

Developing 4D-Var for Strongly Coupled Data Assimilation Using a Coupled Atmosphere–Ocean Quasigeostrophic Model

MICHAEL GOODLIFF^{a,b} AND STEPHEN G. PENNY^{c,a,b}

^a *Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado*

^b *NOAA/Physical Sciences Laboratory, Boulder, Colorado*

^c *Sofar Ocean Technologies, San Francisco, California*

(Manuscript received 14 September 2021, in final form 21 May 2022)

ABSTRACT: Four-dimensional variational (4D-Var) data assimilation (DA) is developed for a coupled atmosphere–ocean quasigeostrophic application. Complications arise in coupled data assimilation (CDA) systems due to the presence of multiple spatiotemporal scales. Various formulations of the background error covariance matrix (\mathbf{B}), using different localization strategies, are explored to evaluate their impact on 4D-Var performance in a CDA setting. 4D-Var requires access to tangent linear and adjoint models (TLM/AM) to propagate information about the misfit between the forecast and observations within an optimization window. In practice, particularly for coupled models, the TLM and adjoint are often difficult to produce, and for some models are nonexistent in analytic form. Accordingly, a statistical data-driven alternative is also employed and evaluated to determine its feasibility for a 4D-Var CDA system. Using experiments conducted with a coupled atmosphere–ocean quasigeostrophic model, it is found that ensemble generation of flow-dependent error covariance statistics can increase the accuracy of 4D-Var CDA. When observing all variables, the hybrid climatological/flow-dependent \mathbf{B} constructions outperform either independently. The use of a hybrid \mathbf{B} matrix combined with a rapid updating ensemble transform Kalman filter (RU-ETKF) using either strongly or weakly CDA resulted in lower overall RMSE. The ocean component achieved its lowest RMSE when using a fully flow-dependent \mathbf{B} matrix generated using 4D-ETKF and using weakly CDA. These results show the importance of time scales and analysis update frequencies. The use of a statistically derived TLM/AM generated from the ETKF ensemble perturbations produces results similar to cases using the analytical coupled TLM/AM in 4D-Var.

KEYWORDS: Coupled models; Data assimilation; Ensembles


1. Introduction

Coupled data assimilation (CDA) is quickly growing in importance as operational prediction centers around the world transition to the use of fully coupled forecast models, with the intention of transitioning to more seamless prediction between time scales ranging from weather (days to weeks) to seasonal (weeks to months). The primary initialization strategy used by most prediction centers for coupled models is to initialize each model component (e.g., the atmosphere or ocean) independently. As the CDA paradigm matures, a key feature of CDA will be improved utilization of information across the model components (e.g., across the air–sea interface).

Coupled data assimilation can be implemented in many ways, typically spanning from *weakly coupled data assimilation* (WCDA), in which data assimilation is applied independently for each coupled model component, to strongly coupled data assimilation (SCDA), in which the coupled system is treated as a single dynamical system so that the analysis can be formed by using data across coupled model components (Penny et al. 2017). Penny et al. (2019) reviewed and summarized previous

literature on SCDA and then investigated a number of open foundational questions using a coupled quasigeostrophic (QG) model. In general, 4D-Var has been one of the most effective DA techniques used in practice for operational numerical weather prediction (NWP) (Janisková et al. 1999; Rabier et al. 2000; Gauthier et al. 2007; Rawlins et al. 2007). However, Penny et al. (2019) found that when observations are sparser the canonical strong-constraint 4D-Variational method (4D-Var) generally underperforms due to its use of a climatological error covariance to represent cross-domain atmosphere–ocean interactions. This is in comparison to the Ensemble Kalman Filter (EnKF), which estimates the forecast error covariance dynamically. Hence, the results of Penny et al. (2019) motivate further study to understand and improve the capabilities of 4D-Var in the context of SCDA, particularly in the presence of multiscale dynamics.

The background error covariance matrix is a key mechanism of DA methods such as 4D-Var or the EnKF for information transfer across the component models in SCDA (Smith et al. 2017, 2018). The background error covariance matrix (typically denoted \mathbf{B}) was traditionally calculated using errors from short forecasts sampled over a long historical period. This “Climatological” background error covariance matrix \mathbf{B}_c provides a time-average statistical relationship between errors in each variable, but as a result cannot represent transient state-dependent relationships. An alternative approach is to resolve the “errors of the day” (Kalnay et al. 1997; Corazza et al. 2003) by dynamically estimating the background error

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Michael Goodliff, Michael.Goodliff@noaa.gov

DOI: 10.1175/MWR-D-21-0240.1

© 2022 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

covariance matrix \mathbf{B}_f using an ensemble forecast (Kalnay 2019; Buizza 2019). Due to limitations in using such statistical correlations at long distances, localization is an important tool that can reduce or eliminate spurious correlations in covariance matrices. This can also be used to limit the transfer of inaccurate information across component models in CDA. The current state-of-the-art in operational NWP is to combine flow-dependent errors derived from an online ensemble forecast with supplemental climatological statistics, generating a hybrid covariance matrix \mathbf{B}_h (Hamill and Snyder 2000; Lorenc et al. 2015). This hybrid formulation increases the rank of the dynamic estimate, which is usually limited due to the use of a relatively small sample of ensemble perturbations.

An additional technical constraint in the use of 4D-Var for SCDA is that large-scale coupled systems typically do not have a corresponding tangent linear model (TLM) and adjoint model for the coupled atmosphere–ocean dynamics. These tools are required by the most commonly used optimization methods in order to minimize the 4D-Var cost function. Developing the TLM and adjoint for a large-scale numerical model of the atmosphere or ocean can be a challenging and tedious endeavor, and already serves as a barrier to applying 4D-Var in practice. Even if each independent model has an associated TLM and adjoint, in most cases the coupled Earth system model does not. State-of-the-art component models of the atmosphere and ocean are usually coupled through a software interface in order to form modern Earth system models, adding further difficulty to producing a unified TLM and adjoint for use in SCDA.

To enable the application of 4D-Var for general cases in which an analytic or software-based TLM and adjoint cannot be easily developed or maintained, we additionally consider an alternative surrogate model based on statistical linear regression (Lermusiaux and Robinson 1999; Pelc et al. 2012). To achieve this, we use the multiple realizations of the system trajectory produced by ensemble forecasts to create a statistical ensemble-based TLM (ETLM). This type of approach has been investigated in applications of increasing complexity. Frolov and Bishop (2016) proposed a localized version of such a statistical TLM called the local ensemble tangent linear model (LETLM) applied to a linear wave advection model. This localization helped to address issues of rank deficiency with the ETLM when the ensemble size is insufficient for modeling applications with larger degrees of freedom. Allen et al. (2017) and Bishop et al. (2017) studied the performance of 4D-Var using the LETLM with a shallow-water model, and with an idealized coupled Lorenz model, respectively. These works concluded that the LETLM can be an accurate substitution of the analytic TLM provided that the ensemble size is larger than the dimension of the local state vector. Frolov et al. (2018) applied the LETLM to a more realistic model NAVGEM (the U.S. Navy's global atmospheric prediction model), comparing the ability of the LETLM to predict the difference between two nonlinear trajectories against a conventional software-based TLM. They showed that the LETLM was generally more accurate than the software-based TLM (partially due to the age of the TLM software), though in their implementation the computational cost was more expensive. This was then expanded upon in Allen

et al. (2020), which examined the challenges of applying the LETLM to the high-resolution (1°) NAVGEM. This paper showed that the LETLM could overcome challenges like balancing ensemble size with the computational stencil size, and upper-atmospheric propagation of fast-moving gravity modes. They concluded by showing that the LETLM has higher skill in the lower troposphere (below 700 hPa), but slightly lags in the upper troposphere and stratosphere (700–2 hPa). Payne (2021) proposed a hybrid of the LETLM and the analytical TLM, using the LETLM to fill in for missing physical processes that were not represented by the TLM. This was proven to be cheaper to run than the LETLM, and when using a very large ensemble, performed better than a pure LETLM. Given the context of our simpler quasigeostrophic application, we will consider both the ETLM and a variation of the LETLM as replacements for the TLM and adjoint in our investigation of 4D-Var SCDA.

Here, we extend the work of Penny et al. (2019) to further advance the exploration of 4D-Var in its application for SCDA, applied using the same MAOOAM coupled atmosphere–ocean QG model configuration. We will explore a broader range of options for constructing the background error covariance for use in 4D-Var SCDA. We compare various covariance structures, including climatological, flow-dependent, and hybrid error covariance matrices. Our primary tool for estimating flow-dependent error covariance is the EnKF; however, there are multiple options for implementing the EnKF in conjunction with the 4D-Var analysis cycle. We evaluate two different ensemble generation methods—one using a 4D-ETKF with an analysis window identical to 4D-Var, and another using a rapid-updating 3D-ETKF (RU-ETKF) that is applied multiple times within the 4D-Var analysis window. We also compare various strategies for localization across component models, including none (i.e., SCDA), restriction to component domains (i.e., WCDA), restriction to individual geophysical variables, and restriction to individual modes. We further use two implementations of 4D-Var, one using a numerical implementation of the analytic coupled atmosphere–ocean TLM and adjoint (which are available for our coupled QG model), and another using the regression-based ETLM/LETLM (to address scenarios in which the TLM and adjoint are not available).

This study is intended to provide further understanding to the application of 4D-Var for SCDA. The contents are structured as follows: section 2 describes the methodology. Section 3 outlines the experiment design, and section 4 presents the results. The conclusions are provided in section 5.

2. Methods

a. The Modular Arbitrary-Order Ocean–Atmosphere Model

We use the Modular Arbitrary-Order Ocean–Atmosphere Model (MAOOAM; De Cruz et al. 2016) for all experiments in this study. This model had its origins in the Ocean–Atmosphere–Quasi-Geostrophic–Wind Stress model (OA-QG-WS v1; Vannitsem 2014), which coupled a QG atmospheric model (Charney and Straus 1980) with a QG ocean model (Pierini 2011). This earlier coupled model used 12 atmospheric and

4 oceanic modes. This was then extended to 20 atmospheric modes in OA-QG-WS v2 (Vannitsem and De Cruz 2014) to include model variables from Reinhold and Pierrehumbert (1982) and increase model resolution. This was followed by the VDDG model (Vannitsem et al. 2015), which extended the range for coupled atmosphere–ocean predictions by producing low frequency variability in the atmosphere and ocean. MAOOAM adapts the VDDG model by expanding the dynamical fields to allow for an arbitrary number of basis functions (or modes). The model includes a two-layer QG atmospheric component coupled both thermally and mechanically with a single layer QG ocean component. The coupling between the two components includes wind forcings, radiative and heat exchanges.

The motion for the atmospheric streamfunction fields ψ_a^1 at 250 hPa and ψ_a^3 at 750 hPa, and the vertical velocity equations are defined by

$$\frac{\delta}{\delta t}(\nabla^2 \psi_a^1) + J(\psi_a^1, \nabla^2 \psi_a^1) + \beta \frac{\delta \psi_a^1}{\delta x} = -k'_d \nabla^2(\psi_a^1 - \psi_a^3) + \frac{f_0}{\Delta p} \omega, \quad (1)$$

$$\frac{\delta}{\delta t}(\nabla^2 \psi_a^3) + J(\psi_a^3, \nabla^2 \psi_a^3) + \beta \frac{\delta \psi_a^3}{\delta x} = +k'_d \nabla^2(\psi_a^1 - \psi_a^3) - \frac{f_0}{\Delta p} \omega - k_d \nabla^2(\psi_a^3 - \psi_o). \quad (2)$$

Here, the Coriolis parameter (f) is linearized around the value f_0 . This is estimated at the $\phi_0 = 45^\circ$. The parameters k'_d and k_d quantify the friction between the atmospheric layers, and between the ocean surface and the atmosphere, respectively.

The motion for the streamfunction of the ocean layer (ψ_o), has the following equations:

$$\frac{\delta}{\delta t} \left(\Delta^2 \psi_o - \frac{\psi_o}{L_R^2} \right) + J(\psi_o, \Delta^2 \psi_o) + \beta \frac{\delta \psi_o}{\delta x} = -r \Delta^2 \psi_o + \frac{C}{\rho h} \Delta^2(\psi_a^3 - \psi_o). \quad (3)$$

The reduced Rossby deformation radius (L_R), the density (ρ), the depth (h), and the friction (r) are represented at the bottom of the active ocean layer. The impact of the wind stress is represented by the right-hand side of the equation. This is modulated by the drag coefficient of the mechanical ocean–atmosphere coupling, $d = c/(\rho h)$. Terms T_a and T_o represent the time evolution of the atmosphere and ocean temperatures, respectively. These are given by the following equations:

$$\gamma_a \left[\frac{\delta T_a}{\delta t} + J(\psi_a, T_a) - \sigma \omega \frac{p}{R} \right] = -\lambda(T_a - T_o) + \epsilon_a \sigma_B T_a^4 - 2\eta_a \sigma_B T_a^4 + R_a, \quad (4)$$

$$\gamma_o \left[\frac{\delta T_o}{\delta t} + J(\psi_o, T_o) \right] = -\lambda(T_o - T_a) - \sigma_B T_o^4 + \epsilon_o \sigma_B T_a^4 + R_o. \quad (5)$$

In these equations, the heat capacities of the atmosphere and active ocean layers are given by γ_a and γ_o . The atmospheric

barotropic streamfunction is given by $\psi_a = (1/2)(\psi_a^1 + \psi_a^3)$. The ocean–atmosphere interface heat transfer coefficient is given by γ , with σ representing the (constant) static stability of the atmosphere. A more comprehensive description on the MAOOAM model can be found in De Cruz et al. (2016).

Here, we symbolize the coupled dynamical system with the mapping:

$$\mathbf{x}_t = \mathcal{M}_t(\mathbf{x}_{t-1}), \quad (6)$$

where \mathcal{M}_t is the propagation of the dynamics over one model time step dt , and \mathbf{x}_t is a state vector of time-dependent coefficients, both defined at time t . We use the configuration of Vannitsem and Lucarini (2016), which has spectral truncation levels of 2×2 for the atmosphere, and 2×4 for the ocean. The coupled drag coefficient for this configuration produces fairly strong low frequency oscillations in the coupled dynamics. The state vector with this configuration has $N_a = 20$ and $N_o = 16$ spectral modes for the atmosphere and ocean, respectively. Both N_a and N_o comprise streamfunction (θ) and temperature (ψ) modes. One model time unit (MTU) corresponds to about 2.7 h, while our model time step $dt = 0.37152$ corresponds to about 1 h.

b. Background error covariances

We implement 4D-Var with three different background error covariance estimates: 1) the traditional static climatological background error covariance \mathbf{B}_c , 2) a fully flow-dependent background error covariance \mathbf{B}_f estimated from a cycled ETKF running in tandem with the 4D-Var, and 3) a hybrid background error covariance \mathbf{B}_h defined as

$$\mathbf{B}_h = \alpha \mathbf{B}_c + (1 - \alpha) \mathbf{B}_f. \quad (7)$$

The static \mathbf{B}_c can be recovered using $\alpha = 1$, while the fully flow-dependent \mathbf{B}_f can be recovered using $\alpha = 0$. In our hybrid experiments, we use $\alpha = 0.5$, which is in line with Penny et al. (2015) and Bonavita et al. (2015).

We generate our climatological \mathbf{B}_c matrix from a 30-yr data assimilation “reanalysis” using the RU-ETKF (this will be discussed in section 2e). The ensemble perturbations are averaged over the time domain in the following form:

$$\mathbf{B}_c = \langle \mathbf{X}_b^T(t) \cdot \mathbf{X}_b(t) \rangle, \quad (8)$$

where the columns of $\mathbf{X}_b(t)$ contain the ensemble perturbations produced by the RU-ETKF 6-hourly forecasts at each time t , and the statistics are averaged over all t .

c. Localization

Localization methods are designed to reduce or omit spurious correlations in the background error covariance matrix \mathbf{B} . Typically, it is assumed that distant correlations in space are less accurate when using small sample sizes, and improve as the sample size increases. However, the concept of localization can be applied more generally to manually reduce correlations that are either known or expected to have lower accuracy. Such approaches have been applied implicitly in the design of WCDA systems.

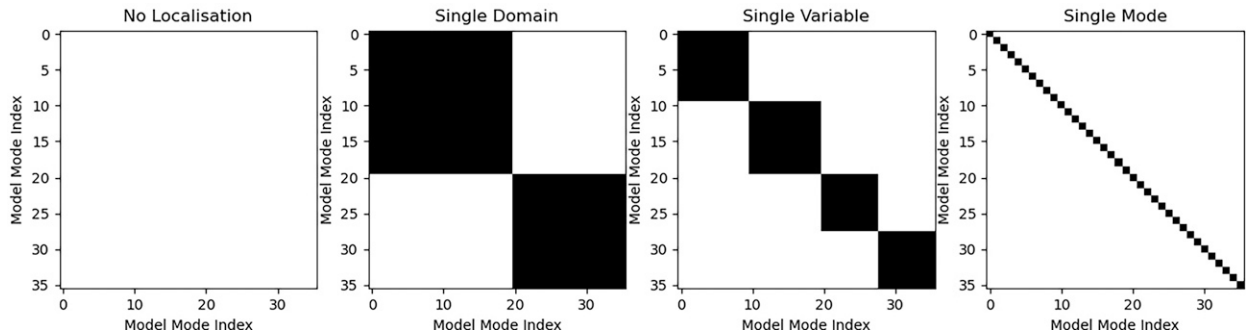


FIG. 1. Each subplot shows how each localization structure looks. “No Localization” treats the coupled model as a single dynamical system (i.e., SCDA). Localization within a “Single Domain” separates the atmosphere and ocean, while cross-domain interactions are ignored (i.e., WCDA). “Single Variable” focuses only on temperature or streamfunction within each domain. “Single Mode” focuses on individual model modes. Focus cells are shown in black.

Localization is applied to the background error covariance matrix as

$$\mathbf{B} = \mathcal{L} \circ \mathbf{B}_x, \quad (9)$$

where \mathcal{L} is the localization matrix, $x \in \{c, f, h\}$, and \circ represents the Schur (element-wise) product. We use the following localization approaches (see Fig. 1 for a visual representation):

- None: We implement SCDA by omitting any localization of the \mathbf{B} matrix, keeping all atmosphere–ocean cross covariances.
- Single domain: We implement WCDA by applying localization to the \mathbf{B} matrix so that the atmosphere–ocean cross covariances are set to zero.
- Single variable: In this case, we localize the \mathbf{B} matrix so that all the cross covariances between variable types (i.e., $\psi_a, \theta_a, \psi_o, \theta_o$) are set to zero.
- Single mode: In this case, we localize the \mathbf{B} matrix so that it is diagonal, with all covariances between individual variables set to zero.

d. Four-dimensional variational data assimilation (4D-Var)

Four-dimensional variational (4D-Var; Talagrand and Courtier 1987) DA attempts to fit an optimal trajectory to all observations within a short analysis window by adjusting the initial conditions of a numerical forecast model. The 4D-Var algorithm is formulated by minimizing a cost function that is defined over a short time window. Its individual terms are weighted using the estimated error characteristics of the forecast dynamics and the observations. Its “strong constraint” form assumes a perfect model representation of the dynamics, which we use in conjunction with the incremental formulation of 4D-Var (Courtier et al. 1994). The basic formulation of the incremental 4D-Var cost function is implemented with the initial state increment $\delta\mathbf{x} = \mathbf{x}_0 - \mathbf{x}_0^b$ as the control variable:

$$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x})^T \mathbf{B}^{-1}(\delta\mathbf{x}) + \frac{1}{2} \sum_{i=0}^N [\mathbf{H}(\delta\mathbf{x}) - \mathbf{d}_i]^T \mathbf{R}_i^{-1} [\mathbf{H}(\delta\mathbf{x}) - \mathbf{d}_i]^T. \quad (10)$$

Here, \mathbf{x}_0^b represents the initial background state (i.e., a forecast initialized from the previous analysis cycle), while \mathbf{B} is the background error covariance, and \mathbf{R}_i is the observation error covariance given at time $t = t_i$; $\mathbf{d}_i = \mathbf{y}_i - \mathcal{H}_i(\mathbf{x}_i^b)$ is the innovation (or departure), and \mathbf{H} is the linearized observation operator. For each time t_i , the nonlinear model operator \mathcal{M}_i propagates the state \mathbf{x}_i from time $t = t_0$ to $t = t_i$, such that $\mathbf{x}_i = \mathcal{M}_i(\mathbf{x}_0)$. The observation vector \mathbf{y}_i , which models the spectral modes directly, and the observation operator \mathcal{H}_i , are both given at time $t = t_i$.

Because 4D-Var is a nonlinear optimization, it is typically solved via nested loops. In the outer loop, a nonlinear model trajectory is generated, and the innovation $\{\mathbf{y}_i - \mathcal{H}_i[\mathcal{M}_i(\mathbf{x}_{i-1}^b)]\}$ at each observation time within the analysis window is calculated, where \mathbf{x}^b represents the initial, and updated, background trajectory. In the inner loop, the 4D-Var cost function is minimized using the linearized dynamics in order to find an improved estimate of the initial conditions that allow the forecast model to best fit the observations. In this inner loop, we optimize the cost function $J(\delta\mathbf{x})$ by setting the gradient $\nabla J_{\delta\mathbf{x}}$ equal to zero:

$$\nabla J(\delta\mathbf{x}^L) = \mathbf{B}^{-1}(\delta\mathbf{x}^L) - \sum_{i=0}^N \mathbf{M}_i^T \mathbf{H}_i^T \mathbf{R}_i^{-1} [\mathbf{H}(\delta\mathbf{x}^L) - \mathbf{d}_i]^T = 0, \quad (11)$$

where \mathbf{M} is TLM and \mathbf{M}^T is the adjoint. We use the BIConjugate Gradient Stabilized iterative method (BI-CGSTAB; van der Vorst 1992), which solves Eq. (11) after converting it to the form $\mathbf{A}\mathbf{x} = \mathbf{b}$. Here, $\delta\mathbf{x}^L$ represents the increment between the analysis and background states ($\mathbf{x} - \mathbf{x}^b$) and \mathbf{H} is the linearized observation operator, where L is the outer loop counter. After convergence, the inner loop is complete and the outer loop nonlinear operations are repeated from the new improved estimate of the initial conditions. The outer loop reruns a background trajectory from the new initial condition solution. This is then re-minimized to update the initial condition. Each variational method uses three outer loops.

e. Ensemble transform Kalman filter (ETKF)

The ensemble transform Kalman filter (ETKF) is an ensemble data assimilation method which updates an ensemble

of model trajectories given sequential observations. We use two versions of the ETKF, where one is updated over an assimilation cycle, and the other is sequentially updated. This is investigated to determine whether taking all observations, or just a singular observation, greatly impacts the results in CDA. The first is a 4D-ETKF as presented by (Hunt et al. 2004, 2007; Harlim and Hunt 2007), but without spatial localization. The 4D-ETKF uses an analysis cycle window that matches that of the 4D-Var. This version of the ETKF assimilates all observations over an analysis cycle by forming \mathbf{Y}_b , which concatenates observations throughout the analysis cycle window. We also consider an alternative ETKF, which we call the “rapid update” ETKF (RU-ETKF). This is a 3D-ETKF that performs an analysis every time a new set of observation becomes available, which occurs multiple times within the 4D-Var analysis window.

The analysis weights (\mathbf{w}^a), transform matrix (\mathbf{W}^a), and analysis error covariance matrix (\mathbf{P}^a) are calculated as

$$\mathbf{P}^a = \left[\frac{k-1}{\rho} \mathbf{I} + (\mathbf{Y}^b)^T \mathbf{R}^{-1} \mathbf{Y}^b \right]^{-1}, \quad (12)$$

$$\mathbf{w}^a = \mathbf{P}^a \mathbf{Y}^b \mathbf{R}^{-1} (\mathbf{y}^o - \bar{\mathbf{y}}^b), \quad (13)$$

$$\mathbf{W}^a = [(k-1) \mathbf{P}^a]^{1/2}. \quad (14)$$

The matrix $\mathbf{Y}^b = \mathbf{H}\mathbf{X}^b$ represents the background ensemble perturbations transformed to the observation space, with \mathbf{y}^o representing the vector of observed quantities, and $\bar{\mathbf{y}}^b = \mathbf{H}\bar{\mathbf{x}}^b$ representing the ensemble mean state transformed to the observation space. Matrices \mathbf{I} and \mathbf{P}^a represent the identity and the ensemble analysis covariance matrix, respectively. The parameter $\rho = 0.01$ is a multiplicative inflation factor that can be used to manually increase the estimated background error covariance [in line with Penny et al. (2019)], and k represents the number of ensemble members. The analysis mean and analysis ensemble members are then computed as follows:

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{X}^b \mathbf{w}^a, \quad (15)$$

$$\hat{\mathbf{X}}^a = \mathbf{X}^b \mathbf{W}^a + \bar{\mathbf{x}}^a \mathbf{1}^T, \quad (16)$$

where the columns of $\hat{\mathbf{X}}^a$ provide the new ensemble of analysis state vectors that can be used as initial conditions for a subsequent forecast. Here, the $\mathbf{1}$ refers to a column vector of ones.

f. Statistical tangent linear model

The TLM and its adjoint are important tools used by 4D-Var to formulate a computationally tractable approach for nonlinear optimization. In practice, the need for a TLM and adjoint can be problematic when they are not readily available as software, and due to the suboptimal availability, TLM and adjoint models can be inaccurate. For situations like these, a statistical implementation to the TLM and adjoint may prove to be a useful alternative. Here, we will consider a statistical TLM generated from an online ensemble DA procedure, which we call the ETLM. Many operational weather centers already use ensemble DA methods, either explicitly or for the purpose of generating

hybrid background error covariance statistics, and so such ensemble statistics are typically available in practice.

As a linear transformation exists so that an ensemble of perturbations \mathbf{X}_{t-1} of a nonlinear state \mathbf{x}_{t-1} can be propagated from time $(t-1)$ to time t :

$$\mathbf{X}_t = \mathbf{M}_{(t,t-1)} \mathbf{X}_{t-1}, \quad (17)$$

$$\mathbf{X}_t \mathbf{X}_{t-1}^T = \mathbf{M}_{(t,t-1)} \mathbf{X}_{t-1} \mathbf{X}_{t-1}^T, \quad (18)$$

$$\mathbf{M}_{(t,t-1)} = \mathbf{X}_t \mathbf{X}_{t-1}^T [\mathbf{X}_{t-1} \mathbf{X}_{t-1}^T]^{-1}. \quad (19)$$

The matrices \mathbf{X}_t and \mathbf{X}_{t-1} contain column-wise the ensemble perturbations at times t and $t-1$, respectively. Here, the initial ensemble perturbations are given by taking the ensemble of model states, removing the ensemble mean, and dividing by the standard deviation for each ensemble member (Allen et al. 2017) such that

$$\mathbf{X}_t = \text{diag} \left(\frac{1}{\sigma^{\text{cv}}} \right) (\mathbf{X}_t^b - \bar{\mathbf{x}}_t^b \mathbf{1}^T), \quad (20)$$

where σ^{cv} is the vector of climatological standard deviation for each model variable (see Fig. 2).

Through a reformulation, the ETLM can also be computed by applying the matrix inversion in the ensemble subspace, which may be preferred when the ensemble size is smaller than the system dimension:

$$\mathbf{M}_{(t,t-1)} = \mathbf{X}_t [\mathbf{X}_{t-1}^T \mathbf{X}_{t-1}]^{-1} \mathbf{X}_{t-1}^T. \quad (21)$$

The LETLM was developed and demonstrated by Frolov and Bishop (2016), Bishop et al. (2017), and Allen et al. (2017) as a practical implementation of the ETLM for operational DA systems in which the number of degrees of freedom of the model far exceeds the feasible ensemble size. The difference between the LETLM and ETLM is the localization variable. This changes the coupling dynamics where the localization of the LETLM uncouples the TLM. The method we use here adapts the formulation presented by Allen et al. (2017) to apply to isolated spectral modes:

$$\mathbf{M}_{(t,t-1)}(p) = \mathbf{S}_p \mathbf{X}_t [(\mathbf{S}_p^{\text{vol}} \mathbf{X}_{t-1})^T (\mathbf{S}_p^{\text{vol}} \mathbf{X}_{t-1})]^\dagger (\mathbf{S}_p^{\text{vol}} \mathbf{X}_{t-1})^T, \quad (21)$$

where $\mathbf{M}_{(t,t-1)}(p)$ denotes the LETLM matrix determining the linear propagation of small perturbations from time $t-1$ to t for spectral mode p . The matrix operators \mathbf{S}_p and $\mathbf{S}_p^{\text{vol}}$ determine which spectral mode is selected (which are of the size $1 \times$ model dimension size). The latter additionally determines its influence range, which is done by selecting the mode and influential modes surrounding it. The notation $[\mathbf{A}]^\dagger$ indicates the pseudoinverse of matrix \mathbf{A} . This can be reformulated as

$$\mathbf{M}_{(t,t-1)}(p) = \mathbf{S}_p \mathbf{X}_t (\mathbf{S}_p^{\text{vol}} \mathbf{X}_{t-1})^T [(\mathbf{S}_p^{\text{vol}} \mathbf{X}_{t-1}) (\mathbf{S}_p^{\text{vol}} \mathbf{X}_{t-1})^T]^{-1}, \quad (23)$$

in the same manner as that of Eqs. (19) and (21). This reformulation reduces the matrix dimension for the matrix inversion operation. In Eq. (22), the pseudoinverse is taken of a

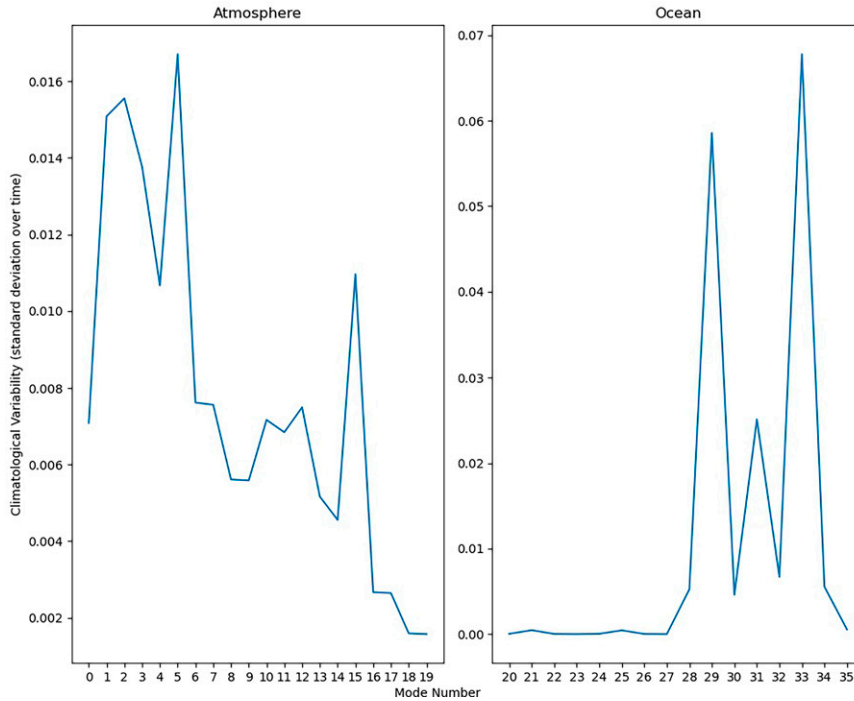


FIG. 2. The climatological variability σ_{cv} of (left) the atmosphere and (right) the ocean at for each variable, calculated over a 30-yr model integration.

$k \times k$ matrix (where k is the ensemble dimension), while Eq. (23) transforms this into a scalar inverse.

To propagate perturbations using the LETLM matrix we use

$$\delta \mathbf{x}_t = \sum_{p=1}^N [\mathbf{S}_p]^T \mathbf{M}_{(t,t-1)}(p) [\mathbf{S}_p^{\text{vol}} \delta \mathbf{x}_{t-1}], \quad (24)$$

where $\delta \mathbf{x}_t$ is a perturbation at time t where the equation is cycled over each mode via a summation. As the notion of spatial locality is eliminated in the spectral formulation of the dynamics, we consider the localization operations \mathbf{S}_p and $\mathbf{S}_p^{\text{vol}}$ as an abstraction of the localization concept, and simply select the current spectral mode without any influence from other spectral modes.

3. Design of experiments

We conduct experiments exploring a range of configurations of the 4D-Var CDA. We use two control experiments, 4D-Var and 3D-Var-FGAT (Lee et al. 2004), both using a climatological background error covariance matrix. The 3D-Var-FGAT is implemented by replacing the analytical TLM in the traditional 4D-Var with the identity matrix. These control experiments provide a point of reference for all further experiments.

For all experiments, we use a 6-h analysis cycle (unless stated otherwise). The motivation behind this analysis cycle length comes from Fisher et al. (2011). This comes from two sides, the first being that 4D-Var relies on long windows lengths so the tangent linear model, and its adjoint, can generate flow dependence between innovations and the initial background error covariance matrix, the second being that this is NOAA's operational analysis

time scale. Each analysis cycle contains observations sampled at 6 separate time steps, one time per hour, based on the model time step of $dt = 0.37152$ corresponding to approximately 1 h. Results are shown using the root-mean-square error (RMSE), as a percentage of the climatological variation (CV) of each model variable (which is calculated as the longtime time standard deviation of the corresponding variable), of the full analysis trajectory against a nature run. Excluding the final experiments, all methods observe all spectral modes directly, with observation errors of $0.1 \times \sigma_{cv}$.

Our investigation commences by examining how the formulation of the coupled atmosphere–ocean background error covariance matrix impacts the 4D-Var CDA. We compare 4D-Var and 3D-Var using climatological background covariance matrices generated from different time periods to evaluate the sensitivity of the variational methods to changes in \mathbf{B}_c . We introduce noise to the climatological background error covariance matrix by generating one estimate over a short time frame (1 year), and the other over a longer time frame (30 years). We also examine flow-dependent and hybrid background error covariance matrices (Rodwell et al. 2016) to investigate how incorporating temporal changes in the background covariances can have an impact on results. The flow-dependent error covariance \mathbf{B}_f is evaluated using two different ensemble generation methods (RU-ETKF, and 4D-ETKF). All ensemble-based experiments use 40 ensemble members, except for a sensitivity study examining the impact of reducing the ensemble size (40, 20, 10). The ensemble generation method and ensemble size affect the generation of \mathbf{B}_f , and as a consequence also affect \mathbf{B}_n . We also examine the impact of localization on the coupled atmosphere–ocean background error

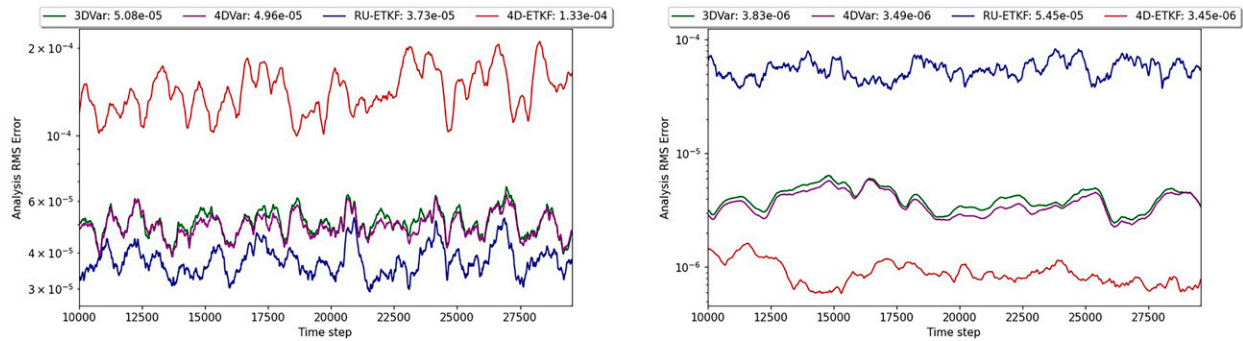


FIG. 3. The analysis RMSE for all (left) atmosphere modes and (right) ocean modes, observing the full coupled state. The following control methods are shown: 3D-Var (green) and analytical 4D-Var (purple) both using the long-range \mathbf{B}_c , the RU-ETKF (blue), and the 4D-ETKF (red). Analysis RMSE (temporal mean shown at the top of each subplot) using a moving average of 442 time steps is shown for the last 20000 model time steps.

covariance matrix by using a range of different localization strategies: 1) none, which emulates SCDA; 2) localizing to only the atmosphere or ocean (ensuring no atmosphere–ocean interaction), which emulates WCDA; 3) by variable, with no interaction between different model variable types; and 4) single mode, with no interaction between modes.

The statistical ETLM and LETLM surrogate models are explored in order to provide a potential alternative for implementing 4D-Var CDA in cases where an analytic or software-based TLM and adjoint are not available. To generate ensemble statistics for the ETLM methods, we use either the cycled RU-ETKF or the 4D-ETKF, both using ensemble forecasts with 40 members.

Our final experiments examine the impact of reduced observation coverage. We consider extreme scenarios where observations are available either only in the atmosphere or only in the ocean. We consider the scenario in which there are only atmospheric observations to be most consistent with the real-world scenario, in which there are plentiful satellite observations of the atmosphere but limited in situ observations of the subsurface ocean. The reverse case is provided for completeness.

4. Results

a. Control experiments

We first examine the control experiments (Fig. 3). We include results from the 3D-Var and 4D-Var, both using the long-range climatological \mathbf{B}_c , as well as the 4D-ETKF and the RU-ETKF ensemble-based methods, compared over 5000 assimilation cycles (6 h per cycle, approximately 3.4 years total) on a strongly coupled system. The 4D-ETKF, which assimilates observations indexed at multiple times within the analysis cycle window, produces the lowest RMSE in the ocean. In contrast, the RU-ETKF, which assimilates observations every time they occur within the analysis cycle window, has the lowest RMSE in the atmosphere. We emphasize that the primary difference between these is the update frequency, and that the lower frequency updates of the 4D-ETKF may favor the slower time scales corresponding to the ocean dynamics,

while the RU-ETKF may favor the faster time scales present in the atmospheric dynamics. The variational methods produce slightly higher RMSE (approximately 33% in the atmosphere and 1% in the ocean) compared to the leading ETKF method in each domain.

b. Impact of background error covariance on 4D-Var CDA

We find that the 4D-Var CDA can be very sensitive to the specification of the background error covariance matrix. Comparisons between 4D-Var using a background error covariance matrix generated from a short training dataset versus another generated from a longer training dataset indicate that the shorter training set is sufficient for the fast-time-scale atmospheric error dynamics. However, this is not sufficient for the ocean, which benefits from the use of the longer training set.

In the atmosphere, the RU-ETKF and 4D-ETKF control methods produce the lowest and highest RMSE as a percentage of CV, respectively (Fig. 4). A summary of the hybrid and flow-dependent experiments can be found in Table 1. It appears that the tuning of the \mathbf{B}_c matrix is particularly important for the slower components of the coupled system. The control methods using \mathbf{B}_c generated from either short or long training datasets have similar (within 10%) RMSE for the atmosphere, while for the ocean the RMSE is reduced by about one order of magnitude due to the use of the longer training dataset. The use of the longer training dataset in the ocean also improves the accuracy of the variational methods to a degree that they outperform the 4D-ETKF.

The hybrid methods are impacted by both the length of the training dataset for \mathbf{B}_c and also the type of ETKF used for dynamic ensemble generation. The use of the RU-ETKF as the dynamic ensemble generator reduces the RMSE in the atmosphere for the variational methods. This applies for both formulations of \mathbf{B}_c .

Overall, the most accurate methods in the atmosphere are the RU-ETKF and the 4D-Var- \mathbf{B}_h^{RU} using the long training method. Though the 4D-ETKF underperforms the RU-ETKF in this case, it still improves the RMSE of the variational methods when used as an ensemble generator when compared to the

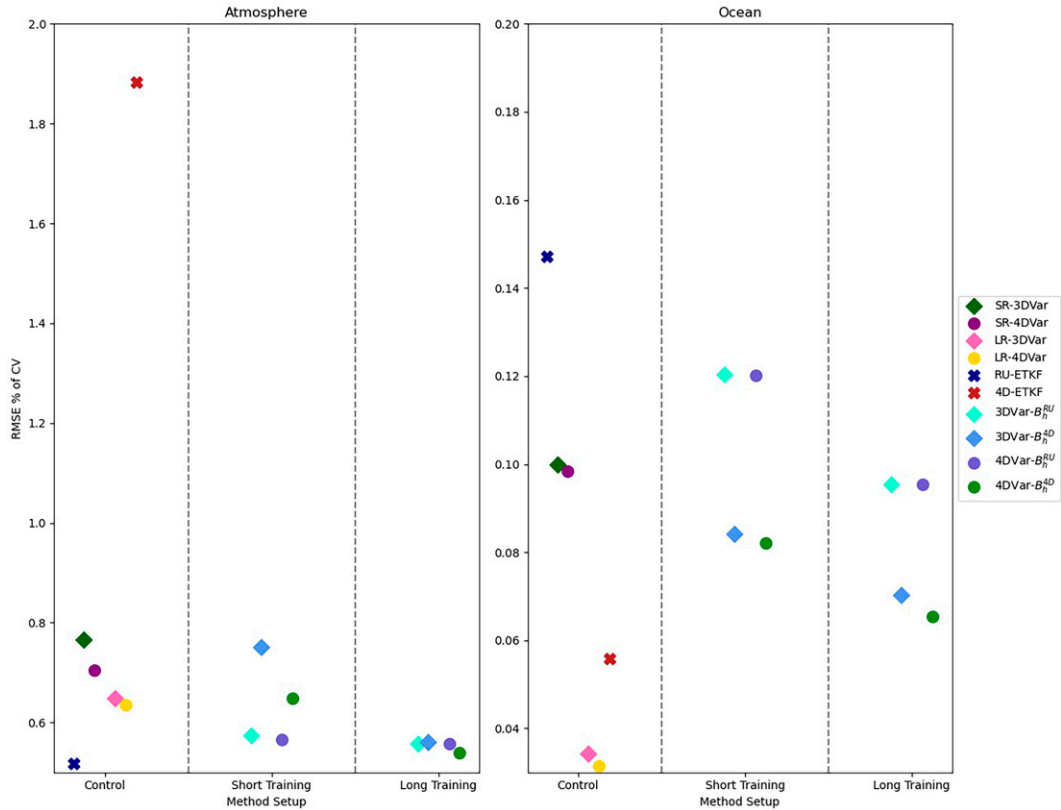


FIG. 4. The (30-yr) analysis RMSE as a percentage of climatological variation for long and short training datasets to generate \mathbf{B}_c for (left) the atmosphere and (right) the ocean, observing the full coupled state. In each plot, the left column contains the control methods, and the second and third columns contain the short and long training methods at time scales of 1 and 30 years, respectively. Here, the long-trained \mathbf{B}_c matrix outperforms the short-trained \mathbf{B}_c in both the atmosphere and the ocean. The 4D-Var methods are given by circles, the 3D-Var methods are given by the diamonds, and the ETKFs are given by the crosses. SR- and LR- reference the short-range and long-range \mathbf{B}_c respectively.

variational control methods which use \mathbf{B}_c . An example of how the ensemble background covariance matrices compare with the climatological background correlation matrix can be found in Fig. 5. Here, from left to right we show the correlation matrices given by the RU-ETKF, the 4D-ETKF, and the climatology. This shows the differences in how the RU-ETKF and 4D-ETKF generate correlations (and thus the \mathbf{B} matrix).

As the 4D-ETKF is more accurate for the ocean, the hybrid methods using the 4D-ETKF as an ensemble generator have a lower RMSE than those using the RU-ETKF in the ocean.

Though the hybrid methods do outperform the control methods using \mathbf{B}_c trained on the shorter training dataset, the methods that perform best for the ocean are the 3D-Var and 4D-Var controls using \mathbf{B}_c trained on the longer dataset, and the 4D-Var- \mathbf{B}_h^{4D} hybrid method. As the \mathbf{B}_c matrices computed from the short and long training datasets perform similarly in the atmosphere, but the longer training datasets produce more accurate results for the ocean, all remaining experiments will use the longer training set for determining the climatological error covariance \mathbf{B}_c .

TABLE 1. The notation for the background error covariance matrix in hybrid and flow-dependent configurations 3D/4D-Var.

Ensemble generator	Data assimilation method			
	3D-Var		4D-Var	
	Background error covariance formulation			
	Hybrid	Flow dependent	Hybrid	Flow dependent
4D-ETKF	3D-Var- \mathbf{B}_h^{4D}	3D-Var- \mathbf{B}_f^{4D}	4D-Var- \mathbf{B}_h^{4D}	4D-Var- \mathbf{B}_f^{4D}
RU-ETKF	3D-Var- \mathbf{B}_h^{RU}	3D-Var- \mathbf{B}_f^{RU}	4D-Var- \mathbf{B}_h^{RU}	4D-Var- \mathbf{B}_f^{RU}

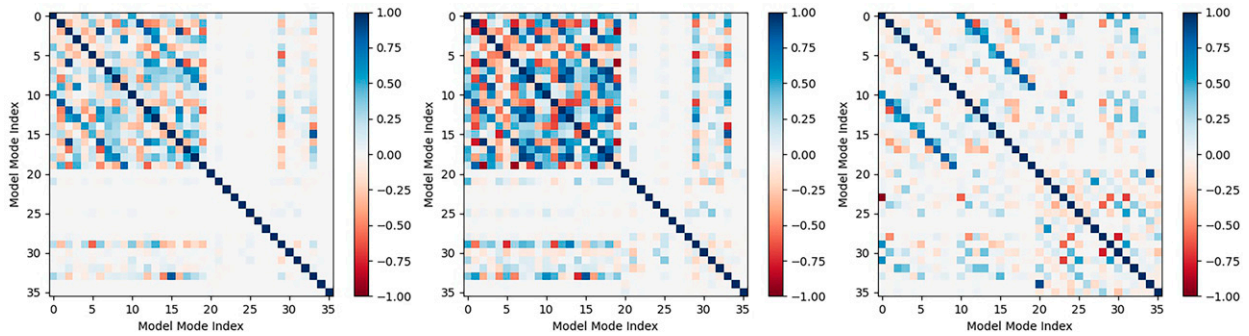


FIG. 5. Example background error correlation matrices after 100 assimilation cycles generated by three different methods: (left) the ensemble method using the RU-ETKF, (center) the ensemble method using the 4D-ETKF, and (right) the climatological background error covariance.

c. Impact of ensemble size on 4D-Var CDA

Generally, a larger sample size of ensemble members will improve the accuracy of ensemble-based DA methods. However, running large ensembles of complex numerical forecast models is computationally expensive. It is typically only feasible to run an ensemble of numerical forecast integrations whose size is only a small fraction of the model dimension. In

Fig. 6, we compare the ensemble generators RU-ETKF and 4D-ETKF in our variational hybrid methods with 10, 20, and 40 ensemble members. As noted by Vannitsem and Lucarini (2016), the number of nonnegative Lyapunov exponents of this MAOOAM configuration is about 20, which implies 10 members is insufficient for the ETKF methods, 20 members should be sufficient [as shown by

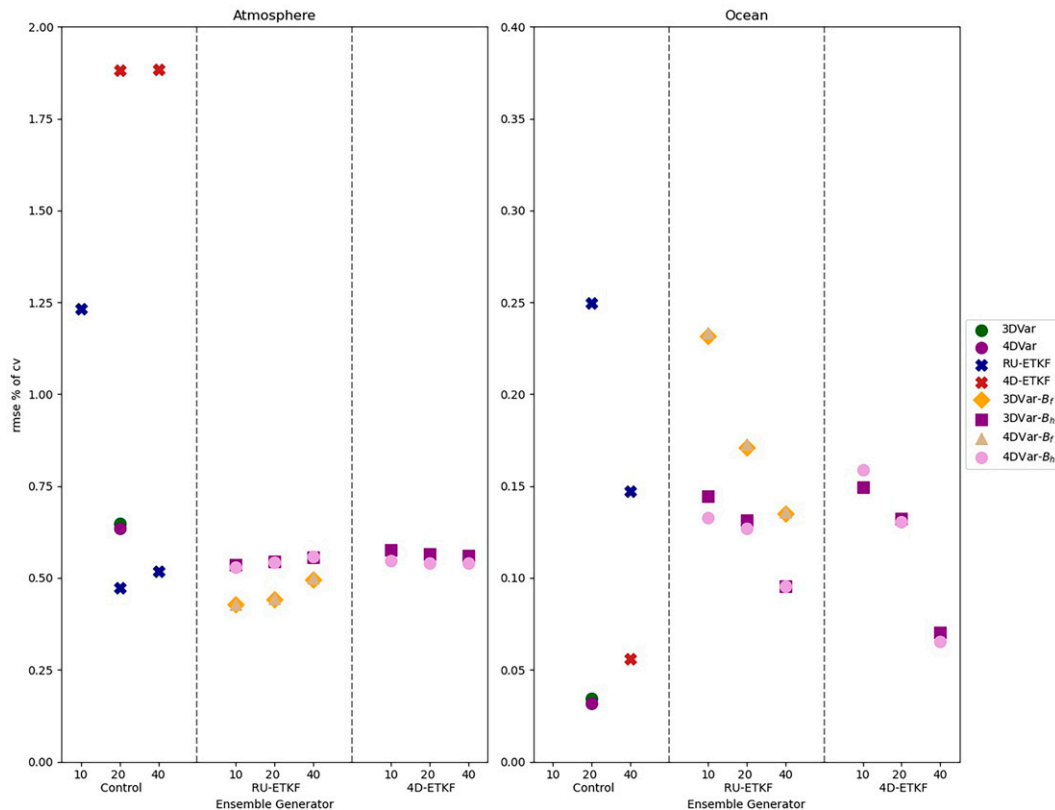


FIG. 6. The analysis RMSE as a percentage of climatological variation for 3D/4D-Var using the RU/4D-ETKF with different ensemble sizes for (left) the atmosphere and (right) the ocean, observing the full coupled state. The B_f methods when generated by the 4D-ETKF diverge, while the other methods decrease in RMSE as the ensemble size decreases. Only the best performing methods are shown in the plot to highlight the differences in the most accurate methods.

Penny et al. (2019)], and 40 members should guarantee that the ETKF is nondivergent. Here, our variational methods use either a climatological \mathbf{B}_c , the hybrid \mathbf{B}_h , or a fully flow-dependent \mathbf{B}_f .

Comparing the control ETKFs, the RU-ETKF has a lower RMSE in the atmosphere, and the 4D-ETKF has a lower RMSE in the ocean. In the atmosphere, the RU-ETKF has higher accuracy than both the 3D/4D-Vars when using either 20 or 40 ensemble members. We also notice the 4D-ETKF diverges with 10 ensemble members. In the ocean, the 4D-ETKF is the most accurate ETKF with 40 ensemble members, followed by the RU-ETKF with 40 ensemble members. Comparisons between the atmosphere and the ocean for ensemble based CDA shows the atmosphere needs only 20 members for accurate results, while the ocean further benefits from a larger ensemble size.

Comparing the variational methods in the atmosphere, we first note that the 3D/4D-Var using \mathbf{B}_f^{4D} generated with the 4D-ETKF have a large RMSE in comparison to the other methods (over 10% of the climatological variation). When instead using the RU-ETKF as the ensemble generator, we see that the 3D/4D-Var using a flow-dependent \mathbf{B}_f^{RU} has a lower RMSE than the 3D/4D-Var using the hybrid \mathbf{B}_f^{RU} . This \mathbf{B}_f^{RU} more accurately describes the system dynamics compared to when the \mathbf{B}_c was introduced within a hybrid covariance. The performance of 3D/4D-Var in the atmosphere is not noticeably different across any of the ensemble sizes tested.

In the ocean, where the 4D-ETKF ensemble generator outperforms the RU-ETKF, we see the 3D/4D-Var using the hybrid \mathbf{B}_f^{4D} results in the lowest RMSE out of all the hybrid methods. The 3D/4D-Var using flow-dependent \mathbf{B}_f^{4D} still significantly underperforms compared to all other methods, which implies that the 4D-ETKF needs the information from the climatological \mathbf{B}_c , even in the ocean in order to incorporate missing system dynamics. Comparing across different ensemble sizes—as expected, using 40 members produces the lowest RMSE for all hybrid variational methods. For consistency, all remaining experiments will use 40 ensemble members for both the RU-ETKF and 4D-ETKF ensemble generators.

d. Impact of localization on 4D-Var CDA

Localization can have a significant impact on the performance of variational DA methods. We have already noted the importance of the formulation of the background error covariance matrix on 4D-Var CDA. Localization further modifies the background error covariance matrix by zeroing out correlations that might likely be the result of statistical noise.

For experiments in the atmosphere, we see no discernible difference between strong (SCDA) and weak (WCDA) localization, but the RMSE increases as the localization becomes stricter. There appears to be no improvement in the atmospheric RMSE due to cross-domain localization (Fig. 7). Given that strong (SCDA) and weak (WCDA) localization produce similar results for variational methods, this implies

that the cross-system covariances have minimal impact on the RMSE in the atmosphere. When introducing more “strict” localization, for example by removing cross-variable or cross-modal covariance, the RMSE increases. Thus, there are cross correlations between variables and between individual modes that are needed to produce the most accurate background error covariance matrix. Comparing these localized variational methods to the control methods, both variational methods are improved by the majority of the localization approaches in conjunction with a hybrid or flow-dependent \mathbf{B} matrix. Though the RU-ETKF is more accurate than 3D/4D-Var using \mathbf{B}_c in the atmosphere, the hybrid methods produce results that are either similar to or outperform the RU-ETKF. Overall, in the atmosphere, the 3D/4D-Var using \mathbf{B}_f^{RU} have the lowest RMSE when using strong (SCDA) or weak (WCDA) localization.

For the ocean experiments, weak (WCDA) localization appears to reduce RMSE compared to strong (SCDA) localization. The stricter localization also results in lower RMSE than the control methods and the strong (SCDA) localization. This implies that, for the given time scales of the analysis cycle and observing frequency, the cross-domain covariances can degrade the accuracy of the ocean state estimate. However, utilizing the cross-variable interactions between the temperature and streamfunction, and the cross-modal interactions, helps to reduce the RMSE. In the ocean, localization is needed for the hybrid methods to improve the RMSE of the control variational methods. Using the 4D-ETKF as an ensemble generator, the hybrid \mathbf{B}_h^{4D} improves on the variational control methods when using weak/variable/single mode localization. For the ocean, this implies that some cross-system covariances degrades performance. Though localization improves the RMSE in the ocean, using a localization that is stricter than the weak (WCDA) case decreases the accuracy of the methods. When generating the ensemble with the RU-ETKF, the RMSE increases due to the RU-ETKF not performing as well in the ocean as it does in the atmosphere. This gives all RU-ETKF based hybrids a higher RMSE than the variational control methods. It can also be seen that the use of a flow-dependent \mathbf{B}_f in the ocean degrades the performance compared to the hybrid methods.

e. Use of ensemble tangent linear models in 4D-Var CDA

Ensemble-based TLMs are shown here to produce results similar to analytical TLMs in 4D-Var CDA. As previous sections have shown, the RU-ETKF and the 4D-ETKF excel in the atmosphere and ocean, respectively. Consequently, using the more accurate ensemble generator provides a better ETLM/LETLM prediction for the respective domain. The 4D-Var implementations using ETLM and LETLM perform well in comparison to the conventional 3D/4D-Var methods (see Fig. 8). When comparing ETLM/LETLMs, in the both the atmosphere and the ocean, both methods perform similarly to the analytical TLMs across all localization parameters. Just as with Fig. 7, where the flow-dependent background error covariance \mathbf{B}_f generated from the 4D-ETKF does not perform well, the corresponding ETLM/LETLM methods perform about as poorly. Outside of this setup, all

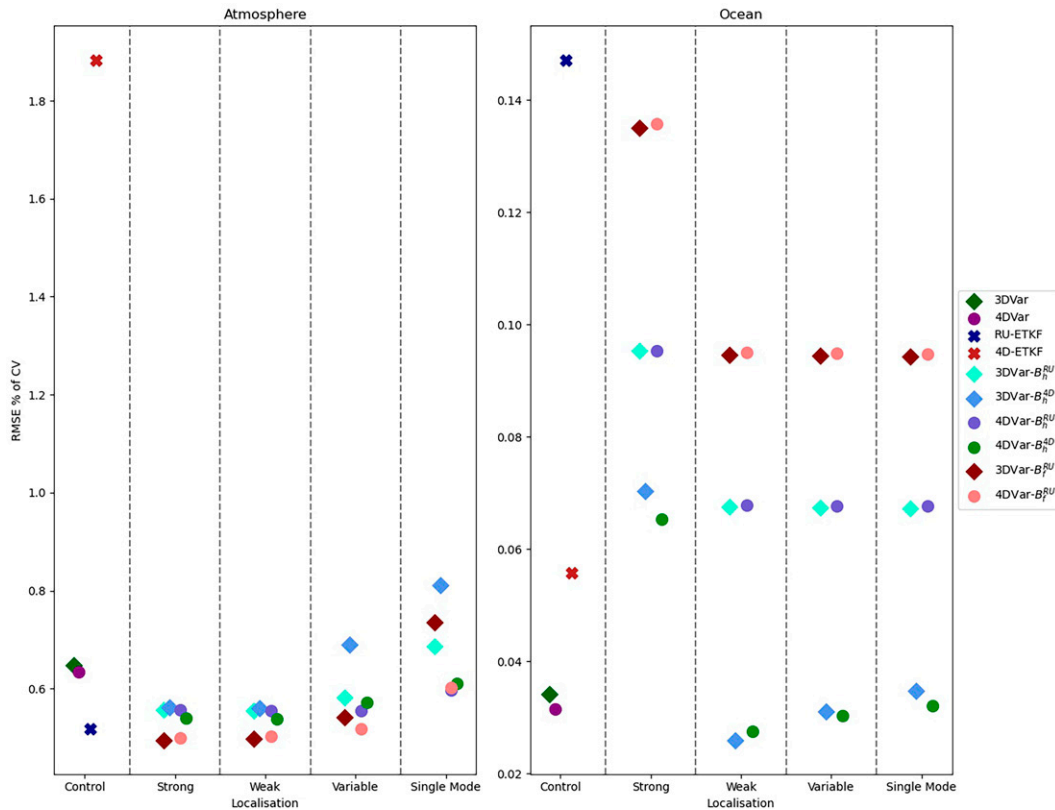


FIG. 7. The analysis RMSE as a percentage of climatological variation for the variational methods under different localization radii for (left) the atmosphere and (right) the ocean, observing the full coupled state. In the atmosphere, the strongly/weakly coupled \mathbf{B}_h^{RU} methods have the lowest RMSE, while when using \mathbf{B}_f^{4D} , the weakly coupled methods have the lowest RMSE. The 4D-Var methods are given by circles, the 3D-Var methods are given by the diamonds, and the ETKFs are given by the crosses. The \mathbf{B}_f^{4D} diverges in the atmosphere and is not shown.

ETLM/LETLM methods give an approximate estimation of the analytical TLMs, giving similar RMSEs when used within 4D-Var. In the atmosphere, the most accurate methods are those using the flow-dependent covariance matrix \mathbf{B}_f^{RU} , under a strong (SCDA) or weak (WCDA) localization setup. In the ocean, the most accurate methods are those using the hybrid background error covariance matrix \mathbf{B}_h^{4D} with weak (WCDA) localization. Overall, the differences in system dynamics have different demands of each method, and so each method has benefits in either system.

f. Impact of sparse observations on 4D-Var CDA

Next we examine the use of a sparser observing system. In practice, there are typically many more observations available of the atmosphere compared to the ocean. Our investigation here will focus on using either only atmosphere observations, as an idealized extreme of the realistic scenario, or only ocean observations. Using this approach, we can test how each CDA method updates the unobserved component of the coupled system. Assimilating only observations in the atmosphere, the RU-ETKF diverges in the ocean, meaning numerical instabilities cause the forecast model solution to “blow-up” (Gottwald and Majda 2013). This in turn has an

effect on all methods that use the RU-ETKF as an ensemble generator (see Fig. 9).

Due to the divergence of the RU-ETKF in the ocean, and the 4D-ETKF being stable and having a lower RMSE, the variational methods in the atmosphere are outperformed by methods which use the 4D-ETKF for ensemble generation. Thus in the atmosphere, 4D-Var methods which use the RU-ETKF have an increased RMSE, this gives some indication of how sensitive the ETLMs and hybrid methods are to the ensemble generator. In terms of localization, atmospheric localization all performed similarly, with weakly coupled localization giving slightly better accuracy. In the ocean, though the 4D-Var implementations using the ETLMs are approximately as accurate as those using the analytical TLM, the results indicate that they are still a viable replacement when a coupled atmosphere–ocean analytical TLM is not available. When only observing atmospheric observations, SCDA underperforms in comparison to the other forms of localization. In general, for this coupled system using only atmospheric observations, using the 4D-ETKF ensemble generator with a weak localization setup can give total system results equivalent to, or better than, the control variational methods.

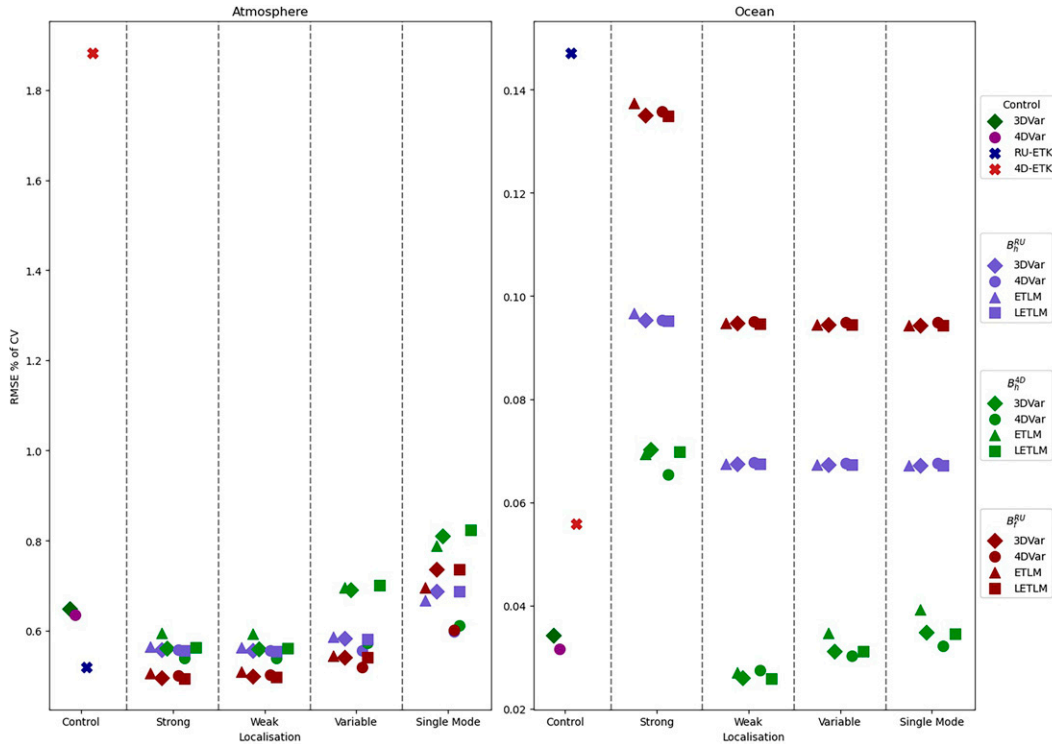


FIG. 8. The analysis RMSE as a percentage of climatological variation for the variational methods and the ETLM/LETLM's estimation of the analytical TLM in a variational setup for (left) the atmosphere and (right) the ocean, observing the full coupled state. The ETLM/LETLMs provide a good estimation of the analytical TLM methods inside all variational setups. The 4D-Var methods are given by circles, the 3D-Var methods are given by the diamonds, and the ETKFs are given by the crosses. The ETLMs are given by the triangle symbols, and the LETLMs are given by the square symbols. Since \mathbf{B}_f^{4D} diverges in the atmosphere, the results are not shown.

We contrast the previous result with an experiment observing only the ocean variables (see Fig. 10). First, there is a significant increase in RMSE in the atmosphere. All methods have a high RMSE (with the 4D-ETKF performing the most accurately), which is approximately one–two orders of magnitude higher than when observing the full coupled state. In the ocean, the RU-ETKF produces a slightly lower RMSE than the 4D-ETKF. The effect of localization on these methods is different in the ocean. The methods with the lowest RMSE are those that use the 4D-ETKF ensemble generator \mathbf{B}_f^{4D} . As localization becomes stricter, the RMSE increases. When using the RU-ETKF, the opposite occurs. As the localization becomes stricter, the RMSE decreases. This points to an important distinction in the application of SCDA versus WCDA (and more strict localization), in which the time scale and frequency of the analysis updates is critically important.

The 4D-Var using the ETLM produces good agreement with the 4D-Var using the analytical TLM in the ocean, but as the atmosphere diverges with both ensemble generators, the ETLM methods do not perform well in the atmosphere. Thus, we reiterate that the performance of the 4D-Var methods using the ETLM is highly dependent on the ensemble generator used to produce the inputs for the calculation of the ETLMs.

To expand on the above results, we examine how the SCDA and WCDA systems that use a flow-dependent \mathbf{B}_f matrix reacted to an increase in observation frequency. By focusing on \mathbf{B}_f , we remove the need to tune a climatological \mathbf{B}_c . Here we increase the analysis frequency from once every 6 h to once every 1.5 h, and also increase the observation frequency from hourly to every 15 min, in order to maintain the same number of observations in each analysis cycle.

In the variational control cases, the \mathbf{B}_c matrix from the previous experiments is used. The 3D/4D-Var controls increase in RMSE with the increased analysis frequency (Fig. 11). This indicates that the \mathbf{B}_c from the previous experiments likely requires retuning for the higher analysis update frequency. In terms of the ETKFs, the RU-ETKF decreases RMSE in the atmosphere, but increases in RMSE in the ocean, while the 4D-ETKF shows only a slight improvement in RMSE.

As we increase the analysis update frequency, the cross atmosphere–ocean dynamics in the background error covariance matrix, due to the use of SCDA, appear to become more important for 4D-Var CDA. All 4D-Var methods that use the 4D-ETKF as the ensemble generator produce improved RMSE in both the atmosphere and the ocean. For WCDA, while all methods show a small decrease in RMSE in the atmosphere, there is a significant increase in RMSE in the ocean, indicating a destabilizing imbalance for the coupled

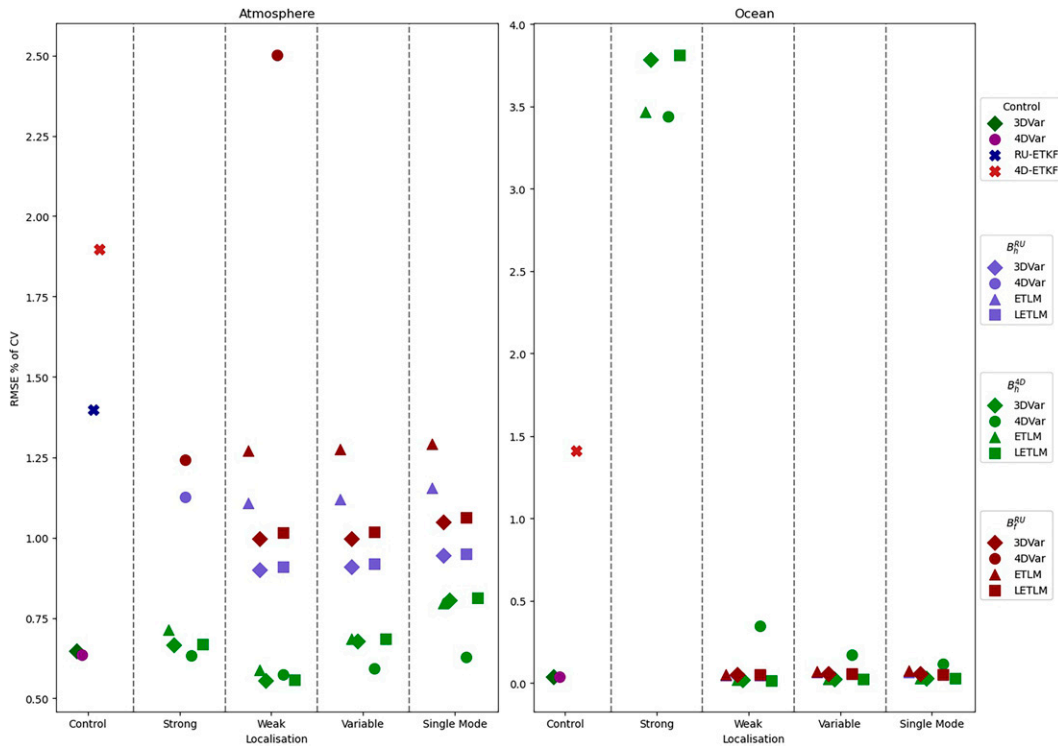


FIG. 9. ETLM/LETLM estimate vs the analytical TLM in the variational setup when observing the atmosphere variables only. In both the atmosphere and ocean, 3D-Var- B_h^{4D} using weak coupling produces the lowest RMSE. The 4D-Var methods are given by circles, the 3D-Var methods are given by the diamonds, and the ETKFs are given by the crosses. The ETLMs are given by the triangle symbols, and the LETLMs are given by the square symbols. Since B_f^{4D} diverges in the atmosphere, results are not shown.

system dynamics (Greybush et al. 2011). These results indicate trade-offs that are present when deciding on the design of a CDA system. The traditional 6-h update intervals used for the atmosphere may not be ideal, particularly for the implementation of SCDA 4D-Var. The results imply that SCDA 4D-Var may be most effective as observations are collected and analyzed more frequently than is the current practice.

5. Conclusions

The 4D-Var data assimilation (DA) method was applied in a coupled atmosphere–ocean quasigeostrophic model setting in order to investigate the challenges in using 4D-Var for strongly coupled data assimilation (SCDA). We investigated a range of configurations to identify sensitivities of the 4D-Var algorithm in this context, including various methods for formulating the background error covariance matrix, different localization strategies, and different observing network designs. For SCDA, a tangent linear model (TLM) and adjoint model are not always readily available [e.g., coastal ocean–biogeochemical (Goodliff et al. 2019), and space weather (Lang et al. 2017)]. Due to the difficulties in producing a TLM and adjoint model for large-scale coupled atmosphere–ocean models, we also considered the use of a statistical regression-based TLM and adjoint within 4D-Var. The background

error covariance matrix appears to play a significant role in 4D-Var SCDA. The flow-dependence of the background error covariance affects the assimilation accuracy. It helps to modify the climatological covariance by applying a hybrid formulation; this led to improvements in analysis RMSE. For the atmosphere, we found that the use of a hybrid background error covariance formulated using a rapid updating 3D-ETKF (RU-ETKF) produced the largest reduction to RMSE. As the atmospheric dynamics operate on faster time scales than the ocean, having the ETKF provide flow-dependent updates to the covariance matrix tended to improve the performance of 4D-Var. For the ocean, generating the hybrid background error covariance using the 4D-ETKF provided a decrease in RMSE over the control methods, but localization was needed. In this case, the cross-system error covariance did not provide a boost in accuracy in the atmosphere, but did decrease the RMSE in the ocean in comparison to the climatological control methods. A localization structure emulating weakly coupled DA provided enough information for the 4D-Var to perform well in both the atmosphere and ocean. Increasing the amount of localization (e.g., by variable or by each single mode) had a negative effect when observing all variables or observing only the atmosphere. This suggests that cross-variable and cross-mode error covariances are important in this CDA application. When observing

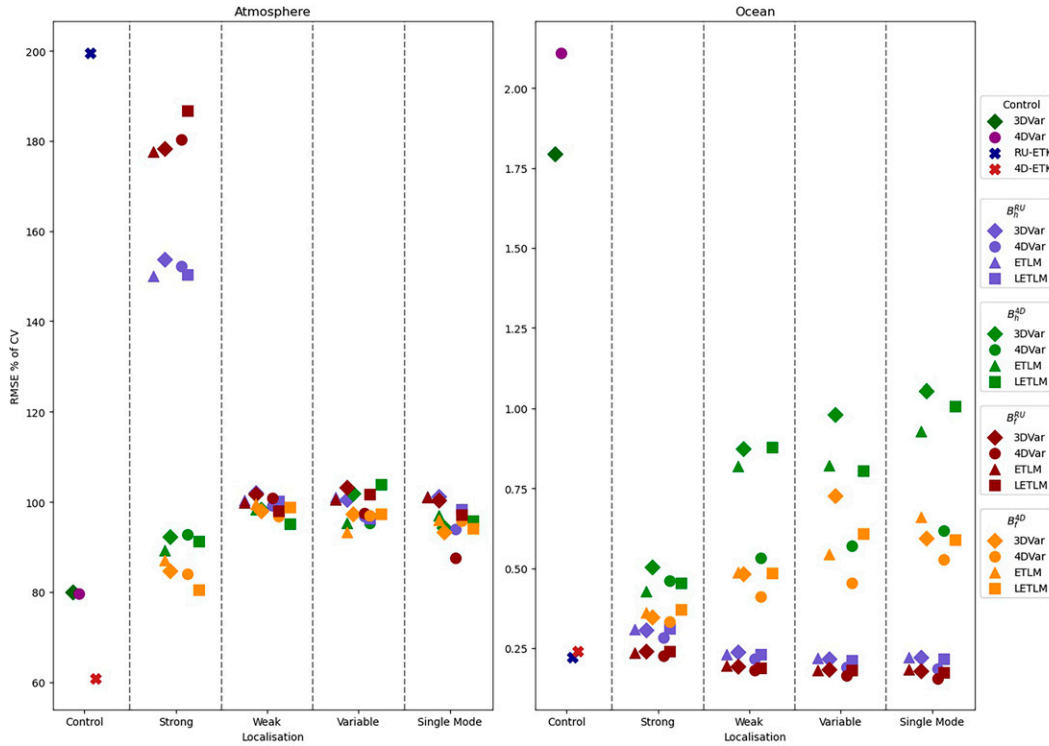


FIG. 10. RMSE when observing only ocean variables. The atmosphere has low accuracy for all methods, while the ocean has the lowest RMSE with the B_f^{RU} method for all localization strategies. The 4D-Var methods are given by circles, the 3D-Var methods are given by the diamonds, and the ETKFs are given by the crosses. The ETLMs are giving by the triangle symbols, and the LETLMs are given by the square symbols.

only the ocean, the stricter localization approaches had the best performance, suggesting that cross-mode covariances were less useful to transfer information from the ocean to the atmosphere, though this information became more useful as the analysis update frequency was increased. In particular, results suggested that a 4D-Var SCDA system may have advantages over a WCDA system when the analysis update frequency is increased. We noted differences in performance of the variational methods that utilized flow-dependent error covariance information. The rapid updating RU-ETKF as an ensemble generator benefited from more cross-domain localization, while the

4D-ETKF benefited from less cross-domain localization. When increasing the analysis update frequency, all SCDA methods showed improvement in RMSE in the atmosphere, with a slight decline in RMSE when using the rapid update ensemble generator. In comparison, all WCDA methods deteriorated at the higher frequency. We emphasize that our results indicate that the time scale of the analysis update frequency needs greater consideration in SCDA applications relative to existing standard practice in NWP. For models without an analytical TLM, it has been shown for an atmospheric case (Allen et al. 2017) that using a localized ETLM (LETLM)

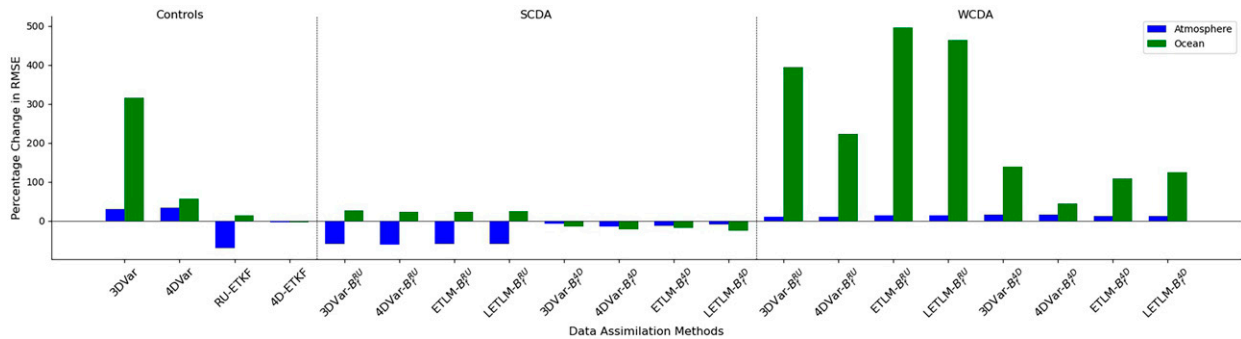


FIG. 11. Change in RMSE when shortening the analysis cycle update interval from 6 to 1.5 h. Negative values indicate improvement. An overall reduction in RMSE is noted for the 4D-Var SCDA methods that use the 4D-ETKF ensemble generator.

could be a good substitute for the analytic or software-based TLM and adjoint when used within 4D-Var. We found that when used in the 4D-Var CDA, the ETLM/LETLM under certain regimes performed as accurately as the 4D-Var configurations that used the analytical TLM. We note that while the ETLM is not appropriate for larger systems due to scalability issues, the LETLM provides a scalable alternative. We showed that if an ensemble generator performs well enough to improve the accuracy of 4D-Var via the introduction of its corresponding flow-dependent \mathbf{B}_f or hybrid \mathbf{B}_h background error covariance matrix, then the same ensemble generator produces reliable statistics for generating the ETLM.

When examining the performance of the ETLM, the flow dependence of the rapid updating RU-ETKF provided more accurate estimates of the atmosphere while the slower-updating 4D-ETKF provided more accurate estimates of the ocean. Both ensemble generation methods were effective for producing hybrid/flow-dependent background error covariance matrices to further increase the accuracy of variational methods compared to the use of a static climatological background error covariance. As most operational weather centers have some form of ensemble-based forecasting capabilities, we expect that a variant of the ETLM may be viable in more realistic applications. Particularly with the increasing adoption of coupled forecast models, new operational models are becoming more complex, which makes the implementation of an analytical TLM a greater challenge. The use of an ensemble-based TLM and adjoint can provide an adaptive and feasible solution. As this requires some form of ensemble-generator, this facilitates the use of flow-dependent and hybrid background error covariances as well, which have been established to have positive impacts on forecast skill.

Acknowledgments. M. Goodliff and S. G. Penny acknowledge support from National Oceanographic and Atmospheric Administration (NOAA) Grant NA18NWS4680048, as well as Office of Naval Research (ONR) Grants N00014-19-1-2522 and N00014-20-1-2580. S.G. Penny is further supported by NOAA Grants NA19NES4320002, NA20OAR4600277, and NA20NWS4680053.

Data availability statement. The software used to produce the data presented in this manuscript is available from the MAOOAM GitHub (<https://github.com/Climdyn/qgs> and <https://zenodo.org/record/5569583#.Yxi76nbMKM9>; Demaeyer et al. 2020), with a spinup of 100 045 000 MTU (approximately 30 000 years).

REFERENCES

- Allen, D. R., C. H. Bishop, S. Frolov, K. W. Hoppel, D. D. Kuhl, and G. E. Nedoluha, 2017: Hybrid 4dvar with a local ensemble tangent linear model: Application to the shallow-water model. *Mon. Wea. Rev.*, **145**, 97–116, <https://doi.org/10.1175/MWR-D-16-0184.1>.
- , S. Frolov, R. Langland, C. H. Bishop, K. W. Hoppel, D. D. Kuhl, and M. Yaremchuk, 2020: Challenges of increased resolution for the local ensemble tangent linear model. *Mon. Wea. Rev.*, **148**, 2549–2566, <https://doi.org/10.1175/MWR-D-20-0016.1>.
- Bishop, C. H., S. Frolov, D. R. Allen, D. D. Kuhl, and K. Hoppel, 2017: The local ensemble tangent linear model: An enabler for coupled model 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **143**, 1009–1020, <https://doi.org/10.1002/qj.2986>.
- Bonavita, M., M. Hamrud, and L. Isaksen, 2015: EnKF and hybrid gain ensemble data assimilation. Part II: EnKF and hybrid gain results. *Mon. Wea. Rev.*, **143**, 4865–4882, <https://doi.org/10.1175/MWR-D-15-0071.1>.
- Buizza, R., 2019: Introduction to the special issue on “25 years of ensemble forecasting.” *Quart. J. Roy. Meteor. Soc.*, **145** (S1), 1–11, <https://doi.org/10.1002/qj.3370>.
- Charney, J. G., and D. M. Straus, 1980: Form-drag instability, multiple equilibria, and propagating planetary waves in baroclinic, orographically forced, planetary wave systems. *J. Atmos. Sci.*, **37**, 1157–1176, [https://doi.org/10.1175/1520-0469\(1980\)037<1157:FDIMEA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1980)037<1157:FDIMEA>2.0.CO;2).
- Corazza, M., and Coauthors, 2003: Use of the breeding technique to estimate the structure of the “analysis” errors of the day. *Nonlinear Processes Geophys.*, **10**, 233–243, <https://doi.org/10.5194/npg-10-233-2003>.
- Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387, <https://doi.org/10.1002/qj.49712051912>.
- De Cruz, L., J. Demaeyer, and S. Vannitsem, 2016: The modular arbitrary-order ocean-atmosphere model: MAOOAM V1.0. *Geosci. Model Dev.*, **9**, 2793–2808, <https://doi.org/10.5194/gmd-9-2793-2016>.
- Demaeyer, J., L. De Cruz, and S. Vannitsem, 2020: qgs: A flexible Python framework of reduced-order multiscale climate models. *J. Open Source Software*, **5**, 2597, <https://doi.org/10.21105/joss.02597>.
- Fisher, M., Y. Trémolet, H. Auvinen, D. G. H. Tan, and P. Poli, 2011: Weak-constraint and long window 4DVar. ECMWF Tech. Memo. 655, 47 pp., <https://www.ecmwf.int/node/9414>.
- Frolov, S., and C. H. Bishop, 2016: Localized ensemble-based tangent linear models and their use in propagating hybrid error covariance models. *Mon. Wea. Rev.*, **144**, 1383–1405, <https://doi.org/10.1175/MWR-D-15-0130.1>.
- , D. R. Allen, C. H. Bishop, R. Langland, K. W. Hoppel, and D. D. Kuhl, 2018: First application of the local ensemble tangent linear model (LETLM) to a realistic model of the global atmosphere. *Mon. Wea. Rev.*, **146**, 2247–2270, <https://doi.org/10.1175/MWR-D-17-0315.1>.
- Gauthier, P., M. Tanguay, S. Laroche, S. Pellerin, and J. Morneau, 2007: Extension of 3DVar to 4DVar: Implementation of 4DVar at the meteorological service of Canada. *Mon. Wea. Rev.*, **135**, 2339–2354, <https://doi.org/10.1175/MWR3394.1>.
- Goodliff, M., T. Bruening, F. Schwichtenberg, X. Li, A. Lindenthal, I. Lorkowski, and L. Nerger, 2019: Temperature assimilation into a coastal ocean-biogeochemical model: Assessment of weakly and strongly coupled data assimilation. *Ocean Dyn.*, **69**, 1217–1237, <https://doi.org/10.1007/s10236-019-01299-7>.
- Gottwald, G. A., and A. J. Majda, 2013: A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks. *Nonlinear Processes Geophys.*, **20**, 705–712, <https://doi.org/10.5194/npg-20-705-2013>.
- Greybush, S., E. Kalnay, T. Miyoshi, K. Ide, and B. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522, <https://doi.org/10.1175/2010MWR3328.1>.

- Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter-3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919, [https://doi.org/10.1175/1520-0493\(2000\)128<2905: AHEKFV>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2905: AHEKFV>2.0.CO;2).
- Harlim, J., and B. R. Hunt, 2007: Four-dimensional local ensemble transform Kalman filter: Numerical experiments with a global circulation model. *Tellus*, **59A**, 731–748, <https://doi.org/10.1111/j.1600-0870.2007.00255.x>.
- Hunt, B. R., and Coauthors, 2004: Four-dimensional Ensemble Kalman filtering. *Tellus*, **56A**, 273–277, <https://doi.org/10.3402/tellusa.v56i4.14424>.
- , E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- Janisková, M., J.-N. Thépaut, and J.-F. Geleyn, 1999: Simplified and regular physical parameterizations for incremental four-dimensional variational assimilation. *Mon. Wea. Rev.*, **127**, 26–45, [https://doi.org/10.1175/1520-0493\(1999\)127<0026: SARPPF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<0026: SARPPF>2.0.CO;2).
- Kalnay, E., 2019: Historical perspective: Earlier ensembles and forecasting forecast skill. *Quart. J. Roy. Meteor. Soc.*, **145**, 25–34, <https://doi.org/10.1002/qj.3595>.
- , and Coauthors, 1997: Data assimilation in the ocean and in the atmosphere: What should be next? *J. Meteor. Soc. Japan*, **75**, 489–496, https://doi.org/10.2151/jmsj1965.75.1B_489.
- Lang, M., P. Browne, P. J. van Leeuwen, and M. Owens, 2017: Data assimilation in the solar wind: Challenges and first results. *Space Wea.*, **15**, 1490–1510, <https://doi.org/10.1002/2017SW001681>.
- Lee, M.-S., D. Barker, W. Huang, and Y.-H. Kuo, 2004: First Guess at Appropriate Time (FGAT) with WRF 3DVAR. NCAR, 4 pp.
- Lermusiaux, P. F. J., and A. R. Robinson, 1999: Data assimilation via error subspace statistical estimation. Part I: Theory and schemes. *Mon. Wea. Rev.*, **127**, 1385–1407, [https://doi.org/10.1175/1520-0493\(1999\)127<1385: DAVESS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<1385: DAVESS>2.0.CO;2).
- Lorenc, A. C., N. E. Bowler, A. M. Clayton, S. R. Pring, and D. Fairbairn, 2015: Comparison of hybrid-4D-EnVar and hybrid-4D-Var data assimilation methods for global NWP. *Mon. Wea. Rev.*, **143**, 212–229, <https://doi.org/10.1175/MWR-D-14-00195.1>.
- Payne, T. J., 2021: A hybrid differential-ensemble linear forecast model for 4D-Var. *Mon. Wea. Rev.*, **149**, 3–19, <https://doi.org/10.1175/MWR-D-20-0088.1>.
- Pelc, J. S., E. Simon, L. Bertino, G. El Serafy, and A. W. Heemink, 2012: Application of model reduced 4D-Var to a 1D ecosystem model. *Ocean Modell.*, **57–58**, 43–58, <https://doi.org/10.1016/j.ocemod.2012.09.003>.
- Penny, S. G., D. W. Behringer, J. A. Carton, and E. Kalnay, 2015: A hybrid global ocean data assimilation system at NCEP. *Mon. Wea. Rev.*, **143**, 4660–4677, <https://doi.org/10.1175/MWR-D-14-00376.1>.
- , and Coauthors, 2017: Coupled data assimilation for integrated earth system analysis and prediction: Goals, challenges and recommendations. WMO Tech. Rep. WWRP 2017-3, WMO, 59 pp., https://library.wmo.int/doc_num.php?explnum_id=10830.
- , E. Bach, K. Bhargava, C.-C. Chang, C. Da, L. Sun, and T. Yoshida, 2019: Strongly coupled data assimilation in multiscale media: Experiments using a quasi-geostrophic coupled model. *J. Adv. Model. Earth Syst.*, **11**, 1803–1829, <https://doi.org/10.1029/2019MS001652>.
- Pierini, S., 2011: Low-frequency variability, coherence resonance, and phase selection in a low-order model of the wind-driven ocean circulation. *J. Phys. Oceanogr.*, **41**, 1585–1604, <https://doi.org/10.1175/JPO-D-10-05018.1>.
- Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.*, **126**, 1143–1170, <https://doi.org/10.1002/qj.49712656415>.
- Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **133**, 347–362, <https://doi.org/10.1002/qj.32>.
- Reinhold, B. B., and R. T. Pierrehumbert, 1982: Dynamics of weather regimes: Quasi-stationary waves and blocking. *Mon. Wea. Rev.*, **110**, 1105–1145, [https://doi.org/10.1175/1520-0493\(1982\)110<1105: DOWRQS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<1105: DOWRQS>2.0.CO;2).
- Rodwell, M. J., S. T. K. Lang, N. B. Ingleby, N. Bormann, E. Hölm, F. Rabier, D. S. Richardson, and M. Yamaguchi, 2016: Reliability in ensemble data assimilation. *Quart. J. Roy. Meteor. Soc.*, **142**, 443–454, <https://doi.org/10.1002/qj.2663>.
- Smith, P. J., A. S. Lawless, and N. K. Nichols, 2017: Estimating forecast error covariances for strongly coupled atmosphere–ocean 4D-Var data assimilation. *Mon. Wea. Rev.*, **145**, 4011–4035, <https://doi.org/10.1175/MWR-D-16-0284.1>.
- , —, and —, 2018: Treating sample covariances for use in strongly coupled atmosphere–ocean data assimilation. *Geophys. Res. Lett.*, **45**, 445–454, <https://doi.org/10.1002/2017GL075534>.
- Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. 1. Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328, <https://doi.org/10.1002/qj.49711347812>.
- van der Vorst, H. A., 1992: Bi-cgstab: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, **13**, 631–644, <https://doi.org/10.1137/0913035>.
- Vannitsem, S., 2014: Dynamics and predictability of a low-order wind-driven ocean-atmosphere coupled model. *Climate Dyn.*, **42**, 1981–1998, <https://doi.org/10.1007/s00382-013-1815-8>.
- , and L. De Cruz, 2014: A 24-variable low-order coupled ocean–atmosphere model: OA-QG-WS V2. *Geosci. Model Dev.*, **7**, 649–662, <https://doi.org/10.5194/gmd-7-649-2014>.
- , and V. Lucarini, 2016: Statistical and dynamical properties of covariant Lyapunov vectors in a coupled atmosphere–ocean model—Multiscale effects, geometric degeneracy, and error dynamics. *J. Phys. A Math. Theor.*, **49**, 224001, <https://doi.org/10.1088/1751-8113/49/22/224001>.
- , J. Demaeyer, L. De Cruz, and M. Ghil, 2015: Low-frequency variability and heat transport in a low-order nonlinear coupled ocean–atmosphere model. *Physica D*, **309**, 71–85, <https://doi.org/10.1016/j.physd.2015.07.006>.