# *Strongly coupled data assimilation in multiscale media: experiments using a quasi-geostrophic coupled model*

AGU100 ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2019MS001652

# Strongly Coupled Data Assimilation in Multiscale Media: Experiments Using a Quasi-Geostrophic Coupled Model

**S. G. Penny[1,2,3], E. Bach[1,4], K. Bhargava[1], C.-C. Chang[1], C. Da[1], L. Sun[1], and T. Yoshida[1,5]**

[1]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, USA, [2]NOAA National Centers for Environmental Prediction, College Park, MD, USA, [3]RIKEN Center for Computational Science, Kobe, Japan, [4]Institute for Physical Science and Technology, University of Maryland, College Park, MD, USA, [5]Climate Prediction Division, Japan Meteorological Agency, Tokyo, Japan

**Key Points:**
- Strongly coupled data assimilation produces more accurate state estimates and forecasts than traditional approaches to data assimilation
- Dynamic background error covariance estimates are essential for strongly coupled data assimilation, especially with limited observations
- Ocean observations can constrain the atmosphere, provided that the analysis-observing frequency and the ensemble size are sufficiently high

**Correspondence to:**
S. G. Penny,
steve.penny@noaa.gov

**Citation:**
Penny, S. G., Bach, E., Bhargava, K., Chang, C.-C., Da, C., Sun, L., & Yoshida, T. (2019). Strongly coupled data assimilation in multiscale media: Experiments using a quasi-geostrophic coupled model. *Journal of Advances in Modeling Earth Systems*, *11*, 1803–1829. https://doi.org/10.1029/2019MS001652

Received 7 FEB 2019
Accepted 28 MAY 2019
Accepted article online 4 JUN 2019
Published online 24 JUN 2019



**Abstract** Strongly coupled data assimilation (SCDA) views the Earth as one unified system. This allows observations to have an instantaneous impact across boundaries such as the air-sea interface when estimating the state of each individual component. Operational prediction centers are moving toward Earth system modeling for all forecast timescales, ranging from days to months. However, there have been few studies that examine fundamental aspects of SCDA and the transition from traditional approaches that apply data assimilation only to a single component, whether forecasts were derived from a coupled model or an uncoupled forced model. The SCDA approach is examined here in detail using numerical experiments with a simple coupled atmosphere-ocean quasi-geostrophic model. The impact of coupling is explored with respect to its impact on the Lyapunov spectrum and on data assimilation system stability. Different data assimilation methods are compared within the context of SCDA, including the 3-D and 4-D Variational methods, the ensemble Kalman filter, and the hybrid gain method. The impact of observing system coverage is also investigated. We find that SCDA is generally superior to weakly coupled or uncoupled approaches. Dynamically defined background error covariance estimates are essential for SCDA to achieve an accurate coupled state estimate as the observing system becomes sparser. As a clarification of seemingly contradictory findings from previous studies, it is shown that ocean observations can adequately constrain atmospheric state estimates provided that the analysis-observing frequency is sufficiently high and the ensemble size determining the background error covariance is sufficiently large.

**Plain Language Summary** To make accurate predictions of weather and climate, scientists develop complex computer models of the Earth that couple smaller models of the atmosphere, oceans, land, etc. Measurements from satellites, ground stations, and ocean buoys are limited, so these give an incomplete picture. Forecasters combine past predictions with new observations to make an informed guess about what the entire Earth looks like at any instant. This guess is used to initialize computer model forecasts. However, since the atmosphere is a chaotic system, small inaccuracies in this initial guess can lead to wildly diverging forecasts further out in time. This study investigates a new method for initializing the Earth models. In the past, physical inconsistencies between the atmosphere and ocean arose because each was initialized independently. We use atmospheric observations to improve our best guess of the ocean conditions, and vice versa, to improve this consistency. This may improve daily weather forecasts and seasonal guidance when adopted by operational weather prediction centers. Given the growing urgency to respond to near-term climate change risks, improved methods to initialize Earth system models provide the potential to generate accurately initialized climate predictions that can better anticipate the coming changes over the next decade.

## 1. Introduction

Coupled data assimilation (CDA) is a growing but largely understudied discipline essential for the next generation of weather, subseasonal to seasonal, and climate prediction efforts. Data assimilation (DA) is used at operational weather and climate prediction centers to initialize numerical weather forecasts, to provide ocean monitoring, and to initialize Earth system models for subseasonal to seasonal prediction. In addition, there is growing interest in the climate modeling community in using DA to initialize coupled models due to the adoption of higher resolutions that prevent long model runs from reaching an equilibrium state (Kinter et al., 2016; Ringler et al., 2018).

Traditionally, operational prediction centers have analyzed the atmosphere and ocean independently. Early developments in CDA were pioneered at the National Centers for Environmental Prediction with the application of the 3-D Variational (3D-Var) method in the Climate Forecast System (Saha et al., 2006) and at the Geophysical Fluid Dynamics Laboratory with an ensemble data assimilation approach (Zhang et al., 2007). Both efforts used a coupled forecast model, with the DA applied independently in the atmospheric and oceanic components—an approach called weakly coupled data assimilation (WCDA). More recently, Laloyaux et al. (2016) implemented a CDA system with independent analyses in the atmosphere and ocean but with interaction between the two domains during the outer loop of the iterative 4-D Variational (4D-Var) minimization. Today, the leading operational centers are steadily moving toward a paradigm of CDA and are closer to analyzing the Earth as a single system—an approach called strongly coupled data assimilation (SCDA). However, basic challenges remain before widespread adoption of SCDA (Penny et al., 2017; Penny & Hamill, 2017).

The benefits of SCDA have been investigated in several studies using either an EnKF (Liu et al., 2013; Lu et al. 2015a, 2015b; Sluka et al., 2016) or 4D-Var data assimilation system (Smith et al., 2015). We investigate the transition from the paradigm applying DA to independent forced atmosphere and ocean models to the new paradigm using a fully coupled forecast model with CDA. To understand the effects of this transition on the dynamics, we compute the Lyapunov spectrum of the individual component models and of the coupled system both without and with DA. We further examine how the unstable-neutral subspace (e.g., as described by Trevisan & Uboldi, 2004, and Trevisan & Palatella, 2011) of the coupled system is resolved and constrained by variational and ensemble data assimilation methods. We investigate the construction of a climatological background error covariance matrix for use in variational and hybrid methods. We probe the observability of the coupled system and investigate simple permutations of the observing network. Finally, we compare forecasts starting from initial conditions derived using a variety of CDA approaches.

All demonstrations are carried out with the simplified Modular Arbitrary-Order Ocean-Atmosphere Model (MAOOAM) developed by De Cruz et al. (2016). The atmospheric component of the MAOOAM is based on Charney and Straus (1980), Cehelsky and Tung (1987) and Reinhold and Pierrehumbert (1982), while the ocean component is based on Pierini (2011) and Barsugli and Battisti (1998). The coupling between the two components includes momentum transfer (wind forcing), as well as radiative and heat exchanges. Examinations of the Lyapunov spectrum of the coupled configuration of the MAOOAM were conducted by Vannitsem et al. (2015), Vannitsem and Lucarini (2016), Vannitsem (2017), and De Cruz et al. (2018). Vannitsem and Lucarini (2016) also examined MAOOAM's covariant Lyapunov vectors.

## 2. Related Work

Foundational principles of SCDA have been investigated using a variety of simplified models and more realistic coupled general circulation models (GCMs). We identify a number of previous studies that have focused specifically on SCDA. There are many other important works that have explored CDA with lesser degrees of coupling that we neglect here for brevity.

Ballabrera-Poy et al. (2009) investigated strongly coupled multiscale assimilation in the context of the two-scale model of Lorenz and Emanuel (1998) using an ensemble Kalman filter (EnKF). Results indicated that using an ensemble size greater than 50 members could simultaneously constrain the large-scale and small-scale fields when the small-scale fields are exhaustively sampled. However, they found that while assimilating only observations of the large-scale dynamics produced accurate results for the large-scale dynamics, it could not adequately constrain the small-scale dynamics. Assimilating the mostly uncorrelated signal from the small-scale dynamics to constrain the large-scale dynamics was prone to the errors associated with spurious covariances that were likely degrading the filter.

Zheng and Zhu (2010) investigated a coupled El Niño-Southern Oscillation (ENSO) model centered on the tropical Pacific and compared the direct assimilation of sea surface temperature (SST) to the unidirectional SCDA of surface wind observations using a deterministic EnKF. The coupled model was composed of a McCreary-type (1981) modal ocean model extended to include a horizontally varying background stratification, as well as 10 baroclinic modes plus a parameterization of the local Ekman-driven upwelling with an embedded SST anomaly model, and a statistical atmospheric model that relates the wind stress to SST anomaly fields (Zhang et al. 2005). In both perfect model and real-data forecast experiments, Zheng and Zhu

(2010) found that assimilating the surface winds improved surface currents and reduced the average error growth rate of SST anomalies during the forecast versus the direct assimilation of SST.

Han et al. (2013) explored a Lorenz (1963) atmosphere coupled to a simple pycnocline ocean model and found that SCDA was only effective when the DA system was presented with highly accurate cross-domain error covariance estimates. Similarly to Ballabrera-Poy et al. (2009), it was found to be more difficult to improve the fast-varying part of the system using observations from the slow-varying part of the system. Han et al. (2013) attributed this to the linear regression formula of the ensemble adjustment Kalman filter (EAKF; Anderson, 2001), since the fast-varying system contributes to the higher-order (nonlinear) terms of the regression at the timescale of the slow-varying system.

Liu et al. (2013) explored a coupled Lorenz-Jin model (Jin, 1997; Lorenz, 1963) with an EAKF, finding that the SCDA that assimilates observations in both the atmosphere and ocean and that employs the coupled covariance matrix outperforms the WCDA alternative. Further, the assimilation of synoptic atmospheric variability was found to be critical for the improvement of not only the atmospheric state but also the oceanic state through the coupled error covariance matrix, especially in the midlatitude system where oceanic variability is driven predominantly by weather noise. It was also found that the assimilation of synoptic atmospheric observations alone improved the coupled state almost as much as assimilating additional oceanic observations, while the assimilation of oceanic observations had little impact on the atmosphere.

Tardif et al. (2014, 2015) explored a low-order coupled Lorenz-Stommel model (Roebber, 1995) with a perturbed observations EnKF (Houtekamer & Mitchell, 1998). Tardif et al. (2014) proposed the assimilation of time-averaged observations to capture stronger signals in the covariance relationship between the atmosphere and ocean at various timescales. The minimum ensemble size required to resolve correlations between the atmospheric variables with the Meridional Overturning Circulation was reduced by assimilating time-averaged observations. Tardif et al. (2015) compared cycling and no-cycling DA experiments with the time-averaged proxy observations to better constrain the slow modes of the system. They found that by assimilating only time-averaged atmospheric observations, the overturning circulation in the ocean could be constrained similarly to when ocean observations were present.

Smith et al. (2015) used an idealized single-column atmosphere-ocean coupled model to compare uncoupled, weakly coupled, and strongly coupled variations of an incremental 4D-Var. They found that SCDA is able to reduce initialization shocks and improve subsequent forecasts compared to uncoupled and WCDA approaches. Smith et al. (2017, 2018) conducted more detailed studies to better estimate forecast error covariance for use with 4D-Var in the SCDA scenario.

Following an investigation with a simple model (Lu et al., 2015a), Lu et al. (2015b) focused on the effect of SCDA restricted to low-level atmospheric temperatures and SST in a coarse-resolution coupled GCM (7.5° × 4.5° atmosphere and 2.8° × 1.4° ocean) using the EAKF (Anderson, 2003) with a small ensemble size (16 members). Similarly to Tardif et al. (2014), they assimilated time-averaged observations, focusing on the low-level atmospheric variables. The procedure computed a cross-domain error covariance using monthly time averaging of daily atmospheric forecasts and instantaneous SST at the end of each month. There was minimal difference between the monthly time-averaged and instantaneous SST with this model configuration. Atmospheric DA was performed daily, assimilating temperature and wind observations, while ocean DA was performed monthly using a sequential update to the SST field that applied the following: (1) the cross-domain analysis increment due to the monthly averaged observed atmospheric surface temperature and (2) the ocean analysis increment due to the monthly averaged observed SST. Results showed reduction in coupled analysis error in midlatitudes when using a lagged averaging of the atmospheric observations versus assimilating the instantaneous atmospheric observations and also indicated some sensitivity to the duration of the lag period.

Sluka et al. (2016) used the intermediate complexity coupled SPEEDY/NEMO GCM at a higher resolution than the studies described above, both for the atmosphere at T30 and for the ocean using a 2° horizontal grid telescoping to 0.25° in the tropics. They applied CDA to assimilate atmospheric observations into the coupled system and compared the impacts of SCDA versus WCDA on the ocean state estimate. This DA system used the local ensemble transform Kalman filter (LETKF; Hunt et al., 2007) with an ensemble size of 40 members updated at a 6-hr analysis cycle to mirror a typical setup for operational numerical weather

prediction. Significant reductions in analysis error were found when using SCDA versus WCDA throughout the global ocean, with some exceptions in the high-latitude deep ocean. Further, the error reduction in the ocean led to improved accuracy in the surface atmospheric fields of temperature, humidity, and zonal wind. Experiments assimilating only ocean observations showed similar improvements across domain boundaries (Sluka, 2018). Sluka et al. (2016) did not show the impact of using SCDA when observations from both the atmosphere and ocean are assimilated.

Yoshida and Kalnay (2018) proposed an offline method to diagnose the uncertainty reduction of SCDA relative to WCDA. This *correlation-cutoff* method utilized the squared background ensemble correlation between the analyzed and the observed variables. Using the nine-variable coupled model of Peña and Kalnay (2004), the partly coupled LETKF analysis guided by the diagnosed correlation outperformed WCDA and the standard SCDA.

More recently, Wada et al. (2019) used a nonhydrostatic atmosphere model coupled with a wave model and a multilayer ocean model, assimilating the Advanced Microwave Scanning Radiometer 2 level 2 SST product with a SCDA-LETKF system. The system used a 6-hr analysis cycle to capture a diurnal cycle in the SST field. In a study of typhoons Kilo, Etau, and the September 2015 Kanto-Tohoku heavy rainfall, they found that compared to using an ocean-based analysis of SST as an external forcing field, using SCDA to assimilate SST data resulted in improved SST analyses compared to in situ moored buoys and an improved lower atmosphere as verified by radiosonde observations and radar-rain gauge rainfall analysis.

## 3. Methods

As indicated above, much of the prior work investigating SCDA has used simple pointwise or single-column models. While recent work has used more sophisticated coupled Earth system models, the cost of these models limits exploration of coupled dynamics and comparison of CDA approaches. Here we use a simple spatially extended three-layer model with truncated quasi-geostrophic (QG) dynamics, for which the multiscale dynamics have been thoroughly investigated (Vannitsem & Lucarini, 2016). The low computational cost allows a comprehensive evaluation of coupling strategies such as SCDA and WCDA, as well as a direct comparison between DA methods such as the EnKF, 4D-Var, 3D-Var, and hybrids.

### 3.1. Model Configuration

We use the MAOOAM three-layer (two-layer atmosphere and one-layer ocean) truncated QG model developed by De Cruz et al. (2016). Our experiments use the configuration examined by Vannitsem and Lucarini (2016), namely, the $2 \times 2$ atmosphere and $2 \times 4$ ocean (referring to the level of spectral truncation). With this configuration, the state vector is composed of $N = 36$ nondimensionalized coefficients of spectral modes, $N_a = 20$ ($N_o = 16$) of which represent the model's atmospheric (oceanic) state. Analysis and forecast errors presented in the results are measured by the Euclidean norm of these nondimensional coefficients. The model has coupled dynamics and coupled thermodynamics. The derivation of the tangent linear model (TLM) is detailed by De Cruz et al. (2016). The atmosphere component uses a zonally cyclic boundary condition, while the ocean uses a basin configuration. This coupled model configuration is designed to represent a midlatitude system. Formally, we describe the coupled dynamical system as

$$\dot{\mathbf{x}}^c = M(\mathbf{x}^c), \tag{1}$$

and decompose this system into its atmospheric and oceanic components as

$$\begin{cases} \dot{\mathbf{x}}_a^c = M_a\left(\mathbf{x}_a^c, \mathbf{x}_o^c\right) \\ \dot{\mathbf{x}}_o^c = M_o\left(\mathbf{x}_a^c, \mathbf{x}_o^c\right). \end{cases} \tag{2}$$

To emulate the uncoupled forced models typically used in operational numerical weather prediction and for Atmosphere Model Intercomparison Project (AMIP) and Ocean Model Intercomparison Project (OMIP) type climate studies, we decouple the above system into two independent dynamical systems by providing

each component with a sequence of prescribed state values $\{\widetilde{\mathbf{x}}_o^c(t)\}$ and $\{\widetilde{\mathbf{x}}_a^c(t)\}$. Thus, we now have the following pair of independent dynamical systems,

$$\begin{aligned}\dot{\mathbf{x}}_a &= M_a\left(\mathbf{x}_a, \widetilde{\mathbf{x}}_o^c\right)\\ \dot{\mathbf{x}}_o &= M_o\left(\widetilde{\mathbf{x}}_a^c, \mathbf{x}_o\right).\end{aligned} \tag{3}$$

For the perfect forcing case, $\widetilde{\mathbf{x}}_a^c$ and $\widetilde{\mathbf{x}}_o^c$ are determined using trajectories generated by the original coupled model. In practice, such forcing fields contain biases that can have a significant influence on the evolution of the forced system. Thus, we also consider a scenario representing errors in the forcing fields by updating the boundary data less frequently, every $t_b$ time units,

$$\widetilde{\mathbf{x}}_a^c(t) = \mathbf{x}_a^c(t_i), \text{and } \widetilde{\mathbf{x}}_o^c(t) = \mathbf{x}_o^c(t_i), \text{for all } t \in [t_i, t_i + t_b). \tag{4}$$

Following De Cruz et al. (2016), we use the second-order Runge-Kutta scheme to integrate the resulting differential equation. Because the model is used to evolve the state forward in time to an intermediate time step, the forcing data must also be used in this intermediate step. For example, for the uncoupled atmosphere,

$$\begin{cases} k_1 = M_a(\mathbf{x}_a(t), \widetilde{\mathbf{x}}_o(t))\\ k_2 = M_a(\mathbf{x}_a(t) + \Delta t k_1, \widetilde{\mathbf{x}}_o(t))\\ \mathbf{x}_a(t + \Delta t) = \mathbf{x}_a + \dfrac{\Delta t}{2}(k_1 + k_2). \end{cases} \tag{5}$$

The integration time step affects the numerical stability of the analysis-forecast cycle. We choose integration time steps either $dt = 0.01$ (as De Cruz et al., 2016), 0.1, or 0.5 nondimensional model time units (MTUs). The shorter time step facilitates varying the coupling frequency, which we use in the investigation of the Lyapunov exponents (LEs) computed from the coupled model. The longer time step allows faster integration, which is useful for ensemble experiments. De Cruz et al. (2018) adopted a time step of 0.2 time units for even higher-resolution versions of the model. Our experiments indicate that the general behavior of the model, for example, as characterized by the Lyapunov spectrum, does not significantly change within this parameter range. However, the model time step does have an impact on the stability of the DA methods. In such cases, we adopt shorter integration time steps when necessary in order to reduce the occurrence of numerical instabilities. The derivation of the TLM for the uncoupled model is presented in Appendix A.

### 3.2. Lyapunov Spectrum Analysis and Synchronization

When using an independent atmosphere or ocean model, even in the hypothetical that the model is perfect and that perfect forcing data are provided, missing bidirectional feedbacks will lead to nontrivial changes in the dynamics versus the coupled system. We use the mathematical theory of synchronization of chaos in order to better understand the error dynamics of the forced models and later to interpret the results of data assimilation applied to the forced and coupled models. Synchronization of chaotic systems is the phenomenon by which systems started at different initial conditions that intermittently exchange information converge to the same trajectory (e.g., Pecora & Carroll, 2015). While it is not possible to achieve perfect synchronization with noisy observations, obtaining conditions close to synchronization is possible and is critical for improving prediction skill. A key requirement for achieving synchronization is to constrain growing error modes, which has been discussed in the context of weather prediction by Pires et al. (1996).

The LE gives a measure of the asymptotic exponential growth rate of a dynamical system. Assuming that $\mathbf{M}$ is a linearization of the system (1), there exists a set of orthonormal vectors $\mathbf{v}_i$ so that the LE $\lambda_i$ characterizes the asymptotic growth rate corresponding to the vector $\mathbf{v}_i$,

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t} \ln\|\mathbf{M}(t, t_0)\mathbf{v}_i\|. \tag{6}$$

For a more detailed introduction in the context of weather forecasting, see Legras and Vautard (1996).

In a coupled driver/response system, we have one-way coupled synchronization if there is some function $G$ so that for the system,

$$\begin{cases} \dfrac{d\mathbf{y}}{dt} = F^{(n)}(\mathbf{y}) \\ \dfrac{d\mathbf{x}}{dt} = F^{(m)}(\mathbf{x}) + G(\mathbf{y}-\mathbf{x}) \end{cases} \tag{7}$$

we have that $\mathbf{x}(t) = \mathbf{y}(t)$ is a locally attractive invariant subspace. A necessary condition for the response system to synchronize with the driver is for the LEs of the response system (also known as the conditional LEs, or CLEs) to be negative (Pecora & Carroll, 1991).

We examine the ability to achieve synchronization between the true coupled driver system with multiple possible response systems, including forced versions of the uncoupled models, as well as versions of both the coupled and uncoupled models with data assimilation applied. Ruiz-Barradas et al. (2017) noted that DA can introduce coupled dynamics into forced atmosphere-ocean models. As described by Carrassi et al. (2008) and Penny (2017), the effect of the DA analysis update is a contraction,

$$\Phi^{+}(t_i, t_{i-1}) = [\mathbf{I}-\mathbf{K}_i\mathbf{H}_i]\Phi^{-}(t_i, t_{i-1}), \tag{8}$$

where $\Phi^{-}$ and $\Phi^{+}$ are the TLM before and after the effect of the DA contraction, $\mathbf{K}_i$ is the Kalman gain, and $\mathbf{H}_i$ is the linearized observation operator, at time $t_i$. The LEs of the DA system are the CLEs and must be controlled to be negative if synchronization is to be achieved between the DA system (response) and the observed signal from nature (driver).

### 3.3. Data Assimilation Methods

We compare various DA solution methods applied to the coupled and forced models using both WCDA and SCDA, including 3D-Var using the control variable transform (Bannister et al., 2008), strong-constraint 4D-Var (Courtier et al., 1994), the ETKF as formulated by Hunt et al. (2007), and hybrid gain DA (Penny, 2014; Hamrud et al., 2014; Bonavita et al., 2015; Houtekamer et al., 2018; Penny et al., 2015). For clarity, a summary of our implementation of each solution method is presented in Appendix B.

### 3.4. Estimating the Coupled Climatological Background Error Covariance

To estimate the climatological background error covariance matrix, we sample a history of ensemble forecast perturbations derived from the ETKF. We use a $k = 37$ member ensemble (one larger than the model dimension $m = 36$) with a fixed analysis cycle interval and full-coverage observing network. We estimate the corresponding background error covariance matrix based on the sample population of ensemble perturbations,

$$\mathbf{B}_0 = \left\langle \mathbf{X}_b \mathbf{X}_b^T \right\rangle. \tag{9}$$

We use a time series of ETKF analyses conducted over a total of $1 \times 10^5$ MTUs, with the first half discarded as spin-up. For the variational methods, a scalar multiplier $a$ is applied to tune the amplitude of the background error covariance matrix,

$$\mathbf{B} = a\mathbf{B}_0. \tag{10}$$

We further consider a blockwise inflation of the $\mathbf{B}$ matrix to separate the atmospheric and oceanic components. To maintain the original structure of the error correlations in the coupled system, we only inflate the variance of the atmosphere and ocean component corresponding to the diagonal elements of the atmosphere and ocean block in the $\mathbf{B}_0$. This procedure can be explicitly written as

$$\mathbf{B} = \mathbf{F}^{1/2}\mathbf{D}_0^{1/2}\mathbf{C}_0\mathbf{D}_0^{1/2}\mathbf{F}^{1/2}, \tag{11}$$

where $\mathbf{D}_0$ and $\mathbf{C}_0$ are the diagonal variance matrix and the symmetric correlation matrix derived from $\mathbf{B}_0$ and $\mathbf{F}$ is a diagonal matrix containing the inflation factor ($a_{atm}\mathbf{I}_{Natm \times 1}$, $a_{ocn}\mathbf{I}_{Nocn \times 1}$).

### 3.5. General Experiment Design

We generate a nature run that acts as truth by integrating the MAOOAM coupled model for $1 \times 10^5$ MTUs (~27.4 years) with the time step $dt = 0.5$ MTUs (about $5.6 \times 10^{-2}$ days or 1⅓ hr). Unless otherwise noted, the

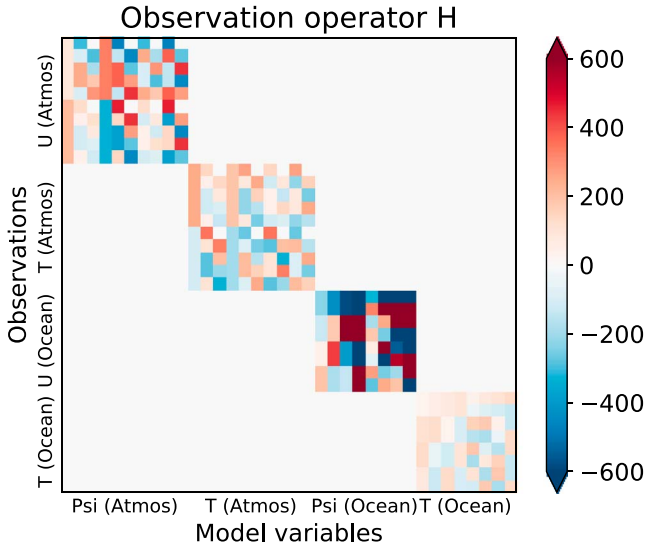


**Figure 1.** Example linear observation operator **H** (used for experiments with the hybrid coupled data assimilation) that converts nondimensional model variables from spectral space into *U* component of wind/current (m/s) and temperature anomaly (K).

analysis is performed every 2.5 MTU (~6 hr). For the 4D-Var, the assimilation window is 2.5 MTU, so that at each analysis cycle the observations over five model integration steps are assimilated. For the 3D-Var and ETKF, only the observations at the analysis step are assimilated. We note that for this configuration, at equal analysis intervals the 4D-Var assimilates 5× more observations than the 3D-Var or ETKF.

#### 3.5.1. Observation Network Configurations

We conduct a series of Observing System Simulation Experiments to assess the impact of different observing network designs on CDA. Observations are sampled from the nature run. We examine cases in which observations are sampled directly from the stream function (for simplicity), as well as observations sampled from the physical grid space. The latter requires the construction of a linearized observation operator **H** (Figure 1).

To simulate observation errors in the subsequent experiments, we choose Gaussian noise to have standard deviation equal to 10% of the climatological variability of the nature run, unless otherwise noted. This error is typically on the order of 1 m/s and 1 K for atmospheric velocity and temperature observations and 0.1 cm/s and 0.1 K for the ocean currents and temperature observations. The climatological variability of the ocean stream function is approximately 4 orders of magnitude smaller than the atmospheric stream function. As a point of reference, Tardif et al. (2014) used 10% of the climatological standard deviation in the atmosphere and approximately 35% in the ocean. In our experiments using a full-coverage observing network indicated that for the ETKF with a 40-member ensemble, the analysis performance degrades significantly as observational noise increases above 30% of the model climatological variability. We assume that the observational error statistics are known accurately for the DA experiments unless otherwise noted. Observations are only available once every analysis, at the end of each analysis window. We note that this design implies that fewer total observations are assimilated in experiments with a larger value of the assimilation window $\tau$.

Our three primary observing network scenarios assimilate observations in the atmosphere only, the ocean only, or both the atmosphere and ocean. Additional experiments begin to explore fixed versus nonstationary observing networks. The fixed observation network configuration for these experiments is shown in Figure 2. The nonstationary observing system evolves like the Brownian motion; namely, each observation station is randomly displaced by a small distance proportional to the time step size.

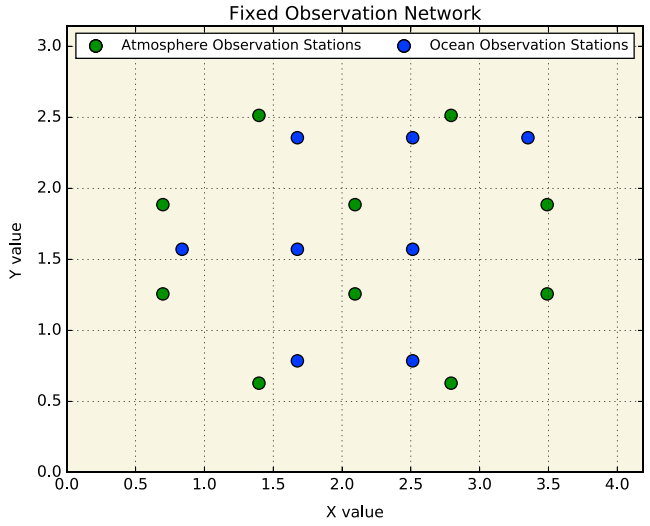


**Figure 2.** Fixed observation network used to assimilate observations in a physical space in comparison to the model spectral space. Each station measures temperature anomaly and stream function for atmosphere (green) and ocean (blue).

For experiments using the hybrid gain method, we use a specific fixed observing network configuration, with the ocean observations emulating a line of moorings. The *U*-wind and atmospheric temperature are observed at 10 collocated points, and the *U*-current and oceanic temperature are observed at 8 collocated points (Figure 3); there are 36 observations in total. The observation operator **H** for this configuration is full rank and well conditioned, with a condition number ~19. We apply independent Gaussian error with standard deviation of 1 m/s (for the wind/current fields) and 1 K (for the temperature fields).

## 4. Results

### 4.1. Lyapunov Spectrum of the Forced Versus Coupled Models

We first compare the Lyapunov spectrum of the coupled model to the uncoupled atmosphere and ocean models under prescribed time-dependent forcing at the interface. We compute the backward LEs (Figure 4) using QR decomposition (see Geist et al., 1990). We verify that as noted by Vannitsem and Lucarini (2016), the fast-decaying and fast-

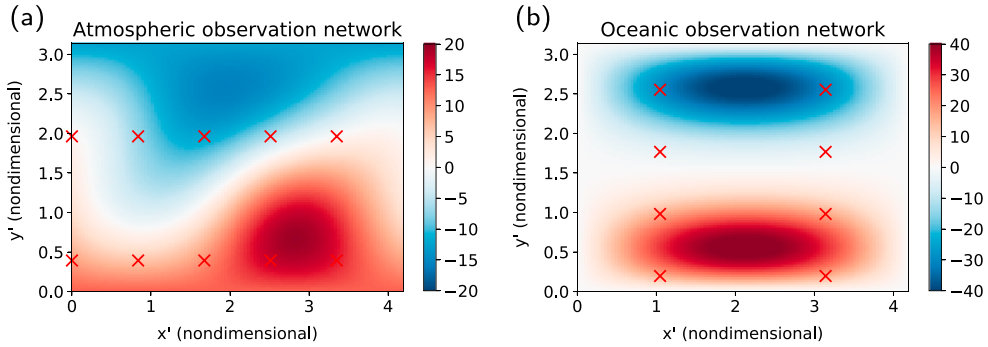


**Figure 3.** Fixed observation network (red crosses) in the physical grid space used for hybrid DA experiments for (a) the atmosphere and (b) the ocean. At each observation location, both the *U* component of the wind/current and the temperature anomaly are measured. The observation positions are chosen asymmetrically so that the resulting observations are linearly independent. The background shadings show an instantaneous anomaly from the spatial mean temperature (K).

growing modes are primarily of atmospheric origin. The decoupling appears to have negligible impact on the atmospheric LEs. However, the near-zero LEs, identified as being of oceanic or coupled origin by Vannitsem and Lucarini (2016), are significantly modified by the decoupling. Vannitsem et al. (2015) demonstrated that low-frequency climate variability in MAOOAM is partially attributable to coupled atmosphere-ocean modes. Goodman and Marshall (2003) also described the importance of neutral singular vectors in coupled atmosphere-ocean dynamics; the atmospheric singular vectors with singular values nearest to zero corresponded to the modes that responded most strongly to changes in surface forcing.

Comparing the near-zero LEs of the coupled model to those of the forced ocean, there is a clear difference in the characteristics of the two spectra (Figure 5). The forced ocean displays a *stepwise* structure with many LEs almost identical, in contrast to a smoother spectrum in the coupled model. Also, the fluctuations of the finite-time LEs are reduced in the forced model. To increase confidence that this difference does not arise solely due to finite-time fluctuations, we compute confidence intervals around each LE calculated with the stationary bootstrap method (a modification of bootstrap sampling for time series data; see Politis & Romano, 1994) using the ARCH Toolbox (Sheppard, 2015). The resampling block sizes are determined using the method of Patton et al. (2009).

We cannot definitively determine through numerical experiment whether the near-zero LEs are positive, negative, or zero for the MAOOAM dynamics. Vannitsem and Lucarini (2016) previously noted the same difficulty with respect to this MAOOAM configuration. Even after long integrations, the confidence intervals for these LEs straddle zero due to the presence of modes with very long timescales. The coupled model has 3 positive, 8 near-zero (of undetermined sign), and 25 negative LEs; the forced atmosphere has 3 positive and 17 negative LEs; and the forced ocean has 8 near-zero and 8 negative LEs. Of the eight near-zero LEs for the coupled model, the first four are numerically positive with the first LE on the scale of $10^{-4}$ and the

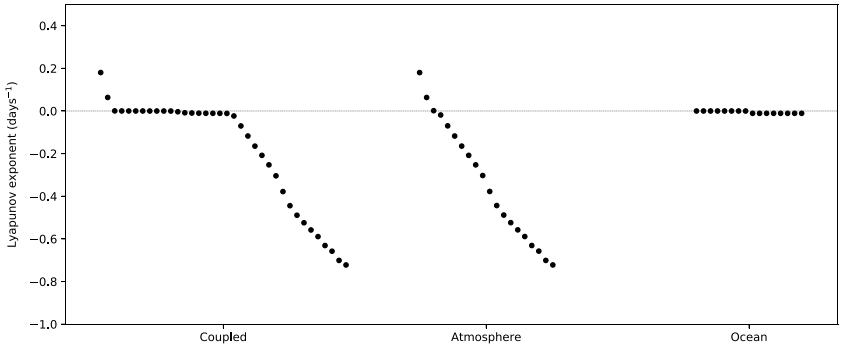


**Figure 4.** The Lyapunov spectrum for the coupled model and forced atmosphere and ocean.

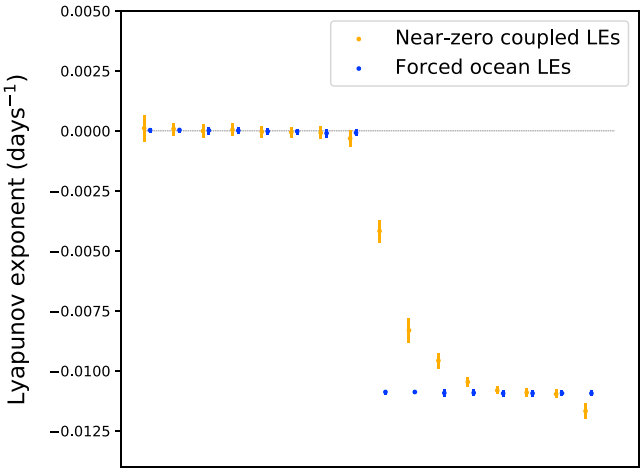


**Figure 5.** The Lyapunov exponents (LEs) nearest to zero for the coupled model along with the LEs for the forced ocean with perfect forcing. The error bars are the 95% confidence intervals. For the forced ocean, the first eight CLEs are between about $10^{-4}$ and $-10^{-4}$.

second to fourth on the scale of $10^{-5}$. The fifth to eighth LEs are numerically negative, with the fifth and sixth on the scale $-10^{-5}$, and the seventh and eighth on the scale $-10^{-4}$. Of the near-zero LEs for the forced ocean, the first three are numerically positive on the scale of $10^{-5}$, the fourth and fifth LEs vary between $10^{-6}$ and $-10^{-6}$, and the sixth to eighth are negative on the scale of $-10^{-5}$.

While it appears that the forced oceanic MAOOAM component differs more from its coupled counterpart than the forced atmospheric MAOOAM component does from its coupled counterpart, these relationships may change if additional processes are taken into account, such as the hydrological cycle. Although Ferrari and Cessi (2003) showed synchronization of a forced atmosphere with oceanic forcing from a coupled run with a simpler ocean-atmosphere model, we see that as indicated by the positive LEs, the forced atmosphere does not synchronize to the trajectory of the coupled one. We find that changing the forcing interval $t_b$ does not noticeably impact the Lyapunov spectrum of the forced atmosphere and ocean models. However, a forcing interval $t_b$ does reduce the degree of synchronization between the trajectories of the forced and coupled models. To illustrate this point, Figure 6 shows the relative error using a model integration time step of $dt = 0.01$ MTU with forcing intervals $t_b = 0.01$, $t_b = 100$, $t_b = 1000$, and $t_b$ = infinite (i.e., constant forcing), as well as the corresponding Lyapunov spectrum for each case. Even when using a forcing interval $t_b$ that is equal to the model integration time step $dt$, and providing the atmosphere or ocean model with perfect forcing, synchronization is not achieved for either case.

### 4.2. Transitioning From a Forced to Coupled Forecast Model

#### 4.2.1. Data Assimilation With Forced Atmosphere and Ocean Models

To compare DA applied to the coupled versus forced models, we estimate the minimum ensemble size required for the ETKF to prevent filter divergence. We compare the forced atmosphere model to the coupled atmosphere model, using an observing network restricted to the atmosphere only (Table 1). We define divergence as the condition that the root-mean-square error reaches the same order of magnitude as a free model run starting from perturbed initial conditions. For the forced atmosphere model, the convergence for a given analysis cycle window $\tau$ depends largely on the forcing interval $t_b$. Using an analysis cycle window shorter than the forcing interval causes the ETKF to diverge. When the data assimilation window is of the same length as the forcing interval, the minimum ensemble sizes are identically the same (i.e., 7), which is greater than the number of unstable-neutral modes (i.e., 3, shown in Figure 4).

When using a coupled forecast model with high-frequency coupling ($t_b = 0.01$), the ETKF converges for the data assimilation windows $\tau = 0.1$, 1, and 10. The minimum ensemble sizes for convergence are smaller (i.e.,

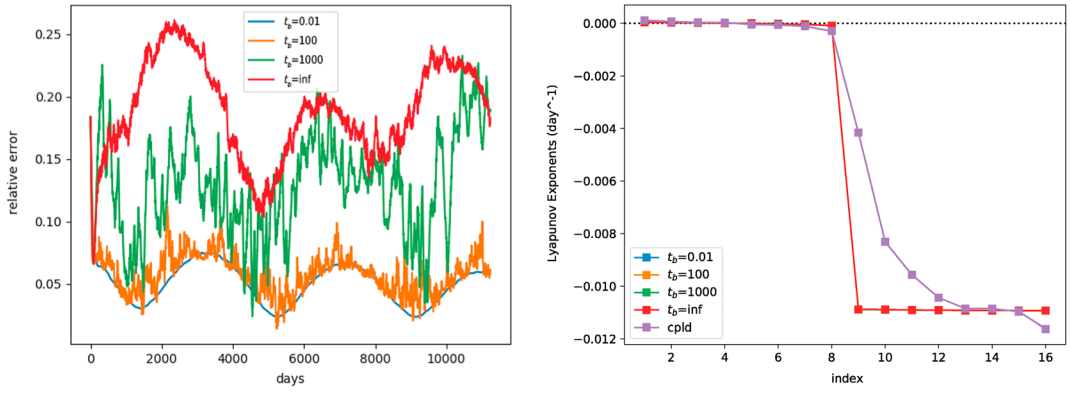


**Figure 6.** (left) The relative error between the forced and coupled trajectory and (right) CLEs of forced ocean model with different forcing window lengths $t_b$.

**Table 1**
*The Minimum Ensemble Size Required to Prevent Filter Divergence for Various Scenarios (Without Inflation) Within the Time $9 \times 10^4$ MTU*

| Experiment type | Forcing/coupling interval | Analysis cycle window (MTU) $\tau = 0.1$ (~16 min) | $\tau = 1.0$ (~2.6 hr) | $\tau = 10.0$ (~1.1 days) |
|---|---|---|---|---|
| **Forced atmosphere** | $t_b = 0.01$ (~1.6 min) | 7 | 7 | 11 |
| | $t_b = 0.1$ (~16 min) | 7 | 7 | 9 |
| | $t_b = 1$ (~2.6 hr) | Diverges with 21 | 7 | 11 |
| | $t_b = 10$ (~1.1 days) | Diverges with 21 | Diverges with 21 | 7 |
| **Coupled system, observed atmosphere** | $t_b = 0.01$ (~1.6 min) | 12 | 9 | 9 |

*Note*. The quantity $t_b$ indicates the forcing frequency, and $\tau$ is the assimilation window.

9–12) compared to the model dimension—roughly 25 to 33% of the coupled model dimension (36) as compared to 35 to 55% of the model dimension (20) for the converging cases of the forced atmosphere system. When using a coupled forecast model the ETKF appears more robust to longer assimilation windows.

#### 4.2.2. Estimating the Coupled Climatological Background Error Covariance Matrix

Before examining SCDA, we must calculate a coupled background error covariance matrix **B** for the variational and hybrid methods. To do this, we use a 40-member ETKF with a fixed analysis cycle window to build the statistics for this matrix. We note that there is larger variability in the short-range forecast errors in the atmospheric component than the ocean. The magnitude of the atmospheric component of **B** is significantly larger than the oceanic component or the cross-domain ocean-atmosphere blocks (Figure 7). The matrix **B** and its corresponding Hessian matrix **S** consequently have large condition numbers on the order of $10^{16}$. After applying the control variable transform (see Appendix B), the condition number of the Hessian matrix $\mathbf{S}_{\mathrm{CVT}}$ with respect to the transformed control variable is 1.0003. Using a forecast length of about 6 hr, the off-diagonal elements of the block matrix (that are by definition zero for WCDA) present nonzero error correlations with magnitudes similar to the off diagonals in the strictly atmospheric or oceanic blocks. These nonzero error correlations indicate that in a CDA context an analysis derived from observing the atmosphere state can meaningfully inform the ocean state, and vice versa.

The character of the **B** matrix is dependent on the timescale of the forecast errors used in its construction. Figure 8 shows the static background error covariance matrices obtained by equation (B12). The left column shows the normalized **B** matrices using a logarithmic color scale, and the center and right panels respectively show eigenvalues and eigenvectors of $\mathbf{B}_0$, arranged in descending order. For $\tau = 1$ MTU in Figure 8, the background error covariance is distributed relatively evenly across every model variable, with the largest variance occurring in ocean temperature (model variables 29–36). The leading eigenmodes (eigenvectors with indices 1–7) at this timescale are also mostly dominated by ocean temperature (Figure 8c). The atmospheric variables exhibit the second strongest error variance, while the ocean stream function is generally weakest. However, because the error variance depends on the error norm (here the Euclidean norm of the nondimensionalized coefficients of each mode), we do not conclude that one component is more uncertain than the others. Instead, we examine the relative amplitude changes with different analysis interval $\tau$.

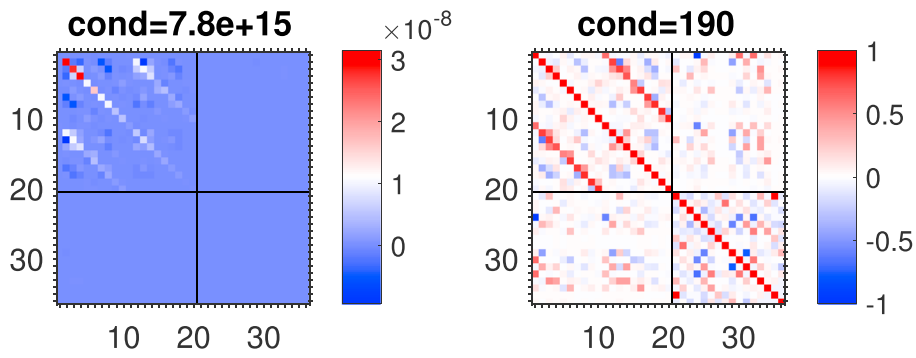


**Figure 7.** The climatological background covariance matrix **B** (left) and its corresponding correlation matrix **C** (right). The atmospheric (1–20) and oceanic (21–36) state vectors are separated by the black lines.

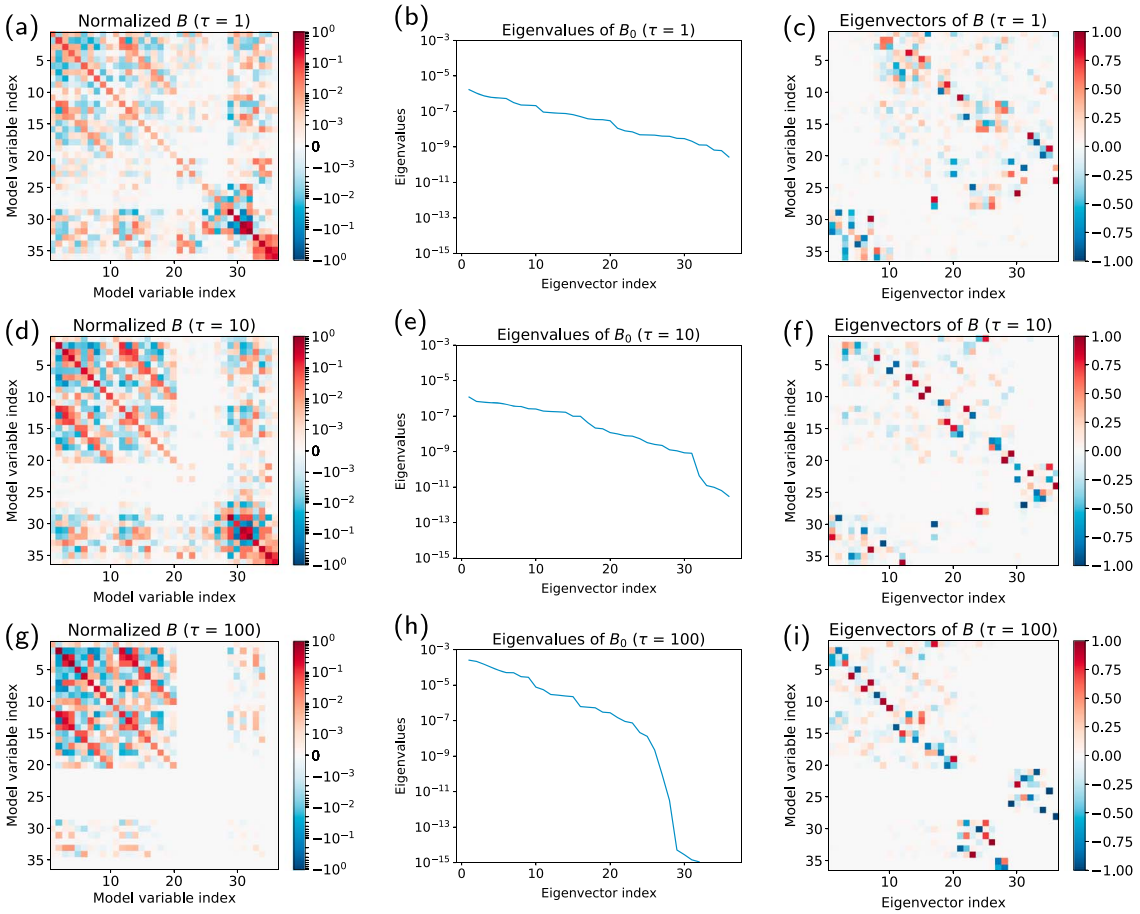


**Figure 8.** Static climatological background error covariance matrices obtained for various forecast times. Each row shows different analysis interval $\tau = 1$, 10, and 100 MTU. The columns show (left) the normalized background error covariance matrices (so that the spectral radius of **B** is unity), (center) the eigenvalues of **B**, and (right) the eigenvectors of **B**.

When we increase the forecast interval, there is a large change in the error covariance structure (panels d and g). As the forecast interval increases to $\tau = 10$ MTU, more error variance is projected toward the atmospheric variables (Figure 8f). This behavior is most apparent in the eigenvectors with $\tau = 100$ (Figure 8i), in which the leading 20 eigenvectors are dominated by atmospheric variables, and the ocean temperature (eigenvectors 21–28) and stream function (eigenvectors 29–36) are clearly separated.

We may further increase the analysis interval $\tau$ so that the atmospheric part of the error is saturated, which is similar to the selection of slower modes by the breeding method (Toth & Kalnay, 1993). However, unlike the breeding method, the EnKF assumes quasi-linearity between the background errors of the analyzed and the observable variables. Hence, the saturated and nonlinear errors often cause filter divergence with larger values of $\tau$ (not shown). Comparing panels (a), (d), and (g), we find that with longer forecast intervals, the ensemble perturbations are more strongly concentrated onto the fastest-growing error modes.

#### 4.2.3. Comparing Degrees of Coupling With 3D-Var

We first compare the performance of 3D-Var under different coupling scenarios using a full observing network. For the uncoupled assimilation experiments, the forward model is driven by three kinds of surface forcing: perfect forcing updated every $t_b = 10$ MTU (~1 day), perfect forcing updated every 30 MTU (~3 days), and constant forcing using the climatological mean. Varying the forcing interval allows us to examine the impact of increasing error in the flux fields. The analysis root-mean-square error (RMSE) for the last $10^3$ MTU ($2 \times 10^3$ steps, ~100 days) out of $10^5$ MTU is shown in Figure 9. All experiments produce similar RMSE for the atmosphere, except the uncoupled atmosphere forced by the climatological mean ocean,

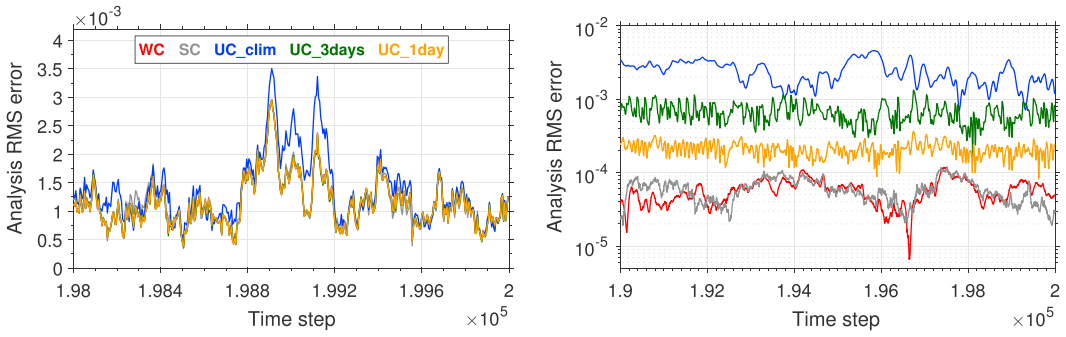


**Figure 9.** The analysis RMSE of the atmosphere (left) and ocean (right) for the strongly coupled (gray), weakly coupled (red), and uncoupled 3D-Var with forcing provided by the climatological mean (blue), or updated forcing from truth every 30 MTU (~3 days, green) and every 10 MTU (~1 day, orange). Results are shown for the last 1,000 MTU for the atmosphere and last 5,000 MTU for the ocean. The WC RMSE for the atmosphere is nearly identical to the UC_1day RMSE.

which exhibits slightly larger RMSE. For the forced ocean model scenarios, improved accuracy can be obtained for the ocean analyses with more frequent updating of the atmosphere forcing.

In general, the RMSE of the CDA systems is smaller than the RMSE of the forced systems. We aggregate the general performance of all coupling scenarios by averaging the analysis RMSE for the last $4 \times 10^4$ MTU (~11.0 years, Table 2). For both atmosphere and ocean, the SCDA 3D-Var achieves the minimum RMSE, which is slightly improved versus the RMSE for the WCDA 3D-Var.

An additional benefit of applying CDA with 3D-Var is that the combined atmosphere-ocean system experiences a more rapid spin-up compared to the uncoupled scenarios. Approximately 300 MTU (~30 days) are required for the uncoupled assimilation system to stabilize the atmosphere RMSE, and about $6 \times 10^4$ MTU (~16.4 years) are required for the uncoupled ocean assimilation system, even when the ocean model is forced daily with true atmospheric forcing. The CDA systems require only $2.5 \times 10^4$ MTU (~6.8 years) to achieve a comparable spin-up for both the atmosphere and ocean, with the SCDA RMSE slightly lower than the WCDA RMSE during the spin-up period.

#### 4.2.4. Comparing Degrees of Coupling With the ETKF

We next investigate the ETKF under different coupling scenarios. We use an ensemble size of 40, which is larger than the degrees of freedom of the coupled model. For the forced model cases, perfect forcing is updated every $t_b = 2.5$, 10, or 30 MTU (or about 6 hr, 1 day, and 3 days). The forced atmosphere ETKF is sensitive to the frequency of the forcing update. Unlike 3D-Var, the ETKF exhibits filter divergence with forcing intervals of 1 day or greater (Figure 10). The ETKF converges using both SCDA and WCDA, though SCDA consistently produces smaller RMSE in the oceanic component than WCDA. Such consistent RMSE difference between WCDA and SCDA was not present when using 3D-Var.

The forced atmosphere ETKF diverges if the forcing from the ocean is updated less frequently than 10 MTU (~1 day). For example, when examining the Lyapunov spectrum for a similar case, the forced model with $t_b =$ 12.5 MTU (~1.4 days) still has a leading CLE larger than zero (Figure 11), indicating filter divergence. The

**Table 2**
*Average Analysis RMSE (Units: $10^{-5}$) of 3D-Var Over the Last $4 \times 10^4$ MTU (~11.0 Years) and of the ETKF After 2,500 MTU (for the Remaining ~26.7 Years)*

| | 3D-Var | | | ETKF | | |
|---|---|---|---|---|---|---|
| | Uncoupled | WCDA | SCDA | Uncoupled | WCDA | SCDA |
| Atmosphere | 116.0 | 116.0 | **115.9** | 41.37 | 42.10 | **40.72** |
| Ocean | 28.72 | 5.516 | **4.915** | 4.510 | 2.784 | **1.318** |

*Note*. The forcing for the uncoupled cases is updated every 10 MTU (~1 day) for 3D-Var and every 2.5 MTU (~6 hr) for the ETKF. The SCDA approach has the smallest RMSE for both DA methods (marked in bold).

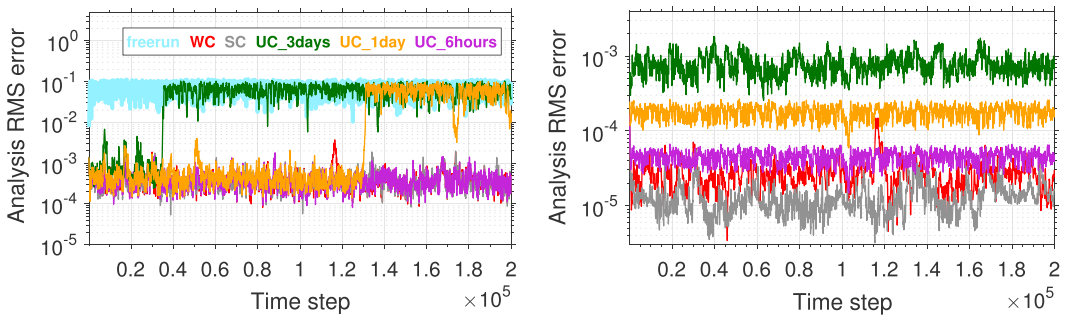


**Figure 10.** The analysis RMSE of the atmosphere (left) and ocean (right) for the weakly coupled (red), strongly coupled (gray), and uncoupled ETKF with updated forcing from truth every 30 MTU (~3 days, green), 10 MTU (~1 day, orange), and 2.5 MTU (~6 hr, purple). A 141-step moving average filter is applied. For the atmosphere, RMSE for the WC and SC cases follows closely with the UC_6hours case but exhibits lower RMSE in the ocean.

CDA ETKF systems generate atmospheric analyses with RMSE similar to the uncoupled ETKF with frequently updated (~6 hr) perfect forcing. We note that inflating the atmospheric background error covariance or perturbing the oceanic forcing at each analysis cycle helped to reduce the occurrence of divergence caused by errors in the surface forcing supplied to the forced atmosphere.

Smith et al. (2015) found that for a 1-D simplified coupled column model using incremental 4D-Var, WCDA was usually comparable to uncoupled assimilation experiments in which the atmosphere and ocean models were forced using accurate surface fluxes. Here we find a similar result and add that the stability is degraded in the forced system only when the forcing accuracy is reduced (e.g., by extending the forcing time window).

The forced ocean ETKF provides lower accuracy analyses than the CDA ETKF even when the forcing is updated with perfect atmospheric conditions as frequently as ~6 hr. Unlike for 3D-Var where the ocean RMSE from SCDA is similar to that from the WCDA after the spin-up period, for the ETKF the SCDA consistently produces higher-accuracy ocean analyses with RMSE about half of the magnitude of the WCDA ETKF even after the spin-up period. Also, the SCDA ETKF requires about 1/10th of the spin-up time compared to the SCDA 3D-Var.

### 4.3. Evaluation of SCDA

#### 4.3.1. Comparing SCDA Methods With a Fully Observed Coupled System

Having explored the transition from a forced to fully coupled forecast model using various CDA strategies, we now focus on comparing various DA methods strictly using SCDA. We first focus on scenarios with a full observing network (Figure 12). We use a 40-member ETKF without inflation and a 20-member ETKF with

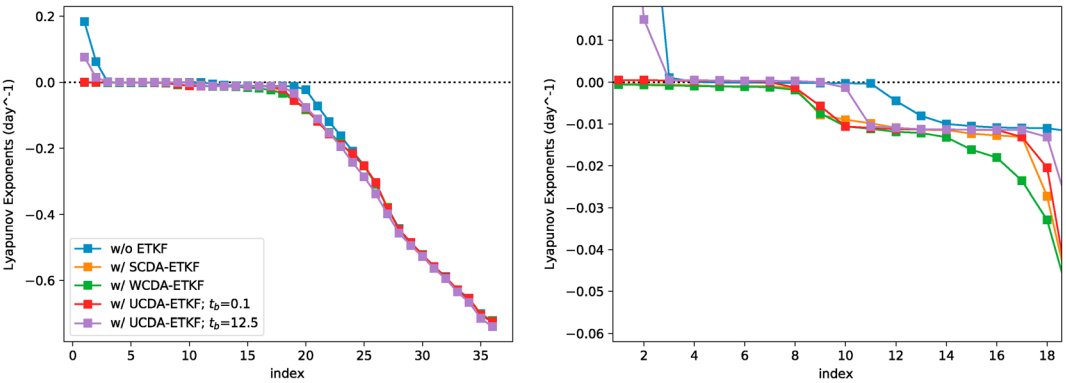


**Figure 11.** Lyapunov spectrum (left) and a magnification of the near-zero LEs (right) for SCDA ETKF, WCDA ETKF, and forced atmosphere and ocean ETKF (UCDA) using forcing intervals of $t_b = 0.1$ (~16 min) and $t_b = 12.5$ (~1.4 days). Each is implemented with an ensemble size of 37, an analysis cycle window of 2.5 MTU, and no inflation.

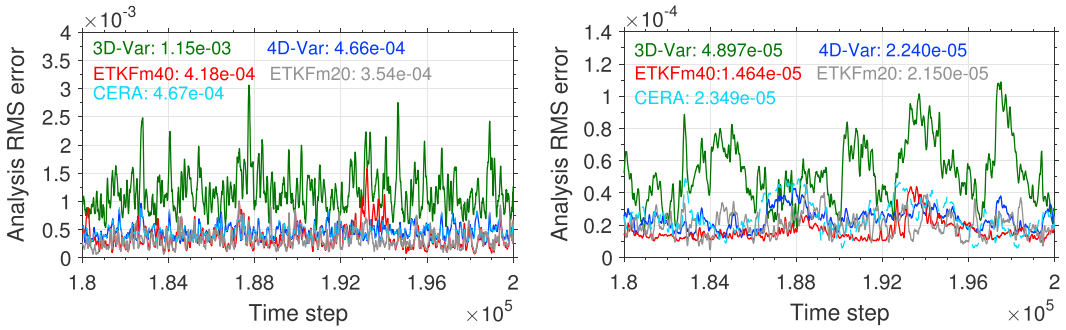


**Figure 12.** The analysis RMSE when using a full-coverage observing system for the atmosphere (left) and ocean (right) with SCDA 3D-Var (green), 4D-Var (blue), 4D-Var/3DFGAT CERA (cyan dash), 40-member ETKF (red), and 20-member ETKF with constant inflation of 1.01 (gray). A 61-MTU running average is applied for both panels. The analysis RMSE for the last $5 \times 10^4$ MTU (~13.7 years) is shown next to the label for each method. For the atmosphere, the RMSEs for 4D-Var and CERA are largely overlapped.

inflation of 1.01. To tune the 3D-Var and 4D-Var, we multiply the normalized background error covariance matrix with a constant factor of 10 and 2, respectively.

For all DA methods, the analysis RMSE for the atmosphere is 1 order of magnitude larger than the RMSE for the ocean. The 3D-Var produces the largest RMSE among all three DA methods. While the atmospheric RMSE is comparable between 4D-Var and ETKF, the RMSE of the ETKF for the ocean component is generally smaller than 4D-Var when the analysis windows is 6 hr. The performance of 4D-Var can be improved by increasing the length of assimilation windows (Smith et al., 2015; Yang et al., 2009), though a larger window was not tested here. The full coupled-state RMSE for the 20-member ETKF is comparable to the RMSE of the 4D-Var and 40-member ETKF. An assimilation approach similar to the CERA configuration (Laloyaux et al., 2016), which integrates the coupled model in two outer loops and uses an incremental 4D-Var analysis for the atmosphere and a 3D-Var using FGAT (First Guess at the Appropriate Time) analysis for the ocean, has accuracy similar to the SCDA 4D-Var.

#### 4.3.2. Comparing SCDA Methods With a Partially Observed Coupled System

Following Sluka et al. (2016), we next consider a reduced observing network by evaluating the performance of SCDA while assimilating only atmospheric observations. The ETKF and 4D-Var produce similar RMSE for the atmosphere, both smaller than the 3D-Var (Figure 13a). However, due to the use of a climatological error covariance, neither 3D-Var nor 4D-Var stabilize the ocean within the length of the experiments (~27.4 years), while after about $4 \times 10^4$ MTU (~10 years) the EKTF does stabilize (Figure 13b).

Next, we expand the investigation to observing networks that cover the atmosphere only, the ocean only, or the full coupled system. The model is run with time step $dt = 0.1$ for $10^4$ MTU (~1,200 days) after a transient spin-up run of $10^7$ MTU. For simplicity, we set the analysis interval and observing interval to be equal. The

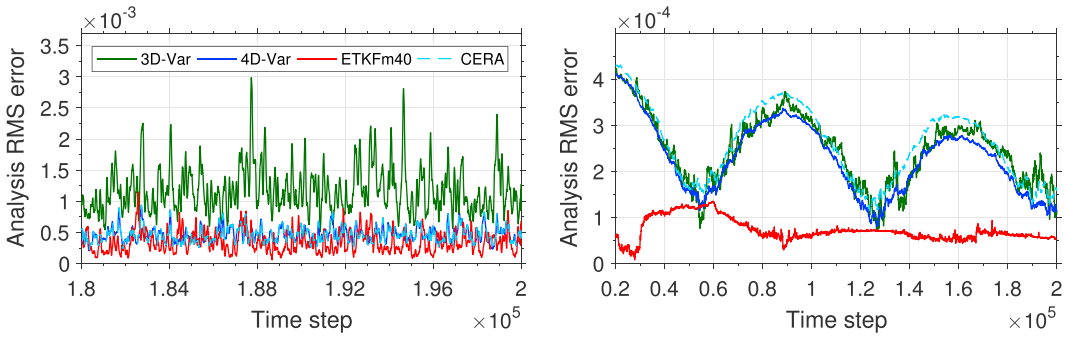


**Figure 13.** The analysis RMSE when observing only the atmosphere for the atmosphere (left) and ocean (right) with strongly coupled 3D-Var (green), 4D-Var (blue), 4D-Var/3D-Var CERA (cyan), and 40-member ETKF (red). A 21-time-step running average filter is applied to the left panel.

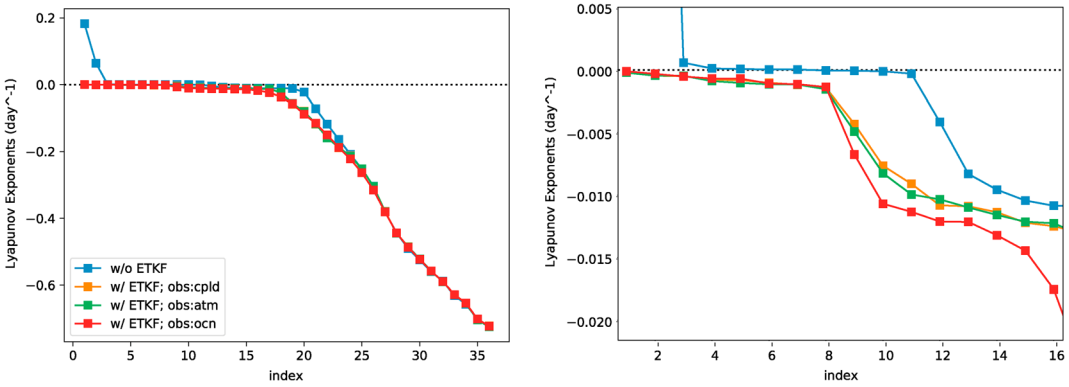


**Figure 14.** The Lyapunov exponents of the coupled model with and without the 37-member ETKF DA, with analysis/observing window $\tau = 0.1$, observing the atmosphere, the ocean, or both domains (left). Magnified view of the same (right).

analysis/observing frequency is varied as $\tau = 16$ min, 1 hr, 2 hr, and 3 hr. The number of ensemble members is varied between 6 and 37.

We attempt to calculate the minimum ensemble size required to prevent filter divergence for each case, using noiseless observations to avoid the introduction of additional sources of nonlinearity to the system. We also fix the multiplicative inflation to $\rho = 1.0$. The minimum ensemble size required to maintain stability in the fully coupled system is generally 16 members when assimilating observations from the full network or the atmosphere only, whether those observations are of the spectral modes or in the physical space. There

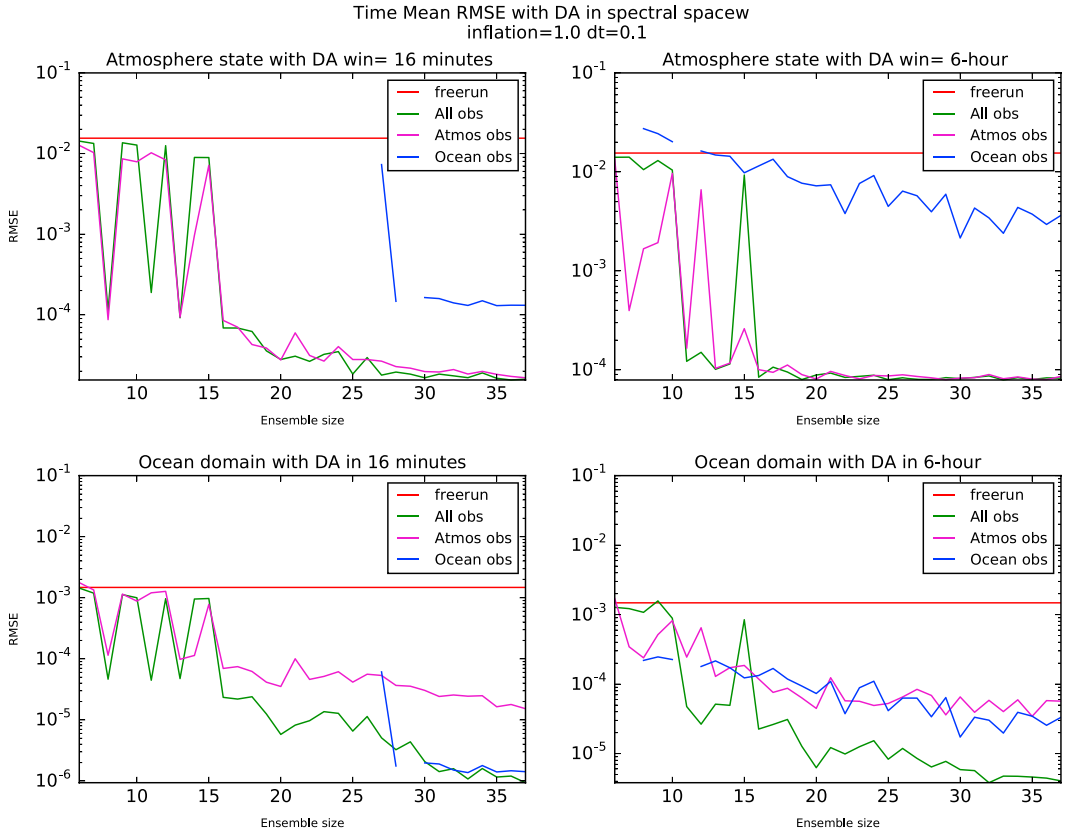


**Figure 15.** Time-mean RMSE for (top) atmosphere and (bottom) ocean, for (left) 16-min and (right) 6-hr observing/analysis frequency. Results are shown for the free run (red), DA using $R = 10\%$ climatology with noisy observations for all observations (green), only atmosphere observations (magenta), and only ocean observations (blue). The analysis window $\tau = 16$ min, model time step $dt = 0.1$ MTU, and inflation $\rho = 1.0$.

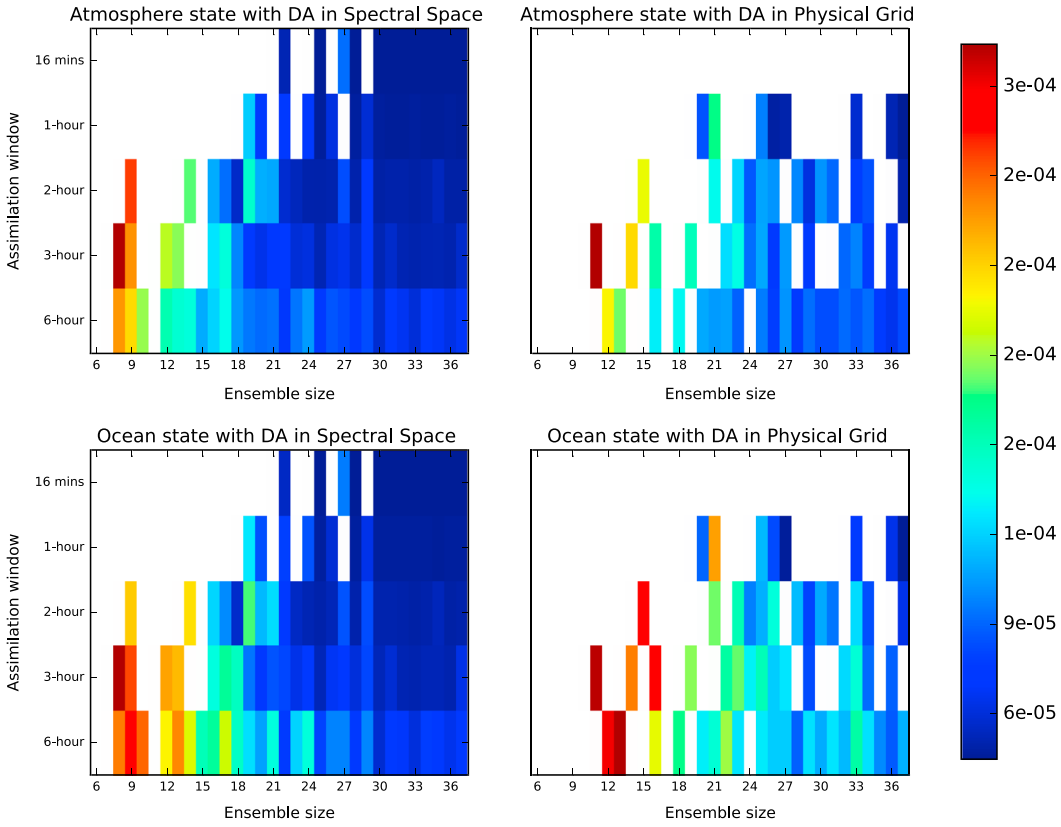


**Figure 16.** Time-mean (1,200 days) RMSE for atmosphere (top) and ocean (bottom) when ocean-only observations are made in the spectral model space (left) or in the physical grid space (right). The observation network is fixed as in Figure 2. White color indicates that experiment encountered numerical overflow. For reference, the free run RMSE is (atm: $1.55 \times 10^{-2}$, ocn: $1.4 \times 10^{-3}$).

are some exceptions. While the ocean state shows asymptotic stability with 16 or more ensemble members, the error grows larger than the free run during the initial few cycles. The SCDA becomes very sensitive when observations are sampled from the ocean only, in which case the minimum ensemble size required to prevent filter divergence in the atmosphere is between 22 and 36 members with no convergence for large analysis cycle windows and between 18 and 30 members for the ocean. The Lyapunov spectrum indicates that the ETKF with 37 members and an analysis cycle of $\tau = 0.1$ MTU stabilizes the DA system (Figure 14).

Next, we return to using an observation error covariance specified at 10% of climatological variability and vary the multiplicative inflation factor as $\rho = 1.0$, 1.05, and 1.10. For the atmosphere, the atmosphere-only observing network and full observing network outperform the ocean-only observing network and have similar RMSEs for stable ensemble sizes (Figure 15). For high-frequency analysis updates (e.g., 16 min) and large ensemble sizes (>27), assimilating only ocean observations, if nondivergent, generates ocean analyses that are almost as accurate as when assimilating a full observing network. As the assimilation frequency decreases, accuracy from assimilating only ocean observations decreases relative to assimilating all observations. For lower frequency analyses with large ensemble sizes, assimilating all observations produces ocean analyses with significantly lower errors than assimilating only ocean observations.

The ETKF has greater sensitivity in the ocean-only observing experiments (Figure 16). Larger ensemble sizes are required to maintain stability compared to the atmosphere-only observing network. The ETKF is generally more stable when sampling observations directly in the spectral model space than observations sampled in the physical grid space, indicating that the latter may require more observations or lower observational noise to achieve similar performance. When the ensemble size is large enough to ensure that the filter is stable, RMSE is reduced by increasing the analysis/observing frequency. That the results due to sampling observations from the native spectral model space were more difficult to replicate when sampling from

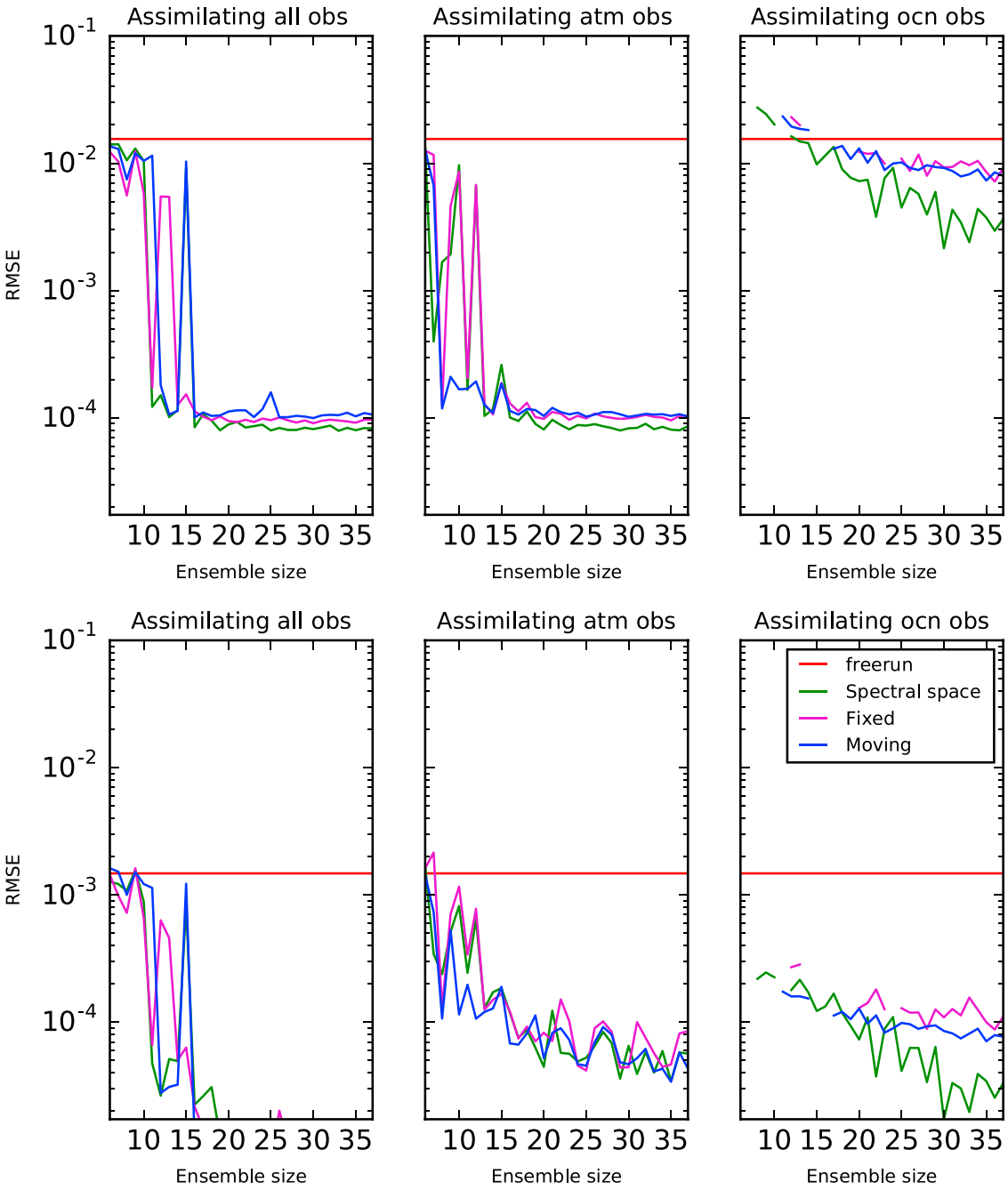


**Figure 17.** Comparing time-mean RMSE for the atmosphere (top) and ocean (bottom) using the ETKF with a dynamic observation network versus a fixed observation network, with analysis/observing window $\tau = 6$ hr, inflation $\rho = 1.0$, and model time step $dt = 0.1$. A reference case is also shown that samples from the native model spectral space.

the physical grid space suggests that new multiscale DA methods may be required to adequately analyze point-sampled ocean observations in a realistic CDA system.

For small ensemble sizes, applying a 5% ($\rho = 1.05$) multiplicative inflation lowers the minimum ensemble size required to prevent divergence. However, as the inflation parameter increases, the RMSE also increases. For the fully observed case, when using multiplicative inflation, the RMSE increases as the ensemble size increases, similarly to what was shown for the Lorenz (1996) model by Penny (2014). When observations are sampled from the opposite model component, the ETKF with inflation diverges to higher RMSE than

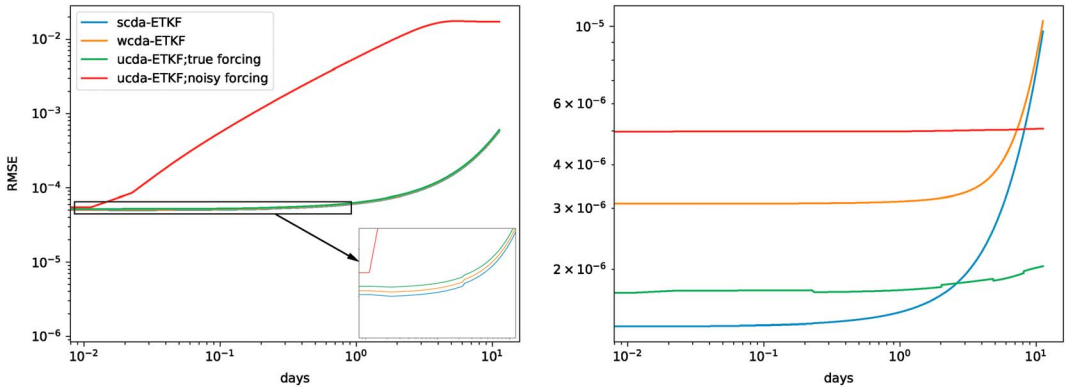


**Figure 18.** The RMSE of forecasts at various lead times in the (a) atmosphere and (b) ocean components for the ETKF using SCDA, WCDA, and the forced ETKF with perfect and noisy forcing. The analysis window is $\tau = 2.5$ MTU (6 hr), and the forcing window of uncoupled ETKF is $t_b = 0.1$ MTU (16 min).

the free run. This indicates that the use of traditional inflation methods may need to be revisited, particularly for SCDA applications. Grudzien et al. (2018a, 2018b) indicate the importance of the choice of inflation method, and a thorough investigation of inflation methods within the context of CDA is a necessary area for future study.

Next we assess the impact of using a fixed versus dynamic observing network, with observations sampled from physical grid space. For reference, we also show the analysis RMSE with observations sampled from the native model spectral space. For large ensemble sizes, when observing either the whole system or the atmosphere only, the ocean analysis is most accurate using a dynamic observing network, while the atmosphere analysis is generally more accurate when using a fixed observing network (Figure 17). When only the ocean is observed, the dynamic observing system generally produces more accurate analyses for both the ocean and atmosphere. The minimum ensemble size required for convergence generally remains unaffected by the type of observations except when the analysis/observing window is 6 hr or greater. For example, when assimilating a dynamic observing system using a 6-hr analysis/observing window, the minimum ensemble size rises from 10 to 16 members.

### 4.4. Forecasts at Various Lead Times

Penny et al. (2017) indicated "For the purpose of seamless prediction, the research community must demonstrate that CDA not only improves initial states but also improves forecasts at all time scales." In this section, we compare the accuracy of forecasts at multiple lead times using DA with different degrees of coupling. For these comparisons we use initial conditions produced from 36,000 DA cycles of a 37-member ETKF using SCDA, WCDA, or uncoupled DA. At each cycle we initialize forecasts with lead times ranging from 0 to 10 days, compute the RMSE, and aggregate the results. For the uncoupled DA, we use a forced model with either perfect forcing or forcing degraded equal to the perfect forcing plus a white noise of 10% of climatological variability. The SCDA forecasts have the highest accuracy, followed by the WCDA, then the uncoupled DA with perfect forcing. For the oceanic states (Figure 18), the forecasts produced by SCDA outperform all the other couplings for the first 2 days of lead time. For both the SCDA and WCDA, the error grows quickly after day 3 to reach almost the same level by day 10. The ocean forecast errors for the uncoupled DA using perfect forcing are generally stable at all lead times, while at shorter lead times the forecasts generated by degraded forcing result in a relatively larger RMSE than all other methods. A more realistic scenario for degraded forcing might allow forcing error to grow with time until it reaches saturation, for example, by relaxing the forcing field to climatology.

### 4.5. Hybrid Gain CDA

We conclude by assessing the performance of the hybrid gain method (Penny, 2014) applied to the coupled model. We provide a comparison using a fixed observing network in the physical grid space (shown in Figure 3) using varying levels of observational coverage. The construction of $\mathbf{B_0}$ and the tuning of its amplitude are repeated for each observing network. The ratio of the diagonals of $\mathbf{B}$ to $\mathbf{B_0}$ is approximately 9.2 for the full observing network case (14 for atmosphere-only and 2.9 for ocean only).

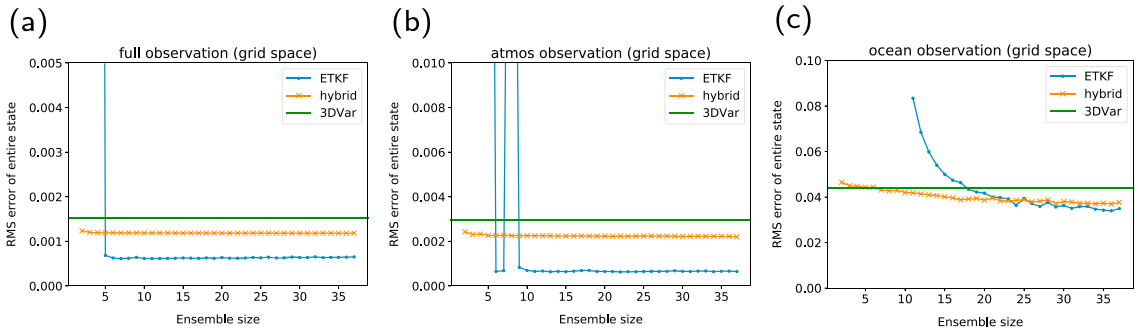


**Figure 19.** Time-mean analysis RMSE of using different DA systems. Observations in the physical grid space (Figure 3) are assimilated from (a) the full observing network, (b) the atmosphere only, and (c) the ocean only. The relaxation-to-prior method (weight 0.6 to the prior perturbation) is applied for inflating the analysis spread for the ETKF; no inflation is applied for the hybrid gain ($\alpha = 0.5$).

When the full observation network is assimilated (Figure 19a), the ETKF with relaxation (Zhang et al., 2004; weight 0.6 to the prior perturbations) achieved the best analysis accuracy with as few as five members. The stability of the ETKF with small ensemble sizes degrades without applying relaxation (not shown). The tuned 3D-Var system achieves less accurate analyses than the ETKF provided that a sufficient ensemble size is used to avoid filter divergence. When combining the ETKF (without relaxation) and 3D-Var via the hybrid gain method ($\alpha = 0.5$), the hybrid gain shows better analysis accuracy than 3D-Var even with the smallest possible ensemble size of two members. Because it acts as a form of additive inflation, the hybrid does not require the application of relaxation (or any other form of covariance inflation) typically required to stabilize the ETKF. The hybrid is more stable than the ETKF at small ensemble sizes. Decreasing $\alpha$, which strengthens the impact of the ETKF in the hybrid analysis, causes the analysis accuracy of the hybrid to become closer to that of the ETKF, but at the expense of its stability for smaller ensemble sizes (not shown).

When only atmospheric observations are assimilated (Figure 19b), the behavior of each system is similar to observing the full network, although the analysis is generally less accurate (and less stable for the ETKF). With the relaxation scheme, the ETKF is stable for ensemble sizes of 9 or greater; the filter is only marginally stable with an ensemble size of 6 to 8. When only ocean observations are assimilated, the analysis accuracy of the entire system is significantly degraded. The hybrid achieved more accurate and more stable analyses than the ETKF for a wide range of ensemble sizes under these conditions.

The performance of the ETKF and the hybrid method assimilating only ocean observations improves gradually as we increase the ensemble size (Figure 19c); this is in contrast to the abrupt transition between the convergence and divergence when full or atmospheric observations are assimilated. It is known that the deterministic ensemble square root filters require an ensemble size large enough so that the perturbations span the unstable and neutral subspaces (e.g., Ng et al., 2011; Trevisan & Palatella, 2011), and many ETKF experiments with simple models have shown such abrupt responses to the ensemble size (e.g., Penny, 2014; Sakov & Oke, 2008). We conjecture that the cause of this gradual response is due to the wide spectrum of near-neutral LEs in the coupled system that result from the slow ocean dynamics that render the unstable-neutral and stable manifolds less distinct.

## 5. Conclusions

We have investigated properties of a coupled atmosphere-ocean model and the application of CDA. The intention is to provide foundational knowledge that can serve as guidance to operational centers currently making the transition from independent atmosphere and ocean DA systems to CDA and more specifically to SCDA. The results have implications for the traditional use of forced atmosphere and ocean models for operational prediction and provide further supporting evidence for the recent trend toward fully coupled forecast models, for example, by the European Centre for Medium-Range Weather Forecasts in June 2018.

Experiments were conducted using the model of De Cruz et al. (2016) with coupled QG dynamics and coupled thermodynamics. We adapted the model to compare CDA with the traditional approach applying DA with forced uncoupled atmosphere and ocean models. The Lyapunov spectrum was computed for the

coupled and uncoupled models, as well as for the DA systems themselves in order to provide a diagnostic of the DA system stability. When applying the ETKF, the CLEs of the forced atmosphere become negative, a necessary condition for the uncoupled atmosphere to synchronize with coupled dynamics. We were not able to show this with the uncoupled ocean. In our experiments, the uncoupled ocean model did not synchronize with the coupled model even when provided with perfect atmospheric forcing. Considering the very long timescales in the ocean, it is possible that synchronization is achievable but requires a much longer integration. We evaluated the error covariance and error correlation matrices of the coupled forecast system. The coupled model contains multiple spatiotemporal scales, which introduces complications for numerical methods.

A number of DA methods were evaluated using uncoupled DA, WCDA, and SCDA. We compared the ETKF, 3D-Var, and 4D-Var. The ETKF and 4D-Var produced comparable accuracy when using a full observing network. When the observing network is reduced to only the atmosphere, the variational methods have difficulty converging in the ocean component.

A more thorough investigation of SCDA was conducted using the ETKF to compare the impact of using varying degrees of model coupling, different analysis cycle windows, and multiple observing networks. We also examined forecasts at various lead times from the ETKF analysis initial conditions. Further clarifying results from Smith et al. (2015) who found that "the weakly coupled system was usually comparable to uncoupled assimilations in which the atmosphere and ocean models were forced using the true SST and surface fluxes," we found that the weakly coupled DA system can actually produce larger errors than the forced system, depending on the accuracy of the surface forcing provided. We also clarify results from Ballabrera-Poy et al. (2009) and Han et al. (2013), who found that SCDA has difficulty in using observations from the ocean (i.e., the low-frequency system) to accurately update the atmosphere (i.e., the high-frequency system). While our results agree with this finding, we also found that by reducing the analysis cycle update time that ocean observations assimilated with SCDA can, in fact, produce stable and relatively accurate atmospheric analyses.

A new phenomenon was found when applying a hybrid DA method with CDA when increasing the ensemble size. Rather than producing a jump from divergent to stable filter at a specific ensemble size, as would be indicated by the dimension of the unstable-neutral subspace, there was instead a gradual improvement in analysis accuracy. This result is attributed to the extended spectrum of near-zero LEs in the coupled atmosphere/ocean system.

Based on the findings of this study, it is recommended that CDA is best performed in a strongly coupled framework. In transitioning to operational applications, appropriate localization strategies must be developed until large ensembles can be employed. While we did not conduct a detailed study of inflation methods here, we found that traditional methods such as multiplicative inflation may be wholly inadequate and encourage future work that develops improved strategies to improve forecast model error representation for SCDA. Finally, more work is needed to characterize the presence of multiple scales in coupled models and to determine how to best initialize these multiple scales for targeted prediction lead times.

The MAOOAM model (De Cruz et al., 2016), and particularly the configuration we use, was originally designed to study low-frequency climate variability in the midlatitudes. In this scenario, the ocean model is largely dependent on the atmospheric forcing. We note that there will be more chaotic dynamics in an ocean model with higher horizontal resolution, increased vertical fidelity, improved thermodynamic coupling, and the additional presence of hydrodynamic coupling.

We anticipate that SCDA will become a central area of study for initializing Earth system models for decadal climate prediction. The atmosphere has historically been far better observed than the ocean. SCDA provides the opportunity to use atmospheric observations to more accurately initialize the ocean state, which is essential for achieving skillful predictions beyond the often stated 2-week predictability limit of the atmosphere. Additional study is needed with more sophisticated Earth system models to characterize geographic dependence of the coupled feedbacks in SCDA. Future work is needed to explore the impacts of SCDA in systems that include multiple Earth system model components, such as sea ice, waves, land, aerosols, and biogeochemistry. Due to the complexity of interactions, we anticipate new insights to be gained from studying SCDA applied to the full range of possible combinations of these components and expect that new DA solution methods may be needed to initialize the full coupled Earth system model.

## Appendix A

We present the derivation of the TLM for the uncoupled model. De Cruz et al. (2016) show that the MAOOAM can be described as a system of Ordinary Differential Equations (ODEs) defined by the tensor $T$ such that

$$\frac{d\mathbf{x}_i}{dt} = M_i(\mathbf{x}, T^i) = \mathbf{x}^T T^i \mathbf{x}, \tag{A1}$$

where $T^i$ is a $(N+1)$-by-$(N+1)$ constant matrix depending on the model parameters. In order to simplify the derivation of the forced atmosphere/ocean model, we decompose the matrix $T^i$ and the augmented state vector $\mathbf{x}$ into its atmospheric and oceanic components,

$$T^i = \begin{bmatrix} T^i_{1,1} & T^i_{1,a} & T^i_{1,o} \\ T^i_{a,1} & T^i_{a,a} & T^i_{a,o} \\ T^i_{o,1} & T^i_{o,a} & T^i_{o,o} \end{bmatrix}, \text{and} \, \mathbf{x} = \begin{bmatrix} 1 \\ \mathbf{x}_a \\ \mathbf{x}_o \end{bmatrix}. \tag{A2}$$

In the equation for the forced atmospheric model, we group terms on the right-hand side as either dependent on or independent from the atmospheric variables,

$$\frac{d\mathbf{x}_{a,i}}{dt} = M_{a,i}(\mathbf{x}_a, \widetilde{\mathbf{x}}^c_o, T^i) \tag{A3}$$

$$= \mathbf{x}_a^T T^i_{a,1} + T^i_{1,a}\mathbf{x}_a + \mathbf{x}_a^T T^i_{a,a}\mathbf{x}_a + \mathbf{x}_a^T T^i_{a,o}\widetilde{\mathbf{x}}^c_o + (\widetilde{\mathbf{x}}^c_o)^T T^i_{o,a}\mathbf{x}_a + c^i_o, \text{for } i \in I_a, \tag{A4}$$

where $c^i_o$ is a term independent of the atmospheric state $\mathbf{x}_a$,

$$c^i_o = T^i_{1,1} + T^i_{1,o}\widetilde{\mathbf{x}}^c_o + (\widetilde{\mathbf{x}}^c_o)^T T^i_{o,1} + (\widetilde{\mathbf{x}}^c_o)^T T^i_{o,o}\widetilde{\mathbf{x}}^c_o, \tag{A5}$$

and where $I_a$ denotes the subset of indices corresponding to the atmosphere.

To form the TLM for each uncoupled system, we first derive the Jacobian of the above ODEs by taking derivatives with respect to each of the component variables. Starting with the atmospheric variables, each row of the Jacobian matrix is

$$J_{a,i} = \begin{bmatrix} \frac{\partial M_{a,i}}{\partial x_{a,1}} & \frac{\partial M_{a,i}}{\partial x_{a,2}} & \cdots & \frac{\partial M_{a,i}}{\partial x_{a,N_a}} \end{bmatrix} \tag{A6}$$

$$= \left( \left(T^i_{a,1}\right)^T + T^i_{1,a} \right) + \mathbf{x}_a^T \left( T^i_{a,a} + T^T_{a,a} \right) + (\widetilde{\mathbf{x}}^c_o)^T \left( T^i_{o,a} + \left(T^i_{a,o}\right)^T \right) \text{for} \, i \in I_a. \tag{A7}$$

The differential form of the forced TLM for a small atmospheric perturbation $\delta\mathbf{x}_a$ is

$$\frac{d(\delta\mathbf{x}_{a,i})}{dt} = J_{a,i}\delta\mathbf{x}_a \tag{A8}$$

$$= \left( \left(T^i_{a,1}\right)^T + T^i_{1,a} \right)\delta\mathbf{x}_a + \mathbf{x}_a^T \left( T^i_{a,a} + \left(T^i_{a,a}\right)^T \right)\delta\mathbf{x}_a + (\widetilde{\mathbf{x}}^c_o)^T \left( T^i_{o,a} + \left(T^i_{a,o}\right)^T \right)\delta\mathbf{x}_a. \tag{A9}$$

Similarly, the forced oceanic model is given as

$$\frac{d\mathbf{x}_{o,i}}{dt} = M_{o,i}(\widetilde{\mathbf{x}}^c_a, \mathbf{x}_o, T^i) \tag{A10}$$

$$= T^i_{1,o}\mathbf{x}_o + \mathbf{x}_o^T T^i_{o,1} + \mathbf{x}_o^T T^i_{o,o}\mathbf{x}_o + (\widetilde{\mathbf{x}}^c_a)^T T^i_{a,o}\mathbf{x}_o + \mathbf{x}_o^T T^i_{o,a}\widetilde{\mathbf{x}}^c_a + c^i_a, \text{for} \, i \in I_o, \tag{A11}$$

where $c^i_a$ is a constant term with respect to the ocean state $\mathbf{x}_o$ such that

$$c_a^i = T_{1,1}^i + T_{1,a}^i \tilde{\mathbf{x}}_a^c + \left(\tilde{\mathbf{x}}_a^c\right)^T T_{a,1}^i + \left(\tilde{\mathbf{x}}_a^c\right)^T T_{a,a}^i \tilde{\mathbf{x}}_a^c, \tag{A12}$$

and the differential form of the forced TLM for a small oceanic perturbation $\delta\mathbf{x}_o$ is

$$\frac{d(\delta\mathbf{x}_{o,i})}{dt} = J_{o,i}\delta\mathbf{x}_o \tag{A13}$$

$$= \left(T_{1,o}^i + \left(T_{o,1}^i\right)^T\right)\delta\mathbf{x}_o + \mathbf{x}_o^T\left(\left(T_{o,o}^i\right)^T + T_{o,o}^i\right)\delta\mathbf{x}_o + \left(\tilde{\mathbf{x}}_a^c\right)^T\left(\left(T_{o,a}^i\right)^T + T_{a,o}^i\right)\delta\mathbf{x}_o, \tag{A14}$$

where each row of the Jacobian is

$$J_{o,i} = \left[\frac{\partial M_{o,i}}{\partial x_{o,1}} \quad \frac{\partial M_{o,i}}{\partial x_{o,2}} \quad \cdots \quad \frac{\partial M_{o,i}}{\partial x_{o,N_o}}\right] \tag{A15}$$

$$= \left(T_{1,o}^i + \left(T_{o,1}^i\right)^T\right) + \mathbf{x}_o^T\left(\left(T_{o,o}^i\right)^T + T_{o,o}^i\right) + \left(\tilde{\mathbf{x}}_a^c\right)^T\left(\left(T_{o,a}^i\right)^i + T_{a,o}^i\right) \text{ for } i \in I_o. \tag{A16}$$

## Appendix B: Data Assimilation Methods

### B1. 3D-Var

Given the background model state vector $\mathbf{x}^b$ and the observations $\mathbf{y}^o$ at time $t_i$, we compute the analysis $\mathbf{x} = \mathbf{x}^a$ that minimizes the cost functional,

$$J(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^b\right)^T\mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right) + \frac{1}{2}(\mathbf{y}^o-h(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}^o-h(\mathbf{x})) \tag{A17}$$

where $\mathbf{B}$ and $\mathbf{R}$ are the background and observation error covariance matrices and $h(\cdot)$ is the (possibly nonlinear) observation operator. By linearizing the observation operator $h(\mathbf{x}) \approx h(\mathbf{x}^b)+\mathbf{H}(\mathbf{x}-\mathbf{x}^b)$, and introducing a change of variables $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}^b$ and $\mathbf{d} = \mathbf{y}^o - h(\mathbf{x}^b)$, we can convert (5) to the incremental form

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{d}-\mathbf{H}\delta\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{d}-\mathbf{H}\delta\mathbf{x}) \tag{A18}$$

Direct calculation of $J(\delta\mathbf{x})$ in this form requires computing the inverse of $\mathbf{B}$. This presents challenges for high dimensional systems and particular difficulties for coupled atmosphere-ocean dynamics for which the condition number of $\mathbf{B}$ can be as large as $10^{16}$ due to the wide range of resolved scales.

We find $\mathbf{x}^a$ using an iterative minimization algorithm. The convergence rate is related to the conditioning of the Hessian matrix with respect to the control variable $\delta\mathbf{x}$,

$$\mathbf{S}_{\delta\mathbf{x}} = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}. \tag{A19}$$

To address this conditioning problem, we perform the control variable transform (CVT) as described by Bannister (2008),

$$\mathbf{L}\delta\mathbf{u} = \delta\mathbf{x} \tag{A20}$$

which converts (27) to

$$J(\delta\mathbf{u}) = \frac{1}{2}\delta\mathbf{u}^T\delta\mathbf{u} + \frac{1}{2}(\mathbf{d}-\mathbf{H}\mathbf{L}\delta\mathbf{u})^T\mathbf{R}^{-1}(\mathbf{d}-\mathbf{H}\mathbf{L}\delta\mathbf{u}), \tag{A21}$$

with the gradient,

$$\nabla J_{\delta\mathbf{u}} = \delta\mathbf{u}-\mathbf{L}^T\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{d}-\mathbf{H}\mathbf{L}\delta\mathbf{u}). \tag{A22}$$

The Hessian matrix with respect to the new control variable $\delta\mathbf{u}$ is

$$\mathbf{S}_{\delta\mathbf{u}} = \mathbf{I} + \left(\mathbf{B}^{1/2}\right)^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{B}^{1/2} \tag{A23}$$

with an upper bound for the condition number,

$$\mathrm{cond}(\mathbf{S}_{\mathrm{CVT}}) = 1 + \left\| \mathbf{R}^{-1/2} \mathbf{H} \mathbf{B} \mathbf{H}^T \mathbf{R}^{-1/2} \right\|_\infty \tag{A24}$$

(Haben, 2011) which is generally much smaller than cond(**S**). The use of a CVT is common in operational systems (Bannister, 2008).

### B2. 4D-Var

The cost function of the strong-constraint 4D-Var with initial condition $\mathbf{x}_0$ is

$$J(\mathbf{x}_0) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}_0^b\right)^T \mathbf{B}^{-1} \left(\mathbf{x}_0 - \mathbf{x}_0^b\right) + \frac{1}{2}\sum_{i=0}^{N} \left(\mathbf{y}_i^o - h(M_i(\mathbf{x}_0))\right)^T \mathbf{R}_i^{-1} \left(\mathbf{y}_i^o - h(M_i(\mathbf{x}_0))\right), \tag{B1}$$

where $M_i(\mathbf{x}_0)$ is the model forecast at time $t_i$. We adopt the incremental approach (Courtier et al., 1994) to solve (34) where we seek a series of states $\mathbf{x}_0^{(j)}, j = 0, 1, 2, \ldots, n_j$. For each $\mathbf{x}_0^{(j)}$, we determine the increment $\delta \mathbf{x}_0^{(j)}$ that minimizes

$$\begin{aligned} J\left(\delta\mathbf{x}_0^{(j)}\right) &= \frac{1}{2}\left[\delta\mathbf{x}_0^{(j)} - \left(\mathbf{x}_0^b - \mathbf{x}_0^{(j)}\right)\right]^T \mathbf{B}^{-1} \left[\delta\mathbf{x}_0^{(j)} - \left(\mathbf{x}_0^b - \mathbf{x}_0^{(j)}\right)\right] \\ &+ \frac{1}{2}\sum_{i=0}^{N} \left[\mathbf{d}_i^{(j)} - \mathbf{H}_i \mathbf{M}_i|_{\mathbf{x}_0^{(j)}} \delta\mathbf{x}_0^{(j)}\right]^T \mathbf{R}_i^{-1} \left[\mathbf{d}_i^{(j)} - \mathbf{H}_i \mathbf{M}_i|_{\mathbf{x}_0^{(j)}} \delta\mathbf{x}_0^{(j)}\right] \end{aligned} \tag{B2}$$

where $\mathbf{d}_i^{(j)} = \mathbf{y}_i^o - H\left(M_i\left(\mathbf{x}_0^{(j)}\right)\right)$ and $\mathbf{M}_i|_{x_0^{(j)}}$ is the TLM of $M_i\left(\mathbf{x}_0^{(j)}\right)$. We set the next state $\mathbf{x}_0^{(j+1)} = \mathbf{x}_0^{(j)} + \delta\mathbf{x}_0^{(j)}$ and repeat the iteration in an outer loop, until the algorithm reaches a minimum error threshold. As for 3D-Var, we apply the CVT so the cost function becomes

$$\begin{aligned} J\left(\delta\mathbf{u}_0^{(j)}\right) &= \frac{1}{2}\left[\delta\mathbf{u}_0^{(j)} - \mathbf{u}_0^{b(j)}\right]^T \left[\delta\mathbf{u}_0^{(j)} - \mathbf{u}_0^{b(j)}\right] \\ &+ \frac{1}{2}\sum_{i=0}^{N} \left[\mathbf{d}_i^{(j)} - \mathbf{H}_i \mathbf{M}_i|_{\mathbf{x}_0^{(j)}} \mathbf{L}\mathbf{u}_0^{(j)}\right]^T \mathbf{R}_i^{-1} \left[\mathbf{d}_i^{(j)} - \mathbf{H}_i \mathbf{M}_i|_{\mathbf{x}_0^{(j)}} \mathbf{L}\mathbf{u}_0^{(j)}\right] \end{aligned} \tag{B3}$$

where

$$\mathbf{L}\delta\mathbf{u}_0^{b(j)} = \left(\mathbf{x}_0^b - \mathbf{x}_0^{(j)}\right) \tag{B4}$$

and the gradient of $J$ is

$$\nabla J_{\delta\mathbf{u}_0^{(j)}} = \left[\delta\mathbf{u}_0^{(j)} - \delta\mathbf{u}_0^{b(j)}\right] + \mathbf{L}^T \sum_{i=0}^{N} \mathbf{M}_i^T|_{\mathbf{x}_0^{(j)}} \mathbf{H}_i^T \mathbf{R}_i^{-1} \left[\mathbf{d}_i^{(j)} - \mathbf{H}_i \mathbf{M}_i|_{\mathbf{x}_0^{(j)}} \mathbf{L}\delta\mathbf{u}_0^{(j)}\right]. \tag{B5}$$

### B3. Ensemble Transform Kalman Filter (ETKF)

Ensemble forecasts are used to form a dynamic estimate of the background error covariance matrix for the ETKF (Bishop et al., 2001; Hunt et al., 2007),

$$\mathbf{P}^b = \frac{1}{k-1} \mathbf{X}^b \left(\mathbf{X}^b\right)^T, \tag{B6}$$

where each column of $\mathbf{X}^b$ is an ensemble perturbation and $k$ is the total number of ensemble members. Assuming that the analysis error covariance matrix is of the same form, we seek a transform matrix $\mathbf{W}$ applied to $\mathbf{X}^b$ so that $\mathbf{X}^a = \mathbf{X}^b\mathbf{W}$. Given that $\mathbf{P}^a = [\mathbf{I} + \mathbf{P}^b\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}]^{-1}\mathbf{P}^b$ (Hunt et al., 2007, equation (11)), we have

$$\mathbf{P}^a = \frac{1}{k-1}\mathbf{X}^b\mathbf{W}\mathbf{W}^T\left(\mathbf{X}^b\right)^T = \left[\mathbf{I} + \frac{1}{k-1}\mathbf{X}^b\left(\mathbf{X}^b\right)^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\frac{1}{k-1}\mathbf{X}^b\left(\mathbf{X}^b\right)^T. \tag{B7}$$

Further letting $\mathbf{Y}^b = \mathbf{H}\mathbf{X}^b$,

$$\frac{1}{k-1}\mathbf{X}^b\left[\mathbf{I} + \frac{1}{k-1}\left(\mathbf{Y}^b\right)^T\mathbf{R}^{-1}\mathbf{Y}^b\right]\mathbf{W}\mathbf{W}^T\left(\mathbf{X}^b\right)^T = \frac{1}{k-1}\mathbf{X}^b\left(\mathbf{X}^b\right)^T. \tag{B8}$$

We can simplify so that

$$\left[\mathbf{I} + \frac{1}{k-1}\left(\mathbf{Y}^b\right)^T\mathbf{R}^{-1}\mathbf{Y}^b\right]\mathbf{W}\mathbf{W}^T = \mathbf{I}. \tag{B9}$$

Because $(\mathbf{Y}^b)^T\mathbf{R}^{-1}\mathbf{Y}^b$ is symmetric and positive semidefinite with at least one eigenvalue of 0, $\mathbf{I} + \frac{1}{k-1}\left(\mathbf{Y}^b\right)^T\mathbf{R}^{-1}\mathbf{Y}^b$ is symmetric and positive definite and so is invertible,

$$\mathbf{W}\mathbf{W}^T = \left[\mathbf{I} + \frac{1}{k-1}\left(\mathbf{Y}^b\right)^T\mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1}. \tag{B10}$$

We compute the symmetric square root,

$$\mathbf{W} = \left[\mathbf{I} + \frac{1}{k-1}\left(\mathbf{Y}^b\right)^T\mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1/2}, \tag{B11}$$

and then update the ensemble perturbations as $\mathbf{X}^a = \mathbf{X}^b\mathbf{W}$ and the analysis ensemble mean as

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \mathbf{X}^b\left[(k-1)\mathbf{I} + \left(\mathbf{Y}^b\right)^T\mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1}\left(\mathbf{Y}^b\right)^T\mathbf{R}^{-1}\left(\mathbf{y}^o - H\left(\overline{\mathbf{x}}^b\right)\right). \tag{B12}$$

Considering the small size of the model under investigation, we do not use localization.

### B4. Hybrid Gain Coupled Data Assimilation (CDA)

Penny et al. (2017) recommended the investigation of multiscale methods for CDA. As a simple type of multiscale DA approach, we test the hybrid gain algorithm of Penny (2014), which forms a combination of climatological and dynamically estimated gain matrices,

$$\mathbf{K} = \beta_1\mathbf{K}^P + \beta_2\mathbf{K}^B - \beta_3\mathbf{K}^B\mathbf{H}\mathbf{K}^P. \tag{B13}$$

We consider hybrid methods as basic multiscale methods because the ensemble-derived error covariance represents short-timescale errors (with respect to the analysis increment) valid at a specific time, while the climatological error covariance uses a long-time average of these errors valid over a long integration of the model. Hybrid DA algorithms can improve the stability and the accuracy of a filter, particularly in the presence of model bias (Bonavita et al., 2015; Penny, 2014; Penny et al., 2015; Penny, 2017). Following Bonavita et al. (2015), we use $\beta_1 = 1 - \alpha$, $\beta_2 = \alpha$, and $\beta_3 = 0$ and set the relative weight between the ETKF and variational method as a constant $\alpha = 0.5$. In the variational component of the hybrid, the climatological background error covariance matrix is tuned for each observation network.

**Acknowledgments**
This work was conducted by the students of the Spring 2018 graduate course AOSC 658E at the University of Maryland. L. S. and E. B. carried out the Lyapunov and synchronization analysis of the forced and coupled models. T. Y. carried out the analysis of background error covariance and hybrid gain experiments. C. D. and C.-C. C. compared the performance of 3D-Var and ETKF with different coupling approaches and the performance of different SCDA methods. K. B. investigated the impact of using different observation network configurations in the CDA systems. L. S. performed the investigation of forecasts at various lead times using each CDA approach. S. G. P. acknowledges support from the NOAA Climate Program Office (CPO) Modeling, Analysis, Predictions, and Projections (MAPP) program under Grant NA16OAR4310140, the NOAA Next-Generation Global Prediction System program under Grant NA18NWS4680048, and the NOAA National Environmental Satellite, Data, and Information Service (NESDIS) SPSP_OODA_17. L. S. is supported by NOAA NESDIS via the NASA ROSES program SPSP_OODA_17. C. D. is supported by NASA Headquarters under the NASA Earth and Space Science Fellowship Program 80NSSC18K1403. E. B. is supported by the University of Maryland Flagship Fellowship and Monsoon Mission II funding (grant IITMMMIIUNIV MARYLANDUSA2018INT1) provided by the Ministry of Earth Science, Government of India. T. Y. is supported by the Japanese Government Long-term Overseas Fellowship Program. The authors acknowledge the University of Maryland supercomputing resources (http://hpcc.umd.edu) made available for conducting computational experiments reported in this study.

## References

Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, *129*(12), 2884–2903. https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2

Anderson, J. L. (2003). A local least squares framework for ensemble filtering. *Monthly Weather Review*, *131*(4), 634–642. https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2

Ballabrera-Poy, J., Kalnay, E., & Yang, S.-C. (2009). Data assimilation in a system with two scales—Combining two initialization techniques. *Tellus*, *61A*, 539–549.

Bannister, R. N. (2008). A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, *134*(637), 1971–1996. https://doi.org/10.1002/qj.340

Barsugli, J. J., & Battisti, D. S. (1998). The basic effects of atmosphere-ocean thermal coupling on midlatitude variability. *Journal of the Atmospheric Sciences*, *55*, 477–493. https://doi.org/10.1175/1520-0469(1998)055<0477:TBEOAO>2.0.CO;2

Bishop, C. H., Etherton, B., & Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly Weather Review*, *129*(3), 420–436. https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2

Bonavita, M., Hamrud, M., & Isaksen, L. (2015). EnKF and hybrid gain ensemble data assimilation. Part II: EnKF and hybrid gain results. *Monthly Weather Review*, *143*, 4865–4882. https://doi.org/10.1175/MWR-D-15-0071.1

Carrassi, A., Ghil, M., Trevisan, A., & Uboldi, F. (2008). Data assimilation as a nonlinear dynamical systems problem: Stability and convergence of the prediction-assimilation system. *Chaos*, *18*(2), 023112. https://doi.org/10.1063/1.2909862

Cehelsky, P., & Tung, K. K. (1987). Theories of multiple equilibria and weather regimes-a critical reexamination. Part II: Baroclinic two-layer models. *Journal of the Atmospheric Sciences*, *44*, 3282–3303.

Charney, J. G., & Straus, D. M. (1980). Formdrag instability, multiple equilibria and propagating planetary waves in baroclinic, orographically forced, planetary wave systems. *Journal of the Atmospheric Sciences*, *37*, 1157–1176.

Courtier, P., Thépaut, J.-N., & Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, *120*(519), 1367–1387. https://doi.org/10.1002/qj.49712051912

De Cruz, L., Demaeyer, J., & Vannitsem, S. (2016). The Modular Arbitrary-Order Ocean-Atmosphere Model: MAOOAM v1.0. *Geoscientific Model Development*, *9*, 2793–2808. https://doi.org/10.5194/gmd-9-2793-2016

De Cruz, L., Schubert, S., Demaeyer, J., Lucarini, V., & Vannitsem, S. (2018). Exploring the Lyapunov instability properties of high-dimensional atmospheric and climate models. *Nonlinear Processes in Geophysics*, *25*, 387–412. https://doi.org/10.5194/npg-25-387-2018

Ferrari, R., & Cessi, P. (2003). Seasonal synchronization in a chaotic ocean–Atmosphere model. *Journal of Climate*, *16*, 875–881. https://doi.org/10.1175/1520-0442(2003)016<0875:SSIACO>2.0.CO;2

Geist, K., Parlitz, U., & Lauterborn, W. (1990). Comparison of different methods for computing Lyapunov exponents. *Progress in Theoretical Physics*, *83*, 875–893. https://doi.org/10.1143/PTP.83.875

Goodman, J. C., & Marshall, J. (2003). The role of neutral singular vectors in midlatitude air-sea coupling. *Journal of Climate*, *16*, 88–102. https://doi.org/10.1175/1520-0442(2003)016<0088:TRONSV>2.0.CO;2

Grudzien, C., Carrassi, A., & Bocquet, M. (2018a). Asymptotic forecast uncertainty and the unstable subspace in the presence of additive model error. *SIAM/ASA Journal on Uncertainty Quantification*, *6*(4), 1335–1363. https://doi.org/10.1137/17M114073X

Grudzien, C., Carrassi, A., & Bocquet, M. (2018b). Chaotic dynamics and the role of covariance inflation for reduced rank Kalman filters with model error. *Nonlinear Processes in Geophysics*, *25*, 633–648. https://doi.org/10.5194/npg-25-633-2018

Haben S. A. (2011). Conditioning and preconditioning of the minimisation problem in variational data assimilation (PhD thesis). University of Reading, 2011.

Hamrud, M., Bonavita, M., & Isaksen, L. (2014). EnKF and hybrid gain ensemble data assimilation. *ECMWF Tech. Rep.*, *733*, 1–34, available at. https://www.ecmwf.int/sites/default/files/elibrary/2014/9766-enkf-and-hybrid-gain-ensemble-data-assimilation.pdf

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. *Journal of Climate*, *26*(24), 10,218–10,231. https://doi.org/10.1175/JCLI-D-13-00236.1

Houtekamer, P. L., Buehner, M., & De La Chevrotiere, M. (2018). Using the hybrid gain algorithm to sample data assimilation uncertainty. *Quarterly Journal of the Royal Meteorological Society*, 1–22. https://doi.org/10.1002/qj.3426

Houtekamer, P. L., & Mitchell, H. L. (1998). Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, *126*(3), 796–811. https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2

Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, *230*(1-2), 112–126. https://doi.org/10.1016/j.physd.2006.11.008

Jin, F. (1997). An equatorial ocean recharge paradigm for ENSO. Part I: A conceptual model. *Journal of the Atmospheric Sciences*, *54*(7), 811–829. https://doi.org/10.1175/1520-0469(1997)054<0811:AEORPF>2.0.CO;2

Kinter, J., O'Brien, T., Klein, S., Lin, S. J., Medeiros, B., Penny, S. G., et al. (2016). High-resolution coupling and initialization to improve predictability and predictions in climate models workshop. U.S. Department of Energy, DOE/SC-0183; U.S. Department of Commerce NOAA Technical Report OAR CPO-5. doi:https://doi.org/10.7289/V5K35RNX. https://science.energy.gov/~/media/ber/pdf/workshop%20reports/High_Resolution_Workshop_Report_FINAL_WEB.pdf

Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K., & Janssen, P. (2016). A coupled data assimilation system for climate reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 65–78. https://doi.org/10.1002/qj.2629

Legras, B., & Vautard, R. (1996). A guide to Lyapunov vectors. In *Seminar on Predictability*, (pp. 143–156). United Kingdom: ECMWF, Reading.

Liu, S., Wu, S., Zhang, Y. L., & Rong, X. Y. (2013). Ensemble data assimilation in a simple coupled climate model: The role of ocean-atmosphere interaction. *Advances in Atmospheric Sciences*, *30*, 1235–1248. https://doi.org/10.1007/s00376-013-2268-z

Lorenz, E. N. (1963). Deterministic non-periodic flow. *Journal of the Atmospheric Sciences*, *20*(2), 130–141. https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

Lorenz, E. N. (1996). *Predictability—A problem partly solved. Proceedings of a Seminar Held at ECMWF on Predictability, ECMWF Seminar Proceedings*, (Vol. 1, pp. 1–18). United Kingdom: ECMWF.

Lorenz, E. N., & Emanuel, K. A. (1998). Optimal sites for supplementary weather observations: Simulations with a small model. *Journal of the Atmospheric Sciences*, *55*, 399–414.

Lu, F., Liu, Z., Zhang, S., & Liu, Y. (2015a). Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part I: Simple model study. *Monthly Weather Review*, *143*(9), 3823–3837. https://doi.org/10.1175/MWR-D-14-00322.1

Lu, F., Liu, Z., Zhang, S., & Liu, Y. (2015b). Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part II: CGCM applications. *Monthly Weather Review*, *143*(11), 4645–4659. https://doi.org/10.1175/MWR-D-15-0088.1

McCreary, J. P. (1981). A linear stratified ocean model of the equatorial undercurrent. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *298*, 603–635.

Ng, G.-H. C., Mclaughlin, D., Entekhabi, D., & Ahanin, A. (2011). The role of model dynamics in ensemble Kalman filter performance for chaotic systems. *Tellus*, *63A*, 958–977.

Patton, A., Politis, D. N., & White, H. (2009). Correction to "Automatic block-length selection for the dependent bootstrap" by D. Politis and H. White. *Econometric Reviews*, *28*, 372–375. https://doi.org/10.1080/07474930802459016

Pecora, L. M., & Carroll, T. L. (1991). Driving systems with chaotic signals. *Physical Review A*, *44*, 2374–2383. https://doi.org/10.1103/PhysRevA.44.2374

Pecora, L. M., & Carroll, T. L. (2015). Synchronization of chaotic systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *25*(9), 097611. https://doi.org/10.1063/1.4917383

Peña, M., & Kalnay, E. (2004). Separating fast and slow modes in coupled chaotic systems. *Nonlinear Processes in Geophysics*, *11*, 319–327. https://doi.org/10.5194/npg-11-319-2004

Penny, S. G. (2014). The hybrid local ensemble transform Kalman filter. *Monthly Weather Review*, *142*, 2139–2149. https://doi.org/10.1175/MWR-D-13-00131.1

Penny, S. G. (2017). Mathematical foundations of hybrid data assimilation from a synchronization perspective. *Chaos*, *27*(12), 126801. https://doi.org/10.1063/1.5001819

Penny, S. G., Akella, S., Buehner, M., Chevallier, M., Counillon, F., Draper, C., et al. (2017). Coupled data assimilation for integrated Earth system analysis and prediction: Goals, challenges and recommendations. World Meteorological Organization, WWRP_2017_3. Available online (last accessed 3 April 2018): https://www.wmo.int/pages/prog/arep/wwrp/new/documents/Final_WWRP_2017_3_27_July.pdf

Penny, S. G., Behringer, D., Carton, J. A., & Kalnay, E. (2015). A hybrid global ocean data assimilation system at NCEP. *Monthly Weather Review*, *143*, 4660–4677.

Penny, S. G., & Hamill, T. (2017). Coupled data assimilation for integrated Earth system analysis and prediction. *Bulletin of the American Meteorological Society*, *97*(7), ES169–ES172, July 2017. https://journals.ametsoc.org/doi/full/10.1175/BAMS-D-17-0036.1

Pierini, S. (2011). Lowfrequency variability, coherence resonance, and phase selection in a low-order model of the wind-driven ocean circulation. *Journal of Physical Oceanography*, *41*, 1585–1604.

Pires, C., Vautard, R., & Talagrand, O. (1996). On extending the limits of variational assimilation in nonlinear chaotic systems. *Tellus*, *48*(1), 96–121. https://doi.org/10.3402/tellusa.v48i1.11634

Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, *89*(428), 1303–1313. https://doi.org/10.2307/2290993

Reinhold, B. B., & Pierrehumbert, R. T. (1982). Dynamics of weather regimes: Quasi-stationary waves and blocking. *Monthly Weather Review*, *110*, 1105–1145.

Ringler, T., Penny, S. G., Koch, D., Barrie, D., Mariotti, A., Adcroft, A., et al. (2018). Workshop on the initialization of high-resolution Earth system models. DOE/SC-0195 ■ NOAA Technical Report OAR CPO-6 ■ https://doi.org/10.25923/20NE-2B12

Roebber, P. J. (1995). Climate variability in a low-order coupled atmosphere–ocean model. *Tellus*, *47A*, 473–494.

Ruiz-Barradas, A., Kalnay, E., Peña, M., BozorgMagham, A. E., & Motesharrei, S. (2017). Finding the driver of local ocean–atmosphere coupling in reanalyses and CMIP5 climate models. *Climate Dynamics*, *48*, 2153–2172. https://doi.org/10.1007/s00382-016-3197-1

Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., et al. (2006). The NCEP Climate Forecast System. *Journal of Climate*, *19*(15), 3483–3517. https://doi.org/10.1175/JCLI3812.1

Sakov, P., & Oke, P. R. (2008). Implications of the form of the ensemble transformation in the ensemble square root filters. *Monthly Weather Review*, *136*, 1042–1053. https://doi.org/10.1175/2007MWR2021.1

Sheppard, K. (2015). ARCH Toolbox for Python. Zenodo, https://zenodo.org/record/15681 (Accessed May 9, 2018).

Sluka, T. C. (2018). *Strongly coupled ocean-atmosphere data assimilation with the local ensemble transform Kalman filter*. Ph.D. dissertation. College Park: University of Maryland.

Sluka, T. C., Penny, S. G., Kalnay, E., & Miyoshi, T. (2016). Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophysical Research Letters*, *43*, 752–759. https://doi.org/10.1002/2015GL067238

Smith, P. J., Fowler, A. M., & Lawless, A. S. (2015). Exploring strategies for coupled 4D-Var data assimilation using an idealised atmosphere-ocean model. *Tellus A: Dynamic Meteorology and Oceanography*, *67*(1), 27,025. https://doi.org/10.3402/tellusa.v67.27025

Smith, P. J., Lawless, A. S., & Nichols, N. K. (2017). Estimating forecast error covariances for strongly coupled atmosphere-ocean 4D-Var data assimilation. *Monthly Weather Review*, *145*(10), 4011–4035. ISSN 00270644. https://doi.org/10.1175/MWRD160284.1

Smith, P. J., Lawless, A. S., & Nichols, N. K. (2018). Treating sample covariances for use in strongly coupled atmosphere-ocean data assimilation. *Geophysical Research Letters*, *45*, 445–454. https://doi.org/10.1002/2017gl075534

Tardif, R., Hakim, G. J., & Snyder, C. (2014). Coupled atmosphere-ocean data assimilation experiments with a low-order climate model. *Climate Dynamics*, *43*(5–6), 1631–1643. https://doi.org/10.1007/s00382-013-1989-0

Tardif, R., Hakim, G. J., & Snyder, C. (2015). Coupled atmosphere-ocean data assimilation experiments with a low-order model and CMIP5 model data. *Climate Dynamics*, *45*(5–6), 1415–1427. https://doi.org/10.1007/s00382-014-2390-3

Toth, Z., & Kalnay, E. (1993). Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society*, *74*, 2317–2330. https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2

Trevisan, A., & Palatella, L. (2011). On the Kalman filter error covariance collapse into the unstable subspace. *Nonlinear Processes in Geophysics*, *18*, 243–250. https://doi.org/10.5194/npg-18-243-2011

Trevisan, A., & Uboldi, F. (2004). Assimilation of standard and targeted observations within the unstable subspace of the observation-analysis-forecast cycle system. *Journal of the Atmospheric Sciences*, *61*, 103–113. https://doi.org/10.1175/1520-0469(2004)061<0103:AOSATO>2.0.CO;2

Vannitsem, S. (2017). Predictability of large-scale atmospheric motions: Lyapunov exponents and error dynamics. *Chaos*, *27*(3), 032101. https://doi.org/10.1063/1.4979042

Vannitsem, S., Demaeyer, J., De Cruz, L., & Ghil, M. (2015). Low-frequency variability and heat transport in a low-order nonlinear coupled ocean-atmosphere model. *Physica D: Nonlinear Phenomena*, *309*, 71–85. https://doi.org/10.1016/j.physd.2015.07.006

Vannitsem, S., & Lucarini, V. (2016). Statistical and dynamical properties of covariant Lyapunov vectors in a coupled atmosphere-ocean model—Multiscale effects, geometric degeneracy, and error dynamics. *Journal of Physics A: Mathematical and Theoretical*, *49*(22). https://doi.org/10.1088/1751-8113/49/22/224001

Wada, A., Tsugiti, H., Okamoto, K., & Seino, N. (2019). Air-sea coupled data assimilation experiment for Typhoons Kilo, Etau and the September 2015 Kanto-Tohoku Heavy Rainfall with the Advanced Microwave Scanning Radiometer 2 sea surface temperature. *Journal of the Meteorological Society of Japan*, *97*(3), 553–575. Special Edition on Tropical Cyclones in 2015–2016. https://doi.org/10.2151/jmsj.2019-029

Yang, S.-C., Corazza, M., Carrassi, A., Kalnay, E., & Miyoshi, T. (2009). Comparison of local ensemble transform Kalman filter, 3DVAR, and 4DVAR in a Quasigeostrophic model. *Monthly Weather Review*, *137*(2), 693–709.

Yoshida, T., & Kalnay, E. (2018). Correlation-cutoff method for covariance localization in strongly coupled data assimilation. *Monthly Weather Review*. *146*(9), 2881–2889. https://doi.org/10.1175/MWR-D-17-0365.1

Zhang, F., Snyder, C., & Sun, J. (2004). Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Monthly Weather Review*, *132*(5), 1238–1253. https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2

Zhang, S., Harrison, M. J., Rosati, A., & Wittenberg, A. (2007). System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Monthly Weather Review*, *135*, 3541–3564. https://doi.org/10.1175/MWR3466.1

Zhang, R. H., Zebiak, S. E., Kleeman, R., & Keenlyside, N. (2005). Retrospective El Niño forecast using an improved intermediate coupled model. *Monthly Weather Review*, *133*, 2777–2802.

Zheng, F., & Zhu, J. (2010). Coupled assimilation for an intermediated coupled ENSO prediction model. *Ocean Dynamics*, *60*, 1061–1073. https://doi.org/10.1007/s10236-010-0307-1