Ensemble-based data assimilation for thermally forced circulations

Altuğ Aksoy, Fuqing Zhang, John W. Nielsen-Gammon, and Craig C. Epifanio
Department of Atmospheric Sciences, Texas A&M University, College Station, Texas, USA

Received 20 December 2004; revised 22 March 2005; accepted 1 June 2005; published 24 August 2005.

[1] The effectiveness of the ensemble Kalman filter (EnKF) for thermally forced circulations is investigated with simulated observations. A two-dimensional, nonlinear, hydrostatic, non-rotating, and incompressible sea breeze model is developed for this purpose with buoyancy and vorticity as the prognostic variables. Model resolution is 4 km horizontally and 50 m vertically. Forcing is maintained through an explicit heating function with additive stochastic noise. Pure forecast experiments reveal that the model exhibits moderate nonlinearity. The strongest nonlinearity occurs along the sea breeze front at the time of peak sea breeze phase. Considerable small-scale error growth occurs at this phase for vorticity, while buoyancy is dominated by large-scale error as the direct result of the initial condition uncertainty. In the EnKF experiments, simulated buoyancy observations (with assumed error of 10^{-3} ms^{-2}) on land surface with 40-km spacing are assimilated every 3 hours. As a result of their resolution, the observations naturally sample the larger-scale flow structure. At the first analysis step, the filter is found to remove most of the large-scale error resulting from the initial conditions and the domain-averaged error of buoyancy and vorticity is reduced by about 83% and 42%, respectively. Subsequent analyses continue to remove error at a progressively slower rate and the error ultimately stabilizes within about 24 hours for both variables. At later model times, while mostly large-scale buoyancy errors due to the stochastic heating uncertainty are effectively removed, the filter also performs well at reducing smaller-scale vorticity errors associated with the sea breeze front. This is an indication that observations also contain useful small-scale information relevant at the scales of frontal convergence.

Citation: Aksoy, A., F. Zhang, J. W. Nielsen-Gammon, and C. C. Epifanio (2005), Ensemble-based data assimilation for thermally forced circulations, *J. Geophys. Res.*, *110*, D16105, doi:10.1029/2004JD005718.

1. Introduction

[2] One of the most prominent areas of investigation in modern numerical weather prediction is data assimilation, the process of utilizing the information content of observations to improve initial conditions for numerical models. Approaches to meteorological data assimilation generally fall into two main categories: the variational techniques (3DVAR and 4DVAR) and the ensemble-based sequential techniques (most notably the ensemble Kalman filter, EnKF). The variational methods (especially 4DVAR with its distinct advantage over 3DVAR due to its dynamical error covariance information) is today accepted as the more established method owing to its 20-year history of research at various scales of atmospheric phenomena and also to its current operational status at the European Centre for Medium-Range Weather Forecasts (ECMWF) [*Rabier et al.*, 2000] and Météo-France [*Gauthier and Thépaut*, 2001]. For a detailed discussion of the technique, see, for instance, *Zupanski et al.* [2002], and the references therein.

[3] The EnKF, on the other hand, is a relatively new method that was first proposed in the geophysical literature by *Evensen* [1994]. As a result of its several advantages over 4DVAR, the EnKF has received an increasing amount of research attention in recent years. One such advantage is the fact that tangent linear and adjoint versions of the forecast model are not required, leading to a much simpler implementation process. In addition, the EnKF is naturally integrated within an ensemble-based forecast model, providing sample estimates of very valuable flow-dependent forecast and analysis covariances with much less cost than fully integrating the covariance matrix forward in time. In addition, from a Bayesian perspective, the ensemble-mean state of the analysis provided by the EnKF represents the maximum likelihood state and serves as a direct estimate of the optimal solution to a variance-minimizing estimation problem, which makes the EnKF appealing statistically.

[4] Tests of the standard form of the EnKF have shown that the scheme performs well for large-scale models [*Evensen and van Leeuwen*, 1996; *Evensen*, 1997]. Today, meteorological applications of the EnKF to large-scale flows have nearly reached the point of operational testing [*Mitchell et al.*, 2002; *Houtekamer et al.*, 2004], while assessments have also been performed for massively parallel ocean circulation models [*Keppenne and Rienecker*, 2002]. While results from the above-mentioned studies are

0148-0227/05/2004JD005718$09.00

very encouraging and almost complete at global and synoptic scales, there is much progress to be made before the EnKF is established as a viable assimilation technique for numerical modeling of meteorological phenomena at smaller scales, where applications of the filter have been much more limited due to the implementation challenges introduced by fully three-dimensional flow characteristics, highly nonlinear microphysical processes, and much more complicated boundary condition issues. At regional and mesoscales, *Zhang et al*. [2005] have shown for an explosive winter storm case that the EnKF is very effective in keeping the analysis close to the truth simulation while also showing that the most effective error reduction occurred at larger scales compared to the smaller, marginally resolvable scales where reduction in error was less effective. Only in recent years have there been attempts to assimilate Doppler radar observations into cloud-scale supercell models, albeit with partial success, first through simulated observations [*Snyder and Zhang*, 2003; *Zhang et al.*, 2004], and then subsequently using real observations [*Dowell et al.*, 2004].

[5] Another very important limitation is the fact that all of the synoptic and mesoscale phenomena that have been put under the scrutiny of EnKF research to date can be regarded as a special class of flows exhibiting free dynamics; i.e., they do not proceed under the dominant influence of an external forcing. There is thus a wide range of dynamical systems, commonly known as chaotic forced-dissipative systems that range from the sea breeze circulation at the smaller-scale end of the spectrum to the climate at very large scales, that have not yet been subjected to assimilation experiments with the EnKF.

[6] There are two possible lines of reasoning to suggest that chaotic forced-dissipative systems may exhibit a sufficiently dispersive nature to warrant the application of data assimilation. The first is of an empirical nature: daily experience with local phenomena such as the sea breeze or the mountain breeze circulations suggests that the timing, strength, and penetration properties of the nonlinear frontal structures, as exhibited through cloud formation, convective activity, changes in air quality, and/or strong wind shifts, can differ significantly from day to day even under seemingly similar synoptic conditions [*Simpson*, 1994; *Miller et al.*, 2003]. The second argument is based on theoretical evidence related to the chaotic nature of nonlinear, forced-dissipative systems. At the large-scale end of atmospheric motions with timescales on the order of months to millennia, it has been shown that, even for severely truncated models, multiple stationary solutions exist with varying stability, and transitions between such solutions occur as a result of changes in the forcing parameters. There is in fact a plethora of published material on this topic; for an introductory discussion see, for instance, *Charney and DeVore* [1979], *Vickroy and Dutton* [1979], and *Mitchell and Dutton* [1981]. At the smaller scales, a similar behavior has been recently demonstrated for the sea and land breeze by *Feliks* [2004] with a two-dimensional model. Moreover, due to the strong dependence of the chaotic nature of forced-dissipative systems on the forcing, one might also conjecture that the predictability of such systems should be closely linked to the parameter space that controls the forcing and the system's response to forcing.

[7] The purpose of this paper is to address the feasibility of data assimilation techniques, and specifically of the EnKF, for nonlinear forced-dissipative atmospheric phenomena. To this end, the sea breeze circulation is chosen as the prototype of a strongly forced system, both because of its relatively simple structure and because of its relevance to day-to-day operational forecasting. To keep the problem simple, the sea breeze is modeled as the nonlinear response to a specific oscillating interior heat source in two dimensions, as in the linear theory of *Rotunno* [1983].

[8] The structure of this article is designed as follows: We briefly give some relevant background information on the EnKF in section 2. Section 3 describes the two-dimensional sea breeze model. Section 4 contains our analysis of the perfect-parameter results. We conclude with a summary and conclusions in section 5.

2. Ensemble Kalman Filter

[9] The Kalman filter (KF), in its original form, as developed by *Kalman* [1960], is an optimal (in the linear limit) Bayesian filter. It combines information from observations and flow-dependent forecast covariances to update a given forecast state under the constraint of minimizing variance. As a result of its appealing sequential form, the KF quickly became one of the most widely used and investigated estimation methods in both engineering and science literature [the reader is referred to *Gelb* [1986] for a rigorous yet accessible introduction to the subject]. A critical issue about the original form of the KF is that the time integration of the error covariance matrix requires two successive matrix multiplications. Each of these operations involves matrices of the dimension of the state vector which, for meteorological applications, is usually on the order of 10^6–10^7. With today's computer resources, such a computation is not feasible.

[10] To address this problem, the ensemble Kalman filter (EnKF) was developed as a modification to the KF which uses an ensemble of model forecasts to directly estimate the background error covariance structure. As such, the ensemble can be viewed as a collection of samples that provides a sampled estimate of the mean state and the covariance matrix. The KF equations are modified to reflect the update of the ensemble mean, while the time integration of the covariance matrix is performed implicitly through the time integration of individual ensemble members:

$$\begin{aligned} \overline{\mathbf{x}}^a &= \overline{\mathbf{x}}^f + \mathbf{K}\left(\mathbf{y}^0 - \mathbf{H}\overline{\mathbf{x}}^f\right), \\ \mathbf{K} &= \mathbf{P}^f\mathbf{H}^{\mathrm{T}}\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R}\right)^{-1}. \end{aligned} \tag{1}$$

Here, $\overline{\mathbf{x}}^f$ and $\mathbf{P}^f$ represent the forecast (background) mean state and covariance matrix (sampled directly from the ensemble perturbations), respectively, while $\overline{\mathbf{x}}^a$ represents the updated analysis mean state. Observations are represented by the observation vector $\mathbf{y}^0$ and observational error matrix $\mathbf{R}$. The mapping between model space and observational space is accomplished through the matrix $\mathbf{H}$. In this form, $\mathbf{H}$ represents the first-order linear approximation to a possibly nonlinear observation function $\mathcal{H}$. The

matrix $\mathbf{K}$ is known as the Kalman gain matrix and contains the coefficients of the linear combination of forecast state vector and observations. This "standard" EnKF formulation was first proposed by *Evensen* [1994] and further tested by *Evensen and van Leeuwen* [1996] and *Evensen* [1997] who concluded that the scheme performed well for large-scale models. For more background on the application of the EnKF in the atmospheric sciences, see, for example, *Mitchell and Houtekamer* [2000], *Keppenne* [2000], *Hamill and Snyder* [2000], *Anderson* [2001], *Keppenne and Rienecker* [2002], *Snyder and Zhang* [2003], *Anderson* [2003], and *Zhang et al.* [2004].

[11] Sequential processing of observations leads to considerable simplification of the analysis scheme and is therefore an attractive alternative to assimilating the entire observation vector at once [e.g., *Whitaker and Hamill*, 2002; *Snyder and Zhang*, 2003]. This is possible because observation errors are assumed to be independent so that the order in which observations are processed does not influence the final estimate of the state (as long as the analysis error covariance, $\mathbf{P}^a$, is updated at each step). For a single, scalar observation y^0, the term $\mathbf{P}^f\mathbf{H}^{\mathrm{T}}$ of equation (1) becomes a column vector c of dimension of the state, while the term $\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R}$ becomes a scalar d. Updating of the ensemble mean is performed similar to equation (1) but by replacing the Kalman gain $\mathbf{K}$ by its simpler form as follows:

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{c}\left(y^0 - \mathbf{H}\overline{\mathbf{x}}^f\right)/d\,. \tag{2}$$

[12] In addition to updating the ensemble mean, the covariance matrix also needs to be updated, which is achieved through updating the difference of each ensemble member from the mean. Following the ensemble square-root filter (EnSRF) formulation proposed by *Whitaker and Hamill* [2002], the update for the difference of each ensemble member from the mean (perturbations) is calculated by

$$\mathbf{x}^a - \overline{\mathbf{x}}^a = \left[\mathbf{I} - \beta\left(\hat{\mathbf{c}}/\hat{d}\right)\mathbf{H}\right]\left(\mathbf{x}^f - \overline{\mathbf{x}}^f\right)\,, \tag{3}$$

where $\beta = [1 + (r/\hat{d})^{1/2}]^{-1}$, $\hat{\mathbf{c}}$ and $\hat{d}$ are sample estimates of $\mathbf{c}$ and d, and r is a scalar representing the observational error at the respective observation location. The analysis step defined in this manner is an alternative to perturbing observations in the analysis step of the EnKF [*Burgers et al.*, 1998; *van Leeuwen*, 1999].

[13] In passing we note that the use of the acronym "EnKF", in its strictest sense, refers to the "standard" perturbed-observations formulation of the filter. However, because of the fact that our goal here is not to draw the attention to any differences between perturbed-observation and deterministic filters but merely use the EnSRF due to its both practical and theoretical advantages, we decided to use the acronym "EnKF" throughout the paper to represent the broader class of ensemble-based Kalman filters.

[14] Another modification that is commonly employed within recent EnKF applications pertains to the rank deficiency problem associated with undersampling when using small ensembles. *Houtekamer and Mitchell* [1998] noted that the EnKF analysis could be improved by preventing distant observations from influencing the update of a grid point. They argued that this was mainly due to the fact that small-sized ensembles tended to overestimate covariances between greatly separated grid points and, as a remedy, applied a cutoff radius beyond which covariances were not calculated. Commonly known as covariance localization, *Houtekamer and Mitchell* [2001] have since experimented with a more sophisticated Schur product approach where the ensemble-based covariance estimates are multiplied element by element with a distance-dependent correlation function that varies from 1.0 at the observation location to 0.0 at some predefined cutoff distance (a widely used such correlation function is *Gaspari and Cohn*'s [1999] compactly supported fifth-order function). *Houtekamer and Mitchell* [2001] maintain that "the effect of localization is to increase the effective size of the ensemble". For this study, we have chosen to employ *Gaspari and Cohn*'s [1999] compactly supported fifth-order function as a means to localize the covariance structure with no posterior covariance inflation applied.

3. Sea Breeze Model and Filter Design

3.1. Model Numerics

[15] As in the linear theory of *Rotunno* [1983], an idealized prototype problem is adopted in which the sea breeze circulation is modeled as the response to a specified oscillating heat source in the fluid interior. For simplicity, the flow is taken to be Boussinesq with only hydrostatic and non-rotating disturbances in two dimensions. The equations of motion can then be written in vorticity-stream function form as

$$\begin{aligned}\frac{\partial \eta'}{\partial t} + (\overline{u} + u')\frac{\partial \eta'}{\partial x} + w'\frac{\partial \eta'}{\partial z} + \frac{\partial b'}{\partial x} &= \kappa_\eta \frac{\partial^2 \eta'}{\partial z^2}\,,\\ \frac{\partial b'}{\partial t} + (\overline{u} + u')\frac{\partial b'}{\partial x} + w'\frac{\partial b'}{\partial z} + N^2 w' &= Q + \kappa_b \frac{\partial^2 b'}{\partial z^2}\,,\end{aligned} \tag{4}$$

where $\mathbf{u} = (u', w')$ is the disturbance fluid velocity, $\eta' = \frac{\partial u'}{\partial z}$ is the hydrostatic vorticity, and $b' = g\theta'/\theta_0$ is the Boussinesq buoyancy. Here g is the acceleration of gravity and $\theta = \theta_0 + \theta_B(z) + \theta'(x, z, t)$ is the potential temperature, where θ_0 and θ_B are reference and background potential temperatures, respectively, while $N^2 = (g/\theta_0)(\partial\theta_B/\partial z)$ is the square of the background Brunt-Väisälä frequency.

[16] Dissipation in the model is represented through the vertical diffusion of vorticity and buoyancy, where κ_η and κ_b are the constant eddy diffusion coefficients for η and b, respectively. Free-slip and thermal insulation conditions are applied at the lower boundary of the model domain by setting both η' and $\partial b'/\partial z$ to zero at $z = 0$.

[17] Forcing takes the form of an explicit heating function Q defined by

$$Q = [A_0 \cos\omega t + \varsigma(t)]\left[\left(\frac{1}{2} + \frac{1}{\pi}\tan^{-1}\frac{x}{x_0}\right)e^{-z/z_0}\right]\,. \tag{5}$$

Here, A_0 is a constant heating amplitude; x_0 and z_0 are horizontal and vertical length scales of heating, respectively; ω is the diurnal frequency; and ζ is a stochastic white noise term. In this functional form, the initial time corresponds to a maximum heating phase.

[18] *Walter* [2004] has shown through a scale analysis that, for weak background wind $\overline{u}$, the nonlinearity of the

model is controlled by a non-dimensional amplitude parameter defined by $\varepsilon = A_0(N^2\omega z_0)^{-1}$. His findings indicate that the system begins to demonstrate nonlinear effects (developing density currents) for $\varepsilon > 0.8$; for smaller values of ε, nonlinearity in the system remains weak. Following this reasoning, then, the heating amplitude A_0 becomes the controlling parameter of nonlinearity when the other parameters N^2, ω, and z_0 are kept constant. Moreover, it is important to note that in a viscous system nonlinearity will also be dependent on the strength of diffusion which is measured by the non-dimensional Reynolds number $\mathrm{Re} = \omega z_0^2 \kappa_\eta^{-1}$. Although there is no direct connection between the Reynolds number and the degree of nonlinearity in this formulation, one can in general argue that the smaller the Reynolds number becomes, the larger should be the critical value of ε where effects of nonlinearity begin to be observed. In other words, all other parameters being equal, larger vertical diffusivity would require a larger heating amplitude to force the flow into the nonlinear regime.

[19] The prognostic equations are integrated using leapfrog time differencing with a weak Asselin time filter [*Asselin*, 1972] for all terms except the vertical diffusion. The diffusion terms are integrated using trapezoidal time differencing. All spatial derivatives are computed using second-ordered centered differences, and a weak horizontal filter is included to help stabilize the steep gradients that occur at the fronts. From a theoretical point of view, the leapfrog scheme with the Asselin filter may not be the optimal numerical integration technique for a stochastic dynamical system [see, for instance, *Ewald et al.*, 2004; *Penland*, 1996]. However, our goal here is not to obtain the most accurate numerical solution to the analytic sea breeze equation (4). Rather, we consistently base our truth and forecast integrations on the sea breeze system defined through our numerics, time step, and noise handling.

[20] The horizontal (cross-shore) and vertical dimensions of the forecast domain are 500 km and 3 km, respectively, with the coast located at the center (at $x = 0$) and the land located to the right of the coast. There are two Rayleigh-damping sponge layers to the left and right of the forecast domain, each of width 300 km, and another above the forecast domain with a 2-km depth. Grid spacing is 4 km horizontally and 50 m vertically, thus optimally resolving the mesoscale properties of the sea breeze circulation while marginally resolving the frontal structure.

3.2. Ensemble and Filter Characteristics

[21] A "climatological" initialization scheme was utilized for this study. The key steps of this scheme are as follows: (1) Using the sea breeze model and an initial background state (zero perturbation vorticity and zero perturbation potential temperature throughout the domain), a 15-day hourly time series is produced. (2) To allow for the initial adjustment from the background state, only days 4 through 15 are used for choosing initial states of ensemble members. (3) Initial states are chosen randomly from the available time series with a normal probability distribution centered at the initial time of day (local noon or maximum heating phase for this application) of the ensemble runs. In other words, the maximum heating states within the 15-day time series have the highest probability to be selected into the ensemble and a standard deviation of 8 hours is used for the normal distribution around the central maximum heating state.

[22] In most of the experiments, an ensemble size of 50 is used. In order to obtain a perfect ensemble, truth is taken as one of the ensemble members. Unless noted otherwise in the article, other properties of the filter chosen for this study are: radius of influence of 100 grid points (400 km horizontal and 5 km vertical), and simulated land surface buoyancy observations with error 10^{-3} ms^{-2} and spacing of 40 km.

4. Results

4.1. Model Behavior

[23] To begin, a demonstration of the mean-state behavior of the sea breeze model will be presented for a moderately nonlinear regime. No EnKF analysis is performed during this run. Mean horizontal wind ($\overline{u}$) is set at a weak value of 0.5 ms^{-1} and the Brunt-Väisälä frequency (N) is chosen to be 10^{-2} s^{-1}. For the heating profile, horizontal and vertical length scales (x_0 and z_0) are set at 10 km and 500 m, respectively. The vertical diffusion coefficients of buoyancy and vorticity (κ_b and κ_η) are both chosen to be 0.25 m^{2}s^{-1}, in order to have a Prandtl number $\mathrm{Pr} = \kappa_\eta/\kappa_b = 1$. These settings correspond approximately to a Reynolds number on the order of 50–100 and indicate that the effects of vertical diffusion on nonlinearity should be relatively small. A mean value of 7×10^{-6} ms^{-3} is used for heating amplitude which roughly corresponds to a value of 1.9 for the nonlinearity parameter ε. With this ε, nonlinear features are expected to be locally significant near the coast while the overall nonlinearity is still expected to be moderate [*Walter*, 2004]. At higher heating amplitudes, the model becomes numerically unstable. Finally, the standard deviation of the stochastic heating amplitude is set at 4×10^{-6} ms^{-3}, resulting in an expected range of 0.8–3.0 for the nonlinearity parameter ε (although the actual influence of stochasticity on nonlinearity is smaller than the suggested range because of the effects of time filtering and diffusion).

4.1.1. Diagnosis of Sea Breeze and Land Breeze Cycles

[24] To demonstrate the model behavior, key stages of a 24-hour sea and land breeze cycle are presented in Figure 1. Panels from top to bottom indicate the time evolution of the corresponding sea/land breeze cycle, beginning with the 123-hour forecast time that corresponds to the onset of the sea breeze and continuing with 6-hourly intervals until the 141-hour forecast time that corresponds to the peak land breeze.

[25] The sea breeze is triggered by the differential daytime heating that reaches its maximum on land at the 120-hour (noon, maximum heating phase) forecast time (not shown). At this time, while surface convergence and vertical motion set up at the coast, surface winds do not yet indicate the existence of a sea breeze circulation. At 123 hours (Figures 1a and 1e) the low-level convergence and vertical motion strengthen and the most intense convergence moves inland as weak onshore surface winds develop across the coastline. At 126 hours into the sea breeze cycle (not shown), which coincides with the neutral heating phase, a distinct frontal boundary forms at the leading edge of the sea breeze circulation. The surface buoyancy over land reaches its peak at this time, marking a 6-hour phase difference

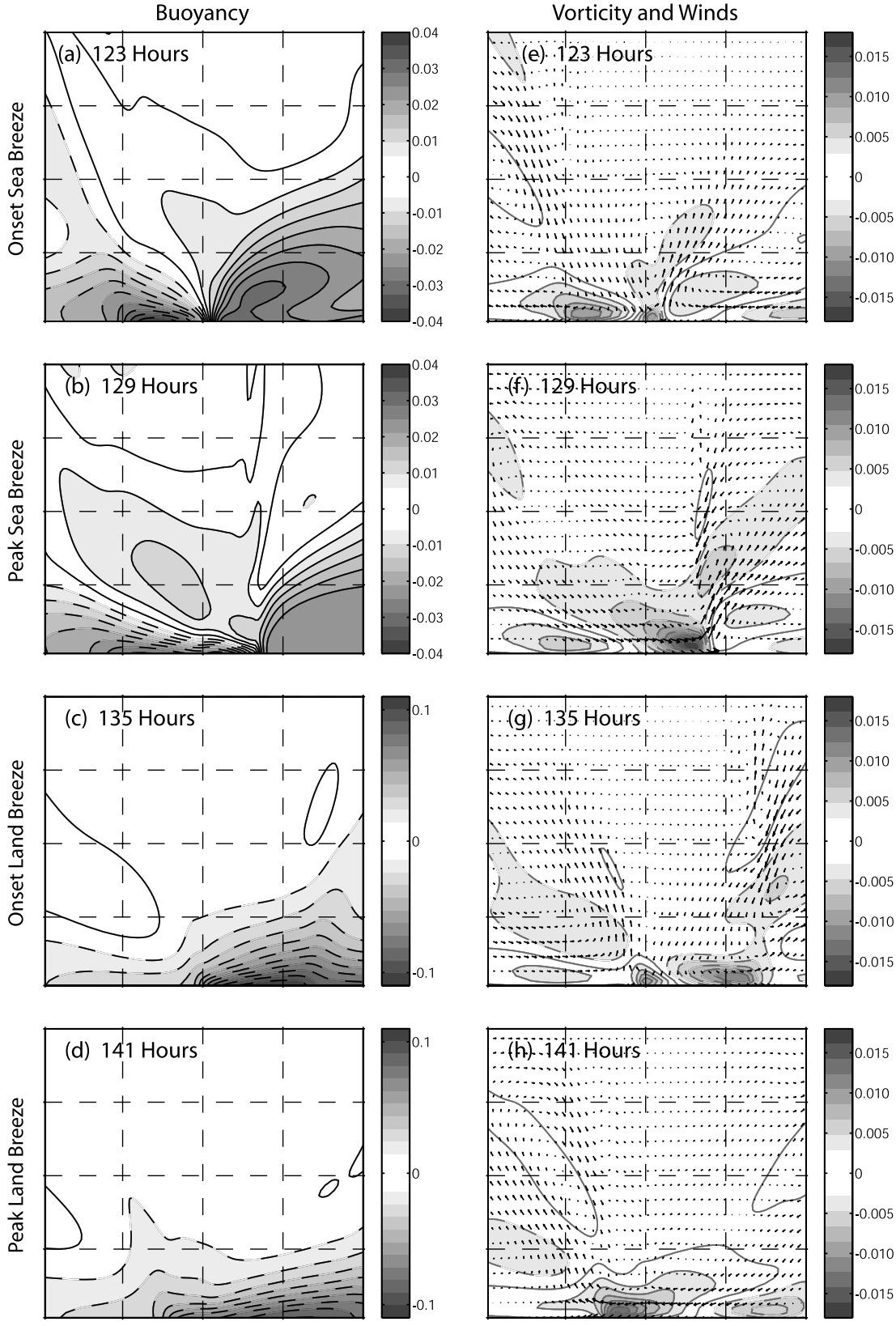

Figure 1. Evolution of the sea and land breeze cycles as observed through the distribution of buoyancy (left panels, ms^{-2}), vorticity (right panels, s^{-1}), and the wind field (right panels, ms^{-1}) within the forecast domain at following forecast times: 123 hours (onset of the sea breeze, Figures 1a and 1e), 129 hours (peak sea breeze, Figures 1b and 1f), 135 hours (onset of the land breeze, Figures 1c and 1g), and 141 hours (peak land breeze, Figures 1d and 1h). Solid (dashed) contours indicate positive (negative) values. Contour intervals are 0.004 ms^{-2} (Figures 1a and 1b), 0.01 ms^{-2} (Figures 1c and 1d), and 0.002 s^{-1} (Figures 1e–1h).

between maximum heating and maximum temperature over land. The associated buoyancy gradient continues to strengthen the circulation well into the cooling phase of the cycle, and both onshore and vertical winds reach their peak values at 129 hours (Figures 1b and 1f). The frontal boundary also matures and penetrates almost 100 km inland. While sloping regions of ascent (over land) and descent (over water) mark the ray paths of gravity waves forced by the heating gradient across the coastline, note also that the nonlinear convergence of the frontal boundary has led to a significant gravity-wave signal extending upwards from the boundary layer (Figures 1f and 1g).

[26] As the cooling phase progresses, the temperature gradient across the coast becomes more negative and the flow at the coast eventually reverses to produce a land breeze. The time of peak land breeze is shown in Figures 1d and 1h. Note that an offshore-propagating front and associated gravity wave signal are again apparent. However, comparison of the Figures 1c and 1d shows that the land breeze front is significantly weaker than the front associated with the sea breeze. This asymmetry in the intensity of the land and sea breezes, which is also commonly observed in real sea/land breezes [*Simpson*, 1994], is apparently the result of the difference in stratification ahead of the respective fronts, with weak stratification ahead of the sea breeze front and essentially unaltered background stratification ahead of the land breeze [*Walter*, 2004].

4.1.2. Structure and Evolution of Ensemble Spread and Covariance

[27] The distributions of standard deviation, as a measure of ensemble spread, of variables buoyancy, vorticity, and vertical motion are shown in Figure 2 as they evolve in the sea and land breeze cycle corresponding to Figure 1. Figure 2 indicates a distinct diurnal variability of the ensemble spread both in space and intensity. Both buoyancy and vorticity exhibit spread that is mostly confined to the surface. For buoyancy, this is primarily a result of the initial condition spread projecting onto horizontally uniform perturbations, which can only weaken slowly through vertical diffusion. Vorticity spread is also influenced by heating through the $\partial b'/\partial x$ term in equation (4) yet has a more localized response with most of the signal concentrated around the frontal boundary. Thus, while the vorticity spread is large near the sea and land breeze fronts, the vertically propagating diurnal gravity waves appear to have very weak signal in the vorticity spread. While the deterministic part of the forcing projects exclusively onto the diurnal frequency, the stochastic part of the forcing projects equally onto all frequencies. Thus, while waves excited by the forcing follow a single ray path pair in the linear limit [*Rotunno*, 1983], waves excited by the stochastic forcing exhibit no such preference. Furthermore, most of the diurnal gravity wave spread associated with initial condition uncertainty of vorticity (not shown) will have propagated out of the domain during the three days before the displayed times.

[28] The spatial distribution of the spread of vertical motion is entirely attributable to the timing and location of the frontal boundary. Additionally, the vertically propagating frontal gravity waves (Figure 1) appear to have their strongest signal in the ensemble spread of vertical motion. This is mainly due to these waves' projecting most of their kinetic energy in the vertical direction as they propagate away from the strong horizontal temperature gradient at the surface.

[29] Corresponding to Figure 2, pointwise (i.e., computed at the same grid point) covariance and correlation coefficient between buoyancy and vorticity are plotted at the different phases of the sea breeze cycle (Figure 3). The covariance structure (Figures 3a–3d) reflects the combined response of buoyancy and vorticity spread and thus is mostly concentrated at the surface and is triggered by the frontal boundary. Meanwhile, a much more intricate structure of the correlation coefficient (Figures 3e–3h) is exhibited throughout the sea breeze cycle. In the linear limit [*Rotunno*, 1983], variations in the heating intensity would lead to a dominantly negative (positive) correlation structure between buoyancy and vorticity over the sea (land) that is independent of the phase of the sea breeze circulation. This simple structure predicted by the linear theory is modified considerably by the nonlinearities in the model. The global evolution of the buoyancy-vorticity correlation structure during the 6th forecast day (Figures 3e–3h) reflects the remnants of the initial-condition correlation structure (not shown). Meanwhile, the local variations near the front are caused by the vertical gravity waves that emanate from the frontal boundary during the peak sea breeze phase (Figures 3f–3h, over land). In contrast, during the land breeze phase, the global structure of the initial-condition correlation is only weakly modified by the presence of the front (Figures 3g–3h, over sea).

4.1.3. Error Dynamics

[30] The 10-day evolution of the ensemble spread (measured by the domain-averaged standard deviation of the ensemble) shows that buoyancy (Figure 4a) and vorticity (Figure 4c) both converge during the first 96 hours, following an initial period of vorticity spread growth. After 96 hours, the buoyancy spread, which is nearly horizontally uniform (Figures 2a–2d), exhibits little or no diurnal variation while the vorticity spread, which is associated primarily with fronts, reaches a diurnal maximum when the sea breeze is strongest. A corresponding asymmetry between the buoyancy and vorticity spread is also apparent in their response to initial phase error and stochastic heating error (not shown).

[31] Because of its horizontally uniform structure, the overall distribution of buoyancy spread remains virtually unaltered by the diurnal sea breeze cycle (compare to Figures 2a–2d). This is both because of the weak advective tendency due to the horizontal uniformity of the buoyancy spread and the weak vertical diffusion of buoyancy. By similar reasoning, then, the localized structure of the vorticity spread that is correlated to the location and intensity of the sea breeze front (compare to Figures 2e–2h) is quickly advected out of the domain through the propagation of the front. Consequently, domain-averaged vorticity spread exhibits a pronounced diurnal variability. In addition, because of the tendency to be advected out of the domain at the timescales of frontal propagation, the memory of initial vorticity spread is not retained within the system beyond 96 hours. Once initial spread information is "forgotten", vorticity spread is then controlled by the stochastic heating. During this period (96–240 hours in Figure 4c), vorticity spread grows rapidly through the intensification of the sea breeze front which, due to the stochastic heating, enhances

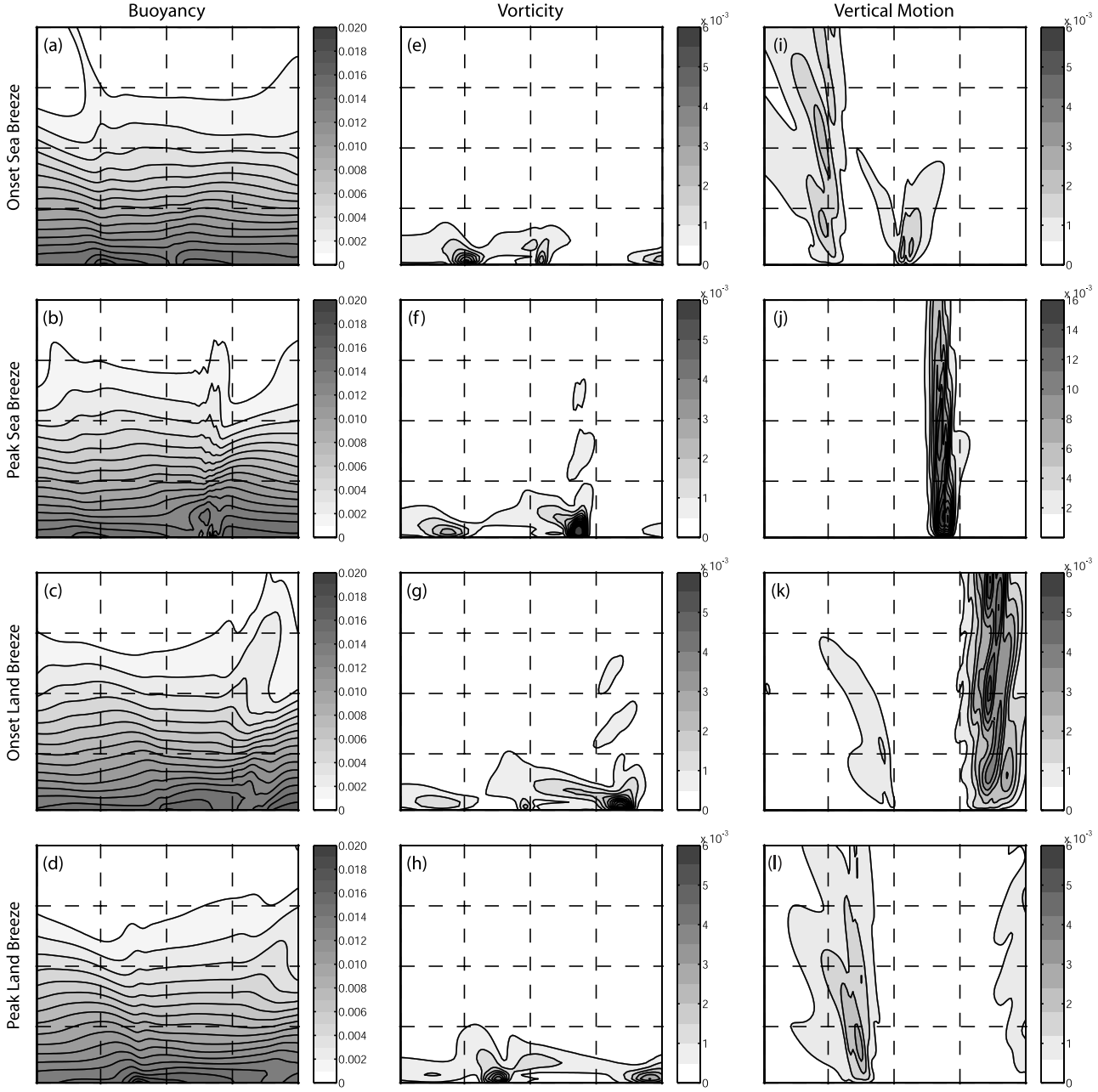

Figure 2. Evolution of the sea and land breeze cycles as observed through the distribution of the standard deviation of buoyancy (left panels, ms^{-2}), vorticity (middle panels, s^{-1}), and vertical motion (right panels, ms^{-1}) within the forecast domain at forecast times same as in Figure 1. Contour intervals are 0.001 ms^{-2} (Figures 2a–2d), 5×10^{-4} s^{-1} (Figures 2e–2h), 6×10^{-4} ms^{-1} (Figures 2i, 2k, and 2l), and 1.6×10^{-3} ms^{-1} (Figure 2j).

location error (compare to Figure 2f) but does not decay as quickly both because of the lingering sea breeze front and the newly formed weaker land breeze front (compare Figures 2f to 2h, the vorticity spread within both the old sea breeze front and the newly formed land breeze front is about half of its value of its peak phase). As a result, the spread fluctuates diurnally about a mean value that linearly depends on the standard deviation of stochastic heating amplitude (not shown).

[32] To further investigate the spatial structure of error and confirm our hypothesis about the differences between buoyancy and vorticity spread, spectra of total power of the difference buoyancy (Figure 4b) and difference vorticity (Figure 4d) are plotted as a function of horizontal wave number at the four different phases of the sea breeze cycle, where differences are determined from the mean for each ensemble member and then the spectra are summed across ensemble members and vertical levels. In order to minimize aliasing due to the "half-wavelength structure" of buoyancy (i.e. the common case of alternating values from one half of the domain to the other), its power spectra are computed after patching the entire forecast domain horizontally. This

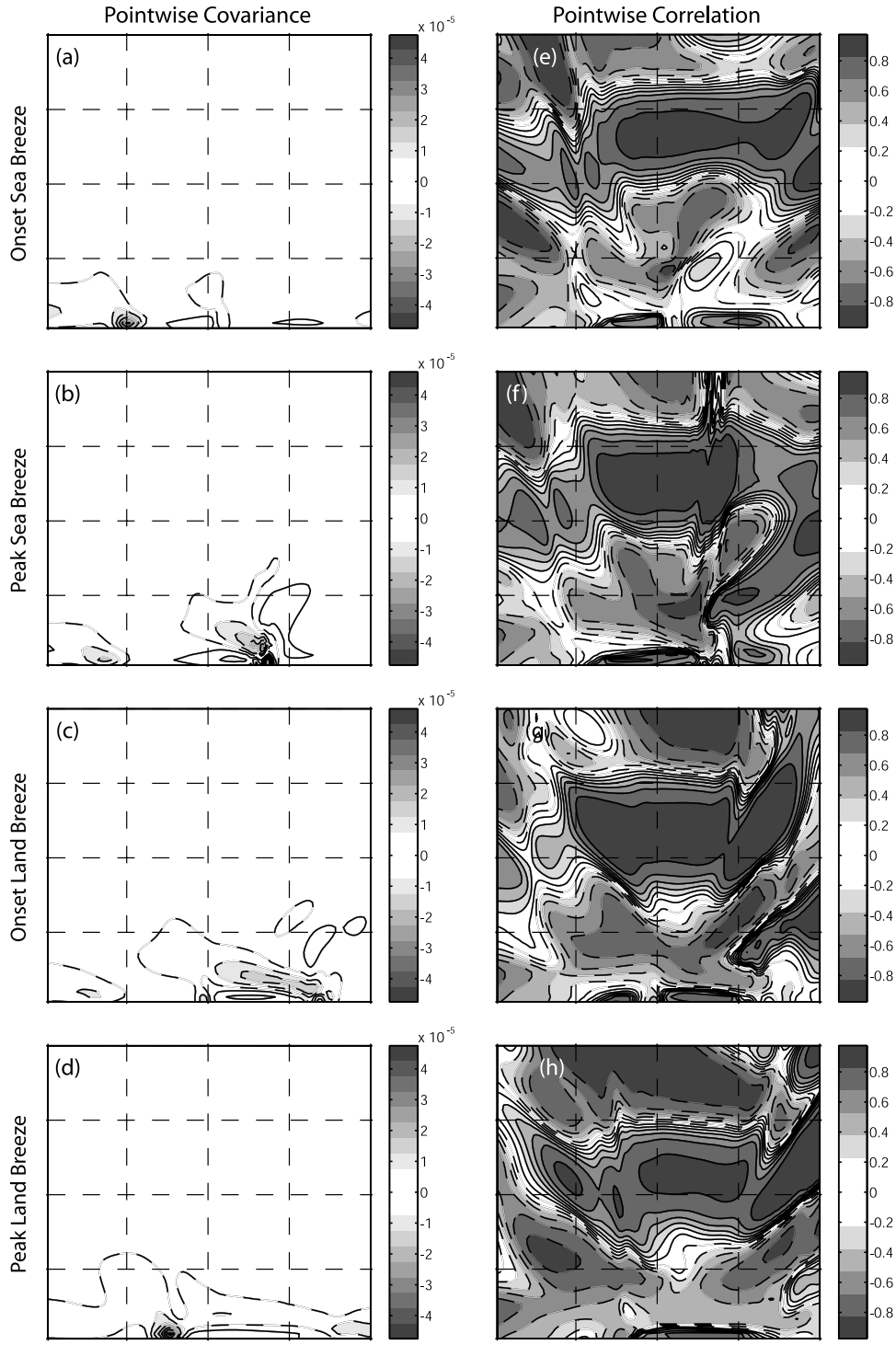

Figure 3. Distribution of pointwise covariance (left panels, ms^{-3}) and pointwise correlation coefficient (right panels) between buoyancy and vorticity at forecast times same as in Figure 1. Solid (dashed) contours indicate positive (negative) values. Contour intervals are 5×10^{-6} ms^{-3} (Figures 3a–3d) and 0.2 (Figures 3e–3h).

gives rise to wave number "1/2" on buoyancy power spectra plots. The large-scale structure of spread is evident for both variables. Yet, as conjectured in the previous paragraph, while both variables lose power selectively at larger scales as the experiment proceeds into the sixth day, buoyancy remains to be dominated by wave numbers 0 and 1/2, which is an indication of retained memory of initial conditions. Vorticity power, on the other hand, appears to have become almost "white" between wave numbers 0–10 and exhibits a much "flatter" spectrum compared to buoy-

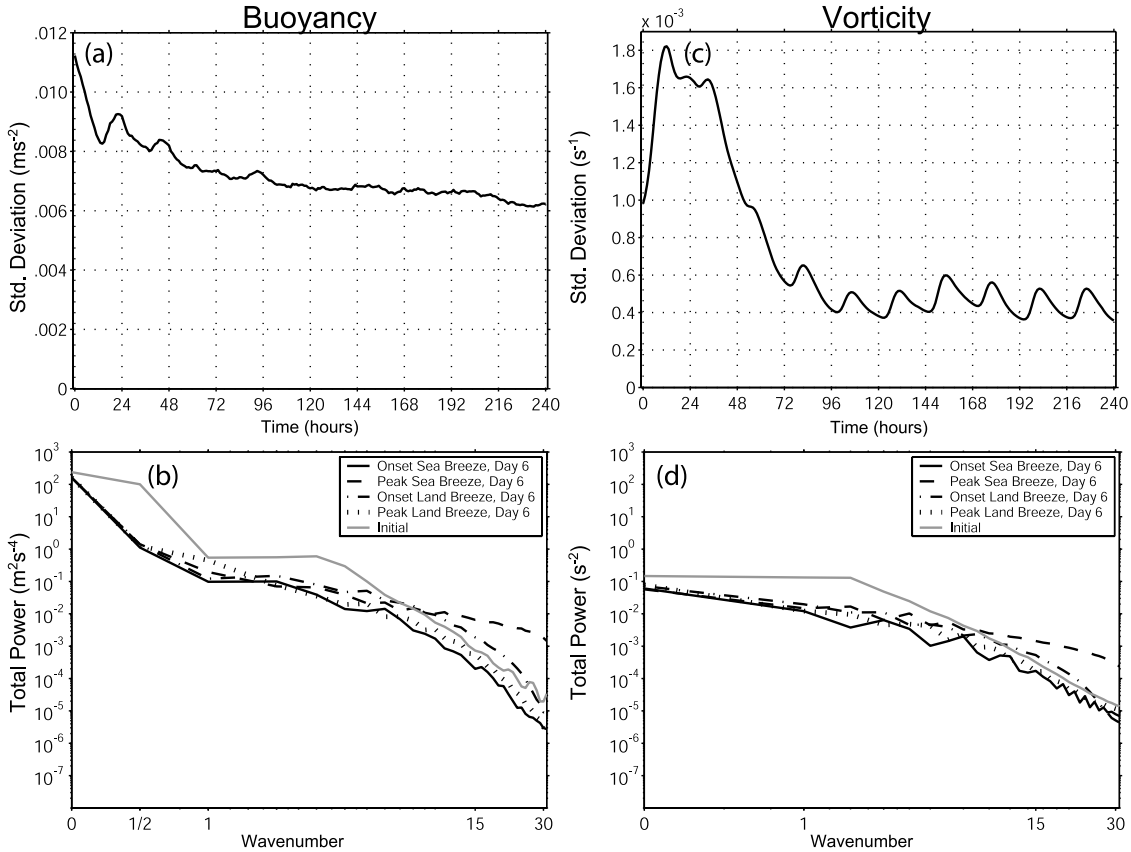

Figure 4. Upper panels: 240-hour evolution of domain-averaged standard deviation of buoyancy (Figure 4a, ms^{-2}) and vorticity (Figure 4c, s^{-1}) from a pure forecast run. Lower panels: Domain-total power spectra of the cumulative difference of each ensemble member and the ensemble mean for buoyancy (Figure 4b, m^2s^{-4}) and vorticity (Figure 4d, s^{-2}) at initial time and forecast times 123 hours (onset of the sea breeze phase, solid), 129 hours (peak sea breeze phase, dashed), 135 hours (onset of the land breeze phase, dashed-dotted), and 141 hours (peak land breeze phase, dotted). Wave number 1 corresponds to a wavelength of 500 km while wave number 31 approximately corresponds to the double Nyquist wavelength of 16 km.

ancy. This confirms our hypothesis that, compared to buoyancy, initial conditions have a much smaller impact on the later evolution of the vorticity spread.

[33] Active frontal dynamics in the sea breeze change the power spectra considerably and introduce smaller-scale structure to both variables. While the power of buoyancy spread increases by about 3 orders of magnitude at wave number 30, the overall spectrum remains "red" with power in wave numbers 0 and 1/2 still larger by about 5 and 3 orders of magnitude, respectively. Consequently, domain-averaged buoyancy spread exhibits no noticeable diurnal signal at later stages of the experiment. On the other hand, the vorticity spread reveals a much more pronounced influence of smaller scales at the peak sea breeze phase so that the power at smaller scales becomes almost comparable in magnitude to the power at larger scales. The relatively large power at smaller scales projects into rapid error growth and results in the distinct diurnal signal in the vorticity spread. Despite this considerable contribution of error at smaller scales during the peak sea breeze phase, the system appears to lack a mechanism to grow smaller-scale errors into larger scales. As a result, beyond day-to-day variability, overall level of error remains relatively stable (i.e., error saturation), not showing signs of long-term growth.

4.2. Filter Performance

[34] Filter evaluation is carried out using the sequential square-root EnKF as explained in section 2 and its basic properties as described in section 3.2. In these tests, simulated land-surface observations (thus spanning only half of the domain at surface) of buoyancy were used with 40-km horizontal spacing and an assumed observational error of 10^{-3} ms^{-2}. Analyses are performed every 3 hours, beginning with the 3-hour forecast. While the entire simulation domain (including the sponge layers) are used to perform analyses and simulate observations, statistics are computed within the main (interior) domain. Thus, the dimension of the state vector is 55000, while the number of observations simulated for each analysis cycle is 13 and the ensemble comprises 50 members. In the following sections, results are presented through the investigation of error structure and filter sensitivity.

4.2.1. Error Characteristics

[35] The evolution of root-mean square (RMS) error and ensemble spread of prognostic variables buoyancy and

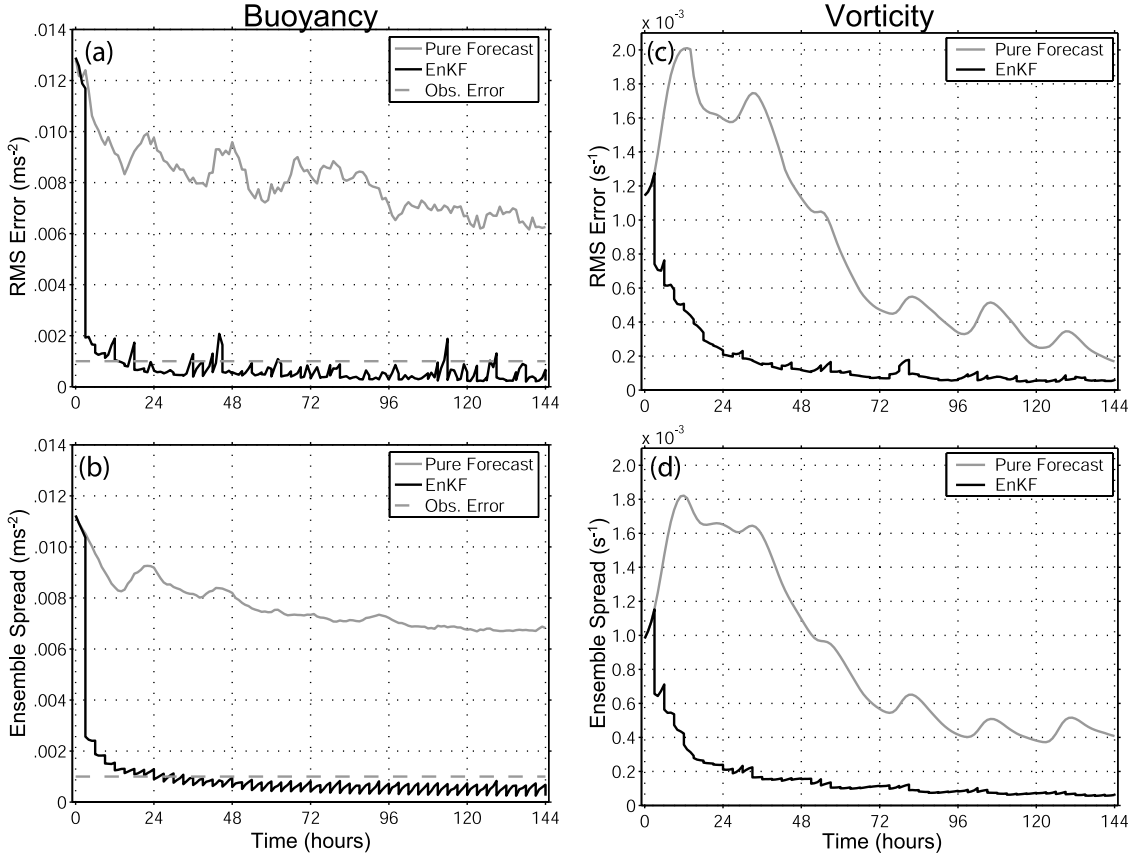

Figure 5. 144-hour evolution of EnKF buoyancy RMS error (Figure 5a, ms^{-2}), buoyancy domain-averaged standard deviation (Figure 5b, ms^{-2}), vorticity RMS error (Figure 5c, s^{-1}), and vorticity domain-averaged standard deviation (Figure 5d, s^{-1}) plotted in solid black. Analyses are performed at 3-hour forecast intervals. In each panel, respective metric from the pure forecast run is plotted in solid gray for comparison.

vorticity from a 144-hour run is shown in Figure 5. During the first analysis cycle, a large reduction in buoyancy error occurs with a correction of about 83% (Figure 5a), while vorticity error reduction is 42% (Figure 5c). Subsequent error reduction between hour 3 and hour 144 analysis is about 90% for both variables. In addition, buoyancy error appears to grow considerably during the 3-hour forecast runs between each analysis cycle although the error introduced in that manner is immediately removed at the following analysis step so that the overall error saturates at a level that is comparable to the observation accuracy (10^{-3} ms^{-2}). Meanwhile, vorticity error growth between analysis cycles has a diurnal nature with strongest error growth occurring during the peak sea breeze phases. Eventually and similar to buoyancy, vorticity error settles to a level that is both controlled by observational accuracy and stochastic uncertainty. The error settling time for both variables is about one day.

[36] To compare how error is reduced during an analysis step, the domain-wide distribution of prior and posterior RMS error of buoyancy and vorticity is shown in Figure 6 at the first analysis step (3 hours) and two peak sea breeze phases (9 and 129 hours). At 3-hour model time (Figures 6a–6d), domain-wide error of both variables is removed considerably in accordance with the large first-analysis error reduction that is observed in Figure 5 (errors in the left quarter of the domain are not reduced due to the constraint of the radius of influence). Further investigation reveals that most of the error removed is larger-scale error that is associated with the ensemble initialization (mostly phase differences between different times that were chosen as ensemble members). Remaining errors are concentrated mostly around the strong temperature gradients which are the focus of nonlinear frontal dynamics.

[37] At 9 hours, which is the third analysis step and coincides with the peak sea breeze phase, the forecast RMS error (Figures 6e, 6g) of both variables exhibit significant small-scale error growth compared to the 3-hour analysis RMS error. For both variables, the strongest error is located at the surface around the front with weaker error extending vertically and showing a distinct gravity wave structure. In addition, a much weaker buoyancy error is found at the surface that extends uniformly over land which is the result of the stochastic heating error. At this time, small-scale errors due to active frontal dynamics dominate over larger-scale errors due to stochastic heating, illustrating that initial frontal errors are not immediately eradicated by the first few analyses. On the other hand, the larger-scale buoyancy error over land surface is completely removed.

[38] At 129 hours, which is the peak sea breeze phase of the sixth day of the experiment, the distribution of forecast RMS error of buoyancy (Figure 6i) shows a distinct large-

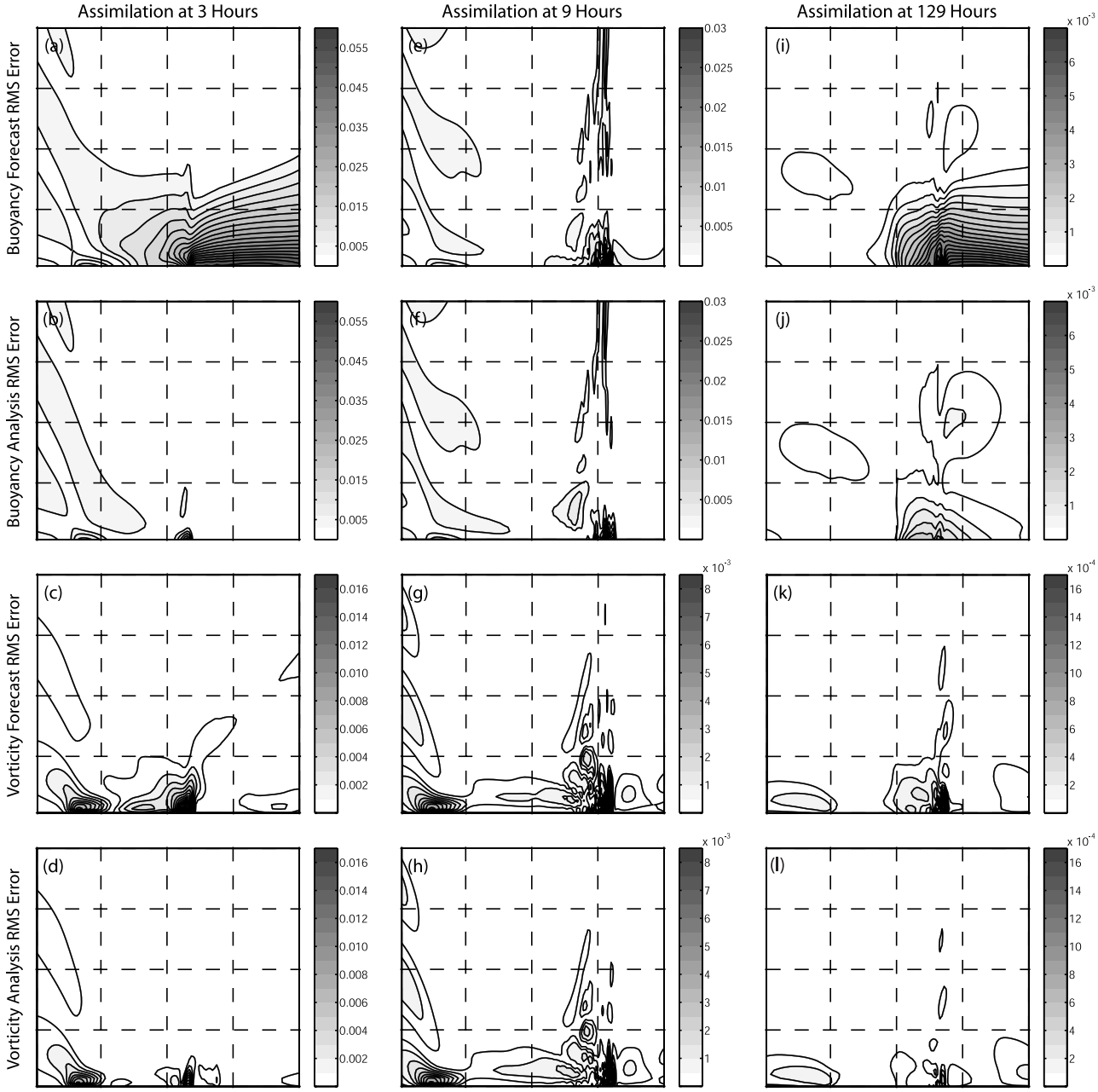

Figure 6. Domain distribution of prior (forecast) and posterior (analysis) RMS error of buoyancy (ms^{-2}) and vorticity (s^{-1}) at 3 hours (first analysis, Figures 6a–6d), 9 hours (peak sea breeze phase of day 1, Figures 6e–6h) and 129 hours (peak sea breeze phase of day 6, Figures 6i–6l) from the same EnKF run as in Figure 5. Contour intervals are 0.003 ms^{-2} (Figures 6a and 6b), 0.001 s^{-1} (Figures 6c and 6d), 0.0015 ms^{-2} (Figures 6e and 6f), 5×10^{-4} s^{-1} (Figures 6g and 6h), 3.5×10^{-4} ms^{-2} (Figures 6i and 6j), and 1×10^{-4} s^{-1} (Figures 6k and 6l).

scale structure much similar to the forecast error distribution at 3 hours (Figure 6a). Embedded in that structure, a small-scale signal is also discernible at the sea breeze front location which is strong in magnitude yet very localized at the surface and weakly extends upward. Comparison of this structure to the 9-hour distribution (Figure 6e) suggests that repeated 3-hour analyses have removed most of the remaining small-scale error caused by initial conditions, so that forecast error is now predominantly generated through the stochastic uncertainty of heating. In contrast to buoyancy but consistent with Figure 2f, forecast RMS error of vorticity appears to have retained its small-scale structure (Figure 6k). As explained before, this is the result of vorticity's sensitivity to the location error of the front that also enables the error to be advected out of the forecast domain. Similar to previous times, the filter at this time successfully removes most of the large-scale error of buoyancy (Figure 6j), leaving some buoyancy error behind the sea breeze. Similarly, considerable reduction also occurs in vorticity error (Figure 6l) and most of the error within the

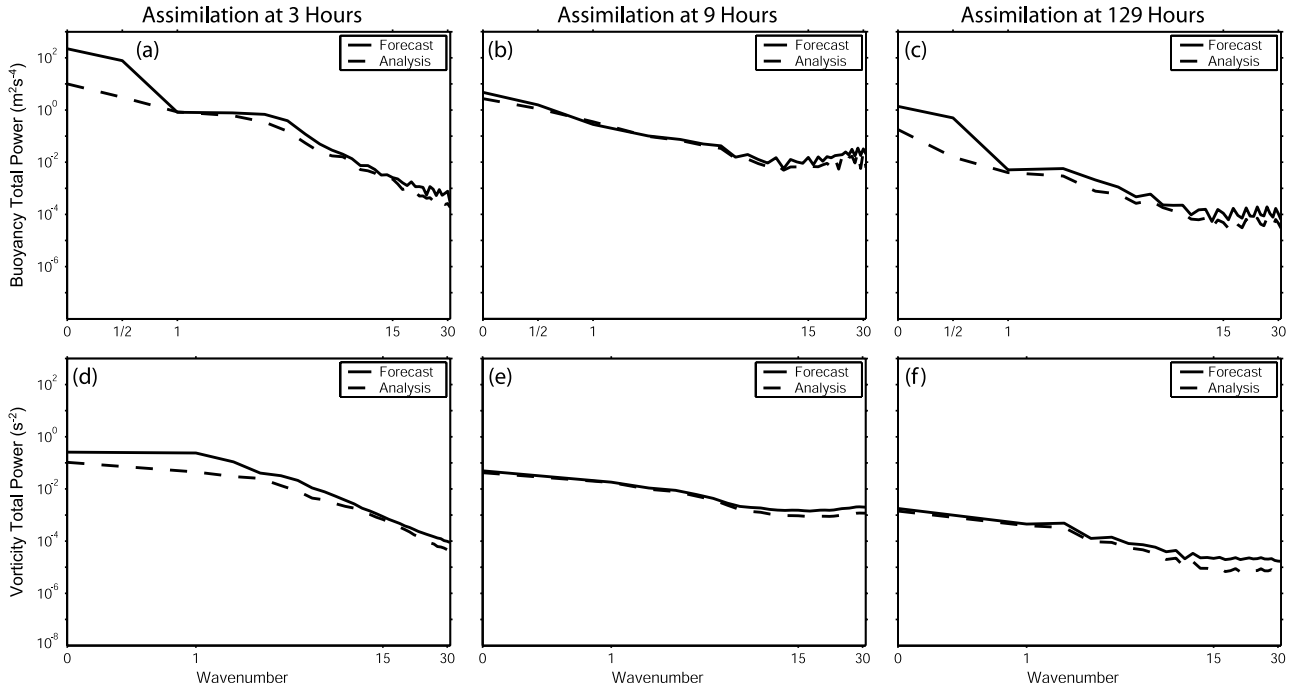

Figure 7. Domain-total power spectra of the cumulative difference of each ensemble member and the ensemble mean for buoyancy (Figures 6a–6c, m^2s^{-4}) and vorticity (Figures 6d–6f, s^{-2}) at forecast (prior, solid lines) and analysis (posterior, dashed lines) steps of 3 hours, 9 hours, and 129 hours.

frontal and gravity-wave structure is eliminated. From this visual analysis, surface buoyancy observations with 40 km spacing, in addition to their information content at their natural large-scale resolution, appear to contain valuable information relevant at the scales of frontal convergence. As a result, the filter not only effectively removes large-scale errors due to stochastic heating but performs also well at reducing small-scale error due to nonlinearities at the front.

[39] To further examine the scale-sensitive behavior of the filter, power spectra of the two variables are plotted at 3, 9, and 129 hours (Figure 7). At the first analysis step of the model (Figures 7a and 7d), both variables exhibit the greatest error reduction at large scales primarily as a result of the error content of initial conditions being predominantly large-scale. At 9 hours, because of the fact that considerable small-scale error still exists in the forecast domain (compare to Figures 6b and 6h) that further intensifies during the active sea breeze phase, power spectra of both variables exhibit significantly larger energy at small scales compared to the initial time (Figures 7b and 7e). This small-scale buoyancy error growth naturally occurs faster than the large-scale error growth normally induced by stochastic heating so that the increase in the power at wave numbers 0, 1/2, and 1 is smaller for the 9-hour forecast. Consequently, the larger-scale information content of buoyancy observations at 9 hours contributes very little to the EnKF analysis and thus reduction in large-scale power is small (almost nonexistent for vorticity). At 129 hours, as the small-scale error becomes saturated after repeated 3-hourly analyses, stochastic heating uncertainty becomes dominant and generates large-scale buoyancy error which is then more effectively removed by the filter (Figure 7c). On the other hand, vorticity error still continues to be mainly concentrated at the front and retains its small-scale structure. As a result, reduction in power occurs mostly at smaller scales (Figure 7f) and its magnitude becomes comparable to that of the large-scale reduction at the 3-hour assimilation (Figure 7d). This confirms our hypothesis that surface buoyancy observations with 40-km spacing contain sufficient small-scale information so that the filter performs well at reducing especially the small-scale vorticity errors associated with the nonlinear frontal processes.

4.2.2. Sensitivity to Observation Accuracy, Ensemble Size, Analysis Frequency, Radius of Influence, Observation Spacing, and Type of Observations

[40] Experiments are also carried out to investigate the extent and nature of the sensitivity of the filter to various filter-related parameters. The results are presented in the form of comparative RMS error plots of the unobserved variable vorticity (Figure 8). In these plots, comparison is made between the parameter value in the control experiment (discussed in detail in section 4.1) and two other parameter values, one larger and one smaller than the control parameter value. In general, vorticity error is found to be sensitive to all parameters especially during the first 24 hours of the assimilation experiments. In all of the cases, errors begin with a relatively large difference between the experiments and, except observation accuracy, converge toward each other at later stages of the experiments so that error differences become progressively smaller. An interesting behavior is observed with the sensitivity to observation accuracy: While more accurate observations (10^{-4} ms^{-2}, one order of magnitude smaller than the control experiment) have almost no positive impact on vorticity error, loss of observation accuracy (10^{-2} ms^{-2}) results in considerably larger error (a more

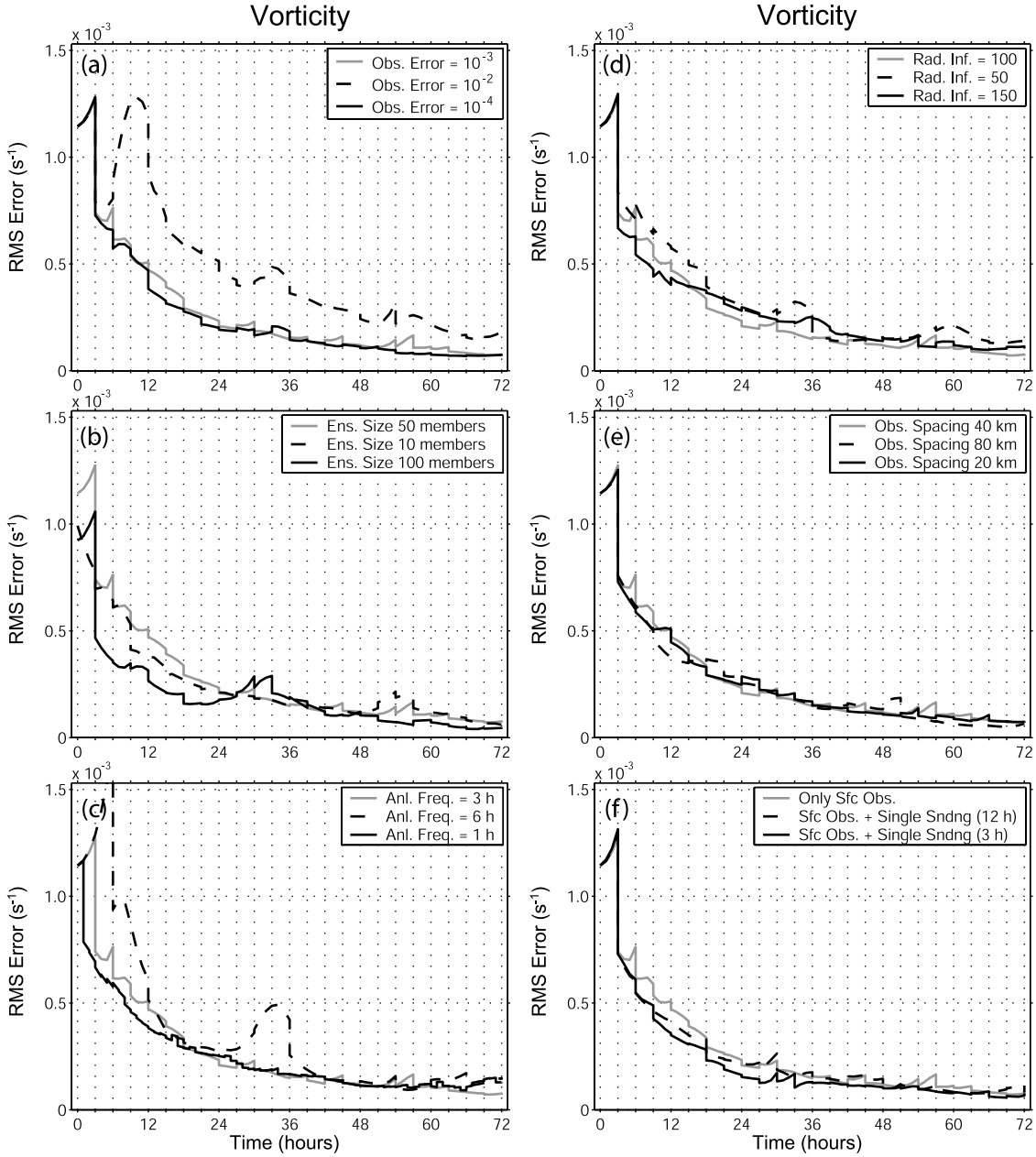

Figure 8. Sensitivity of RMS error (s^{-1}) of the unobserved variable vorticity to (a) observation accuracy, (b) ensemble size, (c) analysis frequency, (d) radius of influence, (e) observation spacing, and (e) assimilation of a single additional sounding.

proportional response to observation accuracy is exhibited by buoyancy error which is not shown here).

[41] Finally, one common behavior is believed to be linked to the large-scale nature of buoyancy. Parameters that exhibit similar characteristics in this respect are observation spacing (Figure 8e) and assimilation of an additional buoyancy sounding observation located about 90 km inland from the coast (Figure 8f). In both cases, vorticity error is not particularly sensitive to the variations in the chosen parameter. Apparently, the scale dependence of the information content of buoyancy observations remains mostly unchanged so that the smaller-scale features of the flow are not sampled more effectively by either an additional sounding or a reduction of observation spacing to 20 km.

5. Summary and Conclusions

[42] This article documents the effectiveness of the ensemble Kalman filter for a thermally forced nonlinear two-dimensional sea breeze model with forcing that is maintained through an explicit spatially and diurnally varied heating function with an added stochastic component. Pure forecast experiments reveal that the model exhibits moderate levels of overall nonlinearity. Ensemble spread of both buoyancy and vorticity is found to remain

stabilized (error saturation) at a certain level and does not show signs of large-scale growth over a 10-day period. Strongest nonlinearity coincides with the peak sea breeze phase of the circulation in timing and with the sea breeze front spatially. Considerable small-scale error growth occurs at this phase, which is most pronounced in the vorticity field. Nevertheless, at other phases of the sea breeze, the model tends to diffuse and advect vorticity errors out of the domain and lacks a mechanism to translate small-scale errors produced during the peak sea breeze phase into larger-scale errors at later times. As a result, the overall vorticity error appears stabilized and fluctuates in a diurnal manner with distinct day-to-day variability. However, buoyancy retains its memory of large-scale initial-condition error for a much longer duration because of its horizontally uniform distribution and weaker diffusion. This results in a dominantly large-scale error structure for buoyancy so that it exhibits no diurnal variability. EnKF control experiments are performed with simulated surface buoyancy observations on land that are placed 40 km apart and sequentially assimilated at 3-hour intervals. At the first analysis step, the filter is observed to successfully remove most of the large-scale phase-difference errors resulting from the initial conditions. At this step, domain-averaged error for buoyancy and vorticity is reduced by about 83% and 42%, respectively. Subsequent analyses continue to remove error at an increasingly slower rate and error ultimately saturates within about 24 hours at a level that is proportional to observation accuracy.

[43] Surface buoyancy observations with 40-km spacing are found to selectively resolve the large-scale features of the buoyancy field. As a result, the filter performs much better at larger scales for buoyancy. This becomes especially important at later stages of an assimilation experiment when most of the large-scale initial-condition error is already removed from the forecast domain. Stochastic heating then consistently results in large-scale errors in buoyancy which are then effectively removed by the filter at each analysis step. Meanwhile, small-scale vorticity errors are mostly induced by nonlinear processes near the front during the peak sea breeze phase. As a result, even though error growth of vorticity becomes more pronounced during the peak sea breeze phases, long-term vorticity error remains small because growth only occurs for a limited duration during the diurnal cycle. Moreover, the filter is observed to perform well in reducing the small-scale vorticity error during such active phases. We believe that 40-km surface buoyancy observations, in addition to their natural large-scale information content consistent with their spacing, contain sufficient small-scale information so that the filter is effective at both removing large-scale errors due to the stochastic heating and small-scale errors due to the nonlinearities of the sea breeze front.

[44] Sensitivity experiments demonstrate that the RMS error of the unobserved variable vorticity is most sensitive to observation accuracy, ensemble size, analysis frequency, and radius of influence. On the other hand, vorticity error reduction by the filter is not sensitive to decreased observation spacing to 20 km or the assimilation of an additional single sounding. Apparently, the scale dependence of the information content of buoyancy observations does not change so that the additional sounding or the reduction of observation spacing do not contribute significantly to the better sampling of the smaller-scale structure of vorticity.

[45] There are a number of limitations of the model worth mentioning briefly, as we believe that they are relevant in terms of sea breeze dynamics and the application of the EnKF. Dynamical simplifications were made to achieve computational efficiency: The Coriolis force is omitted because it does not have a direct impact on the nonlinearity of the sea breeze circulation; the hydrostatic assumption enables vorticity to be independent of the horizontal distance so that the inversion algorithm (between vorticity and stream function) becomes much simpler to implement; and the assumption of dry dynamics eliminates one entire equation and allows the model to be integrated forward much more efficiently. Such simplifications undoubtedly make our sea breeze model less realistic yet more accommodating in terms of understanding and controlling otherwise complicated interactions among model and filter components. Last, an important limitation is related to our choice of representing thermal forcing through an explicit heating function rather than a flux-based scheme that implicitly resolves heating.

[46] We believe that the current model setup offers a sufficient level of simplicity to allow for the preliminary investigation of a novel application of the EnKF to forced-dissipative systems, while retaining most of the key characteristics of sea breeze dynamics such as the inertia-gravity-wave structure and the nonlinear sea breeze front. A major future goal is to apply our improved understanding of the interactions between the forced nature of the sea breeze circulation and the EnKF to an imperfect-model setting. This will allow us to perform simultaneous estimation of the state and the parameters of the model as well as to investigate model error characteristics that arise from parameter uncertainty. Another interesting approach is to estimate source/distribution properties of a scalar tracer through the integration of a chemical concentration model which would facilitate the investigation of connections between the sea breeze circulation and the transport of certain inert chemicals. Both of these studies are already underway and the results will be reported elsewhere in near future.

[47] **Acknowledgments.** The authors are grateful to Chris Snyder, Rich Rotunno, Jeff Anderson, Lee Panetta, and Amy Stuart for their comments on the sea breeze model, ensemble initiation, and filter design. The comments of three anonymous reviewers on an earlier version of the manuscript have helped tremendously in improving both its content and cohesiveness. This research is sponsored by the GTRI/HARC Project H24-2003 and the NSF grant ATM-0205599.

References

Anderson, J. L. (2001), An ensemble adjustment Kalman filter for data assimilation, *Mon. Weather Rev.*, *129*, 2884–2903.

Anderson, J. L. (2003), A local least squares framework for ensemble filtering, *Mon. Weather Rev.*, *131*, 634–642.

Asselin, R. (1972), Frequency filter for time integrations, *Mon. Weather Rev.*, *100*, 487–490.

Burgers, G., P. J. van Leeuwen, and G. Evensen (1998), Analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.*, *126*, 1719–1724.

Charney, J. G., and J. G. DeVore (1979), Multiple flow equilibria in the atmosphere and blocking, *J. Atmos. Sci.*, *36*, 1205–1216.

Dowell, D., F. Zhang, L. Wicker, C. Snyder, and A. Crook (2004), Wind and temperature retrievals in the 17 May Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments, *Mon. Weather Rev.*, *132*, 1982–2005.

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.

Evensen, G. (1997), Advanced data assimilation for strongly nonlinear dynamics, *Mon. Weather Rev.*, *125*, 1342–1354.

Evensen, G., and P. J. van Leeuwen (1996), Assimilation of Geosat altimeter data for the Agulhas Current using the ensemble Kalman filter with a quasigeostrophic model, *Mon. Weather Rev.*, *124*, 85–96.

Ewald, B., C. Penland, and R. Temam (2004), Accurate integration of stochastic climate models with application to El Niño, *Mon. Weather Rev.*, *132*, 154–164.

Feliks, Y. (2004), Nonlinear dynamics and chaos in the sea and land breeze, *J. Atmos. Sci.*, *61*, 2169–2187.

Gaspari, G., and S. E. Cohn (1999), Construction of correlation functions in two and three dimensions, *Q. J. R. Meteorol. Soc.*, *125*, 723–757.

Gauthier, P., and J.-N. Thépaut (2001), Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Météo-France, *Mon. Weather Rev.*, *129*, 2089–2102.

Gelb, A. (Ed.) (1986), *Applied Optimal Estimation*, 374 pp., MIT Press, Cambridge, Mass.

Hamill, T. M., and C. Snyder (2000), A hybrid ensemble Kalman filter-3D variational analysis scheme, *Mon. Weather Rev.*, *128*, 2905–2918.

Houtekamer, P. L., and H. Mitchell (1998), Data assimilation using an ensemble Kalman filter technique, *Mon. Weather Rev.*, *126*, 796–811.

Houtekamer, P. L., and H. Mitchell (2001), A sequential ensemble Kalman filter for data assimilation, *Mon. Weather Rev.*, *129*, 123–137.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen (2004), Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations, *Mon. Weather Rev.*, *133*, 604–620.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, *82D*, 95–108.

Keppenne, C. (2000), Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter, *Mon. Weather Rev.*, *128*, 1971–1981.

Keppenne, C., and M. M. Rienecker (2002), Initial testing of a massively parallel ensemble Kalman filter with the Poseidon isopycnal ocean general circulation model, *Mon. Weather Rev.*, *130*, 2951–2965.

Miller, S. T. K., B. D. Keim, R. W. Talbot, and H. Mao (2003), Sea breeze: Structure, forecasting, and impacts, *Rev. Geophys.*, *41*(3), 1011, doi:10.1029/2003RG000124.

Mitchell, H., and P. L. Houtekamer (2000), An adaptive ensemble Kalman filter, *Mon. Weather Rev.*, *128*, 416–433.

Mitchell, H., P. L. Houtekamer, and G. Pellerin (2002), Ensemble size, balance, and model-error representation in an ensemble Kalman filter, *Mon. Weather Rev.*, *130*, 2791–2808.

Mitchell, K. E., and J. A. Dutton (1981), Bifurcations from stationary to periodic solutions in a low-order model of forced, dissipative barotropic flow, *J. Atmos. Sci.*, *38*, 690–716.

Penland, C. (1996), A stochastic model of IndoPacific sea surface temperature anomalies, *Physica D*, *96*, 534–558.

Rabier, F., H. Jarvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons (2000), The ECMWF operational implementation of four-dimensional variational assimilation: I. Experimental results with simplified physics, *Q. J. R. Meteorol. Soc.*, *126*, 1143–1170.

Rotunno, R. (1983), On the linear theory of land and sea breeze, *J. Atmos. Sci.*, *40*, 1999–2009.

Simpson, J. E. (1994), *Sea Breeze and Local Wind*, 234 pp., Cambridge Univ. Press, New York.

Snyder, C., and F. Zhang (2003), Assimilation of simulated Doppler radar observations with an ensemble Kalman filter, *Mon. Weather Rev.*, *131*, 1663–1677.

Van Leeuwen, P. J. (1999), Comment on "Data assimilation using an ensemble Kalman filter technique," *Mon. Weather Rev.*, *127*, 1374–1377.

Vickroy, J. G., and J. A. Dutton (1979), Bifurcation and catastrophe in simple, forced, dissipative quasi-geostrophic flow, *J. Atmos. Sci.*, *36*, 42–52.

Walter, K. R. (2004), *The Nonlinear Dynamics of the Sea Breeze*, 112 pp., M.S. thesis, Texas A&M Univ., College Station.

Whitaker, J. S., and T. M. Hamill (2002), Ensemble data assimilation without perturbed observations, *Mon. Weather Rev.*, *130*, 1913–1924.

Zhang, F., C. Snyder, and J. Sun (2004), Impacts of initial estimate and observations availability on convective-scale data assimilation with an ensemble Kalman filter, *Mon. Weather Rev.*, *132*, 1238–1253.

Zhang, F., Z. Meng, and A. Aksoy (2005), Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation, *Mon. Weather Rev.*, in press.

Zupanski, D., M. Zupanski, E. Rogers, D. F. Parrish, and G. J. DiMego (2002), Fine-resolution 4DVAR data assimilation for the Great Plains tornado outbreak of 3 May 1999, *Weather Forecasting*, *17*, 506–525.

A. Aksoy, C. C. Epifanio, J. W. Nielsen-Gammon, and F. Zhang, Department of Atmospheric Sciences, Texas A&M University, College Station, TX 77843-3150, USA. (fzhang@tamu.edu)