

## A Low-Order Model Investigation of the Analysis of Gravity Waves in the Ensemble Kalman Filter

LISA J. NEEF\*

*Department of Physics, University of Toronto, Toronto, Ontario, Canada*

SAROJA M. POLAVARAPU

*Environment Canada, Downsview, Ontario, Canada*

THEODORE G. SHEPHERD

*Department of Physics, University of Toronto, Toronto, Ontario, Canada*

(Manuscript received 9 August 2007, in final form 8 November 2008)

### ABSTRACT

The behavior of the ensemble Kalman filter (EnKF) is examined in the context of a model that exhibits a nonlinear chaotic (slow) vortical mode coupled to a linear (fast) gravity wave of a given amplitude and frequency. It is shown that accurate recovery of both modes is enhanced when covariances between fast and slow normal-mode variables (which reflect the slaving relations inherent in balanced dynamics) are modeled correctly. More ensemble members are needed to recover the fast, linear gravity wave than the slow, vortical motion. Although the EnKF tends to diverge in the analysis of the gravity wave, the filter divergence is stable and does not lead to a great loss of accuracy. Consequently, provided the ensemble is large enough and observations are made that reflect both time scales, the EnKF is able to recover both time scales more accurately than optimal interpolation (OI), which uses a static error covariance matrix. For OI it is also found to be problematic to observe the state at a frequency that is a subharmonic of the gravity wave frequency, a problem that is in part overcome by the EnKF. However, error in the modeled gravity wave parameters can be detrimental to the performance of the EnKF and remove its implied advantages, suggesting that a modified algorithm or a method for accounting for model error is needed.

### 1. Introduction

Four-dimensional (4D) data assimilation is now operationally implemented at the European Centre for Medium-Range Weather Forecasts (ECMWF), Météo France, the UK Met Office, the Japan Meteorological Agency, and the Canadian Meteorological Centre (Rabier 2005) and is in the process of being extended to the upper stratosphere and mesosphere, where observations are both new and typically sparse. Gravity

waves, generally considered as unimportant noise in the context of numerical weather prediction, play an important role in the upper stratosphere and mesosphere, especially in driving the quasi-biennial oscillation (QBO) and meridional circulation, as well as in the dissipation of the tides. In these regions gravity waves represent a significant component of the flow in both observations and models (Koshyk et al. 1999). This presents a new challenge for data assimilation: the separation and correct representation of (fast) gravity waves and (slow) balanced motion when both flow components are present in the true state.

Because 4D assimilation methods evolve error covariances in time, they have the potential to improve on the traditional method of statistical or optimal interpolation (OI; Bergman 1979; Daley 1991). Of the two principal 4D assimilation algorithms, four-dimensional variational data assimilation (4D-Var) and the ensemble

---

\* Current affiliation: Chemistry and Climate Division, Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands.

---

*Corresponding author address:* Dr. T. G. Shepherd, Dept. of Physics, University of Toronto, 60 St. George Street, Toronto, ON M5S 1A7, Canada.  
E-mail: tgs@atmosp.physics.utoronto.ca

Kalman filter (EnKF), 4D-Var is currently more widely used for operational deterministic weather forecasting, although the EnKF is used for the ensemble prediction system at the Canadian Meteorological Centre (Houtekamer and Mitchell 2005) and represents a promising algorithm for future use because of its relative ease of implementation. Variational methods are able to analyze unbalanced flow by choosing balanced and unbalanced variables as uncorrelated control variables and estimating their covariances (e.g., Parrish and Derber 1992; Derber and Bouttier 1999; Gauthier et al. 1999). Note that the balance can also be nonlinear so that the unbalanced variables are defined as departures from the nonlinear balances (Fisher 2003).

However, here we focus on the ability of sequential (nonvariational) 4D methods to separate the “slow” balanced vortical motion from relatively “fast” unbalanced motion. Our specific focus is on the EnKF relative to the simpler OI. Because application of the EnKF is still largely in a discussion/testing phase (Lorenz 2003; Houtekamer and Mitchell 2005), the ability of this algorithm to capture unbalanced motion is still poorly understood. Szunyogh et al. (2005) report a single case in which an EnKF-type assimilation method captured a gravity wave that was present in reality but not in the model estimate, thereby suggesting that flow-dependent covariance models can potentially capture unbalanced motion better than stationary ones. It is unclear, however, whether the additional cost of developing an ensemble of states can generally be expected to make it easier to recover different time scales from a set of observations, or whether a static covariance model is sufficient or even preferable.

In regimes in which modeled flows are expected to be balanced, it has been found that 4D data assimilation can cause the excitation of spurious unbalanced motion, essentially because of the development of unphysical correlations (Polavarapu et al. 2000; Lea et al. 2002; Houtekamer and Mitchell 2005; Neef et al. 2006, hereafter NPS06). A general solution to this problem involves imposing balance constraints on the analysis (Courtier and Talagrand 1990; Dee 1991; Todling and Cohn 1994; Polavarapu et al. 2000; Kepert 2004) or simply filtering out fast waves. In the upper stratosphere and mesosphere, excessive filtering can eliminate important gravity waves and either remove or amplify the tides (Sankey et al. 2007). There also appears to be a significant unbalanced component to flow in the tropical stratosphere, which makes it similar in many respects to the mesosphere (Nezlin et al. 2009).

It is also unclear how the frequency and type of the available observations affect the recovery of an unbalanced state. Observed quantities project differently onto

vortical and inertia–gravity wave modes; moreover, gravity waves have time scales (from tens of minutes up to the inertial period) that are similar to or shorter than data assimilation intervals (usually 6 or 12 h). If observations are assimilated roughly once (or less) in a fast period, it will be difficult to glean the wave’s magnitude and frequency.

Here we address these issues using a simple model of a chaotic vortical mode coupled to a linear gravity wave. This work follows in the vein of NPS06, which examined balance (and the loss thereof) in sequential 4D data assimilation in the case of a balanced truth. It was shown there that the EnKF is generally able to conserve balance in the analysis, given a large enough ensemble, because it retains nonlinearity in the estimation of error covariances. Using the same model, and motivated by the current challenge of assimilation in the mesosphere and the tropical stratosphere, we now address regimes of imbalance and time scale overlap.

The model and the EnKF equations are described in section 2. In section 3, the regimes of convergence and divergence of the EnKF are defined in the context of the present model. Section 4 examines numerically how the physical characteristics of the gravity wave that is present in the true state affect the EnKF’s ability to capture the dynamics of both time scales from observations of the partial state. A generalization to imperfect models is made in section 5. Conclusions and some discussion are provided in section 6.

## 2. Methodology

### a. Model

As in NPS06, we use the model derived by Lorenz (1986) and extended by Wirosuetisno and Shepherd (2000):

$$\frac{d\phi}{dt} = w, \quad (1)$$

$$\frac{dw}{dt} = -\frac{C}{2}\sin(2\phi + 2\epsilon bx), \quad (2)$$

$$\frac{dx}{dt} = -\frac{z}{\epsilon}, \quad (3)$$

$$\frac{dz}{dt} = \frac{x}{\epsilon} + \frac{bC}{2}\sin(2\phi + 2\epsilon bx), \quad (4)$$

$$C(t) = a_0 + a_1 \cos(\gamma t). \quad (5)$$

Equations (1)–(5) describe a chaotic vortical mode in  $\phi$  and  $w$ , coupled to a linear gravity wave in  $x$  and  $z$ . The four variables are the spectral coefficients of potential vorticity  $\phi$  and  $w$  (note that  $\phi$  is actually related to the phase of two

potential vorticity coefficients from the original derivation), geostrophic imbalance  $z$ , and divergence  $x$ . The slow vortical mode varies on a nondimensional time scale of about six time units, and the gravity wave has frequency  $\epsilon^{-1}$  (where typically  $\epsilon \ll 1$ ) and period  $T_{GW} = 2\pi\epsilon$ . If the slow time scale is assumed to correspond to typical weather-system time scales of 36–48 h, we can say that one time unit corresponds to 6–8 h; thus, for  $\epsilon = 0.1$ , the gravity wave period is about 4 or 5 h. The two modes are coupled by the parameter  $b$ , which corresponds to a rotational Froude number. Chaos in the slow mode is controlled by the periodic variation of  $C(t)$  with  $a_0 = 1$ ,  $a_1 = 0.8$ , and  $\gamma = 0.92$ . Further description and a derivation of this system are given in appendix 1 of NPS06 and chapter 2 of Neef (2007).

Wirosoetisno and Shepherd (2000) showed that the gravity wave variables have components that are “slaved” to the slow mode and are given [to  $\mathcal{O}(\epsilon^2)$ ] by the asymptotic balance relations

$$U_x(\phi; \epsilon) = -\frac{\epsilon}{2}Cb \sin 2\phi + \mathcal{O}(\epsilon^3) \quad \text{and} \quad (6)$$

$$U_z(\phi, w; \epsilon) = \epsilon^2(Cbw \cos 2\phi + \frac{C'}{2}b \sin 2\phi) + \mathcal{O}(\epsilon^3), \quad (7)$$

where  $C'$  is the time derivative of  $C(t)$ . If the model is initialized with  $x(0) = U_x$  and  $z(0) = U_z$ , the free gravity wave will be suppressed, meaning it will have an amplitude  $\sim \epsilon^3$ , with a growth rate that is exponentially slow in  $\epsilon$  (Wirosoetisno and Shepherd 2000). For a general state of the system, the free gravity wave is the unslaved part of  $x$  and  $z$ ; its magnitude can be approximated [to  $\mathcal{O}(\epsilon^2)$ ] as  $\tilde{I} = \sqrt{\tilde{x}^2 + \tilde{z}^2}$  and its phase as  $\theta = \tan^{-1}(\tilde{z}/\tilde{x})$ , where  $\tilde{x} = x - U_x$  and  $\tilde{z} = z - U_z$ . In principle, the separation between fast and slow components can be made more accurate by extending the balance relations (6) and (7) to higher order in  $\epsilon$  (Wirosoetisno and Shepherd 2000), but we use the second-order relations here.

*b. The nonlinear Kalman filter*

Sequential, variance-minimizing assimilation systems have the following form:

$$\mathbf{x}_k^f = \mathcal{M}_{k-1}(\mathbf{x}_{k-1}^a), \quad (8)$$

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k^f), \quad (9)$$

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (10)$$

(e.g., Kalnay 2003, sections 5.3 and 5.4). Equation (8) represents the forward evolution of the state estimate  $\mathbf{x}_{k-1}^a$  by the forecast model  $\mathcal{M}_{k-1}$ . In (9), the resulting

forecast  $\mathbf{x}_k^f$  is compared to a vector of observations  $\mathbf{z}_k$  (made at a discrete time  $k$ ) and adjusted proportionally to the difference between them (where the operator  $\mathbf{H}_k$  maps the forecast state to observation space). The so-called observation increment  $\mathbf{z}_k - \mathbf{H}_k\mathbf{x}_k^f$  is weighted by the optimal gain matrix  $\mathbf{K}_k$ , a function of the observation error covariance matrix  $\mathbf{R}_k$  and the forecast error covariance matrix  $\mathbf{P}_k^f$ .

OI (Bergman 1979) is a three-dimensional (3D) data assimilation algorithm because the error covariance matrix  $\mathbf{P}_k^f = \mathbf{P}^f$  is defined in the model space but is static in time. In OI, (9) is cycled with (8) to produce a series of forecasts and analyses that sequentially approximate the true state. In the Kalman filter (Kalman 1960; Ghil et al. 1981; Miller et al. 1994),  $\mathbf{P}_k^f$  is instead estimated dynamically by evolving it forward in time using a model, then updating it to reflect the information brought in by observations in (9).

For nonlinear models, both the evolution and analysis update of the covariance matrix depend on higher-order moments of the forecast error distribution, and a closure approximation must be made. The extended Kalman filter (EKF) estimates the evolution of  $\mathbf{P}^f$  by linearizing the model about the state estimate at each discrete point in time. Whereas NPS06 compared the EnKF to the EKF, here we focus entirely on the EnKF, in light of the fact that the EKF algorithm has been shown to be highly unstable (e.g., NPS06; R. Todling 2007, personal communication). The EnKF (Evensen 1994) estimates forecast and analysis error covariances according to the ensemble averages

$$\mathbf{P}_k^f = \langle (\mathbf{x}_{i,k}^f - \langle \mathbf{x}_{i,k}^f \rangle) (\mathbf{x}_{i,k}^f - \langle \mathbf{x}_{i,k}^f \rangle)^T \rangle \quad \text{and} \quad (11)$$

$$\mathbf{P}_k^a = \langle (\mathbf{x}_{i,k}^a - \langle \mathbf{x}_{i,k}^a \rangle) (\mathbf{x}_{i,k}^a - \langle \mathbf{x}_{i,k}^a \rangle)^T \rangle, \quad (12)$$

where the subscript  $i$  denotes individual forecasts in an  $N$ -member ensemble. The analysis step (9) is performed  $N$  times, with each ensemble member updated with a randomly perturbed observation [corresponding to the estimated observation error statistics (Burgers et al. 1998)]. Thus, the EnKF produces an ensemble of state estimates, the mean of which is the EnKF best estimate of the analysis state.

*c. Assimilation experiments*

This study is focused primarily on so-called twin experiments, where “true” and “estimated” states are evolved with the same model, which is also used to generate observations. This way, the forecast error at observation times comes from the initial state error, accumulated analysis error, observation error, and the growth of perturbations between observation times, but not from model error.

Numerical results shown, except for single examples, will be averages over  $K$  different realizations of the initial conditions. In each experiment (or realization), the initial true state  $\mathbf{x}^t = (\phi^t, w^t, x^t, z^t)^T$  is generated by randomly selecting initial values of  $\phi$  and  $w$ , computing the corresponding slaved components of  $x$  and  $z$  [(6) and (7)], and then generating a free gravity wave of magnitude  $\tilde{I}^t$ . Unless noted otherwise, the parameters that define the gravity wave will have values  $\tilde{I}^t = 1.5$ ,  $\epsilon = 0.1$ , and  $b = 0.71$ . The initial estimated state is generated by randomly perturbing the slow component of the truth ( $\phi$  and  $w$ )  $N$  times, with the perturbations for both variables having variance  $\sigma_0^2 = 0.25$ . Ensemble members are initialized with free gravity wave magnitudes chosen randomly from  $\tilde{I}_i \sim \mathcal{N}(0, \sigma_i^2)$ , where  $\sigma_i = \tilde{I}_i$ . The sensitivity of the assimilation result to this estimated initial gravity wave error is examined in section 3. The gravity wave phases are chosen randomly from a uniform distribution between  $[-\pi, \pi]$ , such that the mean free gravity wave in the initial ensemble mean is zero. Likewise, the initial state estimate in OI experiments is generated with no free gravity wave. The truth and forecast are both evolved forward in time using a fourth-order Runge–Kutta method with a time step of  $\Delta t = 0.01$ .

Observations are generated at intervals  $\Delta t^{\text{obs}}$  by adding random noise  $\mathbf{r}_k \sim \mathcal{N}(0, \sigma_{\text{obs}}^2)$  to the truth. In general, we set  $\Delta t^{\text{obs}} = 0.1$  (corresponding to an assimilation interval of about 40 min) and the ensemble size to 24. The choice of these parameters is justified in the following section.

Two types of observations are considered in the experiments. In general, we will observe the mixed–time scale state given by

$$\mathbf{z}_{\text{MIX}} = \begin{pmatrix} \phi_{\text{obs}} \\ w'_{\text{obs}} \end{pmatrix}, \quad (13)$$

where  $w' \equiv (w - bz)/(1+b^2)$  represents a mixed–time scale quantity related to vorticity. To examine the behavior of the EnKF when the gravity wave is excluded from the observations, we also compare observations of the entirely slow state given by

$$\mathbf{z}_{\text{SLOW}} = \begin{pmatrix} \phi_{\text{obs}} \\ w_{\text{obs}} \end{pmatrix}. \quad (14)$$

Because model variables are directly observed, the observation operator  $\mathbf{H}_k$  is linear. Hereafter the time subscripts on  $\mathbf{H}$  and  $\mathbf{R}$  will be dropped because neither operator changes in time. All experiments are run with the correct observation error statistics—that is, with  $\mathbf{R} = \sigma_{\text{obs}}^2 \mathbf{I}_2$ , where  $\mathbf{I}_2$  represents the  $2 \times 2$  identity matrix, and  $\sigma_{\text{obs}} = 0.1$ .

Whereas the EnKF covariance matrix is given by (11), OI requires an estimate of the stationary covariance

TABLE 1. The estimated correlations of the six variable pairs of the exL86 model, estimated from an ensemble of 50 states run to 500 time units.

	Average correlation	Correlation standard error
$\phi, w$	0.4851	0.0941
$\phi, x$	−0.0005	0.1645
$\phi, z$	0.0001	0.1644
$w, x$	0.0000	0.1530
$w, z$	0.0009	0.1531
$x, z$	0.0000	0.1830

matrix. This requires an estimate of multivariate correlations as well as the average variances. Table 1 shows the six correlations between the components of the model state, estimated from an ensemble of 50 states integrated to 500 time units (with  $\epsilon = 0.1$  and  $b = 0.71$ ). The correlations were computed at each time step and then averaged over all time steps, which amounts to 2.5 million samples of each variable. The associated standard deviations for each correlation term are also shown in Table 1. Note that except for  $\rho_{\phi, w}$ , the standard deviations are much larger than the mean correlations. This indicates that the correlations are strongly time and state dependent, which in turn indicates that a stationary covariance matrix, despite the above optimizations, will not be a good approximation.

It can furthermore be seen that all correlations except  $\rho_{\phi, w}$  average to zero. Although the correlations between fast and slow variables (e.g.,  $\rho_{w, x}$ ) increase with increasing  $\epsilon$  and  $b$  (to be discussed in more detail in section 4), we verified that they are still zero in the time average. We therefore formulate the static OI covariance matrix as

$$\mathbf{P}_0^f = \begin{pmatrix} \sigma_s^2 & \sigma_s^2 \rho_{\phi, w} & 0 & 0 \\ \sigma_s^2 \rho_{\phi, w} & \sigma_s^2 & 0 & 0 \\ 0 & 0 & \sigma_f^2/2 & 0 \\ 0 & 0 & 0 & \sigma_f^2/2 \end{pmatrix}. \quad (15)$$

For the correlation term  $\rho_{\phi, w}$  we take the average value given in Table 1,  $\rho_{\phi, w} = 0.49$ . The variances ascribed to the slow and fast variables ( $\sigma_s$  and  $\sigma_f$ , respectively) were optimized by sweeping over a range of values and selecting the values that minimize error (see appendix). The optimal values of  $\sigma_s$  and  $\sigma_f$  were found to depend on the observation interval, with approximate functional dependencies given by  $\sigma_s^{\text{opt}} = 0.3\Delta t^{\text{obs}}$  and  $\sigma_f^{\text{opt}} = 0.03 + 0.14\Delta t^{\text{obs}} + 0.06(\Delta t^{\text{obs}})^2$ . Given these values, (15) represents the best possible OI implementation for the exL86 model.

Although OI does not update the forecast error covariance matrix, an analysis error covariance matrix can

be computed after the assimilation of the observations (Kalnay 2003, chapter 5):

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T. \quad (16)$$

Here, (16) represents the reduction in uncertainty due to the information gained from the observations and corresponds theoretically to the ensemble-estimated errors produced by the EnKF (12).

### 3. Convergence and divergence of the EnKF with two time scales

#### a. Convergence regimes for the slow mode

Filter divergence is the phenomenon whereby the dynamic errors estimated by the assimilation system become increasingly small relative to actual errors, such that subsequent observations are rejected and the algorithm can no longer correct itself (eventually arriving with high confidence at a poor analysis). In the EnKF, filter divergence means that ensemble members become more tightly clustered around their mean than the distance from the mean to the true state. To examine the treatment of gravity waves in the EnKF analysis, we will first establish the regimes (in the space of assimilation parameters) where the EnKF converges for the slow mode, then examine the analysis of the fast mode.

Convergence of the EnKF is controlled by the size of the ensemble and by the frequency at which observations are made. Figure 1 examines the stability of the slow-mode analysis for different regimes of these two parameters by comparing the ratio between the actual and ensemble-estimated error in  $w$  as a function of  $\Delta t^{\text{obs}}$  for five ensemble sizes. Each line represents an average over 50 experiments where the mixed-time scale state  $\mathbf{z}_{\text{MIX}}$  (13) is observed. The ratio shown is computed from the actual and ensemble-estimated errors averaged in time according to

$$E_w = \langle (w^t - \langle w^a \rangle)_{\Delta T}^2 \rangle_{\Delta T}^{1/2} \quad \text{and} \quad (17)$$

$$E_w^e = \langle \langle (w_i^f - \langle w^a \rangle)_{\Delta T}^2 \rangle \rangle_{\Delta T}^{1/2}, \quad (18)$$

where the inner brackets indicate the ensemble average and the brackets  $\langle \cdot \rangle_{\Delta T}$  indicate the average over all time steps of a time window  $\Delta T \equiv [40, 100]$ . The temporal averaging window is chosen to avoid an adjustment period that will be explained in section 3c.

For  $N = 4$  and 8, the ratio is much larger than unity for all observation intervals. Estimated errors match the true errors consistently for  $N \simeq 20$  members. To minimize ensemble sampling error, we will choose  $N = 24$  as

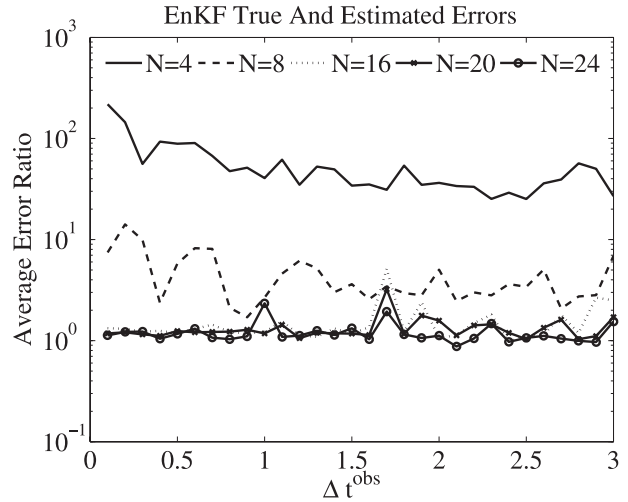


FIG. 1. Averages of the ratio between RMS error in  $w$  and the EnKF-estimated error as a function of the observation interval  $\Delta t^{\text{obs}}$ , for five different ensemble sizes. Each line represents an average over 50 experiments with observations of the mixed-time scale state  $\mathbf{z}_{\text{mix}}$  (13).

the standard ensemble size in all subsequent experiments. [We have verified, but not shown, that the apparently higher average errors in the  $N = 24$  case here are only due to statistical noise. Note also that for small ensemble sizes the ratio between actual and estimated errors is larger for more frequent observations. This happens because the ensemble is updated often but poorly, causing it to cluster around an erroneous state and not spread enough between observation times (NPS06).]

#### b. Fast and slow analysis increments

To understand how the EnKF treats the fast mode, consider the two analysis increments  $\delta w^a = w^a - w^f$  (representing the slow-mode increment) and  $\delta z^a = z^a - z^f$  (representing the fast-mode increment) for observations of either the slow variable  $w$  or the mixed-time scale variable  $w'$ . If  $w$  is observed, then only the balanced components of the fast variables ( $x$  and  $z$ ) can be gleaned from the observations, and the analysis increments become

$$\delta w^a = w^a - w^f = k_{ww}(w_{\text{obs}} - w^f) \quad \text{and} \quad (19)$$

$$\delta z^a = z^a - z^f = k_{zw}(w_{\text{obs}} - w^f), \quad (20)$$

with weights

$$k_{ww} = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_{\text{obs}}^2} \quad \text{and} \quad (21)$$

$$k_{zw} = \frac{c_{zw}}{\sigma_w^2 + \sigma_{\text{obs}}^2}, \quad (22)$$

where  $\sigma_w^2$  is the estimated error variance for  $w$  and  $c_{zw}$  is the estimated covariance between fast  $z$  and slow  $w$ . Because the slow mode and free gravity wave are presumed to be independent,  $c_{zw}$  results entirely from the slaving relationship. The fraction of the observation increment added to  $z$  in (20) is proportional to  $c_{zw}$  because  $w_{\text{obs}}$  contains information about the slaved components of the fast variables. Therefore, the balance relationship must be captured by the estimated fast–slow error covariances (such as  $c_{zw}$ ) in order to correctly update the fast variables with observations of the slow variables. Error in the estimation of these fast–slow covariances will cause a misadjustment of fast variables, which can lead to an excitation of a gravity wave that has nothing to do with the true free gravity wave. Note that the forecast error standard deviations ascribed to the free gravity wave ( $\sigma_{\tilde{x}}$  and  $\sigma_{\tilde{z}}$ ) do not affect the fast-variable increments in this case.

The problem changes if the observed variable is the mixed–time scale quantity  $w'$ , in which case observations also contain information about the free gravity wave, and the analysis increments become

$$\delta w^a = k_{ww'} \left[ w'_{\text{obs}} - \frac{1}{1+b^2} (w^f - bz^f) \right] \quad \text{and} \quad (23)$$

$$\delta z^a = k_{zw'} \left[ w'_{\text{obs}} - \frac{1}{1+b^2} (w^f - bz^f) \right], \quad (24)$$

with weights

$$k_{ww'} = \frac{c_{ww'}}{\sigma_{w'}^2 + \sigma_{\text{obs}}^2} \quad \text{and} \quad (25)$$

$$k_{zw'} = \frac{c_{zw'}}{\sigma_{w'}^2 + \sigma_{\text{obs}}^2}, \quad (26)$$

where

$$c_{ww'} = \frac{1}{1+b^2} (\sigma_w^2 - bc_{zw}) \quad \text{and} \quad (27)$$

$$c_{zw'} = \frac{1}{1+b^2} (c_{zw} - b\sigma_z^2). \quad (28)$$

In this case, fast–slow covariances such as  $c_{zw}$  are needed not only to correctly update the (balanced components of the) fast variables but also to recover information about the slow mode from the fast-variable component of the observations. If estimated correctly, these balanced terms will comprise a small correction to the slow mode analysis. If overestimated, however, they can cause filter divergence in the slow mode.

For the fast variables the analysis step [e.g., (24)] now also depends on estimated fast-variable error variances [ $\sigma_z^2$  in (28), but also  $\sigma_x^2$ ]. These have both balanced and

free gravity wave components, where the free gravity wave components (which have linear evolution) will dominate for an unbalanced truth.

The nature of fast–slow covariances (and how these are explicitly evolved in the EnKF) is illustrated in Fig. 2. Figure 2a shows  $\phi(t)$  and  $x(t)$  for a reference state with a free gravity wave defined by  $\tilde{l}^i = 1.5$ ,  $\epsilon = 0.1$ , and  $b = 0.71$ . The correlation between errors in  $\phi$  and  $x$  is  $\rho_{\phi x} = c_{\phi x} / \sigma_{\phi} \sigma_x$ . The error in  $x$  at each time can be approximated by linearizing the balance relationship (6) about the time-varying state estimate. Then the covariance term  $c_{\phi x}$  becomes

$$c_{\phi x} = \langle e_{\phi} e_x \rangle \simeq \left\langle e_{\phi} \frac{\partial U_x}{\partial \phi} e_{\phi} \right\rangle = -\epsilon C b \cos(2\phi) \sigma_{\phi}^2, \quad (29)$$

where  $e_{\phi}$  represents the error in  $\phi$  at each time. Then the correlation can be approximated as

$$\rho_{\phi x} = -\epsilon C b \cos 2\phi \frac{\sigma_{\phi}}{\sigma_x} \equiv \eta_{\text{LIN}}(t) \frac{\sigma_{\phi}}{\sigma_x}, \quad (30)$$

where

$$\eta_{\text{LIN}}(t) = -\epsilon C b \cos 2\phi \quad (31)$$

approximates the slow–time scale evolution of the covariance between  $\phi$  and  $x$ , and it results from a linearization of the slaving relationship (6). We can evaluate the slow–fast correlations estimated by the ensemble by comparing  $\eta_{\text{LIN}}$  to the quantity  $\eta_{\text{ENS}} = \rho_{\phi x} (\sigma_x / \sigma_{\phi})$ , where  $\sigma_x$  and  $\sigma_{\phi}$  are computed from the ensemble. This is done for an example state in Fig. 2b, where  $\eta_{\text{ENS}}$  is computed from a 24-member EnKF analysis, with mixed–time scale observations (13) assimilated at intervals  $\Delta t^{\text{obs}} = 0.1$ .

Both estimates of  $\eta$  show again that the correlation is strongly state dependent, which implies that a dynamic covariance model will be more useful than a static one for the recovery of both time scales. We have verified (not shown) that the similarity between  $\eta_{\text{ENS}}$  and  $\eta_{\text{LIN}}$  increases with ensemble size. The 24-member EnKF used in this example can at times produce large estimation errors and is contaminated by oscillations with the period of the gravity wave, which result from residual imbalance in the ensemble average. However, it captures the overall variability of  $\eta$ , at least as approximated by linearization.

Figure 3 shows two EnKF analyses for this state, comparing a 24-member ensemble and a six-member ensemble, each with observations of  $\mathbf{z}_{\text{mix}}$  assimilated every  $\Delta t^{\text{obs}} = 0.1$  time units. Figures 3a,b compare the truth and estimates of  $\phi$  and  $x$  for this state. For visual

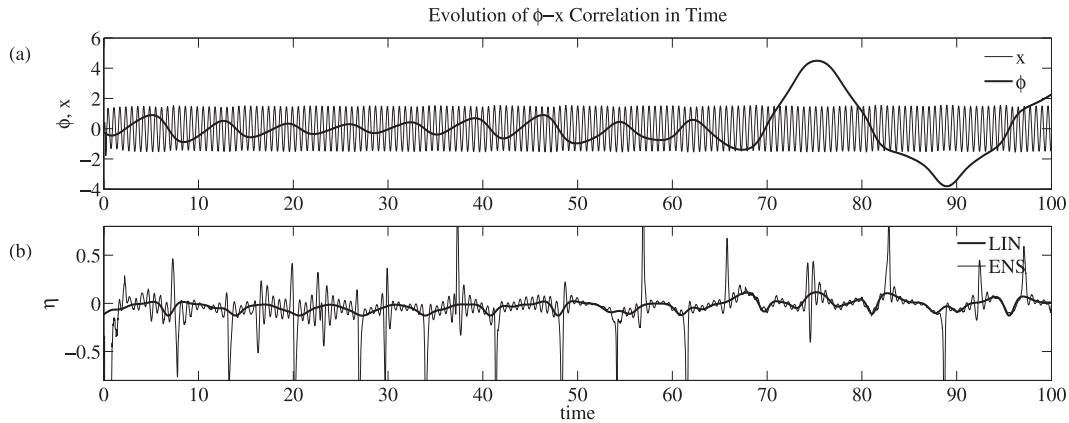


FIG. 2. (a) Evolution of the slow variable  $\phi$  and the fast variable  $x$  over 100 time units, for an example state with a gravity wave of magnitude  $\tilde{I}^t = 1.5$ . In this case,  $\epsilon = 0.1$  and  $b = 0.71$ . (b) The term  $\eta$ , which governs the evolution of the correlation between  $\phi$  and  $x$  for the same state, as given by the linear approximation [(31); thick line] and as computed by a 24-member EnKF (thin line), with mixed-time scale observations (13) assimilated every  $\Delta t^{\text{obs}} = 0.1$  time units.

clarity, only the last 20 time units of the assimilation period are shown. In both plots, the analysis of  $\phi$  is nearly indistinguishable from the truth. Accordingly, the true and EnKF-estimated errors in  $\phi$ , shown in Figs. 3c,d for the respective cases, are similar over the assimilation period. For the six-member EnKF, the analysis of  $x$  (Fig. 3b) has the correct phase but the wrong amplitude. Figure 3f compares actual and estimated error in  $x$  for this case, and the estimated errors are found to quickly become several orders of magnitude smaller than the true error. The EnKF is diverging in the fast mode (with the ensemble tightly locked around a gravity wave of the wrong amplitude), even though it is clearly in a regime of convergence of the slow mode. In the 24-member case, the truth and analysis for  $x$  are also very close (Fig. 3a), with a correspondence between the true and estimated errors in  $x$  (Fig. 3e), indicating that the estimation of fast-slow covariances in Fig. 2b is sufficient for the analyses of both the slow and fast mode and that the estimated fast variances have been adjusted sufficiently. Note also that a slight underestimation of the error in  $x$  is starting to grow by the end of the assimilation period.

*c. Evolution of average slow and fast errors in time*

Figure 4 shows the average time evolution of EnKF-estimated and actual errors. Errors are shown for convergent ( $N = 24$ ) and divergent ( $N = 8$ ) regimes, with  $\Delta t^{\text{obs}} = 0.1$  in both cases. Average errors are subdivided into  $w$ ,  $x$ , and  $\theta$  components, with each line representing an average over 200 experiments. For each ensemble size, solid lines represent actual errors and dashed lines the estimated errors. The initial-time standard devia-

tions in  $w$  [ $\sigma_w(0) = e_s$ ],  $x$  [ $\sigma_x(0) = \tilde{I}^t/\sqrt{2}$ ], and the gravity wave phase  $\theta$  [ $\sigma_\theta(0) = \pi/2$ ] are indicated in the respective panels by a horizontal black line.

The top row of Fig. 4 compares average errors for observations of the mixed-time scale state (13), that is, when both fast-slow covariances and fast-variable variances must be estimated accurately. For both ensemble sizes, the average error in  $w$  is reduced well below the initial standard errors as soon as the first few observations are made, although for  $N = 8$  (as expected) the actual  $w$  errors grow in time whereas the corresponding estimated errors decrease in time. Average errors in  $x$  and  $\theta$  behave quite differently from this for both ensemble sizes, decreasing in time but with the distance between actual and estimated errors increasing in time. Thus, on average there is filter divergence in the fast mode even while there is convergence in the slow mode, for both the small ( $N = 8$ ) and large ( $N = 24$ ) ensembles.

The bottom row of Fig. 4 shows the same quantities but for observations of the slow state (14)—that is, where only the slaved components of the fast variables can be recovered while the free gravity wave is ignored. For  $N = 24$ , average  $w$  errors are not changed significantly by the change from mixed-time scale observations to observations of the full slow mode, suggesting that fast-slow covariances were estimated sufficiently well in the former case to retrieve the slow mode even though it was only partially observed. For  $N = 8$ , the change of observation variable reduces  $w$  error slightly. The extra information about the slow mode that is gained by the change in observation variable is compensated by a loss of information about the fast mode: average  $x$  error does not decrease in time. Estimated  $x$  error, however, does decrease

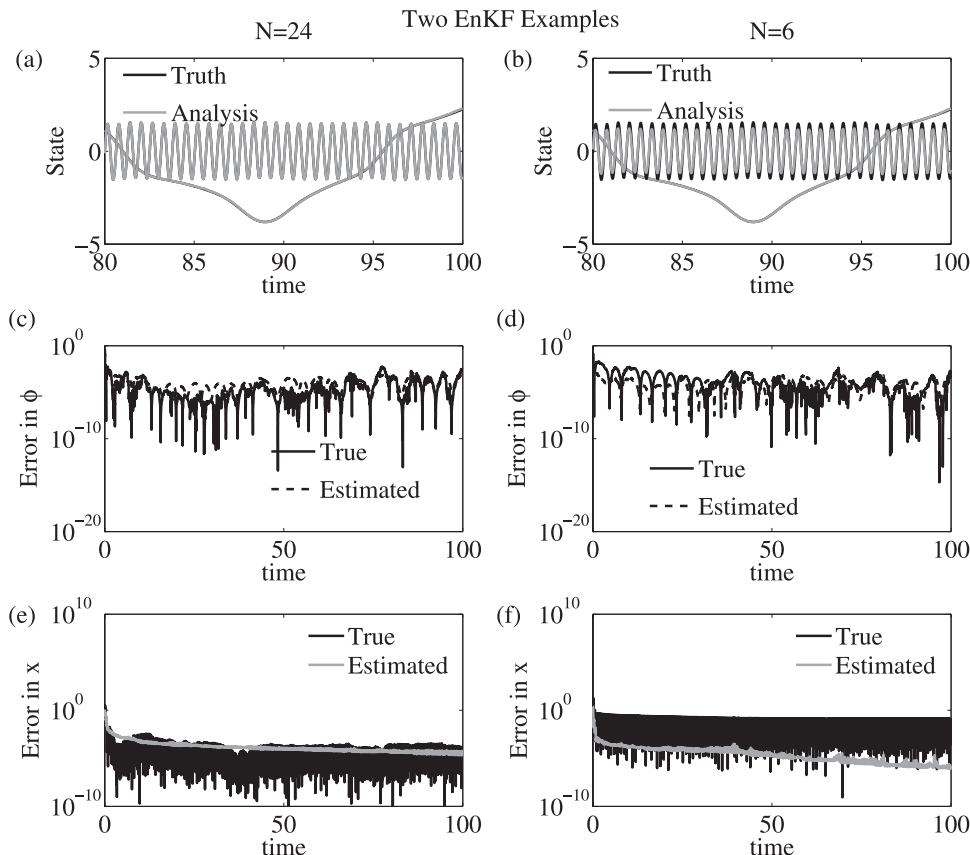


FIG. 3. Two sample assimilations with the EnKF. (a),(b) A comparison of  $\phi$  (the slow chaotic curve) and  $x$  (the fast oscillation) for the true state (black) and the analysis (gray). For clarity, only the last 20 time units of the 100-time-unit assimilation period are shown in each case. In both experiments, the observation interval is  $\Delta t^{\text{obs}} = 0.1$  and the mixed-time scale state (13) is observed. The two experiments differ only in ensemble size: a 24-member ensemble is used in (a), and a six-member ensemble in (b). (c),(d) The true (solid) and estimated (dashed) errors in  $\phi$  over the entire assimilation period, corresponding respectively to (a) and (b). (e),(f) The true (black) and estimated (gray) errors in  $x$  for each case.

in this case. This means that spurious fast–slow correlations are causing the EnKF to reduce estimated error, leading to eventual filter divergence, in the fast mode.

There are two reasons why the EnKF might diverge in the fast mode even while converging in the slow mode. One reason is that the free gravity wave is linear: the ensemble members do not spread in time between observations, so that small estimation errors in the analysis increments accumulate and the ensemble eventually locks too tightly on either an erroneous amplitude (e.g., Fig. 4b) or both the wrong amplitude and phase (e.g., Figs. 4e,f). A second reason is that the fast mode is observed about 10 times less frequently per cycle than the slow mode; this will be examined in the subsequent section.

Despite filter divergence there is still significant error reduction in  $x$  and  $\theta$  (when both modes are observed), even for a small ensemble, while ignoring the fast mode does not significantly impact the slow-mode analysis.

Note also that there exists (in the convergent cases) a spinup period of about 40 time units, after which slow mode error stabilizes. To take this into account, assimilation experiments will hereafter be evaluated by comparing RMS errors for  $w$ ,  $x$ , and  $\theta$ , averaged over the last 60 time units, as in (17), and defining also

$$E_x = \langle (x^t - \langle x^a \rangle)^2 \rangle_{\Delta T}^{1/2} \quad \text{and} \quad (32)$$

$$E_\theta = \langle (\theta^t - \langle \theta^a \rangle)^2 \rangle_{\Delta T}^{1/2}. \quad (33)$$

d. Observation interval

For  $\Delta t^{\text{obs}} = 0.1$  and  $\epsilon = 0.1$  (as above), the gravity wave is observed about six times in a cycle and the slow mode 60 times. Figure 5 examines how the recovery of the fast mode is affected as the observation



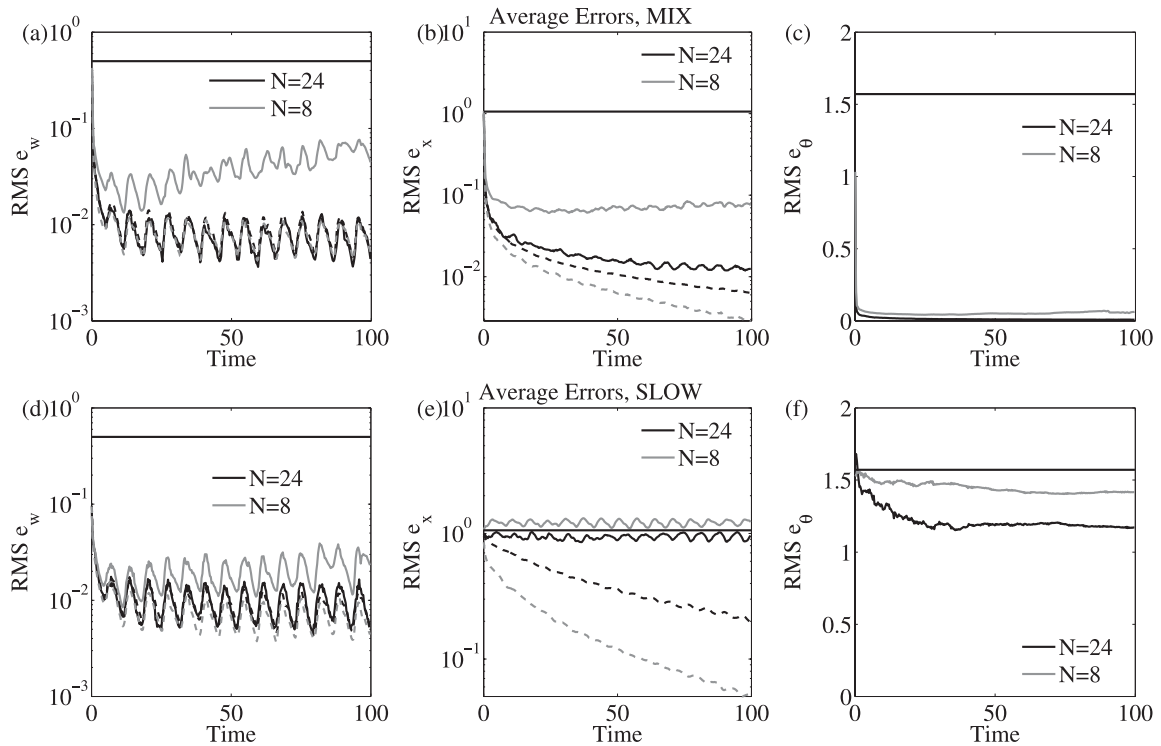


FIG. 4. Convergence/divergence of the EnKF in time. In the top row, average errors are shown for (a) the slow variable  $w$ , (b) the fast variable  $x$ , and (c) the gravity wave phase  $\theta$  for experiments in which MIX observations (13) are assimilated every  $\Delta t^{\text{obs}} = 0.1$  time units. Each line represents an average over 200 experiments, comparing ensemble sizes  $N = 8$  (gray) and 24 (black). Solid lines show actual errors and dashed lines the estimated errors. (d)–(f) As in (a)–(c) but for SLOW observations (14). The horizontal solid lines in each panel represent the average initial-time error.

interval approaches and surpasses the gravity wave period ( $T_{\text{GW}} \simeq 0.6$ ). Average actual and estimated  $E_x$  are compared for a 24-member EnKF and OI, each assimilating mixed-time scale observations (13) with  $\Delta t^{\text{obs}} = 0.1$ . Figure 5a shows that the average true error in the EnKF exceeds the average estimated error at all observation intervals. However, the ratio between actual and estimated  $E_x$  (Fig. 5b) does not increase with  $\Delta t^{\text{obs}}$ , and actual errors are moreover consistently lower in the EnKF than for OI. Thus, despite a tendency to underestimate fast variable error, the EnKF, which dynamically estimates slow–fast correlations, is still able to retrieve the fast mode better than OI, which relies on average correlations.

For both methods there exist peaks, in average actual and estimated error, at  $\Delta t^{\text{obs}}$  equal to multiples of the gravity wave period. These peaks reflect the fact that it is impossible to infer the amplitude of a wave when observing at single phase. Note that the peaks are more pronounced in OI. In the EnKF, estimated error also increases for these special observation intervals, although generally error is underestimated more (as indicated by spikes in the error ratio in Fig. 5b).

*e. Ensemble size*

It is of course also possible to increase the ensemble size beyond what is required to achieve convergence in the slow mode, although this may not be a feasible option in realistic applications. Figure 6 shows average true and estimated  $E_x$  as a function of the ensemble size, again assimilating mixed-time scale observations with  $\Delta t^{\text{obs}} = 0.1$ . Also shown are the initial-time  $E_x$  and the average  $E_x$  that results from 50 OI assimilation runs using mixed-time scale observations with  $\Delta t^{\text{obs}} = 0.1$ . The actual error is reduced from the initial value by over an order of magnitude for as few as six ensemble members; 18–20 ensemble members are required to have errors lower than those returned by OI. In order for estimated errors to match actual errors, around 80 ensemble members are needed—almost 4 times as many ensemble members as are needed to achieve convergence in the slow mode. Nonetheless, even when sampling error is large the assimilation of fast mode observations can be useful, and fewer ensemble members are required to improve on OI than to completely eliminate filter divergence.

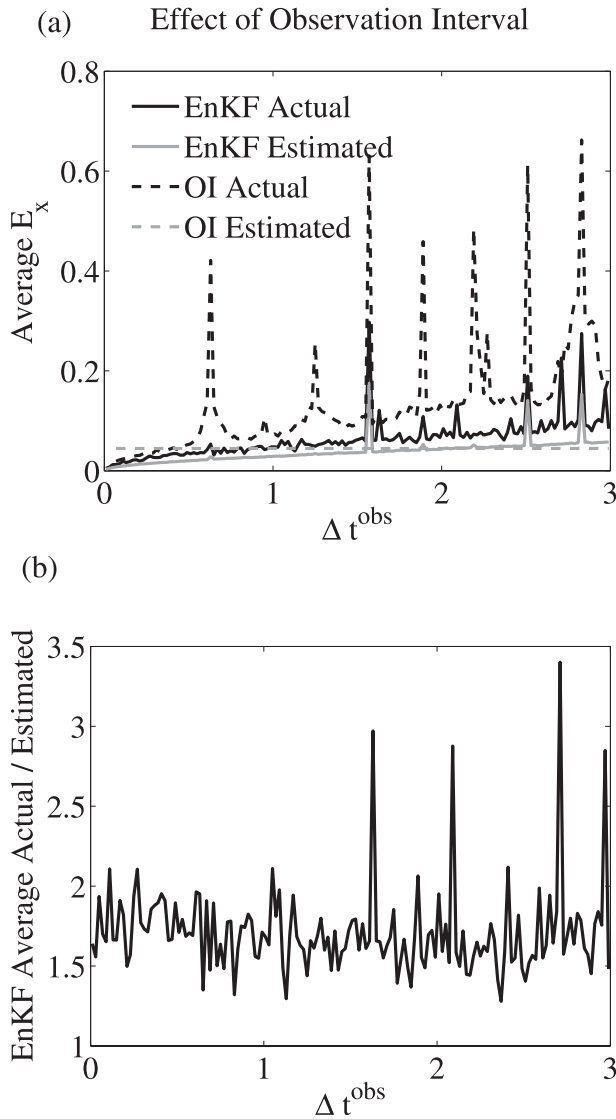


FIG. 5. (a) Average error in  $x$  (32) for a 24-member EnKF, comparing true (black) and estimated (gray) errors as a function of the observation interval. These are further compared to true (black dashed) and estimated (gray dashed) error for OI. (b) The average ratio between true and estimated error in  $x$  for the EnKF. In both the EnKF and the OI, the mixed-time scale state (13) is observed.

#### f. Fast variance of the initial ensemble

Recall that in the case of mixed-time scale observations, the ability to capture the gravity wave also depends on the variance attributed in the ensemble to the fast wave [section 3b, specifically (28)]. Figure 7 examines average true and estimated  $E_x$  over a range of  $\sigma_{\tilde{t}}$ , the standard error from which the gravity waves in the initial ensemble are generated (assimilating mixed-time scale observations, with  $N = 24$  and  $\Delta t^{\text{obs}} = 0.1$ ). The

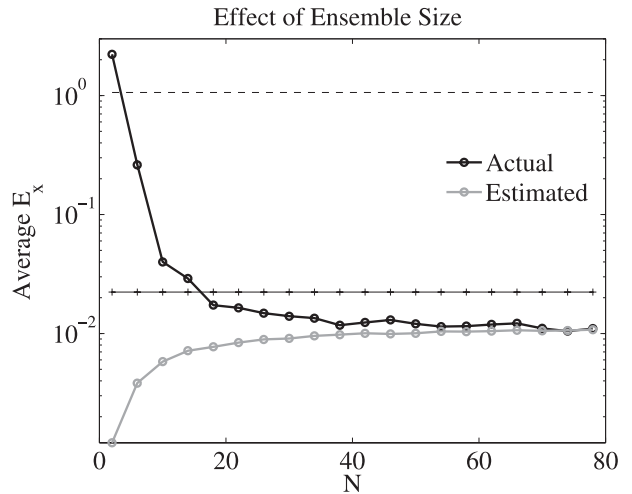


FIG. 6. Average true (black) and EnKF-estimated (gray) error in  $x$  (32) as a function of the ensemble size, computed for 50 realizations (with  $\Delta t^{\text{obs}} = 0.1$  throughout). The initial fast variable error is indicated by a dashed line, and the average OI error is shown by the plus signs. In both the EnKF and the OI, the mixed-time scale state (13) is observed.

errors are again compared to the initial-time error, and the average OI error with similar observations;  $E_x$  decreases as the magnitude of the initial fast variance increases, saturating around  $\sigma_{\tilde{t}} = 1$ . Because there is no increase in error when  $\sigma_{\tilde{t}}(0)$  exceeds the true value ( $\sigma_{\tilde{t}} = \tilde{t} = 1.5$ ), this experiment suggests that it would be prudent, in realistic applications, to overestimate the fast variable error. It is also encouraging that the EnKF can initially underestimate the fast variable standard deviation by about 50% and yet substantially reduce errors from the initial value and improve over OI.

On the other hand, an 80% underestimation of the fast-variable standard deviation [i.e.,  $\sigma_{\tilde{t}}(0) = 0.3$ ] yields analysis errors that are twice the minimum error that can be achieved by increasing  $\sigma_{\tilde{t}}$ . Moreover, actual errors remain consistently larger than estimated errors over all  $\sigma_{\tilde{t}}(0)$ . Thus, Fig. 7 also shows that although it helps to overestimate fast errors, it is also difficult to fully eliminate divergence of the EnKF (in the case where the ensemble is too small to fully converge in the fast mode) by attributing more initial error to the gravity wave.

In summary, we can say that the EnKF is able to retrieve both time scales as long as mixed-time scale observations are made, although with a strong tendency to diverge in the fast mode. Even if the ensemble is not large enough to achieve convergence in the fast mode, however, fast-mode errors are stable in time, generally small compared to the initial-time errors, and may still be lower than OI errors.

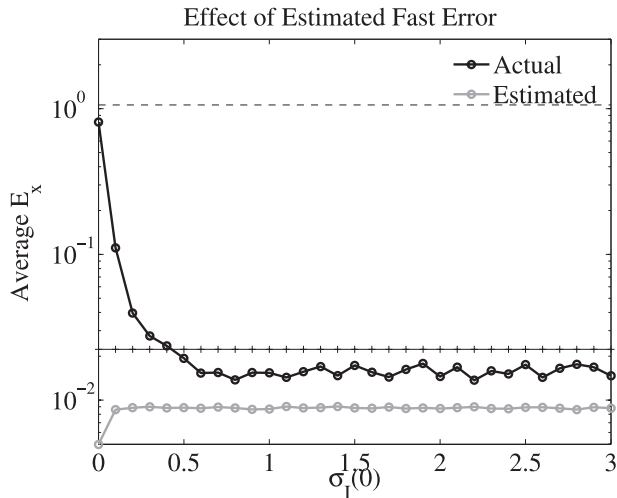


FIG. 7. Average true (black) and EnKF-estimated (gray) error in  $x$  (32), as a function of the standard deviation  $\sigma_f$  from which the fast components of the initial ensemble are generated, computed for 50 realizations (with  $\Delta t^{\text{obs}} = 0.1$  throughout and a 24-member ensemble). The initial-time true fast variable standard deviation is indicated by a dashed line, and the average OI error is shown by the plus signs. In both the EnKF and the OI, the mixed-time scale state (13) is observed.

**4. Gravity wave parameters**

We now examine how the quality of the analysis, the relative value of slow- and mixed-time scale observations, and the EnKF’s possible advantage over OI are affected by the parameters that define the gravity wave. In this set of experiments, we compare three assimilation systems. The first, denoted MIX, is a 24-member EnKF with observations of the mixed-time scale state (13) made at intervals of  $\Delta t^{\text{obs}} = 0.1$ . The second, denoted SLOW, is the same except that the observations are made of the slow-time scale state (14). In the third, mixed-time scale observations are made but assimilated using OI. Comparison of SLOW and MIX shows the relative benefit of observing or ignoring the gravity wave, whereas comparison of MIX and OI shows the possible advantage of time-dependent covariance estimation.

*a. Gravity wave magnitude*

The influence of the gravity wave on the evolution of the full system can be changed by varying its magnitude. This will change the relative value of slow- and mixed-time scale observations because slow observations ignore a greater component of the system as  $\tilde{I}^t$  increases. Because the initial-time ensemble mean state is balanced, a larger true gravity wave magnitude also means that the first few analysis increments in the assimilation period will be larger, thereby amplifying any potential

errors in estimated covariances. Consequently, a larger-magnitude true-state gravity wave, given a balanced initial estimate, could lead to a worse analysis in both modes, in which case it may be better to simply neglect the fast-slow covariances, as in our OI implementation.

Figure 8 shows the three error measures as a function of  $\tilde{I}^t$ , comparing MIX, SLOW, OI, and the corresponding estimated errors for all three. For MIX, both  $E_w$  and  $E_x$  are unaffected by the increased magnitude of the true gravity wave, although the EnKF’s underestimation of fast error remains throughout. For SLOW,  $E_w$  increases slightly as  $\tilde{I}^t$  increases because of the increasingly large source of variability that is now ignored by the observations. For OI, where the fast mode is observed but covariances and variances are kept constant,  $E_x$  increases slightly with  $\tilde{I}^t$ , although we found (not shown) that this increase in error can be removed by slightly increasing  $\sigma_f$  with  $\tilde{I}^t$  in (15). Thus, the advantage of the EnKF over OI does not clearly change with the magnitude of the true gravity wave. However, the value of observing both time scales does increase with increasing gravity wave magnitude.

Note also that  $E_\theta$  decreases with  $\tilde{I}^t$  in the two cases where the free fast mode is observed, MIX and OI. Presumably this happens because increasing the magnitude of the gravity wave decreases the ratio between observation error and the gravity wave signal.

*b. Gravity wave period*

Changing the time scale of the gravity wave also changes the significance of the gravity wave in the evolution of the full system. Increasing  $\epsilon$  means that the balanced components of  $x$  and  $z$  [(6) and (7)] become larger, while the correlations between fast and slow variables become both larger and more nonlinear.

This is illustrated in Fig. 9, which compares  $\eta_{\text{LIN}}$  and  $\eta_{\text{ENS}}$  for a state that is equal to the state shown in Fig. 2a in all respects except that the gravity wave frequency has been changed to  $\epsilon^{-1} = 3$ , corresponding to a gravity wave period  $T_{\text{GW}} \approx 2$ . It can be seen by both approximations that the magnitude of the slow-fast correlation increases relative to Fig. 2b. For OI, this means that the zero correlations used in the covariance matrix, although still representative of the average, will become a poorer estimate. Increased fast-slow correlation also means that error in the estimation of the fast mode more strongly influences the evolution of (true) error in the slow mode, and vice versa. For the SLOW case, increasing  $\epsilon$  will mean that a larger component of the fast variables can be recovered from the observations, while at the same time the lack of observed information about the gravity wave may worsen the analysis of the slow mode. Disagreement between  $\eta_{\text{ENS}}$  and  $\eta_{\text{LIN}}$  also

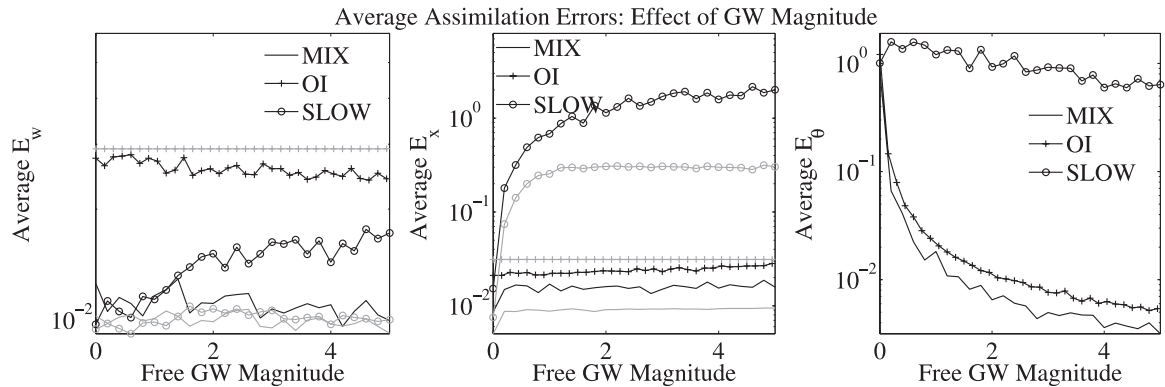


FIG. 8. The three error measures [(17), (32), and (33)], as a function of  $\bar{l}^l$ , the magnitude of the true-state gravity wave. Average errors are compared for the three cases outlined in section 4. The actual errors for each case are indicated by black lines and plot symbols and the estimated errors by gray ones.

increases between Figs. 2 and 9, indicating that the nonlinearity of the balance relationship is increasing, which may make the ensemble less Gaussian and thus increase error in either the SLOW or MIX cases.

Figure 10 compares average errors over a range of  $T_{GW}$  for MIX, SLOW, and OI. The change in fast–slow correlation that comes with changing  $T_{GW}$  does not clearly affect the performance of OI relative to the EnKF (MIX). The MIX case shows only a weak effect of  $T_{GW}$  on average errors (average  $E_\theta$  increases slightly), so we can infer that the increasingly nonlinear balance relationship is not much more difficult to capture by the EnKF cycle. Changing the fast time scale does, however, affect how much information about the fast mode can be recovered from slow observations: both  $E_x$  and  $E_\theta$  decrease with increasing  $T_{GW}$  for SLOW. This means that after a certain point ( $T_{GW} \geq 2$ ), more information can be recovered about the fast mode using the EnKF and slow observations than by using OI and mixed–time scale observations. Note also that  $E_w$  also increases with  $T_{GW}$  for SLOW, reflecting the increasing effect of the unobserved free gravity wave on the slow mode, al-

though average  $E_w$  for SLOW is mostly below average  $E_w$  for OI.

c. Slow–fast coupling

Changing the coupling parameter  $b$  also changes the nonlinearity of the balance relationship and the magnitude of the fast–slow correlations, but now without changing the time scale overlap of the two modes. Figure 11 shows  $\eta_{LIN}$  and  $\eta_{ENS}$  for the same initial condition as in the previous examples, but now with  $b = 3$ . Again, the fluctuations in the correlation term are larger, and the difference between the EnKF-estimated correlation and the linearized value increases, implying greater nonlinearity. Changing  $b$  also changes the information content of observations, with  $w'$  becoming more gravity wave dominated as  $b$  increases. This means that MIX observations as defined in (13) will contain increasingly more information about the gravity wave and less about the slow mode as the coupling between the two modes increases. Changing  $b$  in the exL86 model thus illustrates the effect of observations that project more or less strongly onto the gravity wave manifold.

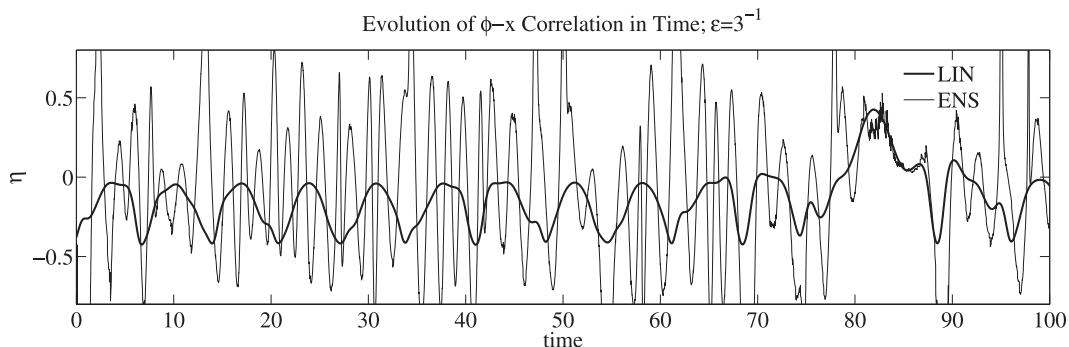


FIG. 9. As in Fig. 2b, but for a state with  $\epsilon = 1/3$ . A 24-member EnKF (with  $\Delta t^{obs} = 0.1$ ) is used to estimate  $\eta_{ENS}$ .

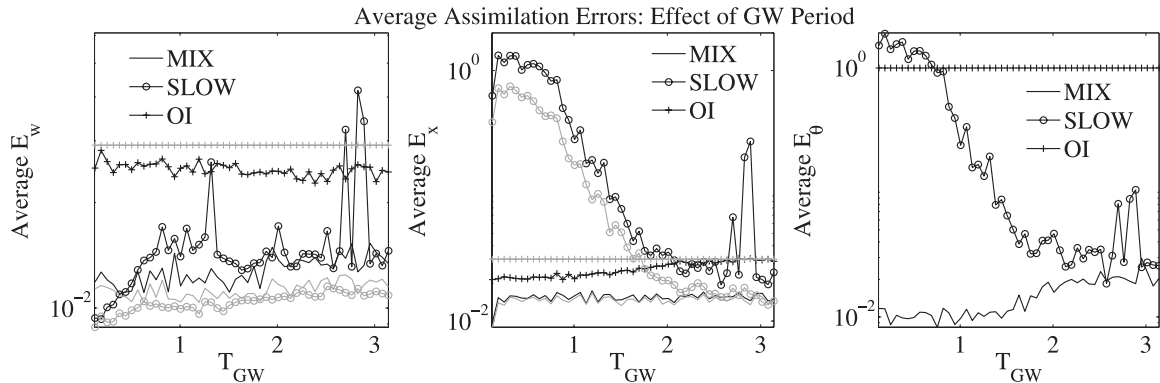


FIG. 10. As in Fig. 8, but as a function of the period of the true-state gravity wave.

Figure 12 compares the effect of changing  $b$  while keeping  $\epsilon$  and  $\bar{l}$  fixed. Here  $E_w$  increases with  $b$  for SLOW and OI but not for MIX, a result similar to that of Fig. 10. In terms of  $E_x$  and  $E_\theta$ , there is an “optimal”  $b$  (roughly  $b = 1$ ) such that MIX and OI give their most accurate recovery of the fast mode, whereas SLOW simply improves in the fast mode with increasing  $b$ . The results can be explained as follows: As coupling increases, SLOW neglects an increasingly relevant source of variability, which increases slow error. At the same time, the additional fast-mode information contained in the observations decreases  $E_x$  and  $E_\theta$ . As coupling increases, OI neglects an increasingly relevant time-varying correlation, which also increases  $E_w$  with increasing  $b$ . In contrast to the previous experiment, however,  $E_x$  and  $E_\theta$  now decrease with  $b$  for OI and MIX when  $b \lesssim 1$ , a decrease that is presumably due to the fact that observations of  $w'$  become more fast-mode dominated as  $b$  increases. The subsequent increase in  $E_x$  and  $E_\theta$  for MIX and for  $b \gtrsim 1$  suggests that the fast–slow correlation, which helps to recover the fast mode, becomes more important at large  $b$ . By preserving fast–slow correlations the EnKF (MIX) excels over OI in its recovery of the gravity wave.

### 5. Model error experiments

In all experiments shown so far, forecasts and truths were evolved with the same model, such that the EnKF had to estimate the amplitude and phase of the gravity wave but was always perfect in gravity wave frequency. It can be argued that this could be the reason why the EnKF performed so much better than OI in the experiments above. In practice, not all gravity wave frequencies are observed and not all frequencies are modeled. For example, fast gravity waves are less likely to be observed than slow gravity waves simply because fast waves spend less time in an instrument’s observation region (Alexander and Barnet 2007). Modeled gravity wave frequencies depend on a model’s resolution and on generation mechanisms such as convective parameterization, and their dispersion relation may be deformed by the time-integration scheme. In consideration of these limitations, two sets of experiments were performed in which the perfect model restriction is removed by introducing error in the modeled gravity wave parameters  $\epsilon$  and  $b$ .

In the first experiment, the gravity wave frequency assumed by the forecast model, now denoted  $\epsilon^f$ , is

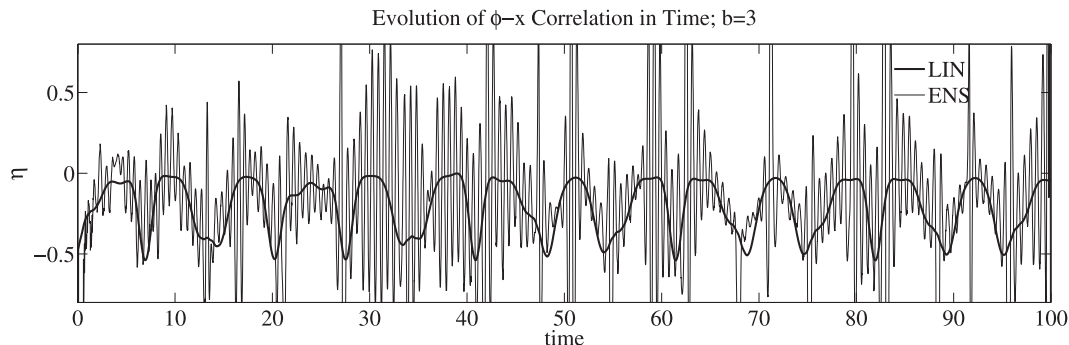


FIG. 11. As in Fig. 2b, but for a state with  $b = 3$ . A 24-member EnKF (with  $\Delta t^{\text{obs}} = 0.1$ ) is used to estimate  $\eta_{\text{ENS}}$ .

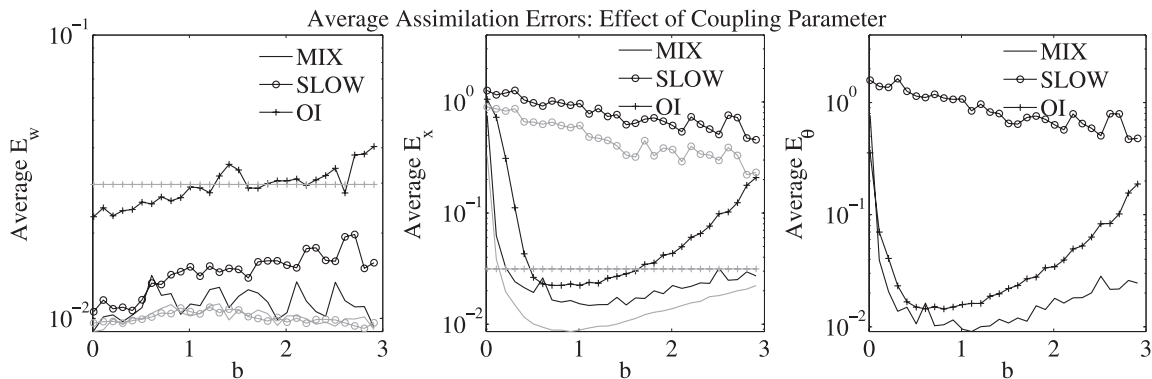


FIG. 12. As in Fig. 8, but as a function of the coupling parameter  $b$ .

changed while the true state is kept at  $\epsilon^t = 0.1$ . Figure 13a shows the three error components as a function of  $\epsilon^f$ . The 24-member EnKF is compared to OI, using mixed-time scale observations (13) in both cases. For the EnKF, misestimation of the gravity wave frequency is detrimental to the recovery of the gravity wave: both  $E_x$  and  $E_\theta$  shoot to the initial error levels for any  $\epsilon^f \neq \epsilon^t$ . Note  $E_x$  even exceeds the initial-state error, indicating the excitation of gravity waves that are not of the same frequency, phase, or magnitude as the true gravity wave and are left uncontrolled by the observations. OI errors in  $x$  and  $\theta$  are a bit lower for  $\epsilon^f \neq \epsilon^t$ , or at least not greater than the initial-time error, which indicates that it is better in these experiments to simply neglect fast–slow covariances. In the slow mode ( $E_w$ ), the EnKF is able to perform at least as well as OI when  $\epsilon^f$  is close to the true frequency (i.e., for  $|\epsilon^f - \epsilon^t| \leq 0.1$ ), but OI clearly yields a better analysis overall. For the EnKF, the strong growth of  $E_w$  for  $\epsilon^f \geq 1.5$  reveals an unfavorable characteristic of the (large-ensemble) EnKF: recovery of the slow mode can be harmed when the gravity wave is not captured at all, in which case it would be better to force fast–slow correlations to zero.

The second case of model error is shown in Fig. 13b, where the estimated coupling parameter, now denoted  $b^f$ , is changed while the true value is kept constant at  $b^t = 0.71$ . The EnKF again returns large errors in all three variables when the estimated  $b$  exceeds the true parameter by a certain amount. Interestingly, underestimating the coupling parameter ( $b^f < 0.71$ ) increases error less than when it is overestimated ( $b^f > 0.71$ ). The difference can be understood from the fact that  $b$  changes the information content of the observation (13) from being purely slow to being mixed. For  $b^f > 0.71$ , the assimilation is assuming that  $w'$  is more time scale mixed than it actually is, resulting in adjustment of the gravity wave away from its true amplitude and phase and not enough adjustment of the slow variables, the

estimates of which have suffered from the misestimated gravity wave. For  $b^f < 0.71$ , the assimilation is assuming that  $w'$  is less mixed; thus, it adjusts the slow mode with more information than is actually contained in observations, while the gravity wave is not adjusted enough. Slow-mode error consequently increases relative to estimated error for  $b^f < 0.71$  but, given less simultaneous adjustment of fast variables, the impact is presumably smaller. Again, although OI behaves similarly, it is more accurate overall.

These admittedly simplistic experiments show that the EnKF's ease in locking onto a linear gravity wave phase will be limited by differences between the resolved frequencies in observations and models. When the ability to capture a gravity wave is lost, the analysis of the vortical mode suffers as well, especially if the correlation between the fast and slow modes is overestimated.

## 6. Summary and conclusions

This study examined the behavior of the ensemble Kalman filter in the physical context in which there is a preponderance of free gravity waves in the truth. This problem differs from the initialization problem of numerical weather prediction but is relevant to data assimilation in the mesosphere and tropical stratosphere (Nezlin et al. 2009). We used a highly simplified model in which the state consists of a vortical mode and a free gravity wave of a relatively fast time scale.

It was found that the EnKF analysis cycle can cause the ensemble to lock onto a gravity wave of the wrong amplitude, causing filter divergence in the analysis of the fast mode, even in regimes where it converges in the slow mode. To prevent this type of filter divergence requires upward of 3 times as many ensemble members as are required for convergence in the slow mode (Fig. 6). When the ensemble size is too small to fully

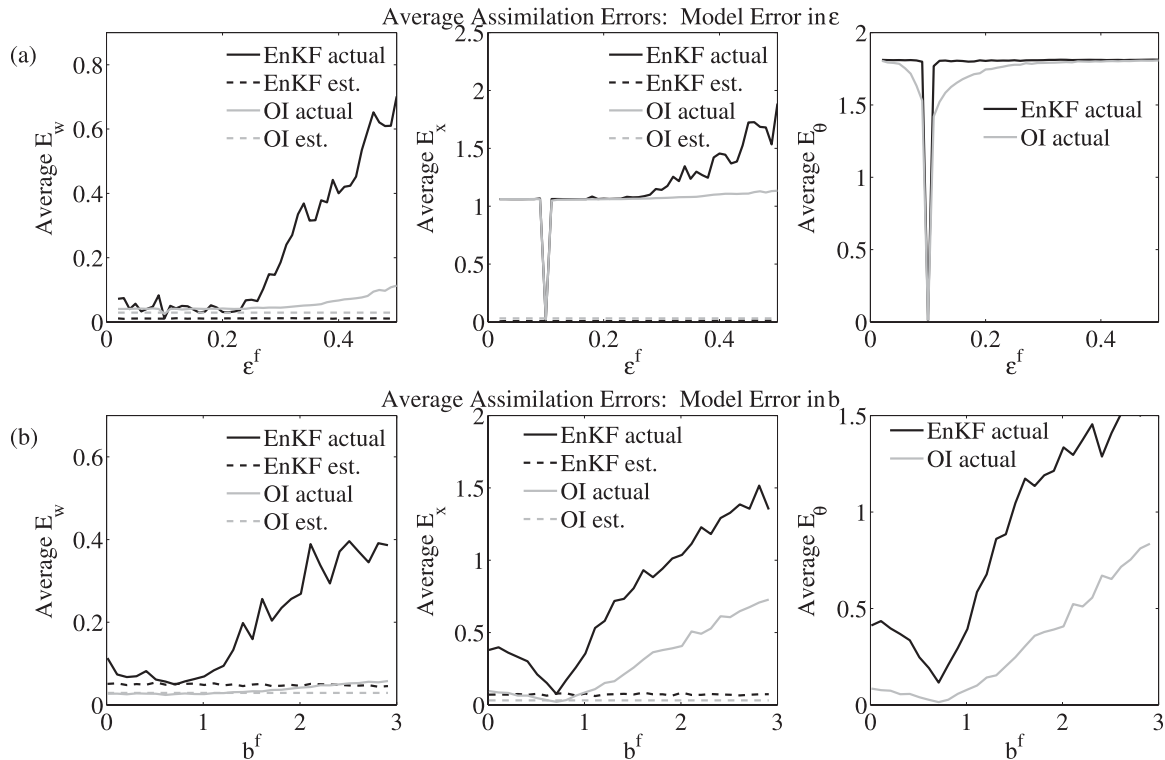


FIG. 13. (a) The three error measures [(17), (32)–(33)] as a function of (a) the gravity wave frequency assumed in the model estimate and (b) the coupling parameter assumed in the model estimate, keeping the true values constant at the reference values,  $\epsilon = 0.1$  and  $b = 0.71$ . The 24-member EnKF (black) is compared to OI (gray), and average actual errors for each case (solid lines) are compared to average estimated errors (dashed lines).

converge in the fast mode, the impact of fast-mode filter divergence on the overall analysis can be lessened by increasing the amount of error initially attributed to the free gravity wave (Fig. 7), although this only alleviates the problem and does not eliminate it. Fast-mode filter divergence comes about because the linear–nonchaotic gravity wave ensemble does not spread between observations, and thus persists even when observations are very frequent (Fig. 5). Because gravity waves in the real atmosphere are also linear to a first approximation, this may indicate a fundamental problem for assimilating unbalanced states with the EnKF.

In realistic applications it is possible that gravity waves that are present in the truth may not be captured in the observations, perhaps because of filtering or averaging of observations. In that case, only the component that is slaved to the slow mode can be controlled by observations, which requires the error covariance estimation cycle to capture the balance relationships between slow and fast model variables. Despite divergence in the fast mode, the EnKF’s covariance model was found to be more accurate than the highly tuned static covariance model used in our OI, resulting in a better recovery of both modes if a mixed–time scale

state is observed, and of the slow mode and the slaved component of the fast mode if slow variables are observed. For sufficiently large ensembles, recovery of the full slow mode from the observations that contain both time scales was found to be nearly as accurate as it would be from the observations where the gravity wave signal is absent (Fig. 4), and the EnKF was able to recover the gravity wave more accurately than OI even for observation periods much longer than the gravity wave period (Fig. 5). The EnKF was also found to more easily overcome the error associated with observing at or near multiples of the gravity wave period (the peaks in Fig. 5) because information is spread forward in time.

Changing the magnitude of the gravity wave, the time scale separation, or the coupling between the two modes changed the relative benefit of observing both time scales versus observing only the slow time scale. Large gravity waves are more difficult to ignore by the EnKF (Fig. 8), and it was found that for gravity waves above a certain magnitude the EnKF can better recover the slow mode if both time scales are observed partially than if the full slow mode is observed alone. On the other hand, the amount of information about the fast mode that can be recovered from entirely slow observations increases

as the correlation between the two modes increases, and we found a case (Fig. 10, center) for which the EnKF returns a better analysis of the fast mode with slow observations than OI does with mixed-time scale observations. Thus, changing the gravity wave parameters (Figs. 8, 10, and 12) also illustrated the flexibility of the EnKF relative to OI.

Model bias is generally large in the middle atmosphere and comes from a variety of sources, such as uncertainty in parameterized processes (Dee 2005; Polavarapu et al. 2005). Model error implies that distances between the truth and forecast will grow faster than estimated forecast errors, leading to increased filter divergence. In a sense, the SLOW experiments shown above can be interpreted as model-error experiments: because one mode is not captured in the analysis, the evolution of the forecast state is systematically different from the truth, resulting in larger analysis increments and the amplification of covariance estimation error. Because the analysis of the slow mode (reflected in  $E_w$ ) does not suffer greatly from the change of observation variable from MIX to SLOW (Figs. 4, 8, 10, and 12), the EnKF seems to be somewhat resistant to small systematic dynamical errors.

However, by introducing an error into the estimated gravity wave frequency or coupling parameter, it was found that capturing the fast mode becomes far more difficult when the gravity wave parameters assumed by the model differ from the true values. The inability to capture a correct gravity wave in this case also harms the slow mode analysis, and no benefit was found in using the EnKF rather than OI. This indicates that other methods, or a method for estimating and accounting for model error, might be necessary when assimilating unbalanced states.

Several issues remain that this study did not address. In the exL86 model, the gravity wave neither propagates away nor dissipates between observation times. Szunyogh et al. (2005) point out that in reality gravity wave events may not last through the entire observation interval and thus may not be captured by either the observations or the assimilation. The propagation of gravity waves and the role of observation spatial density are issues to address in higher-dimensional model studies.

In future studies it will also be important to address the comparative performance of 4D-Var for unbalanced truth states. It was mentioned in the introduction that variational methods can be made to handle unbalanced motion by statistically or analytically prescribing the covariance matrix, which is not evolved explicitly as in the Kalman filter. 3D and 4D variational methods represent a very different approach to the same problem and a detailed comparison between 4D-Var and the

EnKF is beyond the scope of a single paper, although a short comparison of the two algorithms is made in Neef (2007).

This study also did not address currently suggested modifications to the EnKF algorithm, such as inflation of covariances, the addition of stochastic model error, or the use of two ensembles in the EnKF, which are designed to keep the filter from diverging. A brief examination of such modifications is performed in Neef (2007). We also did not look at alternative Kalman filter-type schemes, such as deterministic ensemble filters (Tippett et al. 2003) or the so-called kernel or particle filter (Anderson and Anderson 1999; Xiong et al. 2006). Because the usefulness of these modified algorithms in practical data assimilation problems is still under consideration, future comparisons of these alternative algorithms should include an examination of the effects of different coexisting time scales.

*Acknowledgments.* This work has been supported by the Natural Sciences and Engineering Research Council of Canada and by the Canadian Foundation for Climate and Atmospheric Sciences, through the GCC and C-SPARC projects. LJN was supported by a University of Toronto Blythe Fellowship. The authors thank Ricardo Todling, Chris Snyder, and two anonymous reviewers for helpful comments and suggestions.

## APPENDIX

The optimal variances in the OI forecast error covariance matrix were computed numerically by sweeping

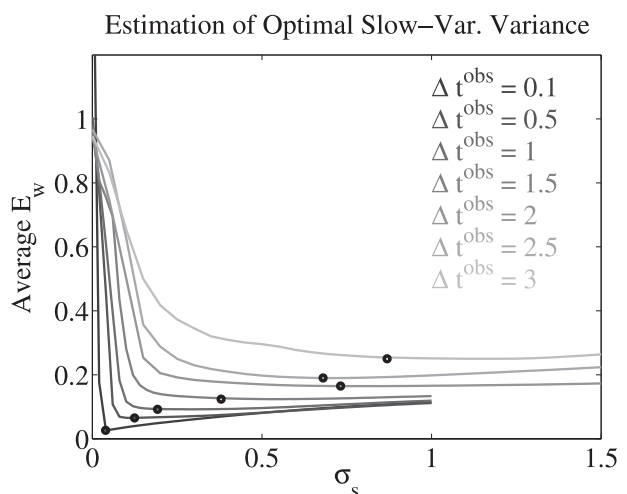


FIG. A1. RMS error in  $w$  as a function of the assumed slow variable ( $\phi$ ,  $w$ ) variance in the static OI covariance matrix for seven different observation intervals. The minima of the curves are indicated by black circles.



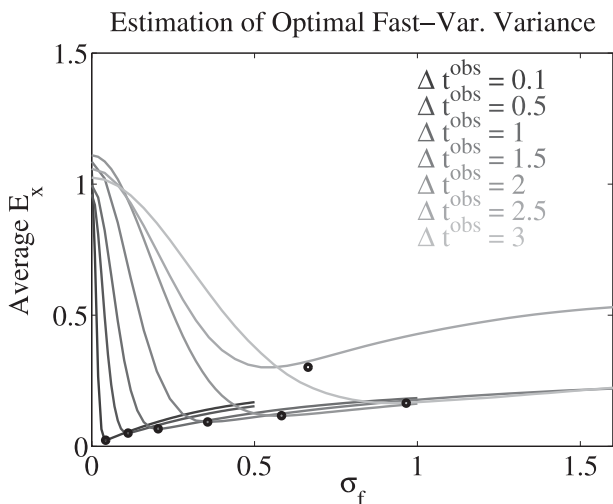


FIG. A2. RMS error in  $x$  as a function of the assumed fast variable variance ( $\sigma_f$ ) in the static OI covariance matrix. The minima of the curves are indicated by black circles.

over  $\sigma_s$  and  $\sigma_f$  for different values of  $\Delta t^{\text{obs}}$  and then performing a polynomial regression to find approximate functional dependencies of  $\sigma_s$  and  $\sigma_f$  on the observation interval.

Figure A1 shows the error measure  $E_w$  as a function of  $\sigma_s$  for seven different observation intervals, each line representing an average over 100 realizations. We performed a polynomial regression on these seven experiment sets and found that the dependence of the optimal  $\sigma_s$  on  $\Delta t^{\text{obs}}$  is approximately linear and is given by  $\sigma_s^{\text{opt}} = 0.3\Delta t^{\text{obs}}$ .

Figure A2 shows average error in  $x$  ( $E_x$ ), averaged over 100 runs for the same seven observation intervals, now as a function of  $\sigma_f$ . Again, the optimal assumed variance is a function of the observation interval, and a polynomial regression yields an approximate functional dependence, given by  $\sigma_f^{\text{opt}} = 0.03 + 0.14\Delta t^{\text{obs}} + 0.06(\Delta t^{\text{obs}})^2$ .

By similar experiments, we also found that the optimal  $\sigma_f$  is not sensitive to the choices of the gravity wave parameters  $\tilde{I}$ ,  $\epsilon$ , and  $b$ .

REFERENCES

Alexander, M. J., and C. Barnett, 2007: Using satellite observations to constrain parameterizations of gravity wave effects for global models. *J. Atmos. Sci.*, **64**, 1652–1665.  
 Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758.  
 Bergman, K. H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev.*, **107**, 1423–1444.  
 Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724.

Courtier, P., and O. Talagrand, 1990: Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus*, **42A**, 531–549.  
 Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.  
 Dee, D. P., 1991: Simplification of the Kalman filter for meteorological data assimilation. *Quart. J. Roy. Meteor. Soc.*, **117**, 365–384.  
 —, 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3323–3343.  
 Derber, J., and F. Bouttier, 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195–221.  
 Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99** (C5), 10 143–10 162.  
 Fisher, M., 2003: Background error covariance modelling. [Available online at [www.ecmwf.int/newsevents/training/meteorological\\_presentations/ppt/DA/Back\\_err.ppt](http://www.ecmwf.int/newsevents/training/meteorological_presentations/ppt/DA/Back_err.ppt).]  
 Gauthier, P., M. Buehner, and L. Fillion, 1999: Background-error statistics modelling in a 3D variational data assimilation scheme: Estimation and impact on the analyses. *Proc. ECMWF Workshop on Diagnosis of Data Assimilation Systems*, Reading, United Kingdom, ECMWF.  
 Ghil, M., S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: Application of estimation theory to numerical weather prediction. *Dynamic Meteorology: Data Assimilation Methods*, L. Bengtsson et al., Eds., Springer-Verlag, 139–224.  
 Houtekamer, P. L., and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **131**, 3269–3289.  
 Kalman, R., 1960: A new approach to linear filtering and prediction problems. *J. Basic Eng.*, **82**, 35–45.  
 Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.  
 Kepert, J. D., 2004: On ensemble representation of the observation-error covariance in the ensemble Kalman filter. *Ocean Dyn.*, **54**, 561–569.  
 Koshyk, J. N., B. A. Boville, K. Hamilton, E. Manzini, and K. Shibata, 1999: Kinetic energy spectrum of horizontal motions in middle-atmosphere models. *J. Geophys. Res.*, **104**, 27 177–27 190.  
 Lea, D. J., T. W. Haine, and R. F. Gasparovic, 2002: Observability of the Irminger Sea circulation using variational data assimilation. *Quart. J. Roy. Meteor. Soc.*, **132**, 1545–1576.  
 Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.  
 Lorenz, E. N., 1986: On the existence of a slow manifold. *J. Atmos. Sci.*, **43**, 1547–1558.  
 Miller, R. N., M. Ghil, and F. Gauthiez, 1994: Advanced data assimilation in strongly nonlinear dynamical systems. *J. Atmos. Sci.*, **51**, 1037–1056.  
 Neef, L. J., 2007: Balance dynamics and gravity waves in four-dimensional data assimilation. Ph.D. thesis, University of Toronto, 200 pp.  
 —, S. M. Polavarapu, and T. G. Shepherd, 2006: Four-dimensional data assimilation and balanced dynamics. *J. Atmos. Sci.*, **60**, 1840–1858.  
 Neelin, Y., Y. J. Rochon, and S. Polavarapu, 2009: Impact of tropospheric and stratospheric data assimilation on the mesosphere. *Tellus*, **61A**, 154–159.

- Parrish, D. F., and J. D. Derber, 1992: The National Meteorological Center spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Polavarapu, S., M. Tanguay, and L. Fillion, 2000: Four-dimensional variational data assimilation with digital filter initialization. *Mon. Wea. Rev.*, **128**, 2491–2510.
- , T. G. Shepherd, Y. Rochon, and S. Ren, 2005: Some challenges of middle atmosphere data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3513–3527.
- Rabier, F., 2005: Overview of global data assimilation developments in numerical weather-prediction centres. *Quart. J. Roy. Meteor. Soc.*, **131**, 3215–3233.
- Sankey, D., S. Ren, S. Polavarapu, Y. Nezlin, and S. Beagley, 2007: Impact of data assimilation filtering methods on the mesosphere. *J. Geophys. Res.*, **112**, D24104, doi:10.1029/2007JD008885.
- Szunyogh, I., E. Kostelich, G. Gyarmati, D. Patil, B. R. Hunt, E. Kalnay, E. Ott, and J. Yorke, 2005: Assessing a local ensemble Kalman filter: Perfect model experiments with the NCEP global model. *Tellus*, **57A**, 528–545.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490.
- Todling, R., and S. E. Cohn, 1994: Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. *Mon. Wea. Rev.*, **122**, 2530–2557.
- Wirosoetisno, D., and T. G. Shepherd, 2000: Averaging, slaving, and balance dynamics in a simple atmospheric model. *Physica D*, **141**, 37–53.
- Xiong, X., I. M. Navon, and B. Uzunoglu, 2006: A note on the particle filter with posterior Gaussian resampling. *Tellus*, **58A**, 456–460.