. Nigam, 1987: On the role of sea surface temperature gradients in forcing low level winds and convergence in the tropics. *J. Atmos. Sci*., **44**, 2418-2436.

Liu, Y., X. Rong, Z. Liu, S. Wu, S. Zhang and R. Jacob, 2012: A random subgrouping scheme for ensemble square root filter (sEnSQF). *Mon. Wea. Rev*., submitted.

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

Luo, J.-J., S. Masson, S. Behera, S. Shingu, and T. Yamagata, 2005: Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. *J. Climate*, **18**, 4474–4497.

Smith, D., M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris and J. M. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. *Science*, 317, 796-799.

Sun C., Z. Hao, M. Ghil and D. Neelin, 2002: Data assimilation for a coupled ocean-atmosphere model: part I: sequential state estimation. *Mon. Wea. Rev.*, **130**, 1074-1099

Tippett, M. K., J. L. Anderson, C. H. Bishop , T. M. Hamill and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485- 1490

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

Zhang, S., A. Rosati, M. J. Harrison, R. Gudgel and W. Stern, 2008: GFDL's coupled ensemble data assimilation system, 1980-2006 coupled reanalysis and its impact on ENSO forecast. *Preprint WCRP3*, Jan. 27 - Feb. 2, 2008, Tokyo.

Zhang, S., A. Rosati and M. Harrison, 2009: Detection of multi-decadal oceanic variability by ocean data assimilation in the context of a ``perfect" coupled model. *J. Geophys. Res.*, **114**, C12018, doi:10.1029/2008JC005261.

Zhang, S., A. Rosati and T. Delworth, 2010: The adequacy of observing systems in monitoring the Atlantic Meridional Overturning Circulation and North Atlantic climate. *J. Climate*, **23**, doi:10.1175/2010JCLI3677.1, 5311-5324.

Zhang, S., 2011a: Impact of observation-optimized model parameters on decadal predictions: Simulation with a simple pycnocline prediction model. *Geophys. Res. Lett.*, **38,** L02702, doi:10.1029/2010GL046133.

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

Zhang, S., Z. Liu, A. Rosati and T. Delworth, 2011. A study of enhancive parameter correction with coupled data assimilation for climate estimation and prediction using a simple coupled model. *Tellus* **A**, **63**, 10963, DOI: 10.3402/tellusa.v63i0.10963

Table 1: Data Assimilation Schemes

Name	Atmos obs.	Ocean obs.	Model	Background Covariance Matrix
CP-AO	yes	yes	coupled	coupled
CP-A	yes	no	coupled	coupled
CP-O	no	yes	coupled	coupled
As-O	yes	yes	1st: atmos. model (forced by SST obs.), 2nd: ocean model (forced by atmos. analysis.)	atmos.-alone ocean-alone
CP-ABOB	yes	yes	coupled	atmos.-alone, ocean-alone
CP-A2OB	yes	yes	coupled	In CP-ABOB, add atmospheric covariance to ocean for oceanic analysis
CP-O2AB	yes	yes	coupled	In CP-ABOB, add oceanic covariance to the atmosphere for atmospheric analysis

Figure Captions

Figure 1. Time series of (a) SST (T) and ocean thermocline depth (h), (b) atmospheric winds (x_1,x_2,x_3) and air temperature (T_a) in the control simulation of the mid-latitude coupled system.

Figure 2. Auto correlations (solid) and cross-correlations (dash) of monthly SST, air temperature and wind in (a) CCSM3.5 North Atlantic average and (b) the mid-latitude coupled system. The wind is the zonal surface wind in (a) and x_2 in (b). The cross-correlations are between SST and the atmospheric temperature and wind, with the positive lags for SST leading the atmosphere.

Figure 3. Same as Fig.1 but for the tropical coupled system.

Figure 4. Same as Fig.2 but for the tropical coupled system.

Figure 5. Analysis RMSE (normalized by the standard deviation of the control run) of all the 6 variables for different assimilation schemes in the mid-latitude coupled system: OB: Observation, coupled schemes CP-A, CP-O, CP-AO (CP-AO and CP-A almost overlap with each other) and the uncoupled scheme As-O (see Table 1). The observational time steps for the atmosphere and ocean are 10 and 40 steps, respectively. The RMSE is calculated as the average of the RMSEs at all the analysis steps.

Figure 6. The ratio of RMSE (accumulated for all time steps) between CP-A and CP-AO as a function of the time steps of atmospheric observation in the mid-latitude system. (a) SST and thermocline depth, (b) air temperature. The oceanic observation time step is fixed at 40 steps. (The ratio of RMSE at the analysis steps are similar).

Figure 7. Analysis RMSE (normalized by the standard deviation of the control run) for As-O (circle), CP-AO (solid dot), CP-ABOB (cross), CP-A2OB (triangle) and CP-O2AB (plus) in the mid-latitude coupled system for h, T, x_2 and T_a. An ensemble of 80 members is performed with the ensemble mean in marks and the ensemble spread (standard deviation) in double bars. The observational time steps for the atmosphere and ocean are 10 and 40 steps, respectively.

Figure 8. Same as Fig.5 but for the tropical system.

Figure 9. Same as Fig.6 but for the tropical system.

Figure 10. Same as Fig.7 but for the tropical system.

Figure 11: Forecast RMSE in the mid-latitude system for h (left) and T (right) initialized in PO (dash) and CP-AO (solid) schemes.

Figure A1. Lagged correlations among all model variables in the mid-latitude system. Each panel represents the pivotal variable that is used for lagged correlation with itself (auto-correlation) and other 5 variables (cross-correlations). The positive lead step is for this pivotal variable leading other variables. Each variable is represented in the same color, blue for h, green for T, red for x_1, cyan for x_2, purple for x_3 and yellow for T_a. For example, in panel (b), the auto-correlation of T is in blue, the cross-correlation between T and h, x_1, x_2, x_3 and T_a are in blue, red, cyan, purple and yellow, respectively.

Figure A2. Same as Fig.A1 but for the tropical system.

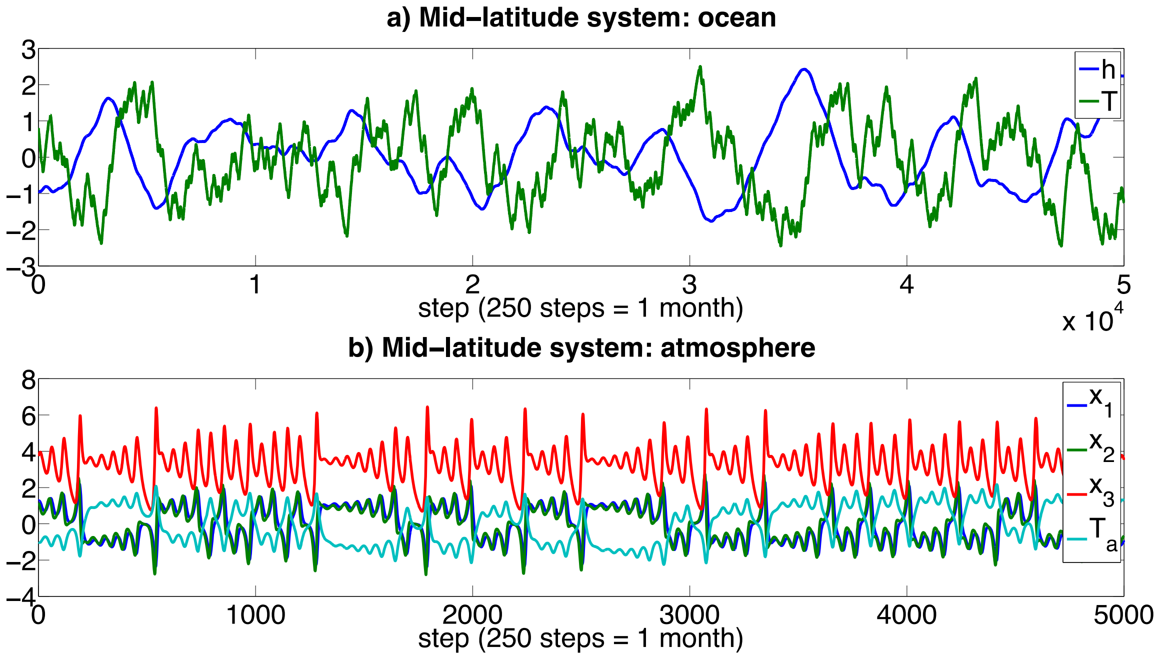

Figure 1. Time series of (a) SST (T) and ocean thermocline depth (h), (b) atmospheric winds (x_1,x_2,x_3) and air temperature (T_a) in the control simulation of the mid-latitude coupled system.

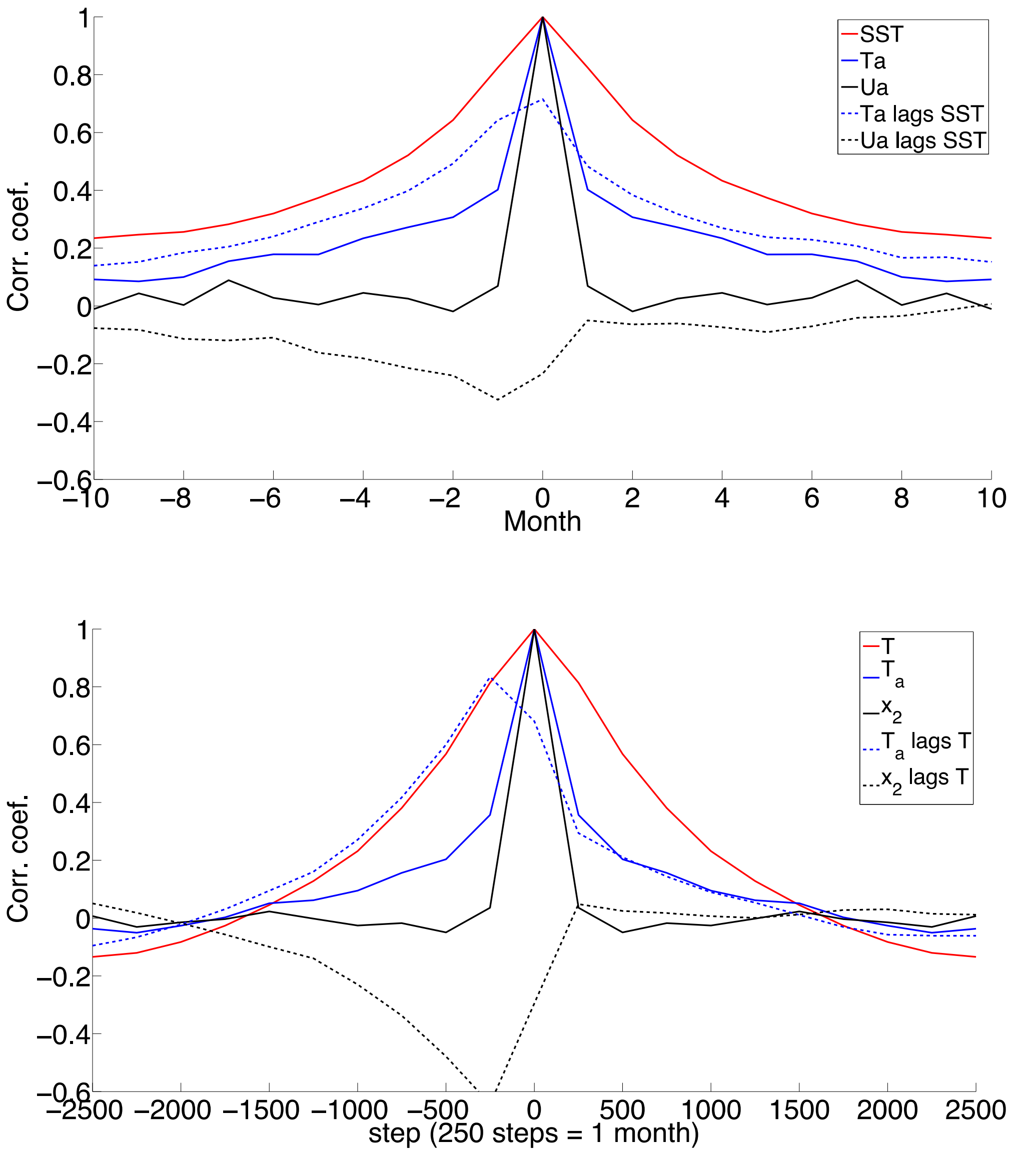

Figure 2. Auto correlations (solid) and cross-correlations (dash) of monthly SST, air temperature and wind in (a) CCSM3.5 North Atlantic average and (b) the mid-latitude coupled system. The wind is the zonal surface wind in (a) and x_2 in (b). The cross-correlations are between SST and the atmospheric temperature and wind, with the positive lags for SST leading the atmosphere.

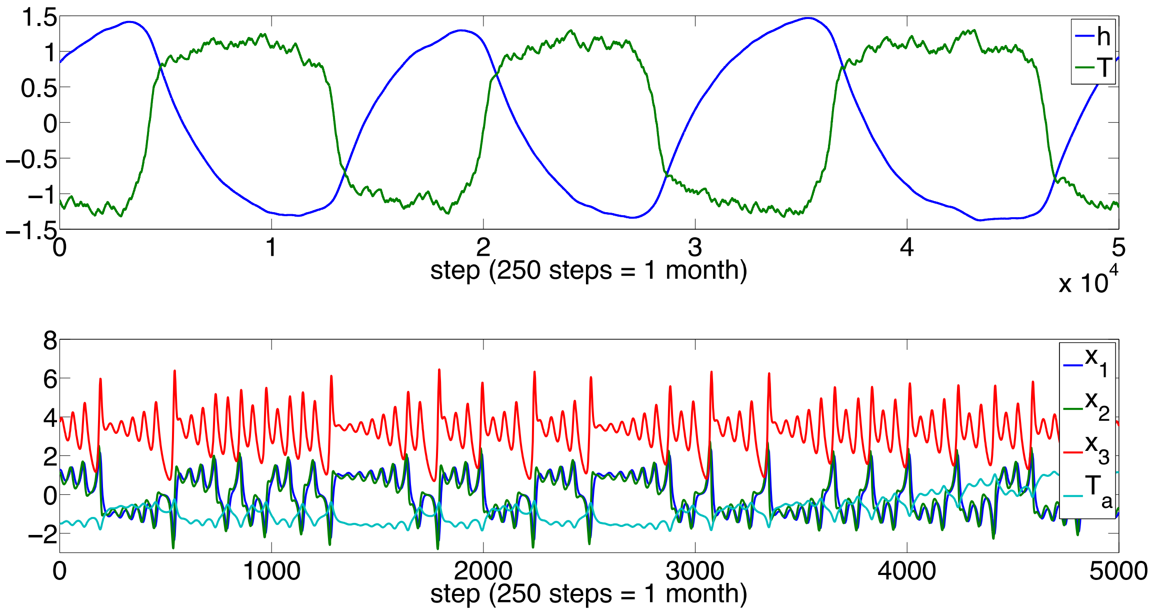

Figure 3. Same as Fig.1 but for the tropical coupled system.

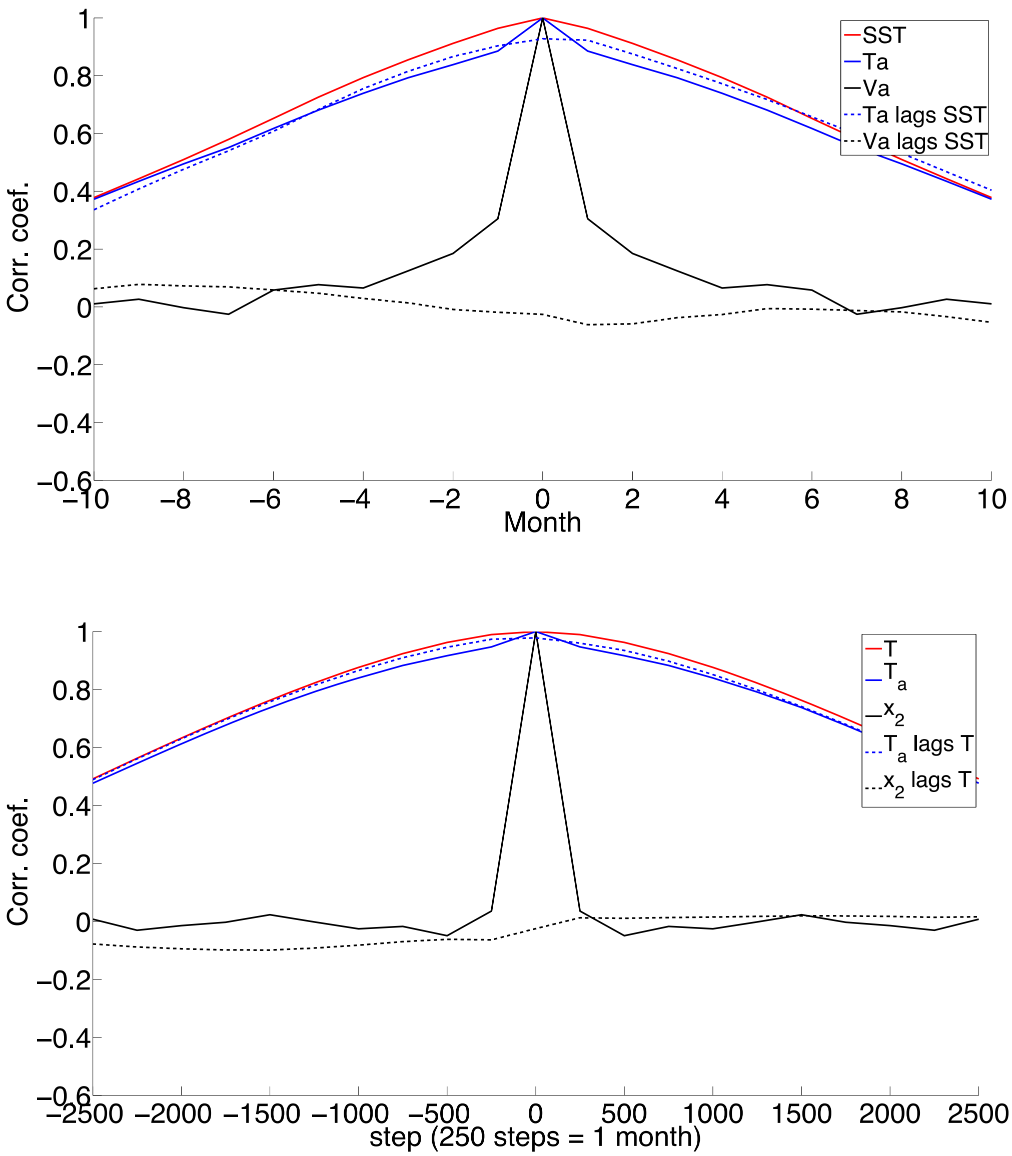

Figure 4. Same as Fig.2 but for the tropical coupled system.

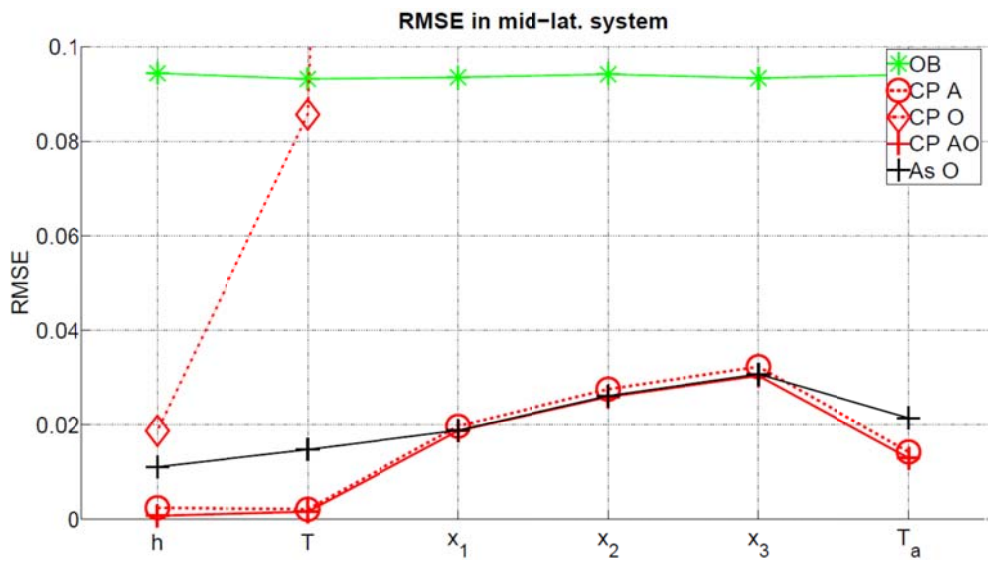

Figure 5. Analysis RMSE (normalized by the standard deviation of the control run) of all the 6 variables for different assimilation schemes in the mid-latitude coupled system: OB: Observation, coupled schemes CP-A, CP-O, CP-AO (CP-AO and CP-A almost overlap with each other) and the uncoupled scheme As-O (see Table 1). The observational time steps for the atmosphere and ocean are 10 and 40 steps, respectively. The RMSE is calculated as the average of the RMSEs at all the analysis steps.

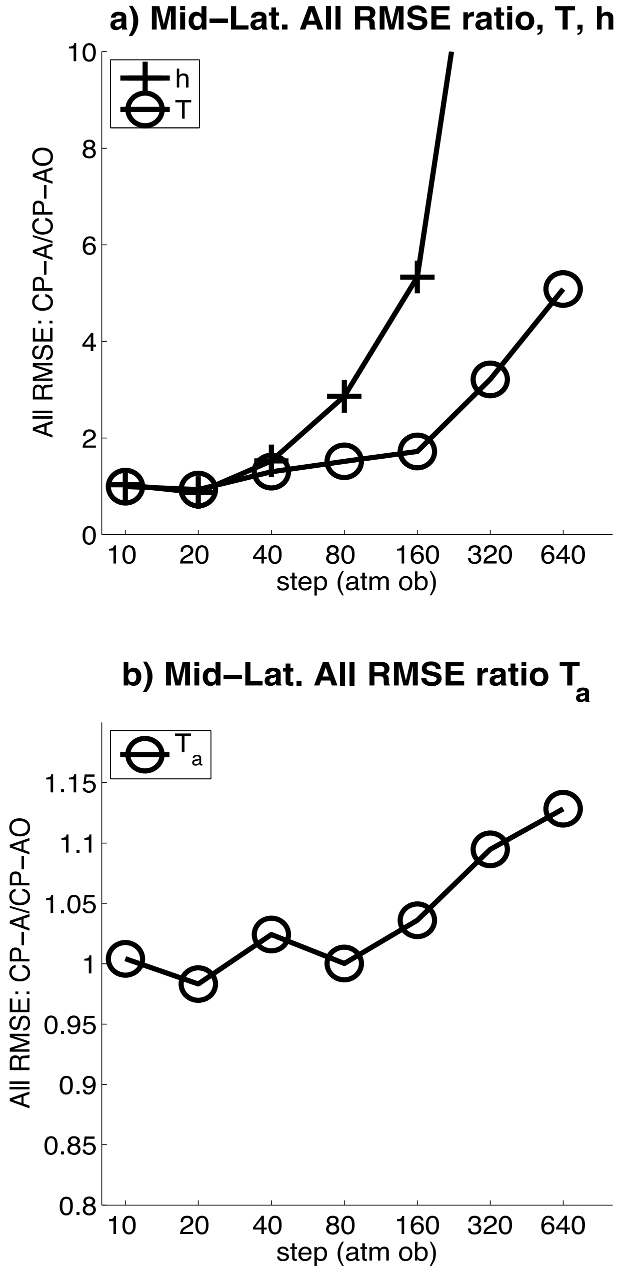

Figure 6. The ratio of RMSE (accumulated for all time steps) between CP-A and CP-AO as a function of the time steps of atmospheric observation in the mid-latitude system. (a) SST and thermocline depth, (b) air temperature. The oceanic observation time step is fixed at 40 steps. (The ratio of RMSE at the analysis steps are similar).

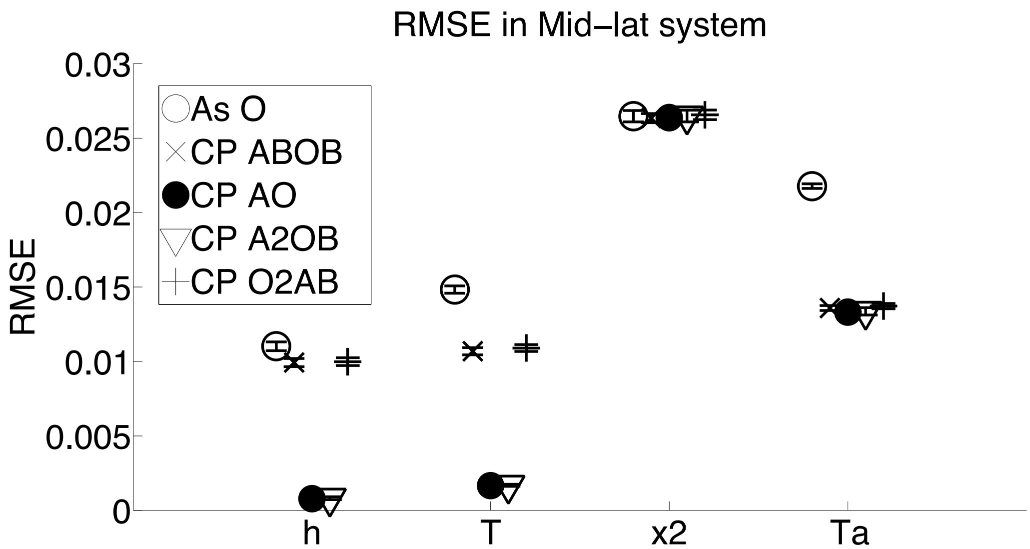

Figure 7. Analysis RMSE (normalized by the standard deviation of the control run) for As-O (circle), CP-AO (solid dot), CP-ABOB (cross), CP-A2OB (triangle) and CP-O2AB (plus) in the mid-latitude coupled system for h, T, x_2 and T_a. An ensemble of 80 members is performed with the ensemble mean in marks and the ensemble spread (standard deviation) in double bars. The observational time steps for the atmosphere and ocean are 10 and 40 steps, respectively.

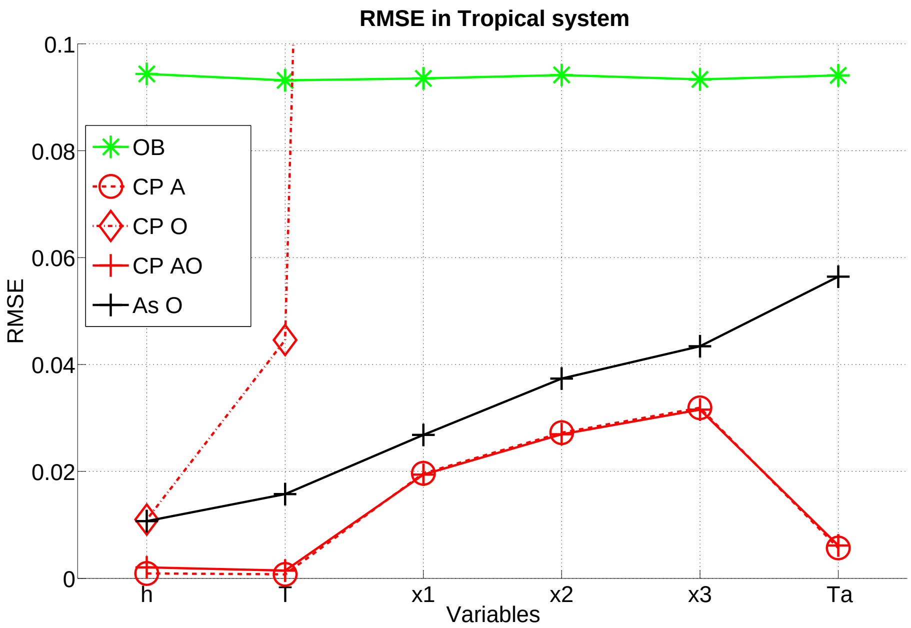

Figure 8. Same as Fig.5 but for the tropical system.

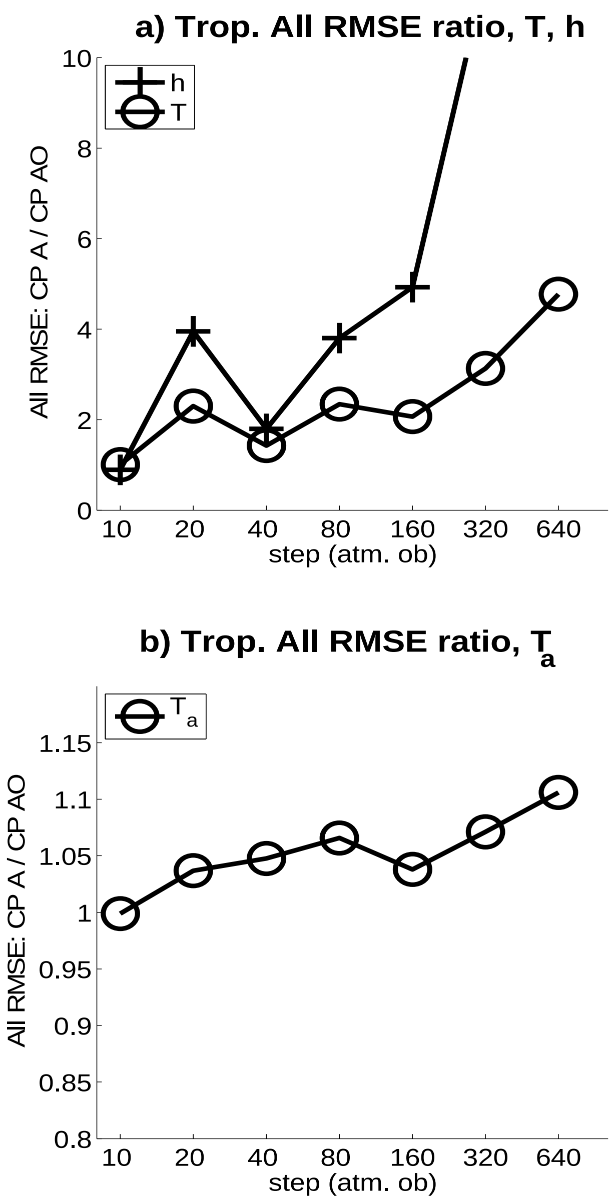

Fig.9: Same as Fig.6 but for tropical system.

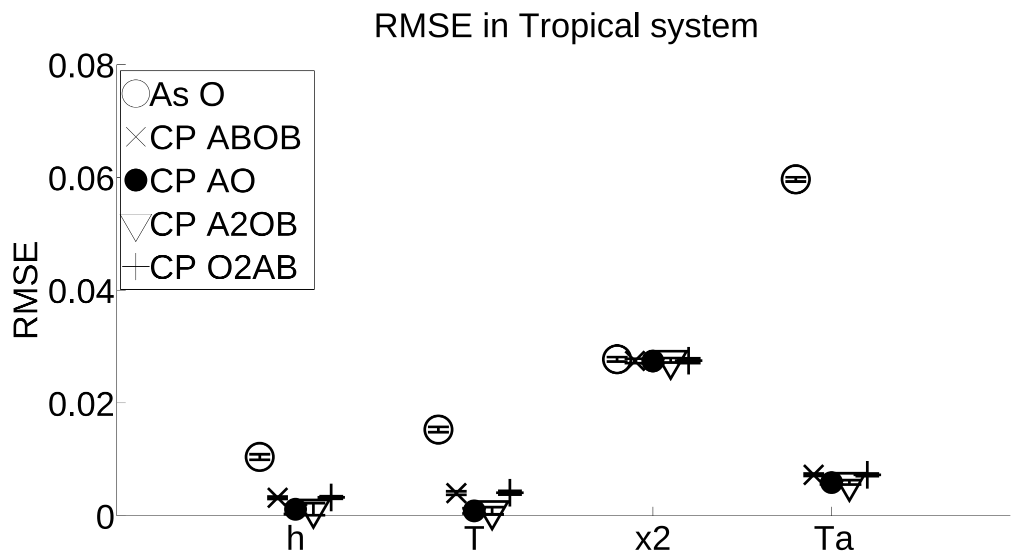

Fig.10: Same as Fig.7 but for the tropical system.

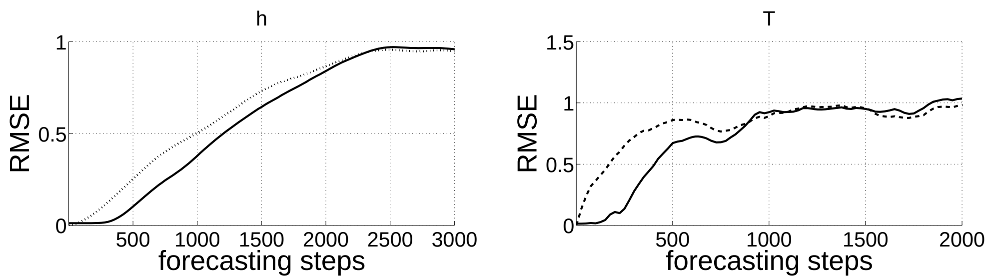

Figure 11: Forecast RMSE in the mid-latitude system for h (left) and T (right) initialized in PO (dash) and CP-AO (solid) schemes.

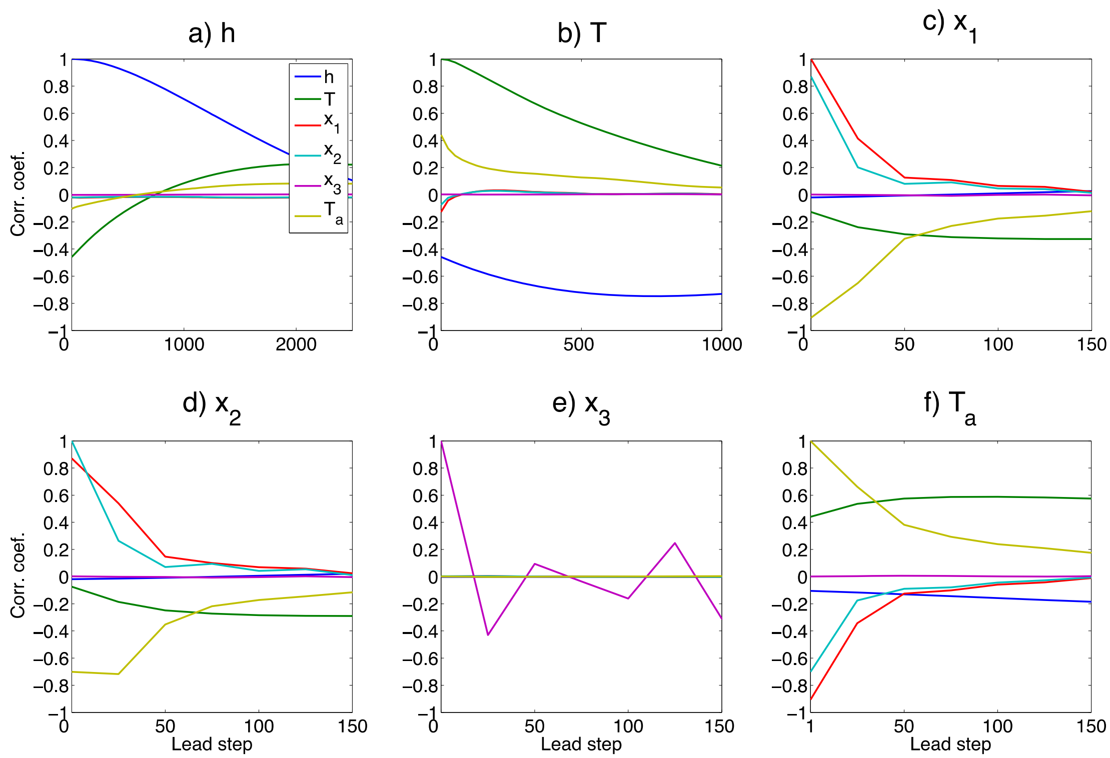

Figure A1. Lagged correlations among all model variables in the mid-latitude system. Each panel represents the pivotal variable that is used for lagged correlation with itself (auto-correlation) and other 5 variables (cross-correlations). The positive lead step is for this pivotal variable leading other variables. Each variable is represented in the same color, blue for *h*, green for *T*, red for x_1, cyan for x_2, purple for x_3 and yellow for T_a. For example, in panel (b), the auto-correlation of T is in blue, the cross-correlation between T and h, x_1, x_2, x_3 and T_a are in blue, red, cyan, purple and yellow, respectively.

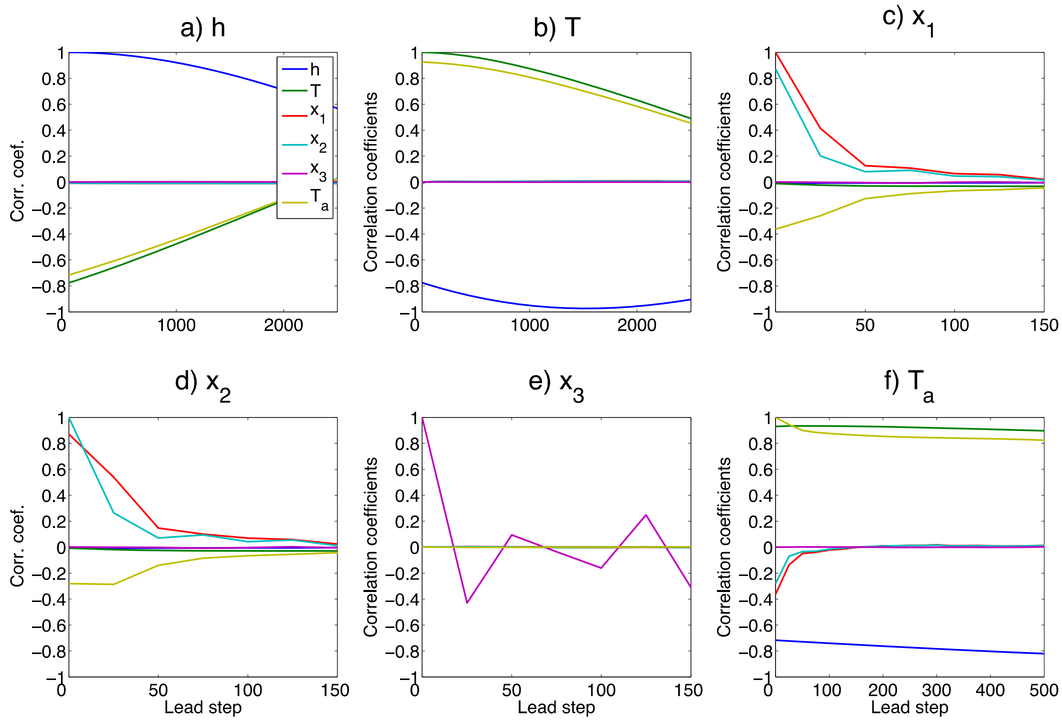

Figure A2. Same as Fig.A1 but for the tropical system.