



Coupled data assimilation and parameter estimation in coupled ocean–atmosphere models: a review

Shaoqing Zhang^{1,2,3} · Zhengyu Liu⁴ · Xuefeng Zhang^{5,6} · Xinrong Wu⁷ · Guijun Han⁵ · Yuxin Zhao⁶ · Xiaolin Yu^{1,2} · Chang Liu⁶ · Yun Liu⁸ · Shu Wu⁹ · Feiyu Lu¹⁰ · Mingkui Li^{1,2} · Xiong Deng⁶

Received: 20 September 2019 / Accepted: 25 April 2020 / Published online: 22 May 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Recent studies have started to explore coupled data assimilation (CDA) in coupled ocean–atmosphere models because of the great potential of CDA to improve climate analysis and seamless weather–climate prediction on weekly-to-decadal time scales in advanced high-resolution coupled models. In this review article, we briefly introduce the concept of CDA before outlining its potential for producing balanced and coherent weather–climate reanalysis and minimizing initial coupling shocks. We then describe approaches to the implementation of CDA and review progress in the development of various CDA methods, notably weakly and strongly coupled data assimilation. We introduce the method of coupled model parameter estimation (PE) within the CDA framework and summarize recent progress. After summarizing the current status of the research and applications of CDA-PE, we discuss the challenges and opportunities in high-resolution CDA-PE and nonlinear CDA-PE methods. Finally, potential solutions are laid out.

Keywords Coupled ocean–atmosphere model · Coupled data assimilation · Coupled model parameter estimation

1 Introduction

Coupled ocean–atmosphere general circulation models, also known as coupled general circulation climate models (CGCMs), are among the most important tools for the prediction of future climate (e.g. Randall et al. 2007; Meehl et al. 2009). However, owing to the imperfect nature of the model physics and numerical schemes, as well as incomplete observations, CGCM simulations and predictions

have been hindered by deficiencies in both the initial conditions and the model itself. It is therefore essential to combine available observations with a coupled model to create the optimal initial conditions for model predictions. This can be done most effectively using data assimilation (DA). Traditionally, a coupled model is initialized using uncoupled DA (Fig. 1a), in which the DA is performed separately in the atmospheric and ocean models, and the coupled model initial conditions are a combination of the

✉ Shaoqing Zhang
szhang@ouc.edu.cn

¹ Key Laboratory of Physical Oceanography, .MOE., Institute for Advanced Ocean Study, College of Ocean and Atmosphere, Frontiers Science Center for Deep Ocean Multispheres and Earth System (DOMES), Ocean University of China, Qingdao, China

² Ocean Dynamics and Climate Function Lab/Pilot National Laboratory for Marine Science and Technology (QNLN), Qingdao, China

³ International Laboratory for High-Resolution Earth System Prediction (iHESP), Qingdao, China

⁴ Department of Geography, Ohio State University, Columbus, OH 43210, USA

⁵ School of Marine Science and Technology, Tianjin University, Tianjin, China

⁶ College of Automation, Harbin Engineer University, Harbin, China

⁷ Ministry of Natural Resources of China, National Marine Data and Information Service, Tianjin 300171, China

⁸ Department of Oceanography, Texas A & M University, College Station, TX 77843, USA

⁹ Nelson Institute Center for Climatic Research, University of Wisconsin-Madison, Madison, WI 53706, USA

¹⁰ Princeton University and NOAA/GFDL, Princeton, NJ, USA

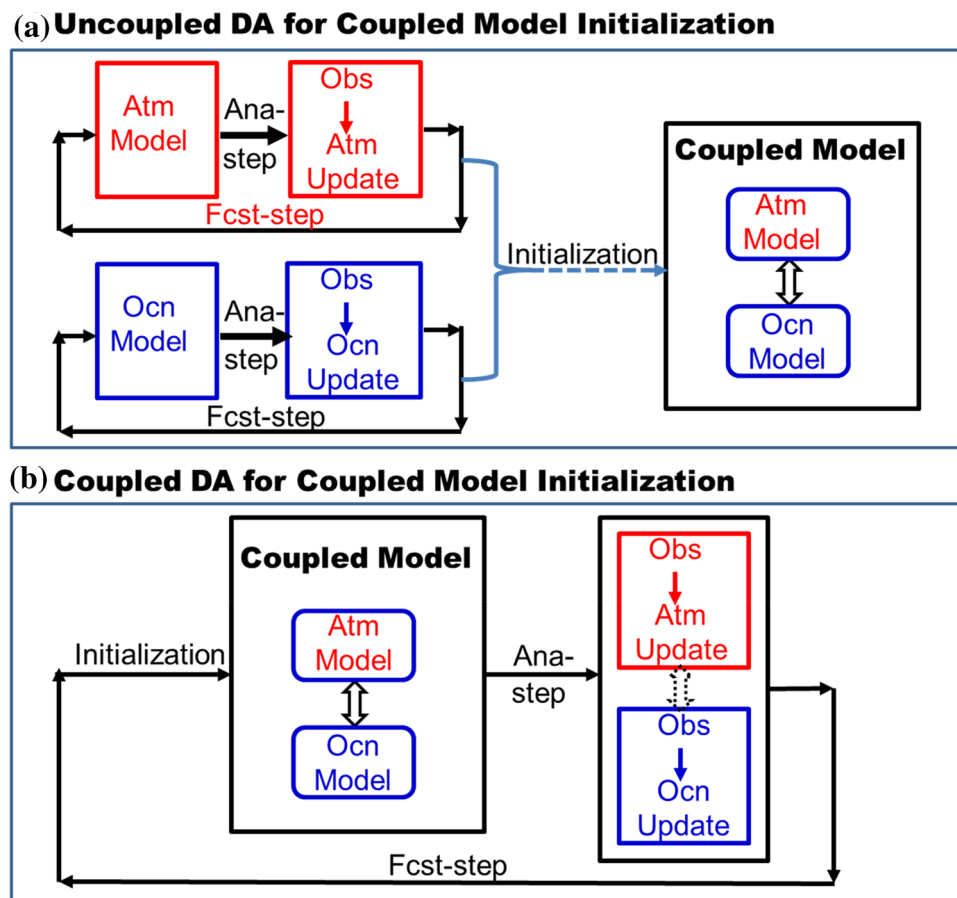


Fig. 1 Schematic illustration of **a** uncoupled data assimilation and **b** coupled data assimilation. While uncoupled data assimilation (**a**) uses the forecasts of uncoupled atmosphere and ocean models as the first-guesses in the atmospheric and oceanic analysis processes (marked by the red and blue arrows), the first-guesses of coupled data assimilation analysis processes (**b**) are the forecasts of the whole coupled model. The initialization of climate prediction is a natural consequence of coupled data assimilation while uncoupled data assimilation takes the combination of separate atmosphere data assimilation and ocean data assimilation results to initialize the coupled model

atmosphere model DA product and/or the ocean model DA product (Derber and Rosati 1989; Ji et al. 1994; Rosati et al. 1997) typically for producing seasonal-interannual climate forecasts (see Saha et al. 2006; Balmaseda and Anderson 2009). However, recent studies have started to employ coupled DA (CDA) (Fig. 1b). In CDA, the DA process is performed within the coupled model directly. Recent studies have shown that interannual climate forecasts from coupled models are improved by using CDA rather than uncoupled DA (e.g., Rienecker 2003; Zhang et al. 2005). Several countries (e.g., the United States, Japan, and the United Kingdom) and Europe have established their own climate analysis and prediction systems based on CDA. The World Meteorological Organization held an international CDA symposium in France

for climate prediction. The dashed double-arrow between the atmospheric and oceanic analysis updates implies that within the coupled data assimilation framework, the atmosphere (ocean) observations can directly impact on ocean (atmosphere) model states through cross-covariance between the atmosphere and ocean states, i.e. implementing strongly coupled data assimilation. Otherwise if only coupled model forecasts are used as the first-guess in the analysis update (without direct observational impact cross the atmosphere and ocean), then weakly coupled data assimilation is implemented

in October 2016. The meeting confirmed that CDA had reached a stage of rapid development, and that CDA would benefit weather prediction at synoptic, intraseasonal to seasonal, and interannual time scales, as well as climate reanalysis (Penny and Hamill 2017). Therefore, CDA has great potential for improving seamless weather and climate predictions (Brunet et al. 2015).

In addition to its application to state estimation and prediction in coupled models, CDA can help improve coupled models and our understanding of physical processes from the perspective of fully coupled models. The CDA framework can be used to optimize model parameters in a fully coupled model system with simultaneous state and parameter estimation. CDA can also be adapted to shed light on the mechanisms of ocean–atmosphere processes, especially

those involving the complex processes of ocean–atmosphere coupling.

This paper provides a review of the CDA approach, with a focus on recent advances and ongoing challenges. The remainder of the paper is organized as follows. In Sect. 2, we review the progress in the implementation of CDA, including topics such as the effect of CDA in reducing initial shock, the development of weakly and strongly coupled DA approaches, and coupled parameter estimation. In Sect. 3, we further review the application of CDA to climate state estimation and climate prediction, as well as its role in improving our understanding of coupled ocean–atmosphere processes in coupled models. Section 4 provides a summary and further discussion, and perspectives on the future challenges and opportunities in CDA.

2 Coupled data assimilation (CDA)

2.1 CDA and its importance

With a coupled model, one can in principle perform data assimilation to combine observations with the model to obtain the optimal initial conditions for both model prediction and the construction of reanalysis datasets for the assessment of climate change and climate variability. Traditionally, and for convenience, one combines the atmosphere and ocean states from existing atmospheric and ocean data assimilation systems together to form the initial conditions of the coupled model forecasts. This process is known as uncoupled data assimilation for coupled model initialization (Fig. 1a). In this approach, atmospheric and oceanic observations are first assimilated in an atmospheric model and an ocean model separately (left part of Fig. 1a). Then, the analysis states are combined to feed into a coupled ocean–atmosphere model as the coupled initial state. Such an approach tends to create an imbalance in the atmospheric and oceanic states because of the mismatch between two separately-achieved analyses using the atmosphere and ocean models respectively, which hinders the prediction skill of coupled models (e.g., Zhang et al. 2005, 2007; Chen and Cane 2008; Zhang 2011a). As the atmosphere and ocean data assimilation stages are conducted separately, such uncoupled data assimilation cannot produce a climate reanalysis that is consistent with the dynamics and physics of the coupled model. The most transparent example is the “twin” experiment forecasts of Zhang (2011b). The author used an extremely simple atmosphere–ocean coupled model by combining the “perfect” atmosphere (ocean) state with perturbed ocean (atmosphere) state as initial conditions compared with CDA initialization and found CDA initialization produces superior forecast quality.

To produce coupled model initial conditions that have balanced and self-consistent multi-component states within the coupled model, recent studies have started to explore the coupled data assimilation (CDA) approach, in which data assimilation is performed within the coupled model framework (Fig. 1b). More generally, as stated in Penny et al. (2017), for data assimilation in multi-component coupled earth system models, coupled model forecasts and potential state estimation are performed jointly so that each model component receives information from observations in other. In that sense, CDA could use the forecast of the whole coupled model as the first guess which is combined with observations to produce the updated analysis using an analysis equation, and the observation-updated coupled model states serve as the initial conditions of the next coupled model forecast cycle. In principle, CDA maintains the interactive processes among the different components of the coupled model during data assimilation, and therefore provides a more natural initialization state for the coupled model forecast. Thus, CDA is expected to produce more coherent and balanced initial conditions and analysis within the coupled system, thus improving climate prediction and advancing the understanding of climate variability (Rienecker 2003; Dee et al. 2014). Moreover, within the CDA framework, observations in one component, for example the atmosphere, can adjust not only the atmospheric model state itself but also other components for example the ocean state through the exchanged flux of the coupled model or coupled covariance (denoted by the double-dashed arrow in Fig. 1b). In general, approaches of CDA can be divided into two groups: weakly coupled DA (WCDA) in which DA is applied to each individual model component and strongly coupled DA (SCDA) in which DA is applied to the coupled Earth system model as a whole (Penny et al. 2017). Specifically, for a coupled ocean–atmosphere model, in WCDA, when the observations in the atmosphere (ocean) are used to adjust its own state, the observational information is transferred to the ocean (atmosphere) dynamically through the exchange of fluxes at the interface between the atmosphere and ocean (as shown in Fig. 1b, excluding the double-dashed arrow). In SCDA, as shown by the double-dashed arrow in Fig. 2b, observations in the atmosphere (ocean) not only adjust the state of the atmosphere (ocean) but also the state of the ocean (atmosphere). In ensemble methodology, SCDA can be implemented by applying the cross-covariance between the atmosphere and ocean to the update equations. In other methods such as variational algorithms, this could also be achieved by introducing a coupled observation operator, or a coupled background error covariance matrix into cost functions. SCDA is the most challenging CDA approach and will be discussed in more detail in Sects. 2.2 and 2.3.

In principle, CDA uses the dynamics and physics of a coupled model to extract observational information,

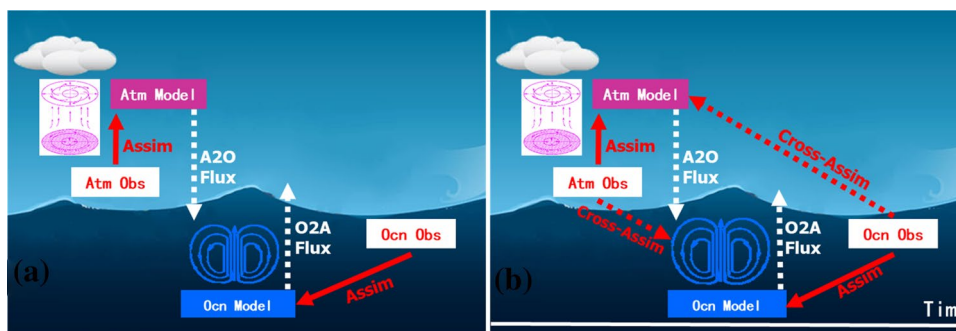


Fig. 2 Schematic of the implementation of **a** Weakly Coupled Data Assimilation (WCDA) vs. **b** Strongly Coupled Data Assimilation (SCDA). **a** WCDA assimilates the observation of the earth system into the corresponding single model component of the coupled model system (solid-red arrows) and transfers the observational information to other model components through the exchange fluxes of the coupled model (dashed-white arrows). **b** Beside the transferring of observational information by the exchange fluxes between coupled model

components, SCDA uses the cross-covariance of the model components to directly assimilate the observation of an earth system component into other coupled model components (dashed-red arrows), denoted by Cross-Assim. The solid-white arrow marks the time direction, suggesting that the SCDA Cross-Assim needs to consider the physical properties of air-sea interaction such as the temporal lag of the ocean response to the atmosphere forcing

allowing any observational information to be incorporated into the coupled system. Thus, CDA is an optimal approach to integrate the Earth observing system in terms of using coupled dynamics and physics. However, there exist two major challenges when moving from single model data assimilation (uncoupled DA) to CDA (see e.g. Yoshida 2019): (1) the design of an efficient computational scheme (e.g., parallelism) to combine the coupled model and analysis update equation together; and (2) the development of an advanced analysis algorithm to deal with multi-timescale variability in a coupled system in which each component may have its own characteristic time scale.

The first challenge has been overcome in recent studies that have implemented WCDA in a variety of applications (e.g., Zhang et al. 2005, 2007; Chen 2008; Fujii et al. 2009; Tardif et al. 2014; Liu et al. 2014a, b; Lea et al. 2015). In WCDA, although atmospheric and oceanic observations are able to impact other components dynamically through an exchange of fluxes, such observational impacts across model components are indirect and therefore may be affected by the imperfect coupled modeling such as parameterized coupling physics, thereby distorting the coupled analysis field. Not only relying on model flux exchanges to transfer observational information between model components, SCDA is able to relax the second challenge. As such, it has to include an observational analysis algorithm that directly adjusts the different components using observational information. The advantage of SCDA has been confirmed in simple model studies and systematic comparisons (e.g. Liu et al. 2013; Smith et al. 2015; Sluka et al. 2016a).

2.2 Implementation of CDA

From fundamentals of data assimilation function and implementation perspective, let's discuss the distinction of different data assimilation approaches when a coupled model is applied. In general, data assimilation incorporates observations into a numerical model to produce the optimal reanalysis and forecast initial conditions. The incorporation of the model and observations consists of two steps: the analysis step and the forecast step. In the analysis step, at the time of the current observation, the observational information is projected onto the model space and is then combined with the model "first guess" or "background" to produce the optimal estimation, or the analysis. In the forecast step, this analysis field is then used to initialize the model for a forecast to the time of next observation, producing the model forecast, or the new "background" or "first-guess". Then, the analysis step is repeated using this new background and new observations.

In a coupled ocean–atmosphere model, when the analysis step is performed within each model component, and the forecast step is performed in the coupled model (Fig. 1b, excluding the double-dashed arrow) to produce the first guess for the next analysis step, this approach implements weakly coupled DA (WCDA). In WCDA, the observational information in one component is transferred to the other component dynamically in the coupled forecast step. The second approach is strongly coupled DA (SCDA), in which both the analysis step and forecast step use the coupled model (Fig. 1b, including the double-dashed arrow). In addition to the use of the coupled model in the forecast stage (as in WCDA), the observational information in one

component of the model is also used to update the other component statistically through the analysis equation. Therefore, in SCDA, the observations are used to update the coupled state statistically in the analysis step and dynamically in both analysis (e.g. Isaksen et al. 2007; Brousseau et al. 2012; Fairbairn et al. (2014) and model forecast (e.g. Zhang et al. 2007). In general, as long as the forecast is made with the coupled model instead of the atmosphere or ocean models separately, the DA process is called CDA (of course, both WCDA and SCDA are types of CDA). Thus, the distinction between uncoupled DA and CDA is in the generation of the first guess of the analysis, whereas the distinction between WCDA and SCDA is in the formulation of the analysis update equation. In principle, SCDA is the best approach to CDA because it assimilates observations in the coupled model statistically and dynamically. The implementation of SCDA, however, involves both the DA algorithm strategy and the understanding of air–sea interaction physics, which is key for modeling the coupling process of the atmosphere and ocean. Thus, SCDA has remained as one of the most important and challenging research topics in CDA, a topic we will revisit in Sect. 2.3.

Current operational CDA mostly uses the WCDA approach. The implementation of WCDA has, to date, used one of four different methods, from simple to complex. In terms of pursuing balanced and coherent coupled initial states, the simplest method to implement the CDA idea is nudging. Within the coupled model framework, one or more model state variables are restored towards the observations or existing reanalysis (e.g., Chen et al. 1995). Nudging usually requires that the observational and/or reanalysis data to which the model states are restored, e.g., satellite sea surface measurements and 3-dimensional reanalysis products, are gridded onto the model space. Keenlyside et al. (2008) applied a nudging method that restored the sea surface temperatures (SSTs) of a coupled climate model towards observations, before the coupled state was used to initialize the coupled model. They reported improved prediction skill on the decadal scale over the North Atlantic and tropical Pacific oceans. Because of the requirement of gridded data, the nudging method, although simple and easy to implement, has very limited application in CDA for assimilating real observations, as the majority of observations are spatially irregular. Moreover, a serious issue with nudging is the lack of uncertainty estimation.

The second method is the 3-dimensional variational method (3D-Var) (e.g., Lorenc 1986), a variant of the optimal interpolation (OI) method (Gandin 1965). It has been applied widely to atmospheric data assimilation (ADA) (e.g. Wu et al. 2002) and ocean data assimilation (ODA) (e.g., Derber and Rosati 1989). OI derives the observational analysis under the constraint of minimal analysis error variance given a background error covariance matrix, whereas

3D-Var realizes this goal by minimizing a normalized distance between the model state and observations using the given background error covariance. It is straightforward to implement the 3D-Var or OI method in a coupled model by inserting the existing 3D-Var ADA and OI ODA into the atmosphere component and ocean component of the used coupled model respectively (e.g., Saha et al. 2010). Then how to implement the multiscale schemes (e.g., Xie et al. 2011; Yu et al. 2018) in CDA is an important research topic.

The third method is the 4-dimensional variational method (4D-Var). This method is in principle an excellent method to generate a DA solution consistent with the model dynamics within a defined observational time window. Based on previous work on ADA (e.g., Courtier et al. 1994) and ODA (Stammer et al. 2002), Sugiura et al. (2008) developed the adjoint model of a fully coupled model. They have done the 4D-Var experiment by using a control vector of low dimensional correction to fluxes with a 9-month time window and reported improved forecast skill at the seasonal to interannual scale. A great advantage of 4D-Var CDA is that once a cost function is properly defined with modeled and observational air–sea interface state variables (e.g., skin sea surface temperature) and/or air–sea interface fluxes (e.g., heat and momentum fluxes), it is feasible to implement SCDA at the air–sea interface. However, the application of 4D-Var to CDA has some disadvantages and challenges. It is difficult for 4D-Var to fully incorporate observational information from different components that have different characteristic timescales (e.g., from the atmosphere at the synoptic timescale to the deep ocean at the decadal timescale) since the tangent linear approximation made for the adjoint would not be valid for a timescale beyond months in state estimation. Moreover, 4D-Var requires the development of an adjoint model for the fully coupled system. Although the issue of adjoint recoding with an updated model is greatly relaxed with sufficient planning and resource allocation (e.g., Giering et al. 2006), the consistent implementation of the coupled model adjoint and minimization is still challenging and demanding work given the widely different time scales between the atmosphere and the ocean. Even with the adjoint model available, the adjoint iteration for minimization is extremely expensive in a fully coupled model. Nevertheless, some of these issues have recently been addressed (e.g., Smith et al. 2015; Fowler and Lawless 2016). Laloyaux et al. (2018) incorporated 4D-Var ADA and 3D-Var ODA into the coupled model of the European Center for Medium-Range Weather Forecasts (ECMWF) to establish a CDA reanalysis and prediction initialization system, in which a multi-loop approach is used. Then, multi-timescale effects in ECMWF CDA system are addressed (Browne et al. 2019).

The fourth method is the ensemble Kalman filter (EnKF). While any other DA methods such as 4D-Var, and 3D-Var (OI) etc. can be indeed derived from Bayes' Theorem,

EnKF uses an ensemble of model integrations to simulate the temporally varying background probability distribution to implement Bayes' Theorem in a straightforward manner (e.g., Evensen 1994; Zupanski 2005; Evensen 2007). EnKF is a convenient method to implement CDA in a complex, multi-component model system, because its forecast stage only involves an ensemble of coupled simulations, and the analysis stage can be carried out by an online subroutine (e.g., Zhang et al. 2005, 2007; Zupanski 2016) or offline program (Nerger et al. 2005; Anderson et al. 2009; Nerger and Hiller 2013). As such, the complexity of the coupled model creates no difficulty in the implementation of the CDA. After successful test experiments in a hybrid coupled model for the forecasting of El Niño–Southern Oscillation (ENSO) (Zhang et al. 2005), Zhang et al. (2007) implemented the ensemble coupled data assimilation system (ECDA) in a fully-coupled atmosphere–ocean general circulation model (CGCM), the Geophysical Fluid Dynamics Laboratory (GFDL) second generation coupled model (CM2). Following the same scheme, Liu et al. (2014a, b) developed the second EnKF CDA system in a CGCM, albeit one with a coarser resolution. The major challenge of EnKF CDA is the computational cost required by the ensemble coupled model integration, which is usually on the order of 10–100 model costs (ensemble members). EnKF CDA also suffers from under-sampling by the finite (both on ensemble integration time and ensemble size) ensemble of the low-frequency background circulation statistics so that it needs localization and inflation techniques to relax the problem. This will be discussed more in Sect. 2.3.

Finally, 4D-Var can be combined with the ensemble method to form a hybrid approach, in which the ensemble statistics are used to update the background error covariance matrix in the cost function of 4D-Var (e.g., Buehner et al. 2017). Such a hybrid method that combines an ensemble scheme with a variational scheme may provide a convenient strategy to utilize the previously available variational schemes while also taking advantage of the ensemble method to overcome the shortcomings of each approach. However, the computational cost could be a major concern when this approach is applied to high-resolution coupled models, for which at least 10 factor of model expenses (ensemble size plus adjoint iterations) are required while integration of the high-resolution coupled model itself costs very high (e.g. Small et al. 2014).

2.3 Strongly coupled DA

Given that SCDA is in principle the optimal CDA approach for reanalysis and prediction initialization, the implementation of SCDA is one of the most important research topics in current studies of coupled model reanalysis and prediction. To clearly explain the principle and implementation

of SCDA, we use Fig. 2 to further illustrate the physical meaning and function of SCDA compared with WCDA. While WCDA (Fig. 2a) assimilates observations to each component model separately (solid red arrows), the observational information is transferred between model components via the surface fluxes (dashed white arrows). Under such a circumstance, the ocean and the atmosphere are coupled dynamically but not statistically, by which the observational adjustment of coupled model states may not be sufficient due to the limitations of the models. In contrast, SCDA (Fig. 2b) uses the statistical relationship between different coupled model components and performs cross-assimilation, with, for example, atmospheric (oceanic) observations used in the ocean (atmosphere) model as well as in the atmospheric (ocean) model itself (dashed red arrows). As such, the ocean and atmosphere are coupled not only dynamically (denoted by white dashed arrows), but also statistically (denoted by dashed red arrows), so that the CDA results better represent the complex air–sea interaction processes in the real world. In principle, SCDA improves on WCDA as it makes full use of observational information in the different coupled system components through the statistical relationship between them. For example, the statistical relationship of air–sea interactions at the air–sea interface. However, because of the different characteristic timescales in the different components of the coupled earth system, as well as sampling errors, the development of SCDA schemes that can make effective use of the reliable (i.e., signal-dominated) statistical relationship across coupled components remains one of the most challenging problems in CDA. Thus, at present, most CDA systems used in more realistic applications are still based on the WCDA framework, as discussed above.

There have been numerous studies of SCDA in simple models, which make efforts to enhance the signal-to-noise ratio of covariance across coupled components (e.g., Han et al. 2013; Liu et al. 2013; Smith et al. 2015; Sluka et al. 2016a; Yoshida and Kalnay 2018; Yoshida 2019). However, a straightforward implementation of SCDA using instantaneous ocean and atmospheric observations with flow-dependent background error covariance in ensemble methodology, referred to as simple SCDA or standard SCDA hereafter, often fails, in which the sample size is limited. One reason for the failure is the large sampling error in the coupled cross error covariance, or coupled cross error correlation (hereafter, both called CCEC), between different model components. This was demonstrated in a 5-variable simple model study (Han et al. 2013), in which the improvement of SCDA over WCDA was realized only when the ensemble size was increased to $\sim 10^4$. Furthermore, it has been found difficult to use observations of the slowly varying components (e.g., the ocean) to improve the state of fast-varying components (e.g., the atmosphere), because the variability in the fast component is dominated by its own timescale

variability. It is relatively easy for the observations of the fast-varying components to improve the state of the slowly varying components (Han et al. 2013; Liu et al. 2013; Sluka et al. 2016a). It's worth to mention that the correlation-cutoff method for covariance localization is also a sound approach that already showed some promising results in low-order models (Yoshida and Kalnay 2018) or training the neural network (Yoshida 2019). Figure 3 gives an example of an improved slow-varying component by SCDA with fast-varying component observations in a simple coupled model in Han et al. (2013).

Thus, in comparison with WCDA, the key to the success of SCDA is the signal-to-noise ratio in CCEC. Under the Gaussian assumption, the coupling intensity between different model components can be a direct function of CCEC (Zupanski 2016). The CCEC can be calculated directly from ensemble samples. The quality of the evaluated CCEC is mainly affected by model errors, coupling errors, sampling errors and the differences in the temporal–spatial scales between different model components, as well as its significant seasonal variability (Smith et al. 2017).

In CGCMs, while some prototype of SCDA is implemented in the ensemble filtering approach (Sluka et al. 2016b), it is impractical to employ a very large ensemble size to sustain significant signal-to-noise ratio in the cross-component covariance. This makes the development of SCDA most challenging. Beside introducing the covariance inflation technique (e.g. Anderson 2007; Liu et al. 2013) that inflates the prior ensemble to increase the ensemble spread,

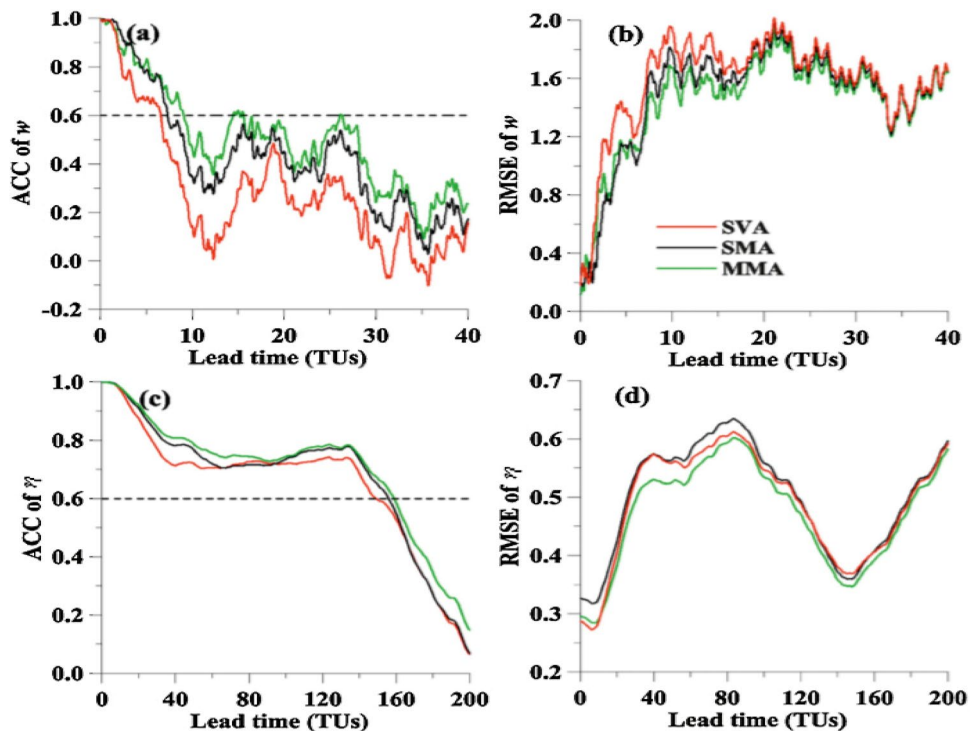
one approach is to make use of the physical mechanism to guide the design of the SCDA, such that the signal-to-noise ratio of the CCEC can be enhanced, as in the following approaches:

(a) Leading averaged coupled covariance (LACC)

Lu et al. (2015a, b) proposed an implementation of SCDA, known as the Leading Averaged Coupled Covariance (LACC) method, to improve the oceanic state directly using atmospheric observations. LACC was designed based on the understanding that ocean–atmosphere temperature cross-correlation in the extra-tropics tends to peak when the atmosphere leads SST by ~ 10 days, because of the dominant role of internal atmospheric variability (Frankignoul et al. 1998) (also see Fig. 4c). Therefore, LACC adjusts the current ocean state using the average of atmospheric state forecasts and observations that lead the current ocean state by ~ 10 days when the same atmospheric observations have been used to adjust the atmospheric states, with the average here being used to further reduce sampling error.

Starting from examination of lag correlation of the ocean to the atmosphere in a low-order coupled model (Fig. 4), the LACC application to a CGCM (Fig. 5) showed that it can effectively reduce the ocean state analysis error to a level below that obtained using WCDA or standard SCDA, whereas the latter uses simultaneous atmosphere–ocean coupled covariance (Lu et al. 2015a) (Figs. 4, 5). The advantages of LACC become even clearer when the ensemble size

Fig. 3 An example of strongly CDA improving forecasting skills in a simple coupled model developed in Zhang (2011a, b). Courtesy to Han et al. (2013): Variation of (left) ACC and (right) RMSE with the forecast lead time of the forecast ensemble means of the **a, b** upper-ocean w and **c, d** deep-ocean h for single variable adjustment (SVA, red curve), single model component adjustment (SMA, black curve), and multiple model component adjustment (MMA, green curve) for an ensemble size of 15000. The dashed horizontal lines mark the ACC value of 0.6



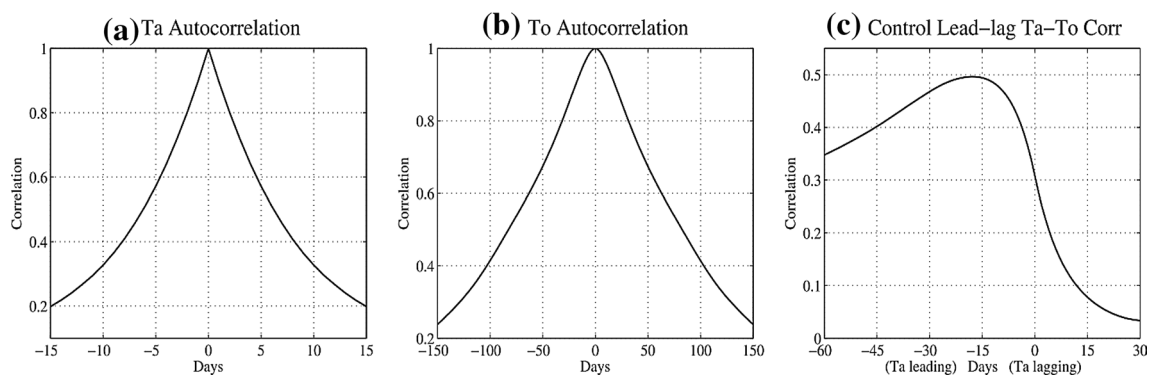


Fig. 4 An example of using lag correlation of the ocean to the atmosphere to improve the atmosphere-ocean cross covariance in strongly CDA. Courtesy to Lu et al. (2015a): **a** Ta (atmosphere temperature)

autocorrelation, **b** To (ocean temperature) autocorrelation, and **c** cross correlation based on the output of a single-member control simulation

is limited and the sampling error is therefore large in both the ocean and atmospheric components. LACC is the first successful implementation of SCDA in a CGCM that shows a significant mitigation of the errors in SST and surface atmosphere temperature when compared with WCDA (Lu et al. 2015b). Then, Yoshida (2019) succeeded to enhance the signal of covariance cross model components by replacing LACC with the error correlation cutoff technique.

(b) Combination of reconditioning and localization

Smith et al. (2018) proposed a reconditioning technique on the original CECC matrix. The new CECC is calculated by modifying the original eigenvectors, so that the signals in the CECC are enhanced. Results in one-dimensional models indicate that reconditioning coupled error correlation coefficient matrixes can avoid losing the modes that are small in magnitude but dynamically significant. This reconditioning method, although able to preserve correlation structures, is unable to eliminate sampling errors. In contrast, a general localization technique of model state space can reduce sampling errors in the CECC, but it tends to lose signals with small coupled error correlation. Smith et al. (2018) combined the two methods and proposed the combined application of the reconditioning and localization, which seems to improve the CECC by relaxing the under-sampling issue described in the paragraph next to the end of Sect. 2.2.

(c) Interface solver

The interface solver method assumes that the correlations outside the atmosphere and ocean boundary layers can be ignored. The atmosphere and ocean components are assimilated separately. When assimilation is carried out in the ocean (atmospheric) component, the assimilation uses observations not only of the ocean (atmosphere) but also

the atmospheric (oceanic) observations that have significant influence on the assimilation (Frolov et al. 2016; Luo and Hoteit 2014). An assimilation test with a simple air–sea coupled model indicated that the interface solver produces a more accurate analysis than that obtained using standard SCDA.

In summary, it has remained difficult to design an effective SCDA between the atmosphere and ocean components in a CGCM simply through traditional statistical data assimilation approaches, such as simple SCDA. However, for the design of SCDA in CGCMs, insight may be gained by utilizing the physical understanding of ocean–atmosphere interactions and evaluating more physically modes based cross-covariance between the atmosphere and ocean. Recent studies have suggested that real world ocean–atmosphere interactions may depend on mesoscale structures that are currently lacking in models (Ma et al. 2016). This poses a further challenge to implementing SCDA in future high-resolution models. A research project led by the first author [sponsored by the State Key Program of Chinese Natural Science Foundation (CNSF)] is currently in progress to study mesoscale air–sea interaction physical processes and associated SCDA.

2.4 Coupled model parameter estimation

The errors in a coupled model can result from errors in the dynamic cores, couplers, numerical schemes, physical parameterization schemes and empirical parameters. Model parameters, however, can be adjusted or optimized using observations. This process is called parameter estimation (PE) or parameter optimization. Similar to state estimation, PE uses the covariance between model states and parameters. Therefore, theoretically, any parameter related to observable state variables could be estimated using a DA method. In practice, parameter estimation is more difficult

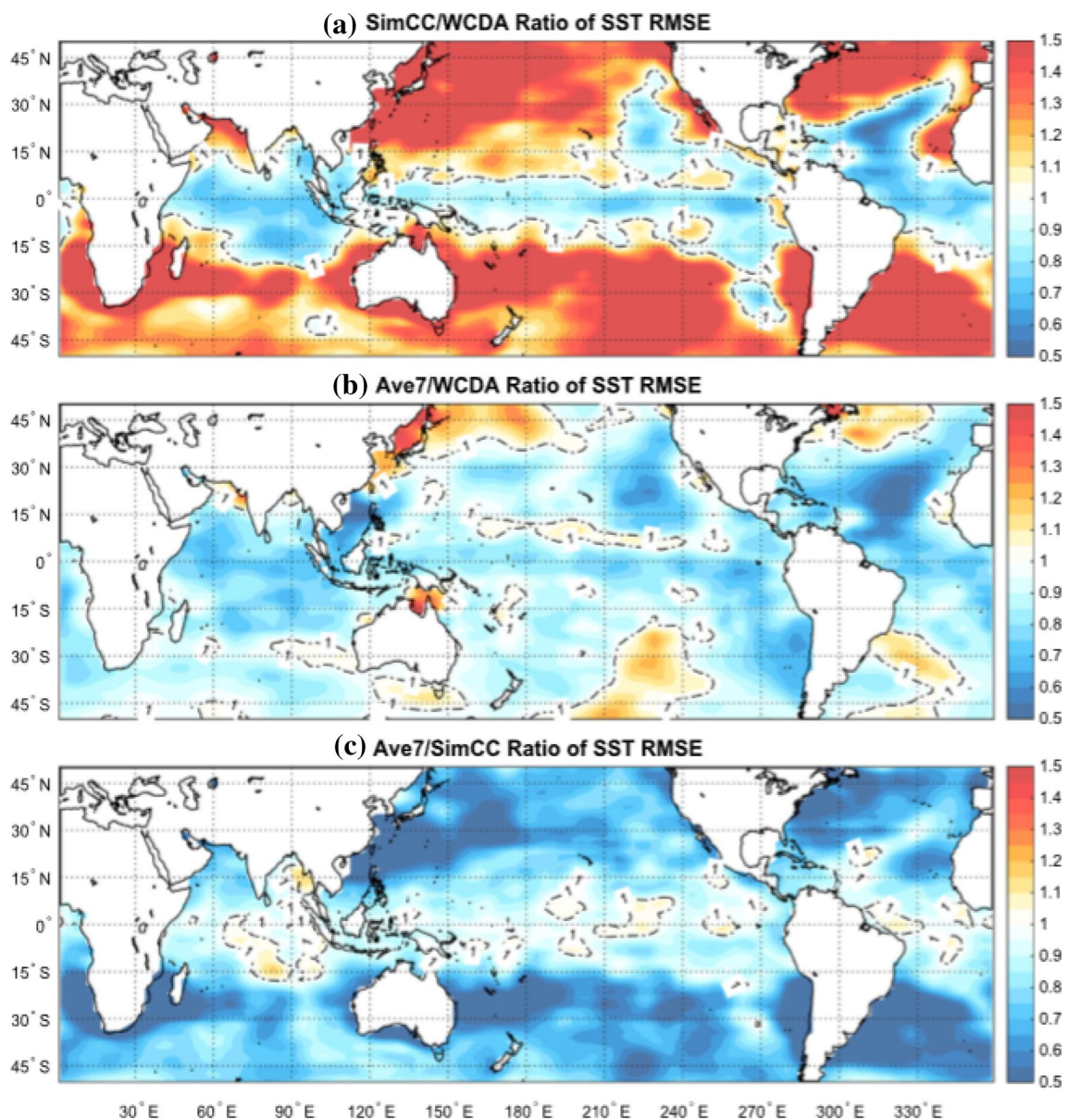


Fig. 5 An example of physical mode-based SCDA to improve SST estimation in a CGCM. Courtesy to Lu et al. (2015b): Spatial distribution of the RMSE of monthly SST from **a** the experiment using the simultaneous coupled covariance (SimCC Exp) (normalized by the

WCDA), **b** the experiment using 7-day atmosphere leading average coupled covariance (Ave7 Exp) (normalized by the WCDA), and **c** the Ave7 Exp (normalized by the SimCC)

than state estimation because of the difficulties in estimating the state–parameter covariance. This covariance is influenced by multiple factors such as model errors, sampling errors, and observational errors, as well as low model sensitivity. In a coupled model, coupled model parameter estimation is even more difficult than the single component model parameter estimation because of the complex model sensitivities associated with the variability of different spatial and temporal time scales in different model components, and the misfitting of air–sea interaction processes. Thus, coupled parameter estimation (CPE) is still in the research stage. To

date, implementation of CPE has mainly involved one of the two methods outlined below.

(a) Objective (Variational) method

First, a cost function is defined to measure the differences between observations and model simulation results obtained from the same initial conditions but different parameter values. The parameters are then estimated with an optimization algorithm. This approach has been used to estimate parameters in land processes. Liu et al. (2005) used CPE in a land–atmos-

phere model to reduce temperature errors and improve model results with the estimated parameters using a variational method. The objective method requires a large number of simulation experiments (depending on the optimization method, the smoothness/convexity of the problem, the accuracy and availability of a gradient etc.) and therefore demands huge computational resources for a CGCM. In addition, the issue of multiple local minima often exists in this method especially in joint parameter/state estimation. To reduce the computational cost and search for globally optimal parameters, attempts have been made to use off-line methods for an objective estimation of parameters. However, this method has only been used in single component atmosphere or ocean models (Posselt and Bishop 2012).

(b) Joint with data assimilation

Based on the state vector augmentation technique, parameter estimation can also be implemented in the data assimilation procedure. During the model integration, it is generally assumed that the parameters are constant and can only be changed by data assimilation. However, before the parameter estimation is carried out, sensitivity analyses should be conducted (Navon, 1997) to ensure that the parameters are identifiable. There are in general two approaches to implement parameter estimation:

- (1) *Four-Dimensional Variational Analysis (4D-Var)* This is accomplished by introducing the parameter estimation term into the objective function (or cost function) and control variables, calculating the gradients from the initial conditions and initial parameters with the adjoint model, and finally obtaining the optimal values for both the initial conditions and the model parameters through an optimization algorithm. Because of the high complexity of adjoint models in CGCMs, research studies into 4D-Var have mainly used simple coupled models (e.g., Han et al. 2015) and coupled climate models of intermediate complexity (e.g., Lu and Hsieh 1998). Research on 4D-Var using climate models mainly focuses on the tropics where air–sea interaction is stronger and parameters are therefore more easily identifiable (Du et al. 2009; Ito et al. 2010; Song et al. 2012).
- (2) *Ensemble Kalman Filter (EnKF)* The EnKF estimates parameters based on the error covariance between observable model states and parameters. It has been widely applied in CPE (e.g., Ruiz et al. 2013; Kang et al. 2011, 2012) because of its direct estimation of the state-parameter error covariance. Development of EnKF-CPE in coupled models of varying complexity has occurred in three phases:
 - (1) Implementation of EnKF or extended EnKF-CPE in intermediate coupled models (Annan et al. 2005; Annan and Hargreaves 2007; Kondrashov et al. 2008).
 - (2) Enhanced parameter estimation: Zhang et al. (2012) showed that the signal-to-noise ratio of the state-parameter error covariance in a coupled model can be significantly improved after the state estimation reaches quasi-equilibrium. Thus, using the observation-constrained states that have reached equilibrium can effectively improve the accuracy of CPE (Zhang 2011a, b). This approach has been extended to estimate multiple parameters depending on different roles (sensitivities) by which parameters play in dynamics and physics (e.g. Zhao et al. 2019).
 - (3) Geographic-dependent CPE: Wu et al. (2012a, b) introduced the spatial distribution of the model state sensitivity to parameters into CPE. In terms of full uses of sensitivity information, the geographic-dependent CPE is similar to the multiple parameter estimation, but it emphasizes the use of spatial distribution information of model parameter sensitivity as well as non-homogeneous nature of observation availability. The geographic-dependent CPE significantly improves the accuracy of climate estimation and prediction compared with single-value PE. Related to this method, Liu et al. (2014a, b) developed a self-adaption spatial average algorithm, which accelerated the convergence of CPE and enhanced the CPE signal-to-noise ratio.

It has been shown that CPE can be used to constrain the uncertainty of parameterization schemes and some model errors caused by the dynamic cores (Han et al. 2014). It can also improve the predictability of ENSO (Wu et al. 2016). CPE has been applied to CGCMs successfully in a perfect model scenario (Liu et al. 2014a, b; Li et al. 2018). Figure 6 shows CPE of convective parameters using EnKF (Li et al. 2018) in the coupled climate model CM2.1 of the Geophysical Fluid Dynamics Laboratory of the National Oceanic and Atmospheric Administration (GFDL/NOAA). The ensemble-based CPE reduced the convective parameter errors and improved the accuracy of the analysis and forecasts for the atmosphere and ocean.

Overall, CPE has been shown to be successful in coupled models of varying complexity. However, it has only been performed in the perfect model scenario. The application of CPE in CGCMs with real observations to improve coupled model reanalysis and prediction still has a long way to go.

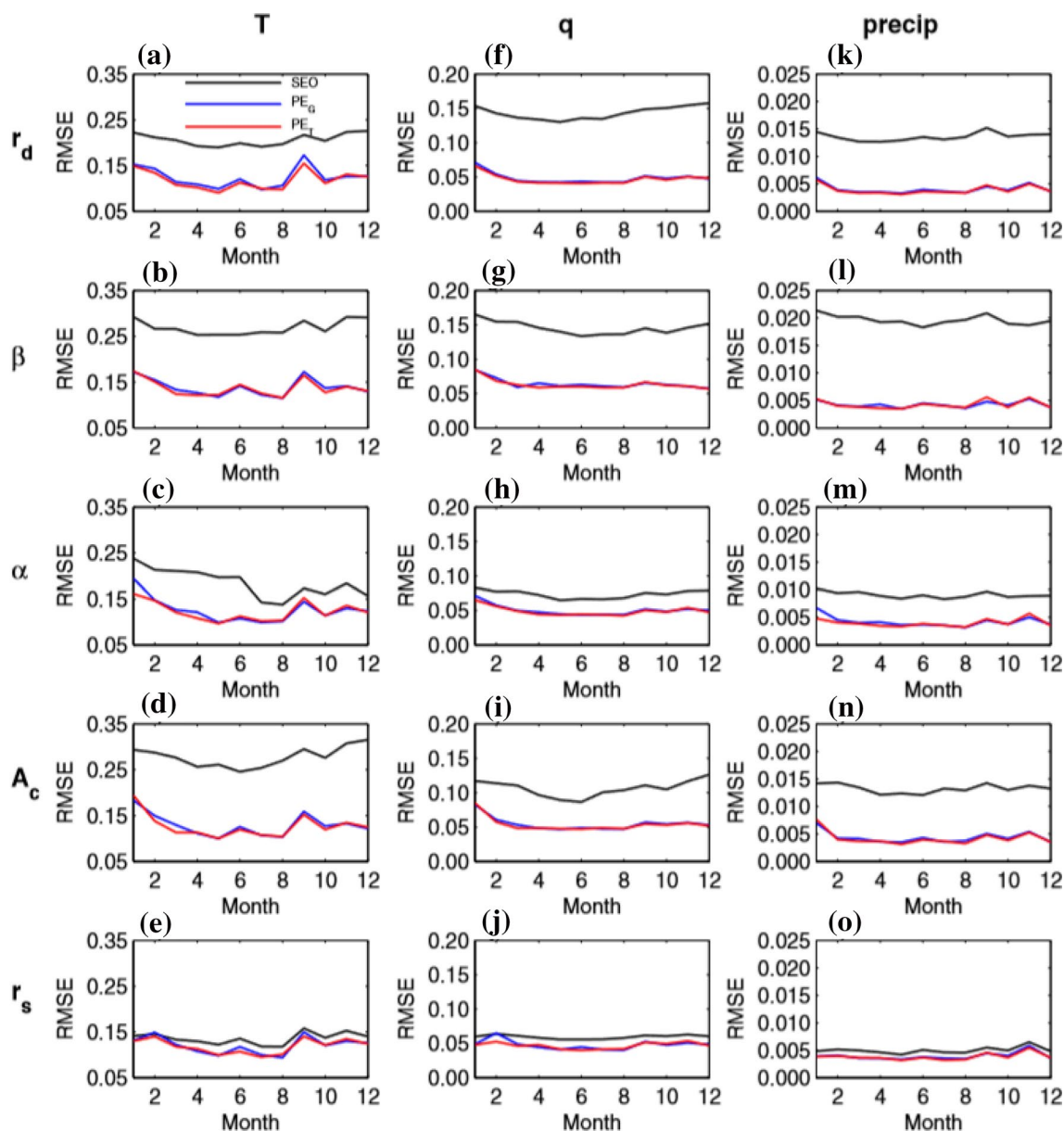


Fig. 6 An example of convection parameter estimation improving temperature and moisture errors. Courtesy to Li et al. (2018): time evolution of RMSEs of **a–e** atmospheric temperature (unit: K), **f–j** specific humidity (unit: g/kg), and **k–o** precipitation (unit: mm/day) in single convection parameter estimation. Each row shows the results

of one parameter, as denoted on the left. The black line represents the RMSE for the state-estimation-only. The blue and the red line represent the RMSEs of parameter estimation with global observation and tropical observation

3 Application of coupled data assimilation

3.1 Minimizing initial shocks in coupled model reanalysis and prediction

Initial shocks refer to the spatial discrepancy in a variable field or imbalance between physical variables in the initial conditions. As discussed in Sect. 2.2, data assimilation is a cycling procedure involving model forecasting and observation-based analysis. Thus, the model initial shock

in the forecasts influences analysis and prediction skill. The reduction of the model initial shock remains an issue of critical importance in numerical weather and climate forecasting. In the last century, limited by sparse observations and a lack of advanced assimilation techniques, climate prediction was based on coupled models initialized with atmosphere and ocean states that were analyzed separately from the atmosphere and ocean component models—i.e., uncoupled DA. This approach can create a mismatch between the atmosphere and ocean dynamics

in their initial conditions, contributing significantly to an initial shock in the coupled climate forecast (Rosati et al. 1997)—i.e., an imbalance between the atmosphere and ocean components. Such initial shocks can also be caused by a mismatch in the atmospheric and oceanic states, such as the heat and momentum budget across the interface, because of the different DA and model systems used to generate the initial conditions. The negative impacts of coupled model initial shocks on climate prediction have been demonstrated in simple coupled models (e.g., Zhang 2011b), intermediate coupled models (Chen and Cane, 2008) and fully coupled general circulation climate models (Mulholland et al. 2015; He et al. 2017).

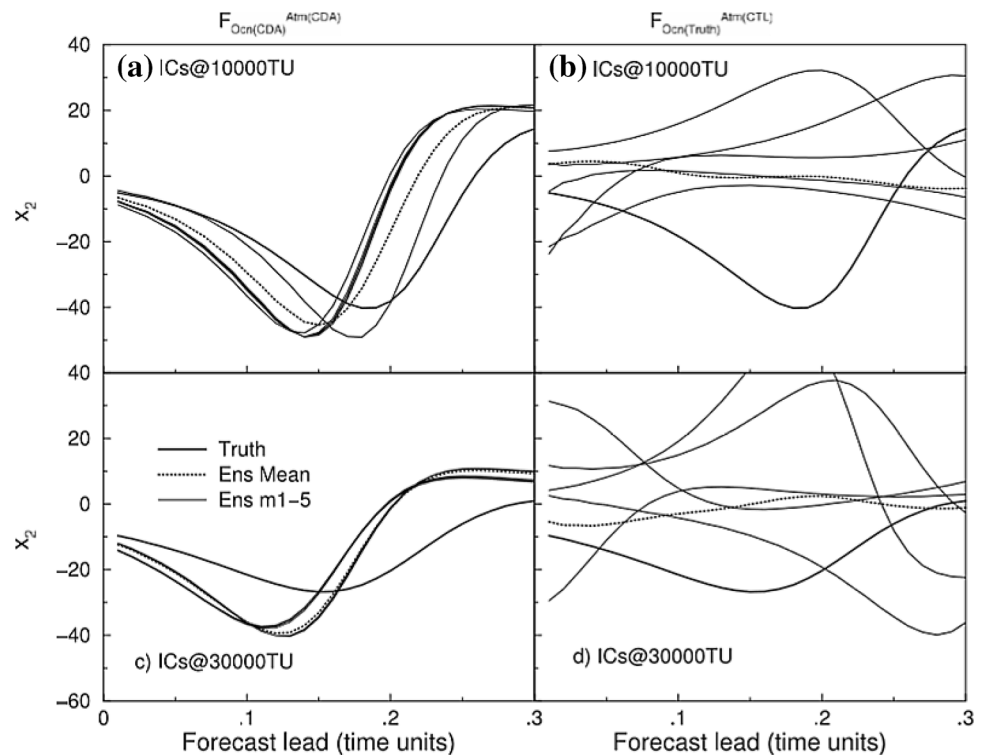
CDA provides an approach to reduce coupled model initial shocks (Singleton 2011). In CDA, observations from the different components of the coupled system are assimilated into a single coupled model system. This method minimizes the initial shock in coupled model predictions by maintaining the dynamic consistency between the different model components (Zhang et al. 2007; Chen and Cane 2008; Sugiura et al. 2008; Liu et al. 2013, 2017b; Mulholland et al. 2015). As a result, CDA has been applied gradually to climate analysis and prediction initialization in CGCMs (e.g., Msadek et al. 2014; Jia et al. 2015). Figure 7 provides an example in a simple coupled model (Zhang 2011a) where the atmospheric forecasts initialized from the CDA coupled states (left panels) show superior skill compared with the forecasts initialized from the perfect ocean state combined with independent atmospheric states (right panels).

Initial shocks in coupled models can also result from other factors. A coupled model with a large bias in its climatology and/or climate variability will generate an initial shock regardless of the initialization method, because the model tends to go back to its climate state.

3.2 Application to climate state estimation

It has been shown that CDA can produce self-balanced and consistent climate estimation (e.g., Zhang et al. 2014a; Feng et al. 2018). With the continuous development of coupled climate models and observational systems, CDA should lead to continuous improvement in climate reanalysis and prediction. Several major operational centers have established their own climate analysis systems based on CDA and CGCM. NOAA/GFDL has built an ensemble coupled data assimilation (ECDA) system for the estimation and prediction of seasonal to interdecadal climate variability by applying the ensemble adjustment Kalman filter and WCDA in CM2 (Zhang et al. 2007). The National Centers for Environmental Prediction (NCEP) has produced a Climate Forecast System Reanalysis (CFSR) from 1979 to the present, primarily for the initialization of seasonal forecasting, with a global atmosphere–ocean–land–sea ice coupled model in which 3D-Var-based WCDA was employed (Saha et al. 2010). ECMWF has established a coupled data assimilation system for the development of global reanalysis products of the 20th century based on an atmosphere–wave–ocean coupled model as well as 4D-Var-based WCDA (Laloyaux et al.

Fig. 7 An example of CDA minimizing initial shocks in coupled model predictions. Courtesy to Zhang et al. 2011a, **b** Time series of X_2 in **(a, c)** $F_{\text{Ocn(CDA)}/\text{Ocn(CDA)}}$ and **(b, d)** $F_{\text{Atm(CTL)}/\text{Ocn(Truth)}}$ as the model is initialized at the **a, b** 10,000th and **c, d** 30,000th time unit. The experiment $F_{\text{Atm(CDA)}/\text{Ocn(CDA)}}$ uses the coupled model ensemble states produced by the coupled data assimilation results as initial conditions, while the forecasts of $F_{\text{Atm(CTL)}/\text{Ocn(Truth)}}$ are initialized from the independent atmospheric ensemble combined with the perfect oceanic states taken from the truth



2016). The Japan Meteorological Agency (JMA) has produced coupled reanalysis products for the period 1940–2006 based on a global atmosphere–ocean coupled model and multivariate 3D-Var-based WCDA, which improved the feedback between sea surface temperature and precipitation (Fujii et al. 2009) as well as surface air and sea surface temperatures (Feng et al. 2018).

3.3 Application to climate prediction

In climate prediction, CDA was first applied in the initialization of forecasts of ENSO, a seasonal-to-interannual air–sea coupled phenomenon, with a coupled atmosphere–ocean model of intermediate complexity. Implementations of CDA include single observation assimilation in ocean (Kleeman et al. 1995), nudging (Chen et al. 1995, 1998), 4D-Var (Lee et al. 2000; Galanti et al. 2003; Gao et al. 2016), and use of the reduced-order Kalman filter (Ballabrera-Poy et al. 2001) and EnKF (Zheng and Zhu 2010). After that, CDA has been gradually applied in CGCMs for ENSO forecasting initialization (Zhang et al. 2005) and decadal-scale climate predictions (Yang et al. 2013, also see Fig. 8).

With the continuous development of high-performance computing capacity and coupled numerical simulation techniques, many countries have established their own climate prediction systems that are based on CDA in realistically configured CGCMs. The United States Naval Research Laboratory (NRL) established an atmosphere–ocean coupled

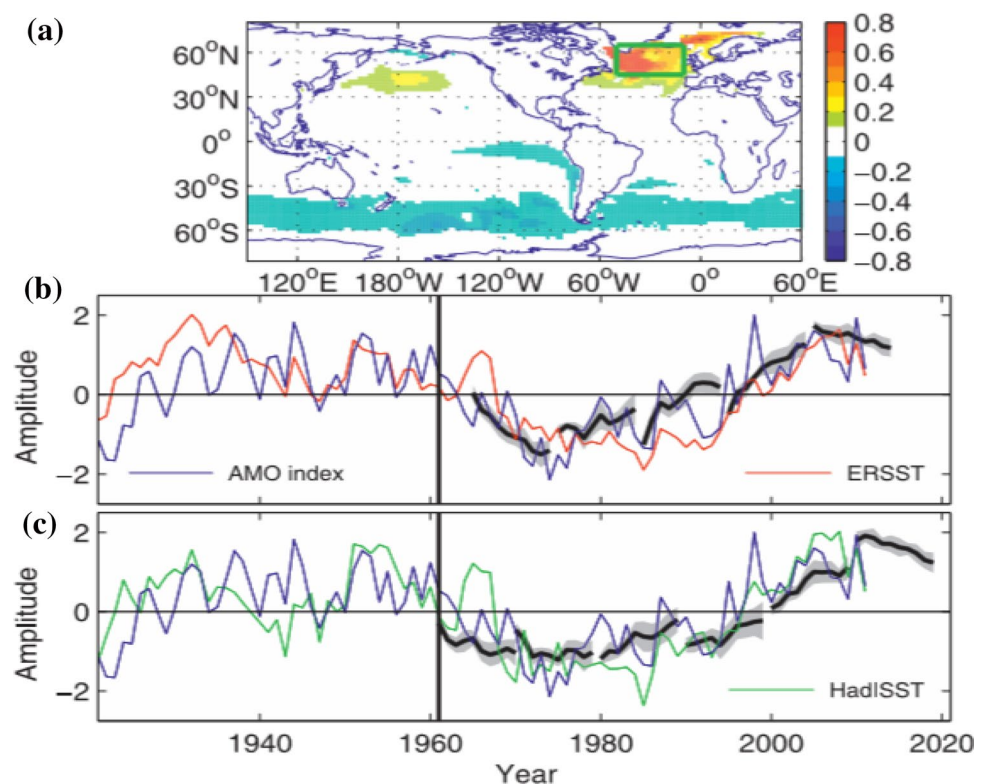
mesoscale ensemble prediction system based on a regional atmosphere–ocean coupled model with 3D-Var-based WCDA (Holt et al. 2011). The Japan Agency for Marine–Earth Science and Technology (JAMSTEC) built an adjoint coupled data assimilation system based on a global atmosphere–ocean coupled model using 4D-Var SCDA, to improve the estimation accuracy and the prediction skill of climate change at seasonal-to-interannual scales (Sugiura et al. 2008). Using a coupled global atmosphere–land–ocean–sea ice model, the UK Met Office set up a WCDA system based on 4D-Var (for the atmosphere) and 3D-Var (for the ocean), which has improved the accuracy of coupled model predictions to the seasonal scale (Lea et al. 2015).

The Institute of Atmospheric Physics, Chinese Academy of Sciences (IAP/CAS) recently established a ten-year climate prediction system based on an independently developed CGCM WCDA with an ensemble optimal interpolation (EnOI) incremental update scheme. The scheme led to improvement in the forecasting skill for the upper 700-m ocean temperature and sea surface temperature anomalies (Wu et al. 2018).

3.4 Application to coupled climate dynamics

CDA can also be used to study climate dynamics. A regional CDA (RCDA) approach, in which observations are assimilated only in a limited region, has recently been proposed to assess and understand the remote climate impact in a full

Fig. 8 An example of CDA implementing decadal climate predictions. Courtesy to Yang et al. (2013): **a** the spatial structure of the component that maximized the average predictability time of SST in the decadal hindcasts, which is called the internal multi-decadal pattern (IMP). **b** The ensemble mean (black solid) and spread (gray shading) time series of the IMP as a function of forecast lead time for the decadal hindcasts initialized on 1 Jan every 10 year from 1965 to 2005, the time series for projecting the ERSST data onto IMP (red solid), and the normalized AMO index (blue solid) from 1920 to 2010. **c** As in (b) but for hindcasts initialized on 1 Jan 1961 and every 10 year from 1970 to 2010. The green line denotes the projected time series of HadISST data onto IMP



CGCM. The application of RCDA to the extratropical climate system has been used to study the impact of extratropical climate on the tropical coupled climate system. These studies show that the extratropical climate can have a significant impact on the development of tropical ENSO in models (Lu et al. 2017a) and in observations (Lu and Liu, 2018). Climate model bias in the tropics can also be influenced by bias in the extratropic region (Lu et al. 2017b).

3.5 Application to numerical weather predictions

Traditional weather forecasts are based on atmosphere-only models and use persistent rather than forecasted sea surface temperature as a forcing at the lower boundary for the atmosphere, which often encounter problems with representing important physical processes at the air-sea interface. As a result, forecast uncertainties during extreme weather events can grow quickly as the forecast lead time increases. To reduce the uncertainties of lower boundary conditions and pursue coherent bottom boundary conditions, the coupled ocean–atmosphere model and corresponding CDA began being used to produce improved weather forecasts in different numerical weather prediction centers (e.g. Skachko et al. 2019; Browne et al. 2019; Guiavarc’h et al. 2019). Such efforts are where most of the developments have been made recently to pursue seamless weather–climate studies with high-resolution coupled models (e.g. Delworth et al. 2012; Small et al. 2014; Roberts et al. 2018) and CDA (e.g. Zhang et al. 2014b, 2015).

4 Summary and discussion

CDA is emerging as a potentially powerful strategy for improving weather and climate reanalysis and prediction. CDA has the following advantages over uncoupled DA:

- (1) It can produce more balanced state estimation for coupled prediction.
- (2) It can significantly improve state estimation in the under-sampled component (for example, sea ice, see e.g. Mahajan et al. 2011; Msadek et al. 2014)
- (3) It can improve coupled models by optimizing model parameters in the coupled framework.
- (4) It can improve our understanding of coupled dynamic processes in the coupled system.

Overcoming the remaining challenges in CDA involves significant multi-disciplinary interactions across the different Earth science disciplines, as well as computing algorithms. Progress in CDA also depends heavily on earth system modeling and data assimilation technology, particularly under the constraint of supercomputing capacity.

Several organizations such as GFDL and NCEP in the United States, JAMSTEC in Japan, as well as ECMWF have independently developed their own CDA techniques and corresponding climate prediction systems in their CGCMs. In recent years, Chinese high-performance computing (e.g., Sunway TaihuLight) and coupled climate models (e.g., the FGOALS-s2 coupled model independently developed by the IAP/CAS) have also developed rapidly, but the development of CDA is still in its initial stages. With the development of coupled models and unconventional observing systems, as well as assimilation technology, CDA has reached an era of rapid development, creating both opportunities and challenges. With the continuous increase of observational types and sources, and increased resolution in coupled models, a high-resolution CDA system that resolves mesoscale eddies in the ocean and tropical cyclones in the atmosphere (i.e., a horizontal resolution of 10 km in the ocean and 25 km in the atmosphere) can enable seamless weather–climate studies at weekly-to-decadal scales.

Many fundamental questions remain to be answered. Although CDA shows advantages over uncoupled DA, there has been no clear evidence that the current CDA implementation schemes have improved operational weather and climate forecasts in a great scope (see e.g., Brassington et al. 2015; Lea et al. 2015; Mulholland et al. 2015). Below, we discuss many ongoing challenges in CDA.

- (1) *CDA with high-resolution coupled models* Seamless weather–climate studies require both high-resolution (HR) coupled model and HR coupled model CDA. Currently, HR coupled models are advanced progressively (Small et al. 2014). However, because of limited computing resources, it is impractical to apply existing ensemble-based CDA methods [for instance, the GFDL-ECDA system (Zhang et al. 2007)] to HR coupled models for which a single model integration is intractable, although the parallelization technique of ensemble data assimilation algorithm progresses well (e.g. Nerger et al. 2005, 2019). How to implement a computationally efficient CDA algorithm that uses a single model integration (Yu et al. 2018) into a HR coupled model is a frontier research topic for seamless weather–climate reanalysis and predictions.
- (2) *Multiscale assimilation* At finer resolutions, more and more spatial–temporal scale phenomena can be simulated by coupled models. How best to consider the multi-scale information of models and observations is another frontier research topic in CDA (Zhang et al. 2014b; Zhao et al. 2017; Yu et al. 2018).
- (3) *SCDA* SCDA is still in its research stage. For example, in ensemble methodology, given the limited ensemble sizes in practice, it is difficult to get the exact coupled error covariance due to more complex model errors

in full CGCMs, disparities in the spatial–temporal scales between each model component, as well as coupled errors and sampling errors. Localizing the coupled error covariance is an efficient approach. Yoshida (2019) has addressed coupled error covariance localization by computing the error correlation for multiple pairs of observation and variable and extending it by training neural networks. The idea of coupled error covariance localization can be further studied in the application of SCDA and CPE in practice.

- (4) *Observation systems* Many existing observation systems that could be beneficial to CDA [e.g., ground-based snow observations (e.g. Moisseev et al. 2017; Lerber et al. 2017)] are not included in the Global Telecommunication System (GTS). Therefore, observations of the Earth system need to be collected more broadly and in a more standard way. Coupling fluxes at the air–sea interface and associated observations (e.g., ocean and atmosphere observations at the same time and place) are very helpful to improve coupled models and associated physics in examination, offset correction, error estimation and data assimilation (Penny and Hamill 2017).
- (5) *Coupled model parameter estimation* How to increase the signal-to-noise ratio in coupled model parameter estimation is still an urgent and challenging research aim. A deep linkage of the model sensitivities to parameter estimation is a viable approach; for example, sensitivity response time scale linking with parameter estimation update frequency (Liu et al. 2017a), sensitivity order linking to the multiple parameters being simultaneously estimated (Zhao et al. 2019). Most CDA parameter estimation studies to date have been performed in the perfect model scenario, where the true parameter is known. The ultimate challenge is the optimization of model parameters for the real world, although it is not clear that an optimal parameter exists. Further studies on this topic are therefore required. Furthermore, given the constraint of computing resources, how to implement HR CDA-PE with high computational efficiency is also an outstanding issue.
- (6) *Nonlinear assimilation* Compared with single models, coupled models have more complex model errors, meaning that existing CDA methods do not satisfy the Gaussian assumption. Therefore, the introduction of nonlinear assimilation techniques, such as the particle filter, into CDA is another important research topic. Currently, particle filters have already been applied to coupled climate models with some success (see e.g. Dubinkina et al. 2011; Dubinkina and Goosse, 2013; Browne and van Leeuwen 2015). Given the particle filter’s excessive computational cost, how to improve its performance in CDA is an ongoing challenging

research topic, but any progress will have good opportunities to improve the quality of CDA. Although quite challenging, how to implement coupled model parameter estimation using particle filters and examining its impact on coupled modeling could be a very interesting research topic.

Acknowledgements This work is supported by the National Key R&D Program of China (2017YFC1404100, 2017YFC1404103 and 2017YFC1404104) and the National Natural Science Foundation of China (Grant No. 41775100, 41830964).

References

- Anderson J, Hoar T, Raeder K, Liu H, Collins N, Torn R, Avellano A (2009) The data assimilation research testbed: a community facility. *Bull Am Meteorol Soc* 90(9):1283–1296. <https://doi.org/10.1175/2009bams2618.1> (ISSN 0003-0007)
- Annan JD, Hargreaves JC (2007) Efficient estimation and ensemble generation in climate modelling filter. *Philos Trans R Soc A* 365:2077–2088
- Annan JD, Hargreaves JC, Edwards NR, Marsh R (2005) Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter. *Ocean Model* 8:135–154
- Ballabrera-Poy J, Busalacchi AJ, Murtugudde R (2001) Application of a reduced-order Kalman filter to initialize a coupled atmosphere–ocean model: impact on the prediction of El Niño. *J Clim* 14:1720–1737
- Balmaseda M, Anderson D (2009) Impact of initialization strategies and observations on seasonal forecast skill. *Res Lett, Geophys.* <https://doi.org/10.1029/2008GL035561>
- Brassington GB, Martin MJ, Tolman HL, Akella S, Balmeseda M, Chambers CRS, Chassignet E, Cummings JA, Drillet Y, Jansen PAEM, Laloyaux P, Lea D, Mehra A, Mirouze I, Ritchie H, Samson G, Sandery PA, Smith GC, Suarez M, Todling R (2015) Progress and challenges in short- to mediumrange coupled prediction. *J Oper Oceanogr* 8(S2):s239–s258. <https://doi.org/10.1080/1755876X.2015.1049875>
- Brousseau P, Berre L, Bouttier F, Desroziers G (2012) Flow-dependent background error covariances for a convective-scale data assimilation system. *Q J R Meteorol Soc* 138(663):310–322. <https://doi.org/10.1002/qj.920>
- Browne B, van Leeuwen P (2015) Twin experiments with the equivalent weights particle filter and HadCM3. *Q J R Meteorol Soc* 141(693 Part B):3399–3414. <https://doi.org/10.1002/qj.2621> (ISSN 00359009)
- Browne PA, de Rosnay P, Zuo H, Bennett A, Dawson A (2019) Weakly coupled ocean–atmosphere data assimilation in the ECMWF NWP system. *Remote Sens* 11(234):1–24. <https://doi.org/10.3390/rs11030234>
- Brunet G, Jones S, Ruti PM (2015) Seamless prediction of the earth system: from minutes to months. *WMO Rep* 1156:483
- Buehner M, Du P, Bedard Joel (2017) A new approach for estimating the observation impact in ensemble-variational data assimilation. *Weather Rev, Mon.* <https://doi.org/10.1175/MWR-D-17-0252.1>
- Chen D, Cane MA (2008) El Niño prediction and predictability. *J Comput Phys* 227:3625–3640
- Chen D, Zebiak SE, Busalacchi AJ, Cane MA (1995) An improved procedure for El Niño forecasting. *Science* 269:1699–1702

- Chen D, Cane MA, Zebiak SE, Kaplan A (1998) The impact of sea level data assimilation on the Lamont model prediction of the 1997/98 El Niño. *Geophys Res Lett* 25:2837–2840
- Courtier P, Thepaut JN, Hollingsworth A (1994) A strategy for operational implementation of 4D-Var using an incremental approach. *Q J R Meteorol Soc* 120:1367–1387
- Dee DP, Balmaseda M, Balsamo G, Engelen R, Simmons AJ, Thépaut JN (2014) Toward a consistent reanalysis of the climate system. *Bull Am Meteorol Soc* 95:1235–1248
- Delworth TL, Rosati A, Anderson W, Adcroft AJ, Balaji V, Benson R et al (2012) Simulated climate and climate change in the gfdl cm². 5 high-resolution coupled climate model. *J Clim* 25(8):2755–2781
- Derber J, Rosati A (1989) A global oceanic data assimilation system. *J Phys Oceanogr* 19:1333–1347
- Du H, Huang S, Cai Q, Cheng L (2009) Studies of variational assimilation for the inversion of the coupled air-sea model. *Mar Sci Bull* 11(2):13–22
- Dubinkina S, Goosse H (2013) An assessment of particle filtering methods and nudging for climate state reconstructions. *Clim Past* 9:1141–1152. <https://doi.org/10.5194/cp-9-1141-2013> (ISSN 18149324)
- Dubinkina S, Goosse H, Sallaz-Damaz Y, Crespin E, Crucifix M (2011) Testing a particle filter to reconstruct climate changes over the past centuries. *Int J Bifurc Chaos* 21(12):3611–3618. <https://doi.org/10.2172/1025774>
- Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res* 99(5):10143–10162
- Evensen G (2007) *Data assimilation: the ensemble Kalman Filter*. Springer Press, Verlag New York, p 187
- Fairbairn D, Pring SR, Lorenc AC, Roulstone I (2014) A comparison of 4dvar with ensemble data assimilation methods. *Q J R Meteorol Soc* 140(678):281–294. <https://doi.org/10.1002/qj.2135>
- Feng XB, Haines K, Boisseson ED (2018) Coupling of surface air and sea surface temperatures in the CERA-20C reanalysis. *Q J R Meteorol Soc* 144(710):195–207
- Fowler AM, Lawless AS (2016) An Idealized study of coupled atmosphere-ocean 4D-Var in the presence of model error. *Mon Weather Rev* 144:4007–4030
- Frankignoul C, Czaja A, L'Heveder B (1998) Air–sea feedback in the North Atlantic and surface boundary conditions for ocean models. *J Clim* 11:2310–2324
- Frolov S, Bishop C, Holt T, Cummings J, Kuhl D (2016) Facilitating strongly-coupled ocean-atmosphere data assimilation with an interface solver. *Mon Weather Rev* 144(1):3–20
- Fujii Y, Nakaegawa T, Matsumoto S, Yasuda T, Yamanaka G, Kamachi M (2009) Coupled climate simulation by constraining ocean fields in a coupled model with ocean data. *J Clim* 22:5541–5557
- Galanti E, Tziperman E, Harrison M, Rosati A, Sirkes Z (2003) A study of ENSO prediction using a hybrid coupled model and the adjoint method for data assimilation. *Mon Weather Rev* 131(11):2748–2764
- Gandin LS (1965) *Objective analysis of meteorological fields*. U.S. Dept. Commerce and National Science Foundation Washington, D.C. (Original in Russian, ob'ektivnyi analiz meteorologicheskikh polei, 1963)
- Gao C, Wu X, Zhang RH (2016) Testing a four-dimensional variational data assimilation method using an improved intermediate coupled model for ENSO analysis and prediction. *Adv Atmos Sci* 33(7):875–888
- Giering R, Kaminski T, Todling R, Errico R, Gelaro R, Winslow N (2006) Tangent linear and adjoint versions of NASA/GMAO's fortran 90 global weather forecast model. In: Bücker M, Collis G, Naumann U, Hovland P, Norris B (eds) *Automatic differentiation: applications, theory, and implementations*. Lecture notes in computational science and engineering, vol 50. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-28438-9_24
- Guiavarc'h C, Roberts-jones J, Harris C, Lea DJ, Ryan A, Ascione I (2019) Assessment of ocean analysis and forecast from an atmosphere ocean coupled data assimilation operational system. *Ocean Sci* 15(5):1307–1326
- Han G, Wu X, Zhang S, Liu Z, Li W (2013) Error covariance estimation for coupled data assimilation using a Lorenz atmosphere and a simple pycnocline ocean model. *J Clim* 26:10218–10231
- Han G, Zhang X, Zhang S, Wu X, Liu Z (2014) Mitigation of coupled model biases induced by dynamical core misfitting through parameter optimization: simulation with a simple pycnocline prediction model. *Nonlinear Process Geophys* 21:357–366
- Han G, Wu X, Zhang S, Liu Z, Navon IM, Li W (2015) A study of coupling parameter estimation implemented by 4D-Var and EnKF with a simple coupled system. In: *Advances in meteorology*, pp 1–16
- He Y, Wang B, Liu M, Liu L, Yu Y, Liu J, Li R, Zhang C, Xu S, Huang W, Liu Q, Wang Y, Li F (2017) Reduction of initial shock in decadal predictions using a new initialization strategy. *Geophys Res Lett* 44:8538–8547
- Holt T, Cummings J, Bishop C, Doyle J, Hong X, Chen S, Jin Y (2011) Development and testing of a coupled ocean–atmosphere mesoscale ensemble prediction system. *Ocean Dyn* 61:1937–1954
- Isaksen L, Fisher M, Berner J (2007) Use of analysis ensembles in estimating flow-dependent background error variance. In *Proc. ECMWF workshop on Flow-dependent aspects of data assimilation*, pages 65–86, 2007
- Ito K, Ishikawa Y, Awaji T (2010) Specifying air-sea exchange coefficients in the high-wind regime of a mature tropical cyclone by an adjoint data assimilation method. *Sola* 6:13–16
- Ji M, Kumar A, Leetmaa A (1994) An experimental coupled forecast system at the national meteorological center: some early results. *Tellus* 46A:398–418
- Jia L, Yang X, Vecchi GA, Gudgel RG, Delworth TL, Rosati A, Stern WF, Wittenberg AT, Krishnamurthy L, Zhang S, Msadek R, Kapnick SB, Underwood SD, Zeng F, Anderson WG, Balaji V, Dixon KW (2015) Improved seasonal prediction of temperature and precipitation over land in a high-resolution GFDL climate model. *J Clim* 28(5):2044–2062
- Kang JS, Kalnay E, Liu J, Fung I, Miyoshi T, Ide K (2011) “Variable localization” in an ensemble Kalman filter: application to the carbon cycle data assimilation. *J Geophys Res*. <https://doi.org/10.1029/2010JD014673>
- Kang J, Kalnay E, Liu J, Fung I, Miyoshi T, Ide K (2012) Estimation of surface carbon fluxes with an advanced data assimilation methodology. *J Geophys Res* 117:D24101. <https://doi.org/10.1029/2012JD018259>
- Keenlyside N, Latif M, Jungclaus J, Kornbluh L, Roeckner E (2008) Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature* 453:84–88
- Kleeman R, Moore AM, Smith NR (1995) Assimilation of subsurface thermal data into a simple ocean model for the initialization of an intermediate tropical coupled ocean-atmosphere forecast model. *Mon Weather Rev* 123:3103–3113
- Kondrashov D, Sun C, Ghil M (2008) Data assimilation for a coupled ocean atmosphere model. Part II: parameter estimation. *Mon Weather Rev* 136:5062–5076
- Lalouaux P, Balmaseda M, Dee D, Mogensen K, Janssen P (2016) A coupled data assimilation system for climate reanalysis. *Q J R Meteorol Soc* 142:65–78
- Lalouaux P, Boisseson E, Balmaseda M, Bidlot JR, Broennimann S, Buizza R, Dalhgren P, Dee D, Haimberger L, Hersbach H, Kosaka Y, Martin M, Poli P, Rayner N, Rustemeier E, Schepers D (2018) CERA-20C: a coupled reanalysis of the twentieth

- century. *J Adv Model Earth Syst* 10:1172–1195. <https://doi.org/10.1029/2018MS001273>
- Lea DJ, Mirouze I, Martin MJ, King RR, Hines A, Walters D, Thurlow M (2015) Assessing a new coupled data assimilation system based on the Met Office coupled atmosphere-land-ocean-sea ice model. *Mon Weather Rev* 143(11):4678–4694
- Lee T, Boulanger JP, Foo A, Fu LL, Giering R (2000) Data assimilation by an intermediate coupled ocean-atmosphere model: application to the 1997–1998 El Niño. *J Geophys Res* 105:26063–26087
- Lerber AV, Moiseev D, Marks DA, Petersen W, Harri A-M, Chandrasekar V (2017) Validation of GMI Snowfall Observations by Using a Combination of Weather Radar and Surface Measurements. *J Appl Meteorol Climatol* 57:797–820. <https://doi.org/10.1175/JAMC-D-17-0176.1>
- Li S, Zhang S, Liu Z, Lu L, Zhu J, Zhang X, Wu X, Zhao M, Vecchi GA, Zhang R, Lin X (2018) Estimating convection parameters in the GFDL CM2.1 model using ensemble data assimilation. *J Adv Model Earth Syst* 10:989–1010. <https://doi.org/10.1002/2017ms001222>
- Liu Y, Gupta HV, Sorooshian S, Bastidas LA, Shuttleworth WJ (2005) Constraining land surface and atmospheric parameters of a locally coupled model using observational data. *J Hydrometeorol* 6:156–172
- Liu Z, Wu S, Zhang S, Liu Y, Rong X (2013) Ensemble Data Assimilation in a Simple Coupled Climate Model: the Role of Ocean-Atmosphere Interaction. *Adv Atmos Sci* 30(5):1235–1248
- Liu Y, Liu Z, Zhang S, Rong X, Jacob R, Wu S, Lu F (2014a) Ensemble-based parameter estimation in a coupled GCM using the adaptive spatial average method. *J Clim* 27:4002–4014
- Liu Y, Liu Z, Zhang S, Jacob R, Lu F, Rong X, Wu S (2014b) Ensemble-Based parameter estimation in a coupled general circulation model. *J Clim* 27:7151–7162
- Liu C, Zhang S, Li S, Liu Z (2017a) Impact of the time scale of model sensitivity response on coupled model parameter estimation. *Adv Atmos Sci* 34:1346–1357
- Liu X, Köhl A, Stammer D, Masuda S, Ishikawa Y, Mochizuki T (2017b) Impact of in-consistency between the climate model and its initial conditions on climate prediction. *Clim Dyn* 49:1061–1075
- Lorenz AC (1986) Analysis methods for numerical weather prediction. *Q J R Meteorol Soc* 112:1177–1194
- Lu JX, Hsieh WW (1998) On determining initial conditions and parameters in a simple coupled atmosphere-ocean model by adjoint data assimilation. *Tellus* 50A:534–544
- Lu F, Liu Z (2018) Assessing extratropical influence on the observed El Niño-Southern Oscillation events using regional coupled data assimilation. *J Clim* (in press)
- Lu F, Liu Z, Zhang S, Liu Y (2015a) Strongly coupled data assimilation using Leading Averaged Coupled Covariance (LACC). Part I: Simple model study. *Mon Weather Rev* 143(9):3823–3837
- Lu F, Liu Z, Zhang S, Liu Y (2015b) Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part II: CGCM applications. *Mon Weather Rev* 143(11):4645–4659
- Lu F, Liu Z, Liu Y, Zhang S, Jacob R (2017a) Understanding the control of extratropical atmospheric variability on ENSO using a coupled data assimilation approach. *Clim Dyn* 48:3139–3160
- Lu F, Liu Z, Zhang S, Jacob R (2017b) Assessing extratropical influence on tropical bias in climate models with regional coupled data assimilation. *Geophys Res Lett* 44:3384–3392
- Luo X, Hoteit I (2014) Ensemble Kalman filtering with a divided state-space strategy for coupled data assimilation problems. *Mon Weather Rev* 142:4542–4558
- Ma X, Jing Z, Chang P, Liu X, Montuoro R, Small RJ, Bryan FO, Greatbatch RJ, Brandt P, Wu D, Lin X, Wu L (2016) Western boundary currents regulated by interaction between ocean eddies and the atmosphere. *Nature* 535(7613):533–537
- Mahajan S, Zhang R, Delworth TL (2011) Impact of the Atlantic meridional overturning circulation (AMOC) on Arctic surface air temperature and sea ice variability. *J Clim* 24:6573–6581. <https://doi.org/10.1175/2011JCLI4002.1>
- Meehl GA, Goddard L, Murphy J, et al. (2009) Decadal prediction. *Bullet Am Meteorol Soc* 90(10):1467–1486
- Moiseev D, von Lerber A, Tiira J (2017) Quantifying the effect of riming on snowfall using ground-based observations. *JGR Atmospheres* 122(7):4019–4037. <https://doi.org/10.1002/2016JD026272>
- Msadek R, Delworth TL, Rosati A, Anderson WG, Vecchi GA, Chang YS, Dixon KW, Gudgel RG, Stern WF, Wittenberg AT, Yang X, Zeng F, Zhang R, Zhang S (2014) Predicting a decadal shift in North Atlantic climate variability using the GFDL forecast system. *J Clim* 27(17):6472–6496
- Mulholland DP, Laloyaux P, Haines K, Balmaseda MA (2015) Origin and impact of initialization shocks in coupled atmosphere-ocean forecasts. *Mon Weather Rev* 143:4631–4644
- Navon IM (1997) Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dyn Atmos Oceans* 27:55–79
- Nerger L, Hiller W (2013) Software for ensemble-based data assimilation systems—implementation strategies and scalability. *Compute Geosci* 55:110–118. <https://doi.org/10.1016/j.cageo.2012.03.026> (ISSN 00983004)
- Nerger L, Hiller W, Schröter J (2005) Pdafthe parallel data assimilation framework: experiences with kalman filtering. In: Zwiefelhofer W, Mozdzyński G (eds) Proceedings of the Eleventh ECMWF Workshop on the use of high performance computing in meteorology. World Scientific, Reading, pp 63–83
- Nerger L, Tang Q, Mu L (2019) Efficient ensemble data assimilation for coupled models with the parallel data assimilation framework: Example of awi-cm. *Geosci Model Dev Discuss*. <https://doi.org/10.5194/gmd-2019-167>
- Penny SG, Hamill TM (2017) Coupled data assimilation for integrated earth system analysis and prediction. *Bull Am Meteorol Soc* 98(7):169–172
- Penny, SG, Akella S, Alves O, Bishop C, Buehner M, Chevallier M, Counillon F, Draper C, Frolov S, Fujii Y, Karspeck A, Kumar A, Laloyaux P, Mahfouf J-F, Martin M, Peña (NOAA/NCEP) M, Rosnay P, Subramanian A, Tardif R, Wang (NERSC) Y, Wu X (2017) Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations. WMO WWRP White paper. https://www.wmo.int/pages/prog/arep/wwrp/new/documents/Final_WWRP_2017_3_27_July.pdf
- Posselt DJ, Bishop CH (2012) Nonlinear parameter estimation: comparison of an ensemble Kalman smoother with a Markov chain Monte Carlo algorithm. *Mon Weather Rev* 140:1957–1974
- Randall DA, Wood RA, Bony S, et al. (2007) Climate models and their evaluation. In: Solomon S, Qin D, Manning M, et al. (eds) Climate change 2007: the physical science basis. Contribution Of working Group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, New York, NY (US), pp 597–599
- Rienecker MM (2003) Report of the coupled data assimilation. Workshop (NOAA/OGP), workshop report, Natl. Oceanogr. and Atmos. Admin., Portland, Oreg., 21–23 Apr
- Roberts MJ, Vidale PL, Senior C et al (2018) The benefits of global high resolution for climate simulation: process understanding and the enabling of stakeholder decisions at the regional scale. *Bull Am Meteorol Soc* 99(11):2341–2359
- Rosati A, Miyakoda K, Gudgel R (1997) The impact of ocean initial conditions on ENSO forecasting with a coupled model. *Mon Weather Rev* 125:754–772

- Ruiz JJ, Pulido M, Miyoshi T (2013) Estimating model parameters with ensemble-based data assimilation: a review. *J Meteorol Soc Jpn* 91(2):79–99
- Saha S, Nadiga S, Thiaw C, Wang J (2006) The NCEP climate forecast system. *J Clim*. <https://doi.org/10.1175/JCLI3812.1>
- Saha S et al (2010) The NCEP Climate forecast system reanalysis. *Bull Am Meteorol Soc* 91:1015–1057
- Singleton T (2011) Data assimilation experiments with a simple coupled ocean-atmosphere model, PhD thesis, Dep. of Atmos. And Oceanic Sci., Univ. of Maryland, College Park, Md
- Skachko S, Buehner M, Laroche S, Lapalme E, Smith G, Roy F, Surcel-Colan D, Bélanger J-M, Garand L (2019) Weakly coupled atmospheric-ocean data assimilation in the Canadian global prediction system (v1). *Geosci Model Dev Discuss* 2019:1–30. <https://doi.org/10.5194/gmd-2019>
- Sluka TC, Penny SG, Kalnay E, Miyoshi T (2016) Strongly coupled enfk data assimilation in coupled ocean-atmosphere models. In: The 96th AMS Annual Meeting, “Earth System Science in Service to Society,” 10–14 January 2016 in New Orleans, Louisiana
- Sluka TC, Penny SG, Kalnay E, Miyoshi T (2016b) Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophys Res Lett* 43:752–759
- Small RJ, Bacmeister J, Bailey D, Baker A, Bishop S, Bryan F, Caron J, Dennis J, Gent P, Hsu HM, Jochum M, Lawrence D, Muñoz E, diNezio P, Scheitlin T, Tomas R, Tribbia J, Tseng YH, Vertenstein M (2014) A new synoptic scale resolving global climate simulation using the Community Earth System Model. *JAMES* 6(4):1065–1094
- Smith PJ, Fowler AM, Lawless AS (2015) Exploring strategies for coupled 4DVar data assimilation using an idealised atmosphere–ocean model. *Tellus A* 67(1):27025
- Smith PJ, Amos SL, Nancy KN (2017) Estimating forecast error covariances for strongly coupled atmosphere-ocean 4D-Var data assimilation. *Mon Weather Rev* 145:4011–4035
- Smith PJ, Lawless AS, Nichols NK (2018) Treating sample covariances for use in strongly coupled atmosphere-ocean data assimilation. *Geophys Res Lett* 45:445–454
- Song JQ, Cao XQ, Zhang WM, Zhu XQ (2012) Estimating parameters for coupled air-sea model with variational method. *Acta Phys. Sin.* 61(11):110401 (**In Chinese**)
- Stammer D, Wunsch C, Giering R, Eckert C, Heimbach P, Marotzke J, Adcroft A, Hill CN, Marshall J (2002) The global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. *J Geophys Res* 107(C9):3118
- Sugiura N, Awaji T, Masuda S, Mochizuki T, Toyoda T, Miyama T, Igarashi H, Ishikawa Y (2008) Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *J Geophys Res* 113:C10017
- Tardif R, Hakim GJ, Snyder C (2014) Coupled atmosphere-ocean data assimilation experiments with a low order climate model. *Clim Dyn* 43:1631–1643
- Wu WS, Purser RJ, Parrish DF (2002) Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon Weather Rev* 130:2905–2916
- Wu X, Zhang S, Liu Z, Rosati A, Delworth TL, Liu Y (2012a) Impact of geographic-dependent parameter optimization on climate estimation and prediction: simulation with an intermediate coupled model. *Mon Weather Rev* 140:3956–3971
- Wu X, Zhang S, Liu Z, Rosati A, Delworth TL (2012b) A study of impact of the geographic dependent of observing system on parameter estimation with an intermediate coupled model. *Clim Dyn* 40:1789–1798
- Wu X, Han G, Zhang S, Liu Z (2016) A study of the impact of parameter optimization on ENSO predictability with an intermediate coupled model. *Clim Dyn* 46:711–727
- Wu B, Zhou T, Zheng F (2018) EnOI-IAU initialization scheme designed for decadal climate prediction system IAP-DecPreS. *J Adv Model Earth Syst* 10:342–356
- Xie Y, Koch S, McGinley J, Albers S, Bieringer PE, Wolfson M, Chan M (2011) A space–time multiscale analysis system: a sequential variational analysis approach. *Mon Weather Rev* 139:1224–1240
- Yang X, Rosati A, Zhang S, Delworth TL (2013) A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *J Clim* 26:650–661
- Yoshida T (2019) Covariance localization in strongly coupled data assimilation. UMD doctoral thesis, <https://drum.lib.umd.edu/handle/1903/25170>
- Yoshida T, Kalnay E (2018) Correlation-cutoff method for covariance localization in strongly coupled data assimilation. *Weather Rev, Mon.* <https://doi.org/10.1175/MWR-D-17-0365.1>
- Yu X, Zhang S, Li J, Lu L, Liu Z, Li M, Yu H, Han G, Lin X, Wu L, Chang P (2019) A multi-time scale EnOI-like high-efficiency filter for coupled model data assimilation. *J Adv Model Earth Syst* 11(1):45–63
- Zhang S (2011a) Impact of observation-optimized model parameters on decadal predictions: simulation with a simple pycnocline prediction model. *Geophys Res Lett* 38(2):L02702
- 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 Clim* 24:6210–6226
- Zhang S, Harrison MJ, Wittenberg AT, Rosati A, Anderson JL, Balaji V (2005) Initialization of an ENSO forecast system using a parallelized ensemble filter. *Mon Weather Rev* 133(11):3176–3201
- Zhang S, Harrison MJ, Rosati A, Wittenberg A (2007) System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon Weather Rev* 135:3541–3564
- Zhang S, Liu Z, Rosati A, Delworth TL (2012) A study of enhance parameter correction with coupled data assimilation for climate estimation and prediction using a simple coupled model. *Tellus* 64:10963
- Zhang S, Chang YS, Yang X, Rosati A (2014a) Balanced and coherent climate estimation by combining data with a biased coupled model. *J Clim* 27:1302–1314
- Zhang S, Zhao M, Lin S-J, Yang X, Anderson W (2014b) Retrieval of tropical cyclone statistics with a high-resolution coupled model and data. *Geophys Res Lett* 41:652–660. <https://doi.org/10.1002/2013GL058879>
- Zhang S, Zhao M, Lin S-J, Yang X, Anderson W, Zhang W, Rosati A, Underwood S, Zeng F (2015) Impact of having realistic tropical cyclone frequency on ocean heat content and transport forecasts in a high-resolution coupled model. *Geophys Res Lett* 42:5966–5973. <https://doi.org/10.1002/2015GL064745>
- Zhao Y, Deng X, Zhang S, Liu Z, Liu C, Vecchi G, Han G, Wu X (2017) Impact of an observational time window on coupled data assimilation: simulation with a simple climate model. *Nonlinear Process Geophys* 24:681–694
- Zhao Y, Deng X, Zhang S, Liu Z, Liu C (2019) Sensitivity determined simultaneous estimation of multiple parameters in coupled models: part I—based on single model component sensitivities. *Clim Dyn.* <https://doi.org/10.1007/s00382-019-04865-3>
- Zheng F, Zhu J (2010) Coupled assimilation for an intermediated coupled ENSO prediction model. *Ocean Dyn* 60:1061–1073
- Zupanski M (2005) Maximum likelihood ensemble filter: theoretical aspects. *Mo. Weather Rev* 133:1710–1726
- Županski M (2016) Data assimilation for coupled modeling systems. In: Park SK, Xu L (eds) *Data assimilation for atmospheric, oceanic, and hydrologic applications, vol III*. Springer International Publishing, Switzerland, pp 55–70