

On the Selection of Localization Radius in Ensemble Filtering for Multiscale Quasigeostrophic Dynamics

YUE YING

Department of Meteorology and Atmospheric Science, and Center for Advanced Data Assimilation and Predictability Techniques, The Pennsylvania State University, University Park, Pennsylvania, and National Center for Atmospheric Research, Boulder, Colorado

FUQING ZHANG

Department of Meteorology and Atmospheric Science, and Center for Advanced Data Assimilation and Predictability Techniques, The Pennsylvania State University, University Park, Pennsylvania

JEFFREY L. ANDERSON

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 7 November 2017, in final form 26 December 2017)

ABSTRACT

Covariance localization remedies sampling errors due to limited ensemble size in ensemble data assimilation. Previous studies suggest that the optimal localization radius depends on ensemble size, observation density and accuracy, as well as the correlation length scale determined by model dynamics. A comprehensive localization theory for multiscale dynamical systems with varying observation density remains an active area of research. Using a two-layer quasigeostrophic (QG) model, this study systematically evaluates the sensitivity of the best Gaspari–Cohn localization radius to changes in model resolution, ensemble size, and observing networks. Numerical experiment results show that the best localization radius is smaller for smaller-scale components of a QG flow, indicating its scale dependency. The best localization radius is rather insensitive to changes in model resolution, as long as the key dynamical processes are reasonably well represented by the low-resolution model with inflation methods that account for representation errors. As ensemble size decreases, the best localization radius shifts to smaller values. However, for nonlocal correlations between an observation and state variables that peak at a certain distance, decreasing localization radii further within this distance does not reduce analysis errors. Increasing the density of an observing network has two effects that both reduce the best localization radius. First, the reduced observation error spectral variance further constrains prior ensembles at large scales. Less large-scale contribution results in a shorter overall correlation length, which favors a smaller localization radius. Second, a denser network provides more independent pieces of information, thus a smaller localization radius still allows the same number of observations to constrain each state variable.

1. Introduction

Covariance localization (Hamill et al. 2001) is a pragmatic procedure for ensemble filters (Houtekamer and Zhang 2016) to remedy sampling errors due to limited ensemble size. In practice, the ensemble size is usually much smaller than the number of state variables and observations, and is not large enough to span all possible directions in phase space for the update, which

is known as the rank problem (Lorenz 2003). If no (or insufficient) localization is applied, the spurious long-distance sample-estimated correlation is not removed and contaminates the filter performance. On the other hand, if too much localization is applied, it not only discards useful observations, but also introduces noise at small scales due to the artificial tapering of the analysis impact. The noise introduced by localization also causes physical imbalance of the analysis (Kepert 2009; Greybush et al. 2011; Lange and Craig 2014). A simple localization function is the Gaspari and Cohn (1999)

Corresponding author: Fuqing Zhang, fzhang@psu.edu

DOI: 10.1175/MWR-D-17-0336.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy \(www.ametsoc.org/PUBSReuseLicenses\)](https://www.ametsoc.org/PUBSReuseLicenses).

fifth-order polynomial with a specified cutoff radius (hereafter referred to as the GC function). This tapering function can be either applied directly to the background error covariance in model space (Houtekamer and Mitchell 1998), or to the analysis increment limiting the impact of observations to nearby state variables (Houtekamer and Mitchell 2001; Hamill et al. 2001). The result of either approach is similar but the effective localization radius is smaller for the latter (Sakov and Bertino 2011; Nerger et al. 2012). The optimal localization radius finds a balance between the removal of sampling noise and the preservation of useful observation information. Operationally, the tuning of localization radius considers multiple factors including fit to observations, balance, computational cost, model resolution, and observation density. Manual tuning of the localization radius by trial and error for a particular model and ensemble filtering system can be very costly. One may desire a more flexible adaptive scheme to determine the localization a priori. However, to design such a scheme requires knowledge of the complex dependence of localization on the underlying correlation scale determined by the model dynamics, the ensemble size, and the observing network.

Anderson (2007) demonstrated that the optimal localization function could be quite different from a smoothed Gaussian. This motivates the exploration of localization functions that adapt to the flow-dependent error covariance (Anderson 2007, 2012; Bishop and Hodyss 2007, 2009). Anderson and Lei (2013) derived an empirical localization function (ELF) based on the information from an observing system simulation experiment (OSSE). The physical intuition behind these adaptive methods is that localization should only reduce the observation impact when the signal-to-noise ratio is low. In practice, localization distance is often tuned to scale with the overall correlation length. The tuned localization distance is $O(1000)$ km for global modeling and data assimilation systems, but a much shorter localization distance of $O(10)$ km is found more suitable for convective weather systems using high-resolution models and observations (Zhang et al. 2009; Sobash and Stensrud 2013). Since the horizontal correlation length scale increases with height, several studies have found that localization radius should also increase with height (Zhu et al. 2013; Houtekamer et al. 2014; Kleist and Ide 2015). Lei et al. (2015) demonstrated that a narrower horizontal but wider vertical localization scale is preferred for precipitating regions. For atmospheric flows with multiple spatial scales (e.g., high-resolution model domains capturing both synoptic- and convective-scale flows), the data assimilation scheme should handle multiple correlation lengths simultaneously. Zhang et al.

(2009) first proposed a successive covariance localization (SCL) approach that localizes the observation impact with a hierarchy of radii to account for different physical length scales. Multiscale localization methods are further explored and found more advantageous than normal single-scale localization in several other studies (Miyoshi and Kondo 2013; Li et al. 2015; Buehner and Shlyueva 2015).

Despite all these efforts to develop a better localization scheme, a comprehensive theory of localization is still lacking. It is still not clear what makes a smaller localization radius more suitable for convective-scale weather systems, whether it is due to the higher model and observation resolution or the shorter overall correlation lengths. Zhen and Zhang (2014) systematically explored the codependence of localization on the underlying physical scales and the ensemble size for a single observation. As a result, an optimal localization scheme was derived and tested in the Lorenz (1996) model framework. Flowerdew (2015) proposed a similar method but allowed the localization function to have a shape other than Gaussian. These methods are yet to be tested in more complex models that contain multiple physical scales. A theory for optimal localization with a dense observing network may be more complicated than the theory for a single observation (Flowerdew 2015). Previous studies have documented the sensitivity of localization to observation density. From a perfect-model OSSE, Anderson (2007) showed that the optimal localization function is broader for regions with dense observations than for an isolated observation. However, in real-data experiments, Dong et al. (2011) and other studies found that a smaller localization radius is necessary to achieve better analysis accuracy for denser observing networks. Kirchgessner et al. (2014) suggested that optimal localization radius is obtained when the effective observation dimension is about equal to the ensemble size for dense observations. Perri nez et al. (2014) derived an optimal localization radius by high-level heuristic arguments assuming a uniform observing network, and they also suggest using a smaller localization radius for denser observations. These studies suggest there may be a more complicated relation between observing network and localization.

In this study, we seek to systematically explore the relative importance of physical correlation lengths, model resolution, ensemble size, and observing networks to the selection of localization radius, which will provide insights on the development of a better localization scheme for multiscale weather systems. A series of sensitivity experiments are conducted using the two-layer quasigeostrophic (QG) model. The QG model, although simple, can capture the essence of multiscale

atmospheric dynamics. GC functions with fixed cutoff radii are investigated in a serial ensemble Kalman filter (EnKF), and the best localization radius for a given scenario is determined by trial and error. Section 2 describes the QG model and EnKF configuration and the design of sensitivity experiments. Section 3 demonstrates the scale dependency of the best localization radius, and is followed by sensitivity experiment results in section 4 that show how the best localization radius varies in response to changes in model resolution, ensemble size, and observing network. Our findings are summarized in section 5.

2. Experimental design

a. Two-layer quasigeostrophic model

The two-layer QG model described in Smith et al. (2002)¹ is adopted to perform numerical experiments in this study. Previous literature has comprehensively documented its dynamical processes (Larichev and Held 1995; Held and Larichev 1996; Salmon 1998; Smith et al. 2002). Harlim and Majda (2010) used this model to investigate the assimilation of a sparse observing network for the atmosphere. The model is defined on a doubly periodic square domain and simulates the large-scale atmospheric flow with baroclinic instability induced from an imposed vertical wind shear. Background streamfunction is defined as $\Psi_1 = -Uy$ for the top layer and $\Psi_2 = Uy$ for the bottom layer, where U is the mean flow. The prognostic equations for perturbations around this background state can be written as

$$\frac{\partial q_1}{\partial t} + J(\psi_1, q_1) + U \frac{\partial q_1}{\partial x} + (\beta + k_d^2 U) \frac{\partial \psi_1}{\partial x} = 0, \quad (1)$$

$$\frac{\partial q_2}{\partial t} + J(\psi_2, q_2) - U \frac{\partial q_2}{\partial x} + (\beta - k_d^2 U) \frac{\partial \psi_2}{\partial x} + r \nabla^2 \psi_2 = 0, \quad (2)$$

$$q_1 = \nabla^2 \psi_1 + \frac{k_d^2}{2} (\psi_2 - \psi_1), \quad (3)$$

$$q_2 = \nabla^2 \psi_2 - \frac{k_d^2}{2} (\psi_2 - \psi_1), \quad (4)$$

where subscript 1 denotes the top layer and 2 denotes the bottom layer; ψ is the perturbation streamfunction and q is the perturbation QG potential vorticity; $J(\psi, q) = \partial_x \psi \partial_y q - \partial_y \psi \partial_x q$ is the Jacobian term representing the nonlinear advection; β is the meridional gradient of the Coriolis parameter; k_d is the Rossby

deformation wavenumber; and r is the strength of linear Ekman drag that removes large-scale energy buildup from the bottom layer. Two characteristic wavenumbers corresponding to the Rossby deformation scale and the Rhines scale are defined as

$$k_d = (L/2\pi)f/\sqrt{g'H} \quad (5)$$

and

$$k_\beta = (L/2\pi)\sqrt{\beta/U_0}, \quad (6)$$

respectively, where f is the Coriolis parameter, g' is the reduced gravity, H is the vertical scale height, U_0 is the horizontal velocity scale, and $L/2\pi$ is a scaling factor. The kinetic energy spectrum of the QG model features an energy injection caused by baroclinic instability near the deformation scale k_d^{-1} . For scales larger than the deformation scale ($k < k_d$), there is an inverse cascade of kinetic energy with a $-5/3$ power law. The cascade halts at a scale $k_{\beta,r}^{-1}$ that is determined by both the β effect and the bottom drag r . At scales smaller than the deformation scale ($k > k_d$), the enstrophy cascades forward and dissipates at the smallest scales, resulting in a -3 power law for kinetic energy. The model mimics this dissipation with an exponential cutoff filter that removes energy buildup at the smallest scales [see appendix B of Smith et al. (2002)].

A baseline configuration used in this study sets the model parameters as $k_d = 20$, $k_\beta = 4$, $U = 0.2U_0$, and $r = 0.5$. Let $L = n dx$ be the domain size, where n is the number of grid points in both zonal and meridional directions and dx is the grid spacing. The model resolution is set to $n = 128$, which resolves $k_{\max} = 63$ modes in each direction. The exponential filter cutoff wavenumber is $k = 40$. These parameters are chosen similarly to the atmospheric case from Harlim and Majda (2010) except that the deformation wavenumber here is larger to produce a wider range of scales with baroclinic instability. The typical deformation length scale for the large-scale atmospheric flow is 1000 km. With this scaling, the model grid spacing dx corresponds to ~ 80 km in real atmospheric models. The average eddy turnover time is ~ 0.15 non-dimensional time units, which corresponds to ~ 2 days.

b. Ensemble filter

The data assimilation method used in this study is a serial version of the ensemble square root filter (Whitaker and Hamill 2002), which is described as follows. Let \mathbf{x} be the state variable vector, and let \mathbf{y}^o be the observation vector. An ensemble of N members is introduced to estimate the flow-dependent background

¹The model code is available at <http://www.cims.nyu.edu/~shafer/tools/>.

error covariance. For each observation indexed with subscript j , the following equations are applied to update the ensemble serially to reach the final analysis:

$$(\bar{\mathbf{x}})_{\text{update}} = \bar{\mathbf{x}} + \rho_j \circ \mathbf{K}_j (\mathbf{y}_j^o - H_j \bar{\mathbf{x}}), \quad (7)$$

$$(\mathbf{x}'_i)_{\text{update}} = \mathbf{x}'_i + \gamma_j \rho_j \circ \mathbf{K}_j (\mathbf{0} - H_j \mathbf{x}'_i), \text{ for } i = 1, 2, \dots, N, \quad (8)$$

where $\bar{\mathbf{x}}$ is the ensemble mean, \mathbf{x}'_i is the ensemble perturbation for the i th member, ρ_j is a localization function, \mathbf{K}_j is the Kalman gain defined as

$$\mathbf{K}_j = \frac{\text{cov}(H_j \mathbf{x}, \mathbf{x})}{\text{var}(\mathbf{y}_j^o) + \text{var}(H_j \mathbf{x})}, \quad (9)$$

and γ_j is a square root modification term defined as

$$\gamma_j = \left[1 + \sqrt{\frac{\text{var}(\mathbf{y}_j^o)}{\text{var}(\mathbf{y}_j^o) + \text{var}(H_j \mathbf{x})}} \right]^{-1}. \quad (10)$$

Here $\text{cov}(H_j \mathbf{x}, \mathbf{x})$ is the background error covariance between observations and state variables, $\text{var}(\mathbf{y}_j^o)$ is the observation error variance, and $\text{var}(H_j \mathbf{x})$ is the associated background error variance. The linearized observation operator \mathbf{H}_j is not used in this study. Instead, the $H_j \bar{\mathbf{x}}$ and $H_j \mathbf{x}'_i$ terms are approximated by $\overline{h_j(\bar{\mathbf{x}})}$ and $h_j(\mathbf{x}'_i) - \overline{h_j(\bar{\mathbf{x}})}$, respectively, where h_j is the nonlinear observation operator and overbars denote ensemble averages.

GC functions with fixed localization radii are adopted as the localization function ρ . Localization radius [radius of influence (ROI)] is defined as the physical distance at which the analysis increments are tapered to zero. The adaptive covariance relaxation method (Ying and Zhang 2015) is also applied to maintain ensemble spread and prevent catastrophic filter divergence.

c. Observing network and assimilation experiments

A control (CNTL) experiment is first performed using the baseline configuration in section 2a. The truth (nature run) is generated by initializing the QG model with white noise and running the model for 50 time units until it reaches a quasi-steady state. A 15-time-units model run (~ 200 days) during the quasi-steady state is taken as the truth. Cycling data assimilation is performed using an ensemble of $N = 64$ members and assimilating synthetic observations simulated from the truth every $\Delta t = 0.05$ time units, which yields 300 cycles in total. Under a perfect-model assumption, the truth model is used as the forecast model during cycling data

assimilation. Using the trial-and-error method, the best-performing ROI that yields the lowest analysis error variance is determined. Analysis errors are also decomposed into spectral components to evaluate the scale dependency in filter performance. Sensitivity experiments are designed to test how the best-performing ROI changes in response to changes in model resolution, ensemble size, and observing network. Table 1 summarizes the model and filter parameters used in each experiment. ROI is defined as a physical distance and expressed in terms of the number of grid points in CNTL (i.e., ROI = 8 means a cutoff distance of $8dx$). Cases with ROI = 8, 16, 32, 64, as well as no localization (ROI = ∞), are tested for each experiment. Some additional values of ROI are tested in the neighborhood of the minimum analysis error to more accurately estimate the best-performing ROI.

Instead of assimilating the state variable ψ directly, a more challenging scenario is considered where temperature θ or horizontal winds (u and v) are observed from the top model layer. Let k_x and k_y be the zonal and meridional wavenumbers, respectively. The nonlinear observation operators that convert ψ to u , v , and θ can be described in spectral space:

$$\hat{u} = -ik_y \hat{\psi}, \quad (11)$$

$$\hat{v} = ik_x \hat{\psi}, \quad (12)$$

$$\hat{\theta} = -\sqrt{k_x^2 + k_y^2} \hat{\psi}, \quad (13)$$

where hats denote the two-dimensional Fourier transform of a variable. Figure 1 shows snapshots of the top-layer θ and u in comparison to the corresponding ψ from the truth simulation. The θ field is overall in phase with ψ but has more small-scale details. On the other hand, u is not in phase with ψ and their peaks are not collocated. Synthetic observations are generated by first converting ψ from the truth simulation to observations according to (11)–(13), and then adding simulated observation errors randomly drawn from a zero-mean normal distribution with an error variance of σ^2 . The observation errors are considered uncorrelated both in space and in time. For CNTL, a uniform observing network of top-layer θ with $\sigma = 3$ is used. Although the bottom layer is not observed, the top-layer observations provide information for both layers because the flow simulated in CNTL is mostly barotropic, especially for the large scales. However, not all model grid points on the top layer are observed. Therefore, the horizontal propagation of information is important for good filter performance. Let Δx be the spacing between observations in both horizontal directions, the number of observations in each direction is $n_o = n dx/\Delta x$. For CNTL,

TABLE 1. Description of model and filter parameters used in control and sensitivity experiments. In each experiment, a range of ROIs is tested and the ROI that minimizes analysis RMSE is determined.

Expt	Model and filter parameters (changes relative to CNTL)
CNTL	$k_{\max} = 63, k_d = 20, k_\beta = 4, U = 0.2U_0, r = 0.5$ (spectral peak at large scale); Ensemble size $N = 64$; Assimilate θ observations every $3dx$ with $\sigma = 3$ every $t = 0.05$
Scale dependence	
M_Scale	$U = 0.5U_0, r = 6$ (spectral peak at intermediate scale)
S_Scale	$k_d = 35, U = 0.6U_0, r = 20$ (spectral peak at small scale)
Sensitivity to model resolution ^a	
LowRes	$k_{\max} = 31$ for both truth and forecast model
LowRes_Model	$k_{\max} = 31$ forecast model but truth is from CNTL
LowRes2	$k_{\max} = 21, k_d = 14$ for both truth and forecast model
LowRes2_Model	$k_{\max} = 21, k_d = 14$ forecast model but truth is from CNTL
Sensitivity to sampling error ^b	
N16, N32, N256, N1024	Ensemble size N changed to 16, 32, 256, and 1024
Sensitivity to observing network ^c	
ObsSparse, ObsDense	Horizontal observation interval changed to $9dx$ and $1dx$, respectively
ObsErrorX3, ObsError/3	Observation error standard deviation changed to $\sigma = 9$ and $\sigma = 1$, respectively ^d

^a The ROI is in physical distance units (number of grid points from CNTL model); for example, ROI = 8 means 8 grid points in CNTL, which corresponds to 4 grid points in LowRes, and only 2.67 grid points in LowRes2.

^b The ensemble size experiments are also repeated for the case where u and v are assimilated instead of θ .

^c The observing network experiments are first performed with regular (horizontally uniform) networks, then repeated with irregular networks (randomly located) that match the density and accuracy of the regular ones.

^d The observation error level is the same for ObsErrorX3 and ObsSparse; they both have 3 times larger RMSE than CNTL.

the observation density is set to $\Delta x = 3 dx$. For a uniform observing network with independent random instrument errors, the spectral variance of its observation error can be exactly calculated. The observation error variance

associated with wavenumber k , $R(k)$, is inversely proportional to the number of observations n_o^2 :

$$R(k) = 2\pi k \sigma^2 / n_o^2. \tag{14}$$

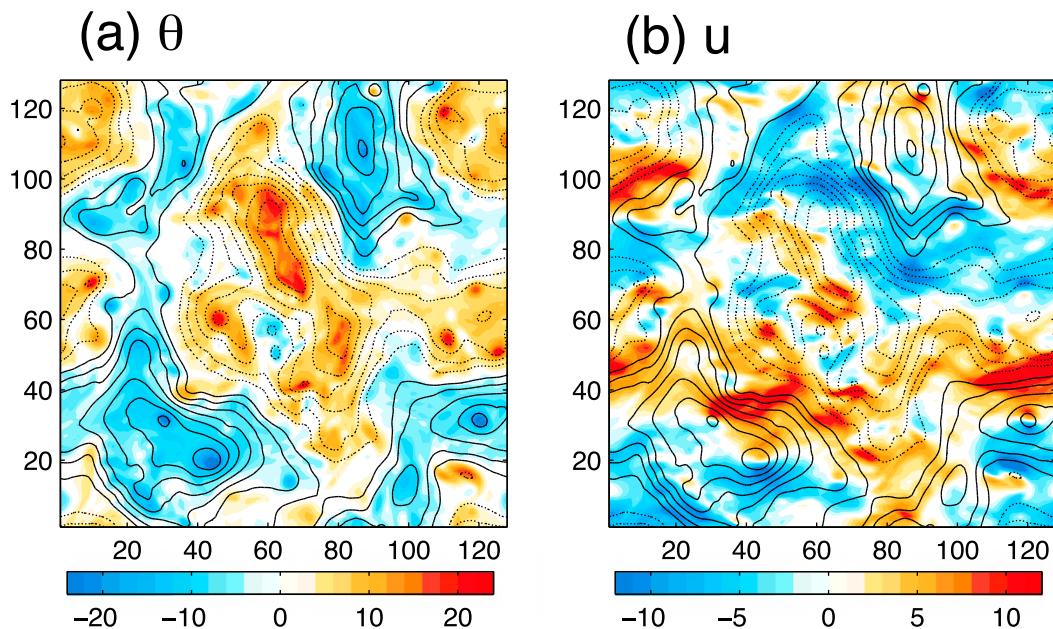


FIG. 1. Snapshots of (a) θ and (b) u in color shadings and state variable ψ in black contours (solid lines for positive and dotted lines for negative values). The ψ , θ , and u variables are nondimensionalized. The scaling of the QG model is described in section 2a.

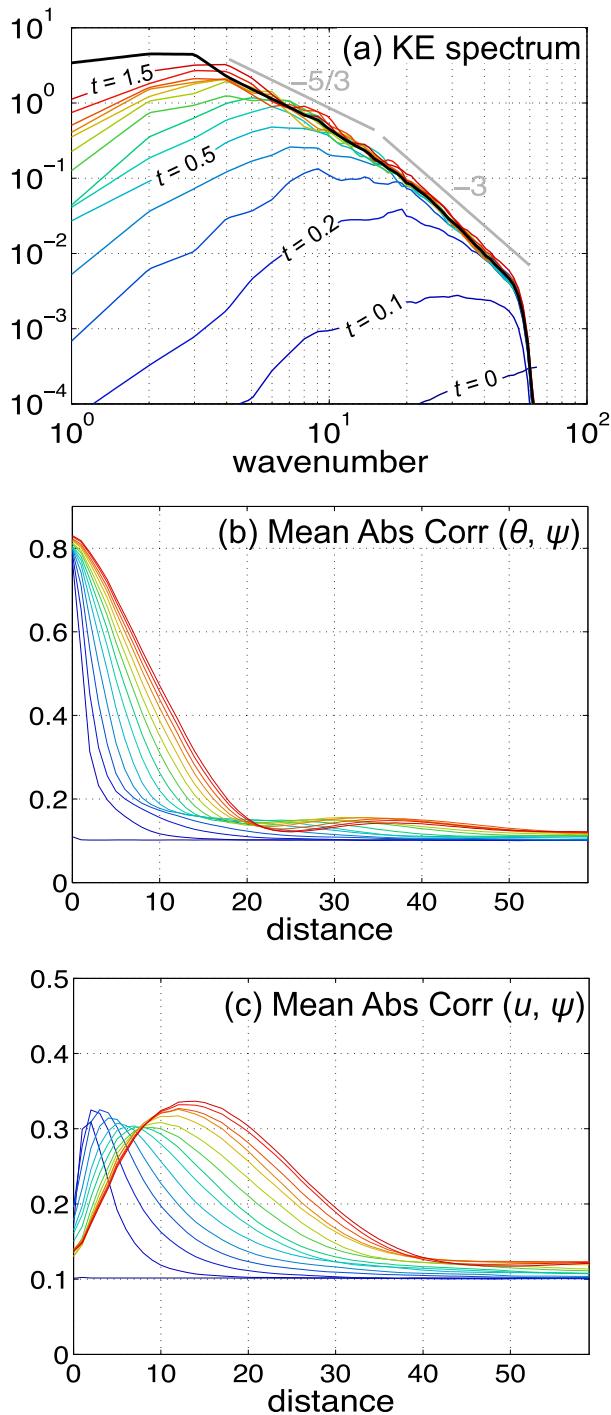


FIG. 2. (a) Spectra of error kinetic energy from ensemble mean (color coded with forecast time) and the time-averaged reference kinetic energy from the truth (black) during an ensemble spinup period of 1.5 time units. (b) Spatially averaged mean absolute correlation (MAC) between θ and ψ plotted as a function of horizontal distance (number of grid points) also color coded with forecast time as in (a). (c) As in (b), but showing correlation between u and ψ .

With fixed instrument error σ , a denser observing network yields lower observation error spectral variance thanks to the larger number of observations. Sensitivity to changes in observing network, including its density, accuracy, and spatial homogeneity, will be tested in section 4c.

To diagnose filter performance, the analysis error (difference between posterior ensemble mean and the truth) variance is averaged over the cycling data assimilation period. When θ is assimilated, the state variables will be converted to θ to calculate the analysis error variance. When u and v are assimilated, the analysis error variance for u and v will be calculated and then combined as an error kinetic energy, which is defined as $(u^2 + v^2)/2$. The errors are further decomposed into spectral components to facilitate the comparison across scales. The analysis error variance associated with the wavenumber k component is defined as

$$E(k) = \sum_{k_x^2 + k_y^2 = k^2} [\hat{\epsilon}(k_x, k_y)]^2, \quad (15)$$

where $\hat{\epsilon}$ is the two-dimensional Fourier transform of the analysis error and $k = \sqrt{k_x^2 + k_y^2}$ is the global wavenumber. Summing $E(k)$ over a range of wavenumbers will yield the mean square error (MSE), or root-mean-square error (RMSE) if the square root is taken, associated with this scale range. In this study, the spectrum is divided into three bands: large (L; $k = 1-4$), intermediate (M; $k = 5-20$), and small (S; $k = 21-63$) scales.

To generate the prior ensemble, random white noise is added to the truth initial condition for each member and an ensemble forecast is run for a spinup period of 1.5 time units until the ensemble spread reaches climatological level. Figure 2a shows the kinetic energy spectrum from the truth (black line) and time evolution of error kinetic energy from the ensemble mean (colored lines) during this spinup period. Error saturates when its kinetic energy reaches the level of the reference kinetic energy. The M scale follows a $-5/3$ power law associated with the inverse energy cascade, where the small-scale errors saturate faster than the large-scale errors and an overall upscale error growth is present. As errors grow upscale, they have more large-scale components, which results in an increasing overall correlation length scale. An estimation of this correlation length is shown in Fig. 2b for θ observations and Fig. 2c for u observations. The mean absolute correlation (MAC) estimated from the ensemble is averaged over the observing network and plotted as a function of horizontal distance. Both MACs for θ and u are broadening over time. The MAC for u has a nonlocal peak because u and ψ maxima are not

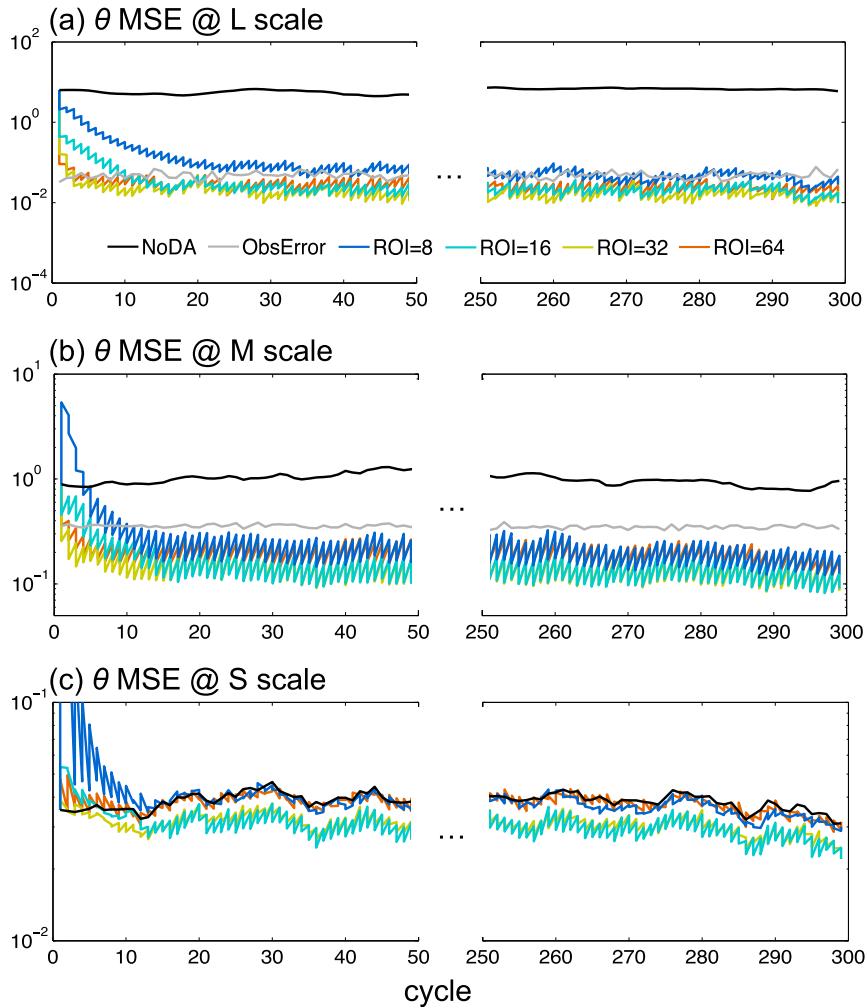


FIG. 3. Time series of θ MSE from CNTL filtered for the (a) L, (b) M, and (c) S scales. Results are shown for errors from free ensemble (NoDA; black), observation (gray), and from analysis mean using ROI = 8, 16, 32, and 64 (colored).

collocated, and the correlation between u and ψ is overall much lower than the correlation between θ and ψ . The ensemble after spinup is used as the prior for the ensuing data assimilation cycles.

3. Scale-dependent best localization radius

Figure 3 shows the time evolution of θ MSE at L, M, and S scales from CNTL using ROI = 8, 16, 32, and 64. The analysis errors undergo an initial adjustment (filter spinup) period of approximately 20 cycles before the filter performance reaches steady state. The cases using smaller ROIs experience longer filter spinup periods than the cases using larger ROIs. Earlier in the spinup period, the larger ROI = 32 and 64 cases perform better than smaller ROIs. However, for the steady state, the ROI = 16 case appears to improve, while the ROI = 64

case is clearly suboptimal. At L scale, ROI = 32 is more favorable. ROI = 16 occasionally performs as well as ROI = 32, but on average is worse. On the other hand, at M scale, the ROI = 16 and 32 cases are competitive at steady state. For the S scale, the ROI = 16 case appears to have better performance. These results indicate a scale dependency for favorable ROIs. The following calculation of time-averaged analysis errors will exclude the first 50 cycles to make sure that results reflect steady-state filter behavior, and the long record (250 cycles) ensures robust statistics not influenced by transient behavior. Figure 4a shows the time-averaged θ spectral error variance $E(k)$ from CNTL. Along with the analysis error variance, the observation error variance $R(k)$ and error variance from a free ensemble without data assimilation (NoDA) are also plotted for reference. The observation error variance follows a +1 power law

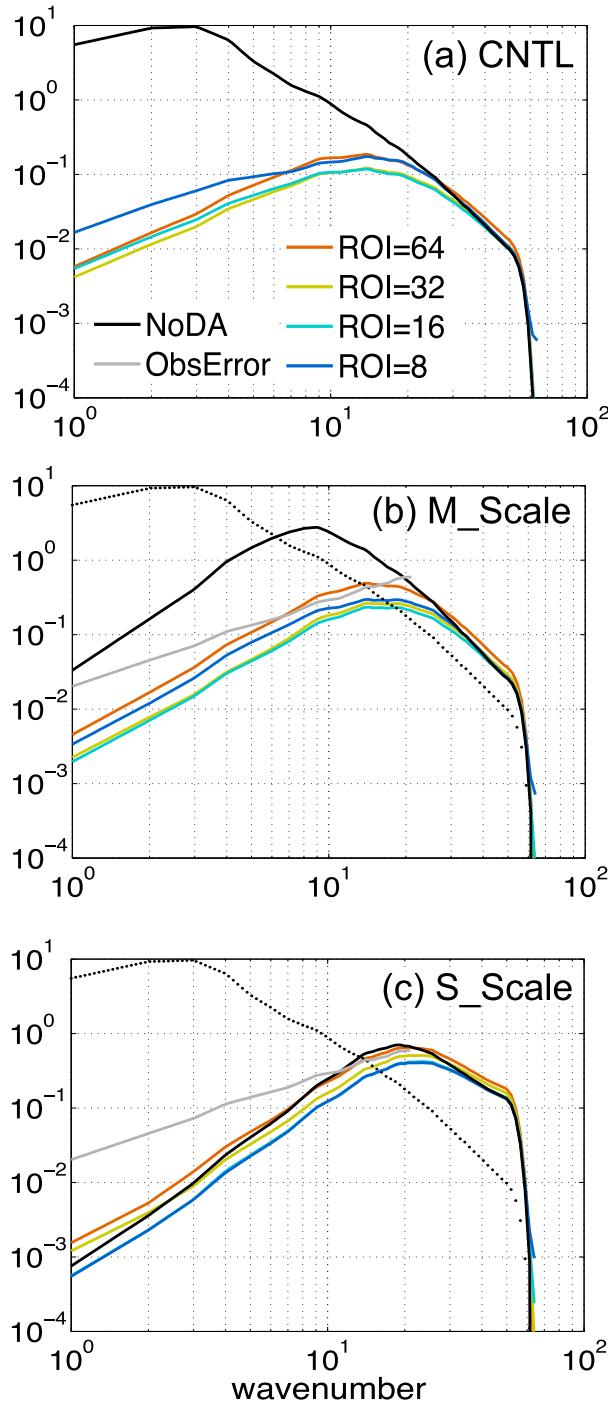


FIG. 4. Temperature (θ) error spectra from (a) CNTL, (b) M_Scale, and (c) S_Scale. Results are shown for errors from a free ensemble (NoDA; black), observation (gray), and from analysis mean using ROI = 8, 16, 32, and 64 (colored). The CNTL NoDA error is shown as dotted lines in (b) and (c) for reference.

associated with white noise, and the NoDA error is fully saturated at all scales and thus resembles the reference spectrum from the truth. Note that the θ

spectrum follows a similar power law to the kinetic energy. From L to S scales, the best ROI that minimizes analysis error variance at that scale appears to shift toward smaller values. The S scale is not observed because of coarser observation resolution than the model grid, thus the analysis errors remain mostly saturated at this scale.

Two experiments, M_Scale and S_Scale, similar to CNTL but with changed reference kinetic energy spectra are conducted to further demonstrate the scale dependency of the best ROI. Figures 4b and 4c show their resulting error spectra. The CNTL reference spectrum has a peak at the L scale ($k = 3$), and this spectral peak is shifted toward M and S scales by changing model parameters r , k_d , and U . As bottom drag r increases, more energy is removed from the large-scale end of the spectrum, and the halting scale $k_{\beta,r}^{-1}$ becomes smaller. The zonal wind shear ($\pm U$) is increased to inject more baroclinic instability to ensure that eddies have similar amplitudes as the CNTL. For S_Scale, the deformation scale (k_d^{-1}) is also shifted to a smaller scale ($k_d = 35$) to allow baroclinic instability to develop. The resulting M_Scale and S_Scale experiments simulate eddies with different sizes from CNTL. M_Scale is dominated by eddies at M scale and S_Scale is dominated by eddies at S scale, which is also reflected in their averaged eddy turnover time (~ 0.1 for M_Scale and ~ 0.07 for S_Scale).

Results confirm that the best-performing ROI decreases as the dominant scale of the system becomes smaller. For a given system, if there is one dominant scale, our results suggest that this scale will determine the best localization radius. A fixed ROI works well when there is only one dominant scale. Figure 5a plots the analysis errors for all scales with respect to ROIs. The overall best ROI that minimizes domain-averaged analysis RMSE is near 24 for CNTL, and it shifts to 16 for M_Scale and 12 for S_Scale. When a wider range of scales is present, a fixed ROI may become insufficient to minimize errors at all scales, and a different ROI should be specified for each scale, as suggested by previous studies (Zhang et al. 2009; Miyoshi and Kondo 2013; Li et al. 2015; Buehner and Shlyueva 2015). Figure 5b illustrates this by plotting the CNTL analysis RMSE filtered for the L, M, and S scales with respect to ROIs. The L-scale component favors ROI = 32, the M scale favors ROI = 24, and the S scale favors ROI = 16.

The relationship between correlation length scale and localization distance is not necessarily one-to-one (Anderson and Lei 2013; Anderson 2016). In this study, we hypothesize that the best localization radius scales with the overall correlation length. One can

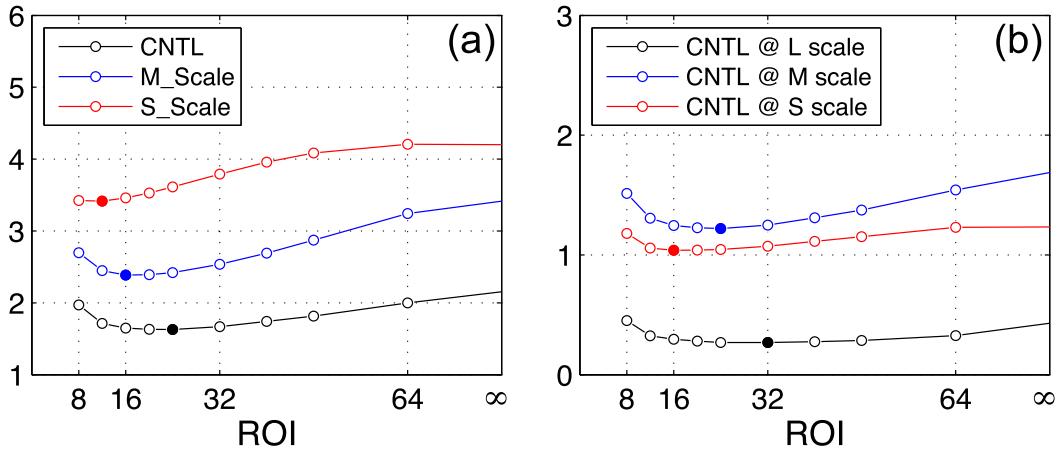


FIG. 5. (a) Temperature (θ) RMSE plotted with respect to ROI from CNTL (black), M_Scale (blue), and S_Scale (red). (b) As in (a), but showing θ RMSE from CNTL filtered for the L (black), M (blue), and S (red) scales. The filled circles indicate the ROI with lowest analysis error.

consider a multiscale data assimilation problem as successively constraining from the large to small scales that have decreasing correlation lengths. Figure 6 plots the MAC functions filtered for L and S scales from CNTL. The shape of correlation functions on average is very different for L and S scales. The correlation function is broader at L scale than at S scale. The magnitude of overall correlation at zero distance is lower at L scale than at S scale. When a larger ROI is used, the averaged correlation remains lower than when a smaller ROI is used, which indicates that larger ROIs constrain the analysis ensemble more than smaller ROIs. The analysis ensemble also tends to be more dispersive for cases using smaller ROIs, which can be inferred from the amount of inflation determined by the adaptive covariance relaxation algorithm. The relaxation coefficient averaged over time for the ROI = 8, 16, 24, 32, and 64 cases are $\alpha = -0.23, 0.17, 0.27, 0.35,$ and $0.57,$ respectively. A larger α means more inflation is applied to the analysis ensemble.

4. Sensitivity experiments

a. Model resolution

Numerical models usually cannot resolve all the scales of an atmospheric flow. In this section, we investigate cases using models that have lower resolutions than the truth model that generates the synthetic observations. For the QG model used in CNTL, the S scale features a forward enstrophy cascade, and the M scale features an inverse energy cascade. A model with resolution reduced to $k_{\max} = 31$ is first investigated. Comparing it to the CNTL model, it has most of its S scale truncated and cannot accurately resolve the forward enstrophy

cascade. Model parameters are selected to ensure that the low-resolution model has a large-scale energy spectrum matching with the high-resolution model in CNTL. Without the forward enstrophy cascade, a stronger enstrophy filter (exponential cutoff at $k = 20$) is applied, and the baroclinic instability is slightly increased. Such model tuning is typically done in real atmospheric models, too, although more sophisticated methods such as parameterization are used to account for processes that these models cannot resolve. The low-resolution model is first tested in a perfect-model scenario in the LowRes experiment, where the observations

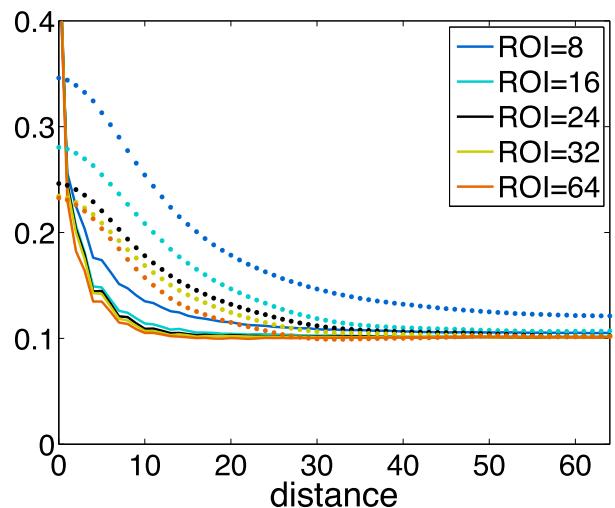


FIG. 6. MAC between θ and ψ filtered for L (dotted) and S (solid) scales estimated by the prior CNTL ensemble using ROI = 8, 16, 24, 32, and 64. The MACs are averaged over space and time and plotted as a function of horizontal distance (number of grid points).

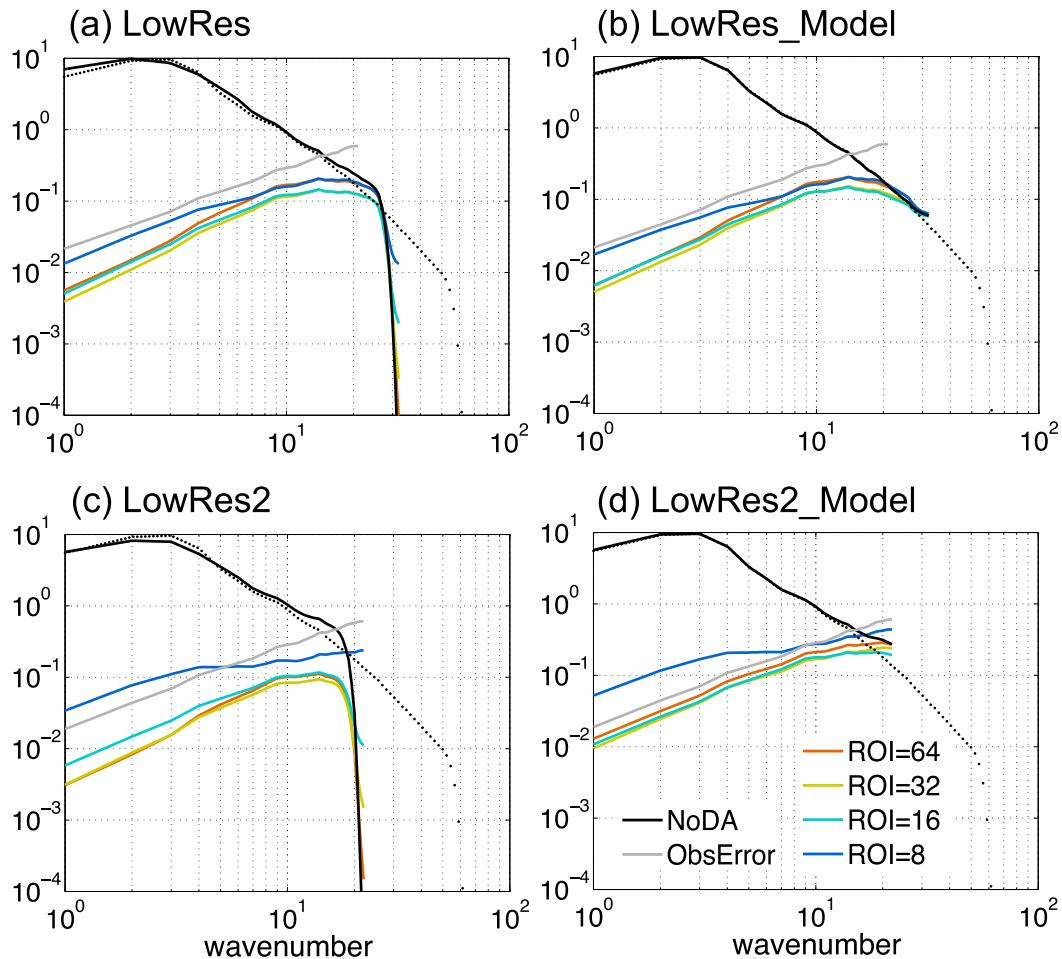


FIG. 7. Temperature (θ) error spectrum of NoDA (black), observation (gray), and analysis ensemble mean using ROI = 8, 16, 32, and 64 (colored) from (a) LowRes, (b) LowRes_Model, (c) LowRes2, and (d) LowRes2_Model experiments. The CNTL NoDA error spectrum is shown as dotted lines for reference. The ROI is in physical distance units (number of grid points from CNTL model); for example, ROI = 8 means 8 grid points in CNTL, which corresponds to 4 grid points in LowRes, and 2.67 grid points in LowRes2).

are generated from a truth run using the same low-resolution model. Figure 7a shows the analysis error spectra from the cases using different ROIs. Other than the slightly elevated smallest-scale energy in LowRes, the results are very similar to the CNTL experiment (Fig. 4a). In LowRes_Model experiment, the low-resolution model is only used in the forecast step of the data assimilation and observations sampled from the high-resolution truth in CNTL are assimilated. The resulting analysis error spectra are shown in Fig. 7b. Figure 8 plots the analysis RMSE filtered for L and M scales with respect to ROIs. Comparing LowRes (red), and LowRes_Model (blue) to CNTL (black), the reduced model resolution does not appear to influence the best ROI at the well-resolved L scale, while the best ROI at M scale slightly decreases because of the representation errors. The time-averaged adaptive

relaxation coefficients are $\alpha = 0.25$ for LowRes (similar to CNTL where $\alpha = 0.27$) and $\alpha = 0.32$ for LowRes_Model, indicating more inflation is applied to the ensemble spread when representation errors are present.

Another model with even lower resolution $k_{\max} = 21$ is tested in LowRes2 (Fig. 7c) and LowRes2_Model (Fig. 7d). In this case, the lack of resolution starts to influence the representation of baroclinic instability near the deformation wavenumber. To fully resolve the baroclinic instability, the deformation wavenumber is changed to $k_d = 14$. Although some tuning is made to match the model climatology with CNTL, the error growth rate is slightly higher at larger scales for this low-resolution model. This is a common situation for real atmospheric models where key dynamical processes cannot be well represented at the small scale and therefore model forcing is biased. In the perfect-model

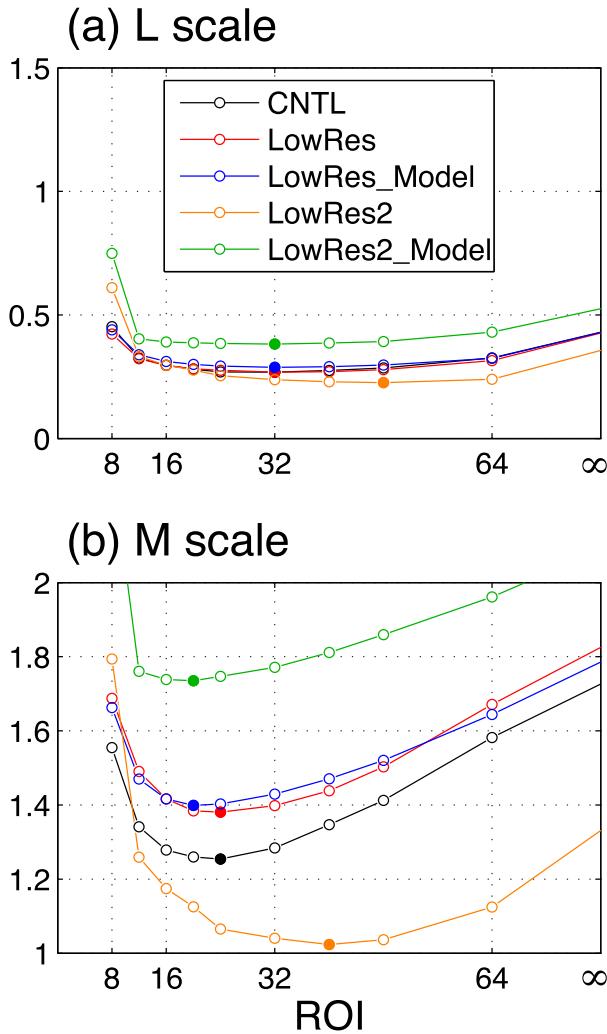


FIG. 8. Temperature (θ) RMSE filtered for (a) L and (b) M scales plotted with respect to ROI. Results are compared for CNTL, LowRes, LowRes_Model, LowRes2, and LowRes2_Model.

scenario, LowRes2 (Fig. 7c) shows that the best ROI becomes larger because of the increased error growth rate at larger scales. Figure 8 also shows this shift in best ROI from CNTL to LowRes2 for both L and M scales. When observations from CNTL are assimilated instead, LowRes2_Model (Fig. 7d and green line in Fig. 8) shows that although the steady-state analysis RMSE is larger because of the presence of larger model errors, the best ROI stays relatively unchanged from CNTL for the L scale. The averaged adaptive relaxation coefficient is larger for LowRes2_Model ($\alpha = 0.58$) than for LowRes2 ($\alpha = 0.23$) because of the presence of a large representation error.

Previous studies (e.g., Aksoy et al. 2012) have demonstrated the negative impact from representation errors when assimilating observations from a high-resolution nature run while the model is at a coarser

resolution. Various methods are proposed to account for these representation errors in data assimilation algorithms to reduce their negative impact (Janjić and Cohn 2006; Bocquet et al. 2011; Hodyss and Nichols 2015; van Leeuwen 2015). Our results indicate that localization is not sensitive to model resolution as long as a dynamical process is well resolved and/or model representation errors are well accounted for in the data assimilation scheme. Note that we define the localization ROI in terms of the physical length, which does not change as model resolution reduces. If the ROI is defined as a number of grid points, it will change as model resolution changes (i.e., fewer grid points for lower resolution). Also note that our conclusions are drawn from QG model experiment results. Whether this can be generalized to other systems is yet to be confirmed with further experimentation.

b. Sampling error due to limited-size ensemble

Given the same observing network and a perfect model, a larger ensemble size provides more accurate sample-estimated error covariances and therefore improves the accuracy of the analysis. The negative impact of sampling error on filter performance can be demonstrated by decreasing the ensemble size N . To test the sensitivity of localization to N , experiments are conducted with $N = 16, 32, 256, 1024$, and compared to $N = 64$ as in CNTL. Figures 9a–c plot the resulting L-, M-, and S-scale analysis RMSEs as a function of ROI. The trend that the ROI that minimizes analysis RMSE at a certain scale becomes larger as N increases is seen for all three bands. These results are consistent with previous studies (Houtekamer and Mitchell 2001; Lorenc 2003; Zhang et al. 2006; Anderson 2007, 2012). When a larger N is used, there is also a wider range of ROIs around the best ROI that produce similar filter performance, indicating less sensitivity to localization scale in this case (Lei and Whitaker 2017). Not surprisingly, the minimum analysis RMSE is achieved using the largest $N = 1024$. This is consistent with the recent study of Kondo and Miyoshi (2016) who demonstrated with their 10240-member ensemble that with large enough ensemble size one can achieve the best filter performance without localization. Figure 10a shows the steady-state MACs estimated from analysis ensembles of different sizes. Correlation decreases as horizontal distance increases. Beyond the decorrelation length scale, the true correlation should be zero on average. Because of sampling error, the limited-sized ensembles have MACs that asymptote to nonzero correlation values, and the asymptotic value increases as N decreases. Since data assimilation constrains the observed scales and reduces error, the analysis error has less large-scale contribution and a shorter overall correlation

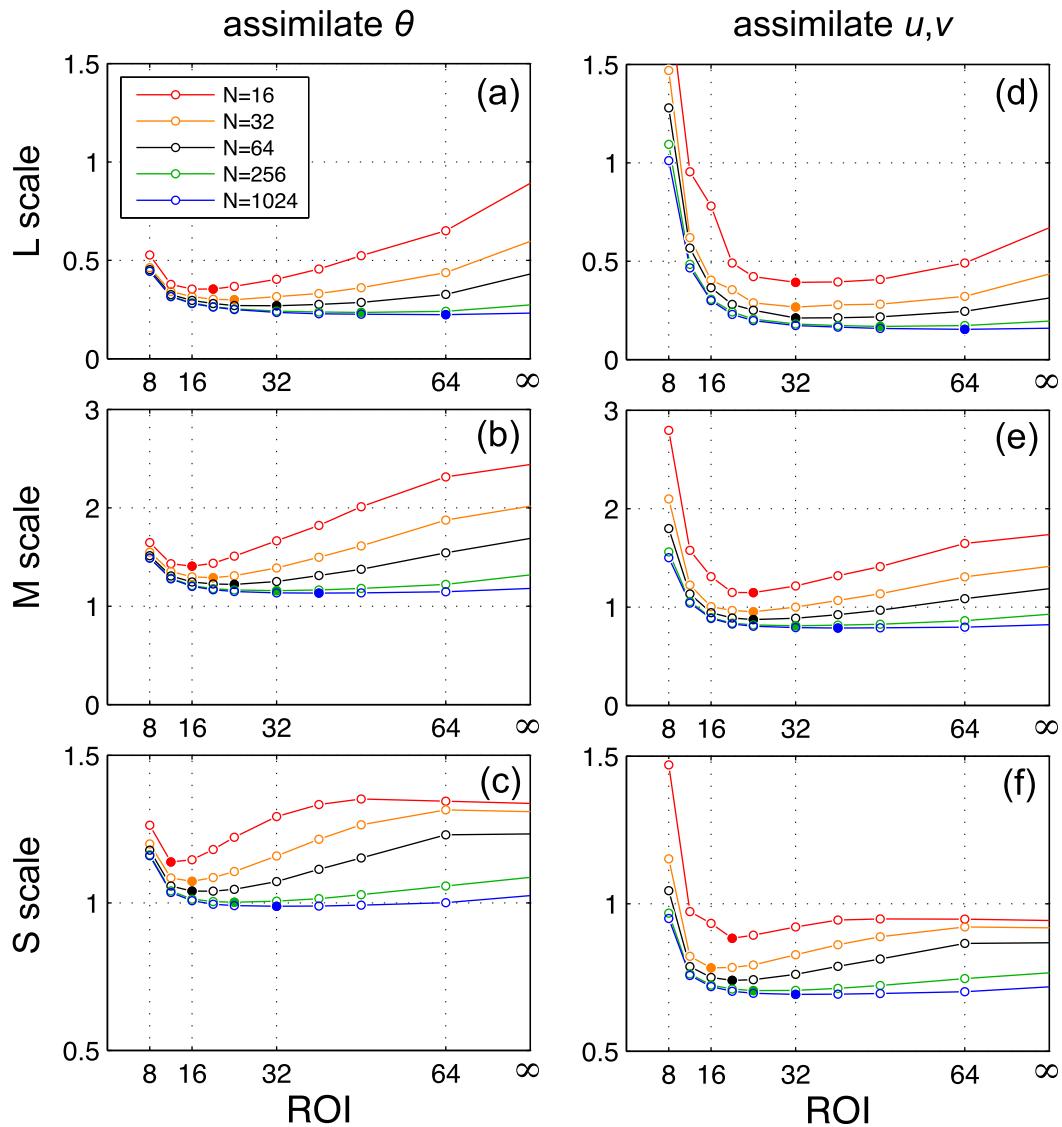


FIG. 9. Temperature (θ) RMSE filtered for (a) L, (b) M, and (c) S scales plotted with respect to ROI compared for the $N = 16, 32, 64$ (CNTL), 256, and 1024 cases. Filled circles indicate the minimum error corresponding to the best ROI. (d)–(f) As in (a)–(c), but showing root-mean-error kinetic energy from the cases assimilating u and v .

length compared to errors from a free ensemble without data assimilation (Yoon et al. 2010). Ensemble forecasts from the analyses will have correlation lengths that increase as errors grow upscale and eventually saturate again as shown in Fig. 10c.

Assimilating u and v observations is more challenging than θ observations since their correlation functions are nonlocal. The sensitivity experiments mentioned above are repeated using u and v as observations, and the resulting analysis RMSEs are plotted with respect to ROIs in Figs. 9d–f. Figure 10b shows the steady-state MACs for each experiment, and the time evolution of a MAC during ensemble forecasts is shown in Fig. 10d. Similar

to the θ observation cases, the asymptotic correlation values are unchanged when u and v observations are assimilated. The overall correlation is much lower compared to θ observations. As forecast errors grow and larger scales contribute more, the nonlocal correlation peak shifts toward longer distances. For larger ROIs, the trend that the best ROI increases as N increases is unchanged. However, an extra penalty is present for smaller ROIs that are too short and exclude the nonlocal correlation peaks that are important sources of information for this case. As a result, as N decreases, the best ROI does not decrease beyond the distance at which these nonlocal correlations peak.

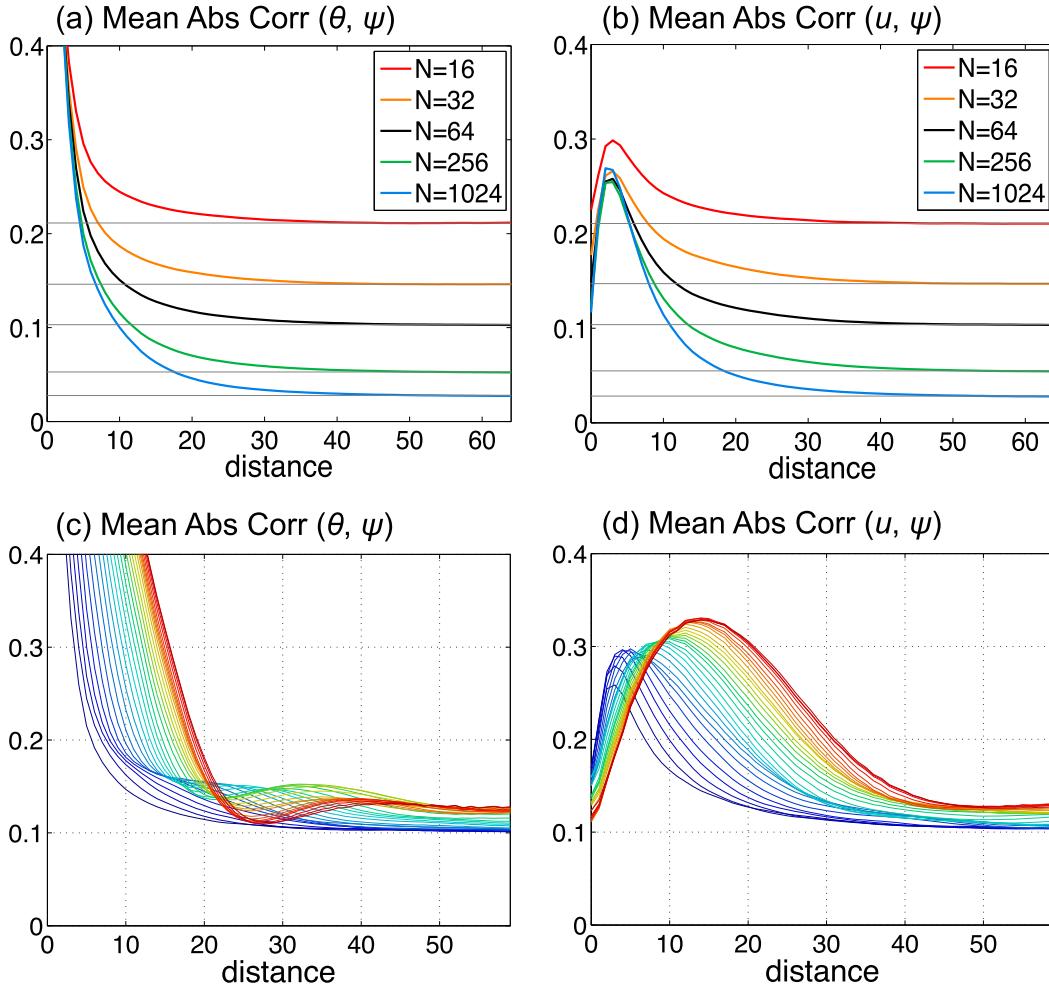


FIG. 10. (a) MAC between θ and ψ from the posterior ensemble with changing ensemble size $N = 16, 32, 64, 256,$ and 1024 using the best ROI. (b) As in (a), but showing MAC between u and ψ from the cases assimilating u and v . (c) Time evolution of MAC between θ and ψ from ensemble forecasts initialized with posterior from the $N = 64$ case, the lines are color coded with forecast time $t = 0$ to $t = 1.5$ every 0.05 time units. (d) As in (c), but showing MAC between u and ψ from the cases assimilating u and v .

c. Accuracy and density of the observing network

The sensitivity of localization to changes in the observing network is tested in this section. In CNTL, we considered a uniform observation grid coarser than the model grid ($\Delta x = 3 dx$). When observation density increases (n_o increases), the observation error spectral variance $R(k)$ will decrease according to (14). In the ObsSparse experiment, the horizontal observation interval is set to $\Delta x = 9 dx$, which yields an $R(k)$ 9 times larger than that from CNTL. Changes in network density have two different effects on filter performance. First, the observation error variance $R(k)$ is changed (“accuracy effect”). Second, the number of independent pieces of information available within the localization scale to constrain each state variable is also changed

(“number effect”). To separate these two effects, another experiment, ObsErrorX3, is conducted in which the observing network is of the same density as CNTL but its observation error standard deviation σ is increased by a factor of 3 resulting in the same $R(k)$ as ObsSparse (i.e., only the accuracy effect is present in ObsErrorX3). In both ObsSparse and ObsErrorX3, the information provided by the observation is less accurate. The other effect of increasing observation accuracy is tested in ObsDense where the horizontal observation interval is set to $\Delta x = 1 dx$, and ObsError/3 where observation density is unchanged but observation error standard deviation σ is decreased by a factor of 3. Figure 11 shows error spectra from these experiments, and Figs. 12a–c plot their analysis RMSEs filtered for L, M, and S scales with respect to ROIs. For ObsSparse

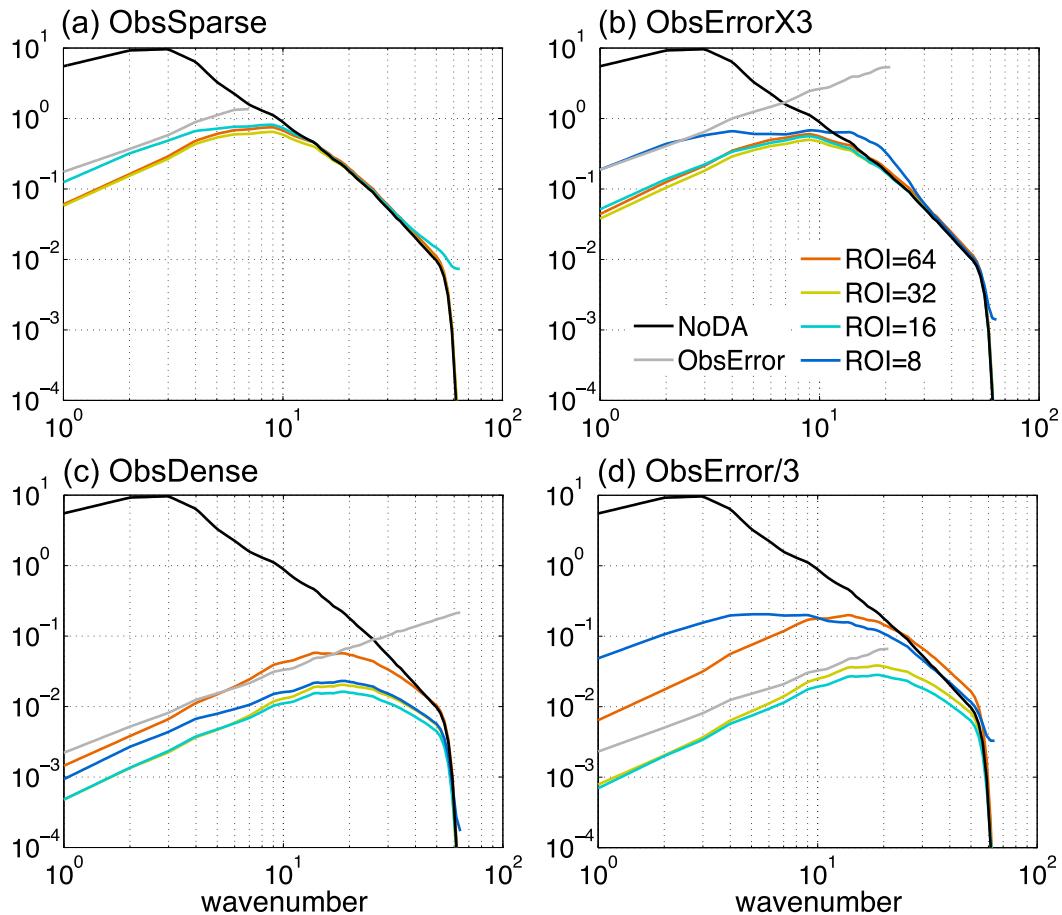


FIG. 11. Error spectra of free ensemble mean (NoDA; black), observation (gray), and analysis mean using ROI = 8, 16, 32, and 64 (colored) from (a) ObsSparse, (b) ObsErrorX3, (c) ObsDense, and (d) ObsError/3 cases.

(Fig. 11a), the ROI = 8 case is excluded because this ROI is smaller than the observation interval that results in only one observation available to constrain each state variable.

With both accuracy and number effects, a sparser observing network favors a larger ROI to achieve the best filter performance, and a denser network favors a smaller ROI (cf. CNTL, ObsSparse, and ObsDense in Figs. 12a–c), which is consistent with findings from several previous studies (Dong et al. 2011; Zhu et al. 2013; Periañez et al. 2014; Kirchgessner et al. 2014; Snook et al. 2015). It is evident that the observation error variance $R(k)$ influences the lowest analysis error variance $E(k)$ achievable, and the best-performing ROI is larger when a less accurate observing network is used. The accuracy effect alone is responsible for some of the sensitivity in localization at L and M scales as $R(k)$ is reduced. A more accurate observing network with smaller $R(k)$ yields a smaller $E(k)$ that has reduced contribution from large scales, which causes its correlation length to be shorter so that a smaller ROI is

favorable. However, for S scale, the accuracy effect does not change the best ROI because prior error is still saturated at this unobserved scale despite a lower $R(k)$. The number effect can be seen by comparing ObsSparse to ObsErrorX3 (also ObsError/3 to ObsDense) in Figs. 12a–c. The former has lower observation density than the latter while they share the same $R(k)$. As fewer pieces of independent information are available within the localization scale (ROI), the range of ROIs with good performance around the best ROI gets narrower. This implies that the sensitivity to localization increases as fewer independent observations are available. Both accuracy and number effects contribute to the sensitivity of localization to changes in the observing network. The accuracy effect appears to be more important when observations are dense. As observations become sparse, the number effect becomes more important.

Spatially inhomogeneous observing networks are more common than uniform observation grids. The irregularity in the horizontal location of observations precludes the calculation of observation error spectral

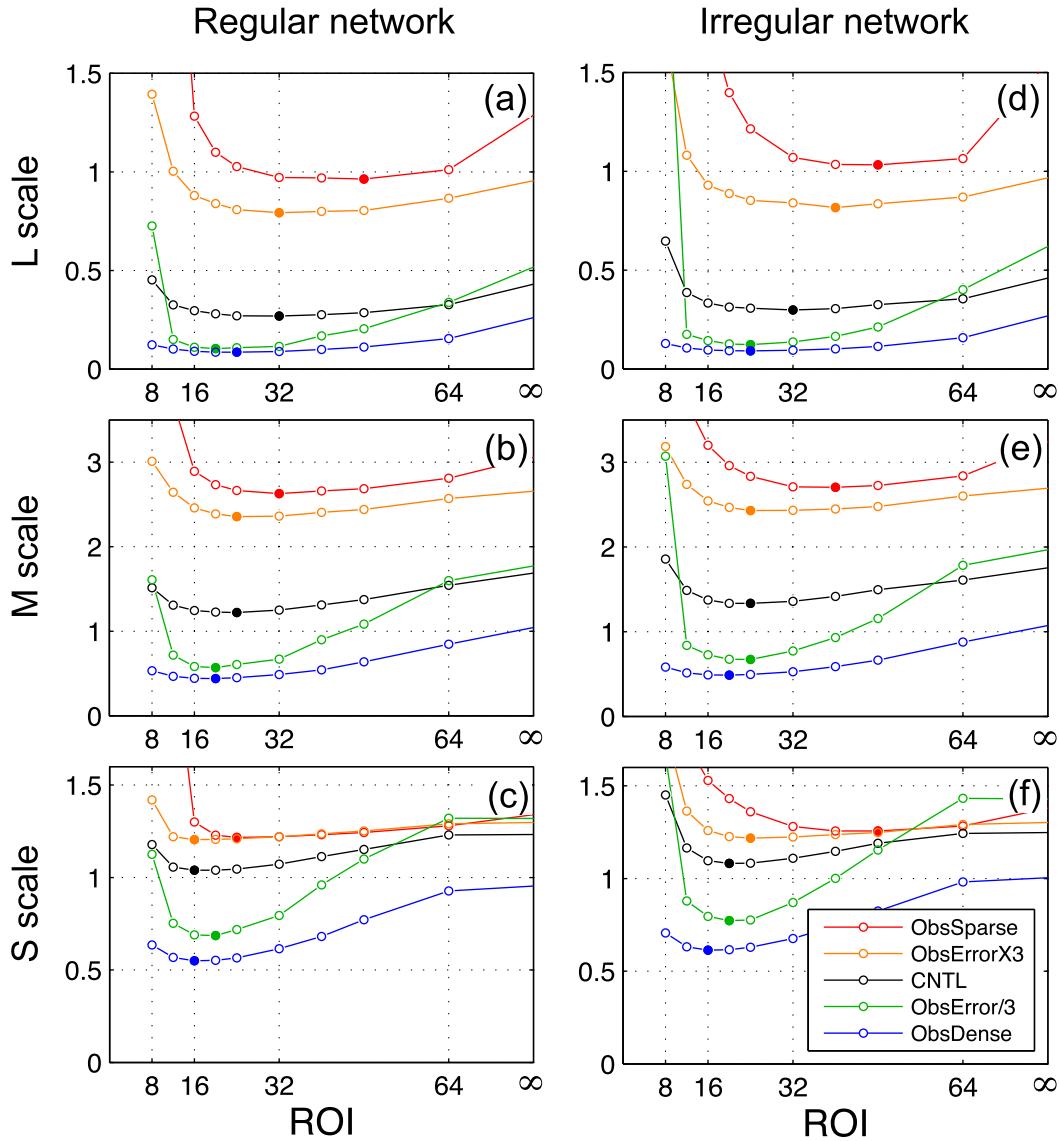


FIG. 12. Temperature (θ) RMSE filtered for (a) L, (b) M, and (c) S scales plotted with respect to ROI for the cases with changing observing networks: ObsSparse, ObsDense, ObsErrorX3, and ObsError/3. Filled circle markers indicate the minimum error corresponding to the best ROI. (d)–(f) As in (a)–(c), but showing results from using irregular observing networks.

variance. We repeated the experiments in this section using observations randomly located in the domain but keeping the number of observations and instrument error unchanged. Figures 12d–f show results from using these irregular observing networks compared to their regular network counterparts (Figs. 12a–c). Overall, the behavior of an irregular network is very similar to a regular one with the same accuracy and density. The only difference is that an irregular network yields larger analysis error, especially for sparse networks where there are occasionally fewer observations within the localization scale. For a sparse irregular observing

network, the best ROI is larger than the best ROI for a regular network with the same number of observations. These results suggest that an irregular observing network is less efficient in reducing analysis error than a uniform network with the same number of observations.

5. Conclusions

In this study, numerical experiments are conducted to test the scale dependency of localization and its sensitivity to several aspects of an ensemble modeling and data assimilation system. The two-layer quasigeostrophic

(QG) model is employed as the forecast model. It is a simple model but captures the essence of large-scale atmospheric dynamics with a realistic spectral energy distribution. Compared to standard test models such as the Lorenz (1996) system, the QG model has better representation of multiscale dynamics and therefore is better suited for testing when scale is the key concern. The data assimilation method considered here is the square root filter (Whitaker and Hamill 2002) with an adaptive covariance relaxation (Ying and Zhang 2015) that ensures filter stability. The Gaspari and Cohn (1999) function with fixed localization radius is applied in the filter, and the localization radius that minimizes analysis error is determined by trial and error. The findings are summarized as follows:

- 1) The best localization radius is scale dependent. For a weather system, the overall best localization radius scales with its dominant correlation length scale. When multiple scales of motion are present simultaneously, the localization that minimizes analysis error at a certain scale also depends on the correlation length for that scale. For example, larger scale favors larger localization radius. A multiscale localization can achieve better filter performance by specifying a scale-dependent localization radius.
- 2) A lower model resolution does not change the best localization radius (defined as a physical distance) for the resolved scales as long as the model representation of the dynamical processes is correct. If the low-resolution model cannot adequately resolve some key dynamical processes, the incorrect model dynamics will give rise to representation errors when assimilating observations. However, if adaptive inflation methods are applied to account for these representation errors, the localization radius is not sensitive to model resolution changes.
- 3) Consistent with previous studies, a decrease in ensemble size is found to cause the best localization radius to shift to smaller values, and the range of localization radii with good performance also becomes narrower. However, this behavior changes when assimilating observations whose correlations with state variables are nonlocal. For correlation functions with peaks at a certain distance (i.e., correlation between u and ψ), further reducing the localization radius will not remedy sampling error as ensemble size decreases.
- 4) Increasing the density of observing networks without changing the instrument error causes the best localization radius to shift to smaller values. Two effects contribute to this behavior. (i) A denser network yields lower spectral observation error variance for

the large scales, which lowers the large-scale prior error. As a result, the overall correlation length scale decreases for the prior ensemble due to less contribution from large-scale errors. (ii) A denser network provides more pieces of independent observation information within a localization radius, and a smaller radius is enough to allow the same number of observations to constrain each state variable. Irregular networks behave similarly to regular ones with the same density in terms of best localization radius, except that a larger radius is favored for irregular networks when observations are too sparse to sample a certain scale.

The particular type of localization used in this study is observation-space localization (Houtekamer and Mitchell 2001; Hamill et al. 2001), which tapers the sample covariance between an observation and state variables and between an observation and other observation priors. Model-space localization (Houtekamer and Mitchell 1998) directly tapers the background error covariance before the observation operator is applied. Although the similarity of the two approaches is proven (Sakov and Bertino 2011; Nerger et al. 2012), these approaches can have different behavior under certain conditions such as small ensemble size and short localization distance (Lei and Whitaker 2015). For example, when assimilating satellite observation impact in the vertical, Campbell et al. (2010) suggested that model-space localization is superior. However, Lei and Whitaker (2015) showed that the opposite can be true for some cases. To develop a more robust localization theory, both types of localization need to be further investigated, especially for cases where nonlocal observation operators involve spatial averaging. Correlated observation errors may also change the behavior of localization. The impact from observation operators and error models are currently being investigated in a follow-up study.

Although the flow simulated by the QG model possesses many scales, its dynamical process (baroclinic instability) is relatively simple. In real atmospheric models, there could be multiple sources of instability at different scales. For example, the addition of moist convective instability may change the error growth rate at small scales, and the scale interaction may be more complicated. Our study motivates further exploration of the behavior of localization in different contexts in pursuit of a more generalized theory.

Acknowledgments. This research was supported by NSF Grant AGS-1305798 and ONR Grant N000140910526. Yue Ying was also sponsored by the

China Scholarship Council (CSC) visitor program. We would like to acknowledge high-performance computing support from Yellowstone provided by NCAR's Computational and Information Systems Laboratory, sponsored by the NSF. The authors would also like to thank Shafer Smith for providing detailed description of the QG model. We also appreciate the insightful comments from three anonymous reviewers and the editor that improved an earlier version of this paper.

REFERENCES

- Aksoy, A., S. Lorsolo, T. Vukicevic, K. J. Sellwood, S. D. Abersson, and F. Zhang, 2012: The HWRP Hurricane Ensemble Data Assimilation System (HEDAS) for high-resolution data: The impact of airborne Doppler radar observations in an OSSE. *Mon. Wea. Rev.*, **140**, 1843–1862, <https://doi.org/10.1175/MWR-D-11-00212.1>.
- Anderson, J. L., 2007: Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D*, **230**, 99–111, <https://doi.org/10.1016/j.physd.2006.02.011>.
- , 2012: Localization and sampling error correction in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, **140**, 2359–2371, <https://doi.org/10.1175/MWR-D-11-00013.1>.
- , 2016: Reducing correlation sampling error in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, **144**, 913–925, <https://doi.org/10.1175/MWR-D-15-0052.1>.
- , and L. Lei, 2013: Empirical localization of observation impact in ensemble Kalman filters. *Mon. Wea. Rev.*, **141**, 4140–4153, <https://doi.org/10.1175/MWR-D-12-00330.1>.
- Bishop, C. H., and D. Hodyss, 2007: Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quart. J. Roy. Meteor. Soc.*, **133**, 2029–2044, <https://doi.org/10.1002/qj.169>.
- , and —, 2009: Ensemble covariances adaptively localized with ECO-RAP. Part 1: Tests on simple error models. *Tellus*, **61A**, 84–96, <https://doi.org/10.1111/j.1600-0870.2008.00371.x>.
- Bocquet, M., L. Wu, and F. Chevallier, 2011: Bayesian design of control space for optimal assimilation of observations. I: Consistent multiscale formalism. *Quart. J. Roy. Meteor. Soc.*, **137**, 1340–1356, <https://doi.org/10.1002/qj.837>.
- Buehner, M., and A. Shlyayeva, 2015: Scale-dependent background-error covariance localization. *Tellus*, **67A**, 28027, <https://doi.org/10.3402/tellusa.v67.28027>.
- Campbell, W. F., C. H. Bishop, and D. Hodyss, 2010: Vertical covariance localization for satellite radiances in ensemble Kalman filters. *Mon. Wea. Rev.*, **138**, 282–290, <https://doi.org/10.1175/2009MWR3017.1>.
- Dong, J., M. Xue, and K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteor. Atmos. Phys.*, **112**, 41–61, <https://doi.org/10.1007/s00703-011-0130-3>.
- Flowerdew, J., 2015: Towards a theory of optimal localization. *Tellus*, **67A**, 25257, <https://doi.org/10.3402/tellusa.v67.25257>.
- Gaspari, G., and S. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, <https://doi.org/10.1002/qj.49712555417>.
- Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522, <https://doi.org/10.1175/2010MWR3328.1>.
- Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790, [https://doi.org/10.1175/1520-0493\(2001\)129<2776:DDFOBE>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2).
- Harlim, J., and A. J. Majda, 2010: Filtering turbulent sparsely observed geophysical flows. *Mon. Wea. Rev.*, **138**, 1050–1083, <https://doi.org/10.1175/2009MWR3113.1>.
- Held, I. M., and V. D. Larichev, 1996: A scaling theory for horizontally homogeneous, baroclinically unstable flow on a beta plane. *J. Atmos. Sci.*, **53**, 946–952, [https://doi.org/10.1175/1520-0469\(1996\)053<0946:ASTFHH>2.0.CO;2](https://doi.org/10.1175/1520-0469(1996)053<0946:ASTFHH>2.0.CO;2).
- Hodyss, D., and N. Nichols, 2015: The error of representation: Basic understanding. *Tellus*, **67A**, 24822, <https://doi.org/10.3402/tellusa.v67.24822>.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, [https://doi.org/10.1175/1520-0493\(1998\)126<0796:DAUAEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2).
- , and —, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137, [https://doi.org/10.1175/1520-0493\(2001\)129<0123:ASEKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2).
- , and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532, <https://doi.org/10.1175/MWR-D-15-0440.1>.
- , X. Deng, H. L. Mitchell, S.-J. Baek, and N. Gagnon, 2014: Higher resolution in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 1143–1162, <https://doi.org/10.1175/MWR-D-13-00138.1>.
- Janjić, T., and S. E. Cohn, 2006: Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Wea. Rev.*, **134**, 2900–2915, <https://doi.org/10.1175/MWR3229.1>.
- Keptert, J. D., 2009: Covariance localisation and balance in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 1157–1176, <https://doi.org/10.1002/qj.443>.
- Kirchgessner, P., L. Nerger, and A. Bunse-Gerstner, 2014: On the choice of an optimal localization radius in ensemble Kalman filter methods. *Mon. Wea. Rev.*, **142**, 2165–2175, <https://doi.org/10.1175/MWR-D-13-00246.1>.
- Kleist, D. T., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-hybrid results. *Mon. Wea. Rev.*, **143**, 433–451, <https://doi.org/10.1175/MWR-D-13-00351.1>.
- Kondo, K., and T. Miyoshi, 2016: Impact of removing covariance localization in an ensemble Kalman filter: Experiments with 10240 members using an intermediate AGCM. *Mon. Wea. Rev.*, **144**, 4849–4865, <https://doi.org/10.1175/MWR-D-15-0388.1>.
- Lange, H., and G. C. Craig, 2014: The impact of data assimilation length scales on analysis and prediction of convective storms. *Mon. Wea. Rev.*, **142**, 3781–3808, <https://doi.org/10.1175/MWR-D-13-00304.1>.
- Larichev, V. D., and I. M. Held, 1995: Eddy amplitudes and fluxes in a homogeneous model of fully developed baroclinic instability. *J. Phys. Oceanogr.*, **25**, 2285–2297, [https://doi.org/10.1175/1520-0485\(1995\)025<2285:EAAFIA>2.0.CO;2](https://doi.org/10.1175/1520-0485(1995)025<2285:EAAFIA>2.0.CO;2).
- Lei, L., and J. S. Whitaker, 2015: Model space localization is not always better than observation space localization for assimilation of satellite radiances. *Mon. Wea. Rev.*, **143**, 3948–3955, <https://doi.org/10.1175/MWR-D-14-00413.1>.
- , and —, 2017: Evaluating the trade-offs between ensemble size and ensemble resolution in an ensemble-variational data

- assimilation system. *J. Adv. Model. Earth Syst.*, **9**, 781–789, <https://doi.org/10.1002/2016MS000864>.
- , J. L. Anderson, and G. S. Romine, 2015: Empirical localization functions for ensemble Kalman filter data assimilation in regions with and without precipitation. *Mon. Wea. Rev.*, **143**, 3664–3679, <https://doi.org/10.1175/MWR-D-14-00415.1>.
- Li, Z., J. C. McWilliams, K. Ide, and J. D. Farrara, 2015: A multi-scale variational data assimilation scheme: Formulation and illustration. *Mon. Wea. Rev.*, **143**, 3804–3822, <https://doi.org/10.1175/MWR-D-14-00384.1>.
- 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, <https://doi.org/10.1256/qj.02.132>.
- Lorenz, E. N., 1996: Predictability—A problem partly solved. *Proc. Seminar on Predictability*, Reading, United Kingdom, ECMWF, 18 pp., <https://www.ecmwf.int/sites/default/files/elibrary/1995/10829-predictability-problem-partly-solved.pdf>.
- Miyoshi, T., and K. Kondo, 2013: A multi-scale localization approach to an ensemble Kalman filter. *SOLA*, **9**, 170–173, <https://doi.org/10.2151/sola.2013-038>.
- Nerger, L., T. Janjić, J. Schroeter, and W. Hiller, 2012: A regulated localization scheme for ensemble-based Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **138**, 802–812, <https://doi.org/10.1002/qj.945>.
- Periáñez, Á., H. Reich, and R. Potthast, 2014: Optimal localization for ensemble Kalman filter systems. *J. Meteor. Soc. Japan*, **92**, 585–597, <https://doi.org/10.2151/jmsj.2014-605>.
- Sakov, P., and L. Bertino, 2011: Relation between two common localization methods for the EnKF. *Comput. Geosci.*, **15**, 225–237, <https://doi.org/10.1007/s10596-010-9202-6>.
- Salmon, R., 1998: *Lectures on Geophysical Fluid Dynamics*. Oxford University Press, 400 pp.
- Smith, K. S., G. Boccaletti, C. C. Henning, I. Marinov, C. Y. Tam, I. M. Held, and G. K. Vallis, 2002: Turbulent diffusion in the geostrophic inverse cascade. *J. Fluid Mech.*, **469**, 13–48, <https://doi.org/10.1017/S0022112002001763>.
- Snook, N., X. Ming, and Y. Jung, 2015: Multiscale EnKF assimilation of radar and conventional observations and ensemble forecasting for a tornadic mesoscale convective system. *Mon. Wea. Rev.*, **143**, 1035–1057, <https://doi.org/10.1175/MWR-D-13-00262.1>.
- Sobash, R. A., and D. J. Stensrud, 2013: The impact of covariance localization for radar data on EnKF analyses of a developing MCS: Observing system simulation experiments. *Mon. Wea. Rev.*, **141**, 3691–3709, <https://doi.org/10.1175/MWR-D-12-00203.1>.
- van Leeuwen, P. J., 2015: Representation errors and retrievals in linear and nonlinear data assimilation. *Quart. J. Roy. Meteor. Soc.*, **141**, 1612–1623, <https://doi.org/10.1002/qj.2464>.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, [https://doi.org/10.1175/1520-0493\(2002\)130<1913:EDAWPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2).
- Ying, Y., and F. Zhang, 2015: An adaptive covariance relaxation method for ensemble data assimilation. *Quart. J. Roy. Meteor. Soc.*, **141**, 2898–2906, <https://doi.org/10.1002/qj.2576>.
- Yoon, Y., E. Ott, and I. Szunyogh, 2010: On the propagation of information and the use of localization in ensemble Kalman filtering. *J. Atmos. Sci.*, **67**, 3823–3834, <https://doi.org/10.1175/2010JAS3452.1>.
- Zhang, F., Z. Meng, and A. Aksoy, 2006: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part I: Perfect model experiments. *Mon. Wea. Rev.*, **134**, 722–736, <https://doi.org/10.1175/MWR3101.1>.
- , Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **137**, 2105–2125, <https://doi.org/10.1175/2009MWR2645.1>.
- Zhen, Y., and F. Zhang, 2014: A probabilistic approach to adaptive covariance localization for serial ensemble square-root filters. *Mon. Wea. Rev.*, **142**, 4499–4518, <https://doi.org/10.1175/MWR-D-13-00390.1>.
- Zhu, K., Y. Pan, M. Xue, X. Wang, J. S. Whitaker, S. G. Benjamin, S. S. Weygandt, and M. Hu, 2013: A regional GSI-based ensemble Kalman filter data assimilation system for the Rapid Refresh configuration: Testing at reduced resolution. *Mon. Wea. Rev.*, **141**, 4118–4139, <https://doi.org/10.1175/MWR-D-13-00039.1>.