Assimilating Observations with Spatially Correlated Errors Using a Serial Ensemble Filter with a Multiscale Approach

YUE YING

Advanced Study Program, National Center for Atmospheric Research, Boulder, Colorado

ABSTRACT

The serial ensemble square root filter (EnSRF) typically assumes observation errors to be uncorrelated when assimilating the observations one at a time. This assumption causes the filter solution to be suboptimal when the observation errors are spatially correlated. Using the Lorenz-96 model, this study evaluates the suboptimality due to mischaracterization of observation error spatial correlations. Neglecting spatial correlations in observation errors results in mismatches between the specified and true observation error variances in spectral space, which cannot be resolved by inflating the overall observation error variance. As a remedy, a multiscale observation (MSO) method is proposed to decompose the observations into multiple scale components and assimilate each component with separately adjusted spectral error variance. Experimental results using the Lorenz-96 model show that the serial EnSRF, with the help from the MSO method, can produce solutions that approach the solution from the EnSRF with correctly specified observation error correlations as the number of scale components increases. The MSO method is further tested in a two-layer quasigeostrophic (QG) model framework. In this case, the MSO method is combined with the multiscale localization (MSL) method to allow the use of different localization radii when updating the model state at different scales. The combined method (MSOL) improves the serial EnSRF performance when assimilating observations with spatially correlated errors. With adjusted observation error spectral variances and localization radii, the combined MSOL method provides the best solution in terms of analysis accuracy and filter consistency. Prospects and challenges are also discussed for the implementation of the MSO method for more complex models and observing networks.

1. Introduction

Data assimilation (DA) combines the information from observations and model forecasts according to their relative uncertainties to obtain an improved estimate of the model state. The ensemble Kalman filter (EnKF; [Evensen 1994](); [Burgers et al. 1998]()) is a popular ensemble filter formulation where prior and observation errors are assumed to follow Gaussian distributions. Serial EnKF formulations ([Anderson 2003](); [Tippett et al. 2003]()), where observation errors are assumed uncorrelated and observations are assimilated one at a time, have good computational scaling and have been successfully implemented in operational prediction systems (e.g., [Whitaker et al. 2008]()). Despite its popularity, the serial EnKF is known to have suboptimal performance when the observation errors are correlated, since the uncorrelated observation error assumption will cause an overfit to the observations.

In this paper, observation errors are considered as including both instrument noises and errors of representativeness, or "representation errors" ([van Leeuwen 2015](); [Hodyss and Satterfield 2017](); [Janjic et al. 2018]()). Representation errors can be caused by errors in the forward operator that maps model state to corresponding observations, scale mismatches between observations and model states, as well as errors introduced during preprocessing of observations. Scale mismatches between observations and model states often give rise to spatial correlations in representation errors ([Stewart 2010](); [Weston et al. 2014](); [Waller et al. 2014]()). Due to these spatial correlations, observations obtained from Doppler radar ([Waller et al. 2016]()), satellite radiances ([Bormann and Bauer 2010](); [Geer et al. 2017](); [Migliorini and Candy 2019]()), and many others cannot be fully utilized in DA. To avoid overfitting to observations with correlated errors, it is common practice to inflate the observation error variances during DA ([Courtier et al. 1998](); [Collard 2004](); [Hilton et al. 2009](); [Bormann et al. 2016]()). However, [Miyoshi et al. (2013)]() pointed out that

Corresponding author: Dr. Yue Ying, mying@ucar.edu

inflating observation error variances to try to compensate for correlated observation errors is suboptimal. Other observation subsampling methods such as thinning (Liu and Rabier 2002, 2003; Dando et al. 2007) and the creation of superobservation, also known as "superobbing" (Berger and Forsythe 2004; Zhang et al. 2009), have been proposed to reduce spatial correlations in observation errors. Hoffman (2018) recently investigated the impact of observation thinning or smoothing on analysis accuracy in a simple one-dimensional setting. Although subsampling the observations will prevent overfitting, doing so will discard small-scale observation information and degrade the overall DA performance.

For high-resolution prediction systems, the need for retaining small-scale information calls for DA methods that can incorporate the observational information at all ranges of scales. Recent studies have investigated DA methods that can account for observation error correlations in simplified (Stewart et al. 2008, 2013; Rainwater et al. 2015; Ruggiero et al. 2016) or realistic scenarios (Bormann and Bauer 2010; Weston et al. 2014; Waller et al. 2016; Bormann et al. 2016; Campbell et al. 2017; Simonin et al. 2019). These DA methods, although more computationally expensive, allow more observation information to be incorporated and thus improves prediction skill. A common approach for ensemble filtering to assimilate observations with correlated errors is to perform a transform of the observations to a space where their errors are no longer correlated (rotating the observation error covariance matrix) and perform DA in that space (e.g., Anderson 2001). For example, if the observation error correlation length is constant in space, the observations can be transformed to spectral space (e.g., Fourier transform) where their errors become uncorrelated. However, reducing sampling noises in spectral space is more difficult, since a distance-based correlation function in physical space will transform to a nonlocal function in spectral space.

Spatial correlations in observation errors change the distribution of error variances in spectral space, effectively changing the relative precision of observation information at different scales (Miyoshi et al. 2013; Rainwater et al. 2015; Fowler et al. 2018). Observation errors with spatial correlation are essentially a red noise with more error variances concentrated at larger scales. Recent ideas of using spatial differences of neighboring observations to extract small-scale information and improve DA performance (Anderson 2016; Bedard and Buehner 2020) inspired a new multiscale approach described in this paper. The new approach decomposes the observations into several scale components and assimilates them sequentially using a serial EnKF. During assimilation of an observation scale component, its corresponding error variance is adjusted individually to account for the effect of spatial correlations.

This paper is organized as follows. Section 2 describes the DA framework used in this study for testing the multiscale approach. Section 3 uses the Lorenz (1996) model to demonstrate the suboptimality due to mischaracterizing observation error correlations and provides a proof of concept of how the multiscale approach reduces such suboptimality. Section 4 further evaluates the multiscale approach using a two-layer quasigeostrophic model. Section 5 discusses further implementation of the multiscale approach in real prediction systems and section 6 summarizes the findings of this study.

2. Methodology

a. The ensemble square root filters

Ensemble filtering is a sequential DA approach that propagates the model state in time and assimilates observations as they become available to update the state. An ensemble of model state realizations is used to provide a flow-dependent representation of its uncertainties. An analysis cycle consists of an filter update step where the prior ensemble from model forecasts is fused with observations to form the posterior ensemble (also known as the analyses) and a forecast step where the posterior ensemble is propagated forward in time to provide priors for the next cycle. This study focuses on a particular variant of ensemble filters known as the ensemble square root filter (EnSRF), which is described as follows.

Let $\mathbf{x}$ denote a random vector of size N_x containing the model state variables. Let N be the ensemble size and subscript n indexes ensemble members. Superscripts b and a denote the prior and posterior state, respectively. The nth prior ensemble member can be written as $\mathbf{x}_n^b = \overline{\mathbf{x}}^b + \mathbf{x}_n'^b$, where

$$\overline{\mathbf{x}}^b = \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}_n^b \qquad (1)$$

is the ensemble mean and $\mathbf{x}_n'^b$ is the nth ensemble perturbation. Let $\mathbf{y}^o$ be a vector of size N_y that contains the observations. H is a forward operator that maps a prior model state to its corresponding observation priors:

$$\mathbf{y}_n^b = H(\mathbf{x}_n^b). \qquad (2)$$

The observation priors can also be written in terms of the sum of mean and perturbations, $\mathbf{y}_n^b = \overline{\mathbf{y}}^b + \mathbf{y}_n'^b$. The ensemble filter derives the posterior states by weighting the prior states and the observations according to their

relative uncertainties. The prior state and observation errors are both assumed to follow Gaussian distributions with zero mean (unbiased) and error covariances $\mathbf{B}$ and $\mathbf{R}$, respectively. The prior ensemble $\mathbf{x}_n^b$ provides a sample estimate of $\mathbf{B}$, while $\mathbf{R}$ is typically specified.

A shorthand notation for sample-estimated error covariance between two vectors $\mathbf{a}$ and $\mathbf{b}$ is defined as

$$\text{cov}(\mathbf{a}, \mathbf{b}) = \frac{1}{N-1}\sum_{n=1}^{N} \mathbf{a}_n'(\mathbf{b}_n')^{\text{T}}, \qquad (3)$$

where n indexes the ensemble member and the prime indicate ensemble perturbations. For the sake of simplicity, the covariance between a vector $\mathbf{a}$ and itself is denoted as $\text{cov}(\mathbf{a})$. Note the notation also holds for scalars [e.g., $\text{cov}(\mathbf{a}, c)$ is the error covariance between a vector $\mathbf{a}$ and scalar c]. The covariance between a scalar c and itself is just its variance, and therefore denoted as $\text{var}(c)$.

The EnSRF (Whitaker and Hamill 2002) update equations can be written as

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \text{cov}(\mathbf{x}^b, \mathbf{y}^b)[\text{cov}(\mathbf{y}^b) + \mathbf{R}]^{-1}(\mathbf{y}^o - \bar{\mathbf{y}}^b), \qquad (4)$$

and

$$\mathbf{x}_n'^a = \mathbf{x}_n'^b - \text{cov}(\mathbf{x}^b, \mathbf{y}^b)\left[\left(\sqrt{\text{cov}(\mathbf{y}^b) + \mathbf{R}}\right)^{-1}\right]^{\text{T}}\left[\sqrt{\text{cov}(\mathbf{y}^b)} + \sqrt{\mathbf{R}}\right]^{-1}\mathbf{y}_n'^b, \qquad (5)$$

where $\mathbf{x}_n^a = \bar{\mathbf{x}}^a + \mathbf{x}_n'^a$ is the nth posterior member. The EnSRF is a deterministic ensemble filter formulation (e.g., Bishop et al. 2001; Anderson 2001) that does not require perturbing the observations and therefore avoids introducing additional sampling noise.

For observations with uncorrelated errors ($\mathbf{R}$ is diagonal), the observations can be assimilated serially, which is known as the "serial EnSRF" (Whitaker and Hamill 2002; Tippett et al. 2003). Let

$$\mathbf{z}_n = [\mathbf{x}_n, H(\mathbf{x}_n)] = [\mathbf{x}_n, \mathbf{y}_n] \qquad (6)$$

be a state vector in a joint state–observation space (Anderson 2003) for the nth ensemble member. The serial EnSRF assimilates one observation at a time to iteratively update the joint space state vectors. Let y_j^o be the jth observation and $\sigma_{o,j}^2$ be its error variance (jth diagonal term of $\mathbf{R}$). Superscript (j) indicates the updated values after assimilating the previous $j-1$ observations. Starting from the prior state $\mathbf{z}_n^{(1)} = \mathbf{z}_n^b$. The update equations for the jth observation are

$$\bar{\mathbf{z}}^{(j+1)} = \bar{\mathbf{z}}^{(j)} + \mathbf{K}_j[y_j^o - \bar{y}_j^{(j)}], \qquad (7)$$

and

$$\mathbf{z}_n'^{(j+1)} = \mathbf{z}_n'^{(j)} - \phi\mathbf{K}_j y_{j,n}'^{(j)}, \qquad (8)$$

where

$$\mathbf{K}_j = \frac{\text{cov}[\mathbf{z}^{(j)}, y_j^{(j)}]}{\text{var}[y_j^{(j)}] + \sigma_{o,j}^2} \qquad (9)$$

is the Kalman gain associated with the jth observation, and

$$\phi = \left(1 + \sqrt{\frac{\sigma_{o,j}^2}{\text{var}[y_j^{(j)}] + \sigma_{o,j}^2}}\right)^{-1} \qquad (10)$$

is a square root modification factor. When all N_y observations are assimilated, $\mathbf{x}_n^a = \bar{\mathbf{x}}^{(N_y+1)} + \mathbf{x}_n'^{(N_y+1)}$ gives the nth posterior ensemble member. A full $\mathbf{R}$ matrix can be specified in the EnSRF to account for error correlations. Section 3a will use numerical experiments to show how the serial EnSRF solution deviates from the EnSRF solution due to mischaracterization of $\mathbf{R}$.

b. Spatially correlated observation errors

This study only considers spatially uniform observing networks. A simple first-order autoregressive process is selected as observation error model. The $\mathbf{R}$ matrix is defined as

$$\mathbf{R}_{ij} = \sigma^2 \exp(-D_{ij}/L), \qquad (11)$$

where σ^2 is the error variance, D_{ij} is the distance between the ith and jth observations, and L is the correlation length scale. Both σ and L are considered to be constant in space and time. Figure 1a shows the spatial correlation function when $L = 0, 1, 2, 5,$ and 10 grid points on a one-dimensional domain. $L = 0$ is the case when observation errors are uncorrelated. The $\mathbf{R}$ matrix can be diagonalized as

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}^o\mathbf{U}^{\text{T}}, \qquad (12)$$

where $\mathbf{\Lambda}^o$ contains the eigenvalues (error spectrum) along the diagonal and $\mathbf{U}$ contains the discrete Fourier bases as eigenvectors; Λ_k^o, the kth diagonal term in $\mathbf{\Lambda}^o$, is the spectral error variance associated with wavenumber $k = 0, 1, \ldots, N_y/2$. The (j, k)th element of $\mathbf{U}$ is given by

$$U_{jk} = \sqrt{\frac{2}{N_y}} \exp\left(2\pi i\frac{kj}{N_y}\right), \qquad (13)$$

where $i = \sqrt{-1}$. Figure 1b shows the error spectra for $\sigma = 1$ and L varying from 0 to 10. The uncorrelated
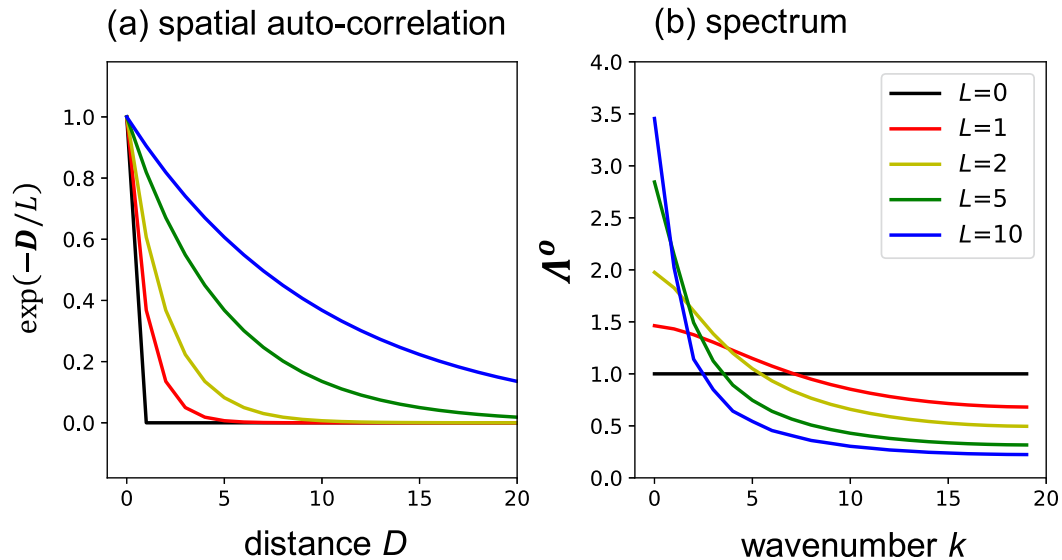
(a) spatial auto-correlation (b) spectrum

FIG. 1. (a) Spatial autocorrelation function and (b) its corresponding spectrum of a one-dimensional error field with 40 grid points. Varying correlation length scales $L = 0, 1, 2, 5$, and 10 grid points are shown in different colors.

errors ($L = 0$) have an even distribution of spectral variances. As L increases, more spectral error variances are distributed in the large scales (small k). When the true observation errors are spatially correlated, assuming uncorrelated observation errors (in the serial EnSRF) is effectively overspecifying observation error variances at small scales and underspecifying them at large scales.

Note that the exact choice of observation error correlation functions is not important for reasons that will become clear later as the multiscale approach is introduced in section 2d. Decomposing the observations into scale components will allow adjustment of spectral error variances to adapt to any arbitrary correlation functions. What is more challenging is the spatially and temporally varying σ and L in real observations, which will be discussed in detail in section 5.

c. Metrics for evaluating filter performance

Let us now define several quantities to measure filter performance. The true model state from a nature run is denoted as $\mathbf{x}^*$. Synthetic observations are generated as $\mathbf{y}^o = H(\mathbf{x}^*) + \boldsymbol{\varepsilon}^o$, where random noises $\boldsymbol{\varepsilon}^o$ are drawn from the true observation error distribution $N(0, \mathbf{R}^*)$ with the true error statistics σ^* and L^*. The observation error covariance specified a priori in the EnSRF is denoted as $\mathbf{R}$ with parameters σ and L. The true and specified observation error spectra are denoted as Λ^{o*} and Λ^o, respectively, according to (12). The error of the prior ensemble mean compared to the truth, $\boldsymbol{\varepsilon}^b = \bar{\mathbf{x}}^b - \mathbf{x}^*$, measures the actual accuracy of the prior model state, while the prior ensemble spread represents the uncertainty

"specified" in the EnSRF which reflects the filter's knowledge of the uncertainty in state estimate. The spectrum of the actual prior error, Λ^{b*}, can be calculated as

$$E(\boldsymbol{\varepsilon}^b \boldsymbol{\varepsilon}^{bT}) = \tilde{\mathbf{U}} \Lambda^{b*} \tilde{\mathbf{U}}^T, \qquad (14)$$

where E denotes the expected value that is evaluated by averaging $\boldsymbol{\varepsilon}^b \boldsymbol{\varepsilon}^{bT}$ over many analysis cycles, $\tilde{\mathbf{U}}$ is similar to $\mathbf{U}$ in (13) but replacing N_y with N_x. The spectrum of prior ensemble spread, Λ^b, can be calculated as

$$E[\text{cov}(\mathbf{x}^b)] = \tilde{\mathbf{U}} \Lambda^b \tilde{\mathbf{U}}^T. \qquad (15)$$

Ideally, the ensemble spread should match the actual error standard deviation for each wavenumber, so that the specified uncertainties match the truth and correct weights are assigned to the observations and prior states. A consistency ratio (CR) is defined as the ratio between Λ^b and Λ^{b*} to evaluate this match. When CR < 1, the ensemble spread underestimates the actual uncertainties in model states (the ensemble is underdispersive), and observations are not given enough weights to adjust the state estimates. In extreme cases, the observations can be completely ignored when the ensemble spread collapses, a phenomenon known as "filter divergence". Note that the same metrics can be calculated for the posterior error and ensemble spread, Λ^{a*} and Λ^a. The overall error and spread can be measured by summing over all wavenumbers. For example, the square root of the trace of Λ^{a*} and Λ^a gives the overall posterior RMSE and posterior ensemble spread, respectively.

Covariance localization (Houtekamer and Mitchell 2001; Hamill et al. 2001) and inflation (Anderson and Anderson 1999) are also implemented to counter the deficiencies in ensemble filter and improve its performance. Localization reduces the sampling noises due to the use of a limited-size ensemble. The Gaspari and Cohn (1999) localization function is used in this study with radius of influence (ROI) defined as the distance at which ensemble estimated covariance is tapered to zero. A multiplicative inflation factor λ is applied to the prior ensemble perturbations before the filter update to maintain filter consistency in the presence of model errors or other unaccounted for error sources.

d. A multiscale ensemble filtering approach

This section introduces a multiscale ensemble filtering approach comprising two methods: multiscale observation (MSO) and multiscale localization (MSL). The MSO method decomposes the observations into several scale components and the filter update is calculated sequentially for each component to update the model state. The uniform observing network considered in this study allows scale decomposition to be performed by fast Fourier transforms. The sth observation scale component can be found by

$$\mathbf{y}_s^o = \mathbf{U}[\mathbf{f}_s \circ (\mathbf{U}^T \mathbf{y}^o)], \qquad (16)$$

where $\mathbf{U}$ contains the Fourier bases from (13), $\mathbf{f}_s$ is a scale-selective response function for the sth scale component, and $\circ$ denotes the element-wise product. Approaches for extracting scale component from nonuniform observing networks will be discussed in section 5. The number of scale components N_s and the response functions are predefined in this study. The MSO method starts from the prior ensemble $\mathbf{x}_n^{(1)} = \mathbf{x}_n^b$ and observation priors $\mathbf{y}_n^{(1)} = \mathbf{y}_n^b$ (subscript n indexes ensemble member), and performs the following steps for $s = 1, \ldots, N_s$ to iteratively update the states and observation priors.

1) Use (16) to compute the sth scale component of the observations and observation priors, $\mathbf{y}_s^o$ and $\mathbf{y}_{n,s}^{(s)}$.
2) Use $\mathbf{x}_n^{(s)}$ as prior ensemble, $\mathbf{y}_s^o$ as observation, and $\mathbf{y}_{n,s}^{(s)}$ as observation priors; use $(\lambda_s^o \sigma)^2$ as observation error variance, where λ_s^o is an adjustment factor for the spectral observation error varainces. Apply the serial EnSRF update equations in (7)–(10) and set $\mathbf{x}_n^{(s+1)}$ as the updated ensemble and $\mathbf{y}_n^{(s+1)}$ as the updated observation priors.
3) If $s < N_s$, let $s = s + 1$ and go to step 1, otherwise the final posterior ensemble is $\mathbf{x}_n^a = \mathbf{x}_n^{(N_s+1)}$.

In this study, the optimal spectral observation error variance adjustment factors are found by $\lambda_s^o = \sqrt{\Lambda_s^{o*}/\Lambda_s^o}$,

where Λ_s is the averaged spectral variances for the sth scale component. Section 5 will discuss how to estimate λ_s^o when the truth (Λ_s^{o*}) is unknown, and how to account for spatially and temporally varying observation error statistics.

Multiscale localization methods have been suggested in the literature to improve DA performance for high-dimensional problems (e.g., Zhang et al. 2009; Miyoshi and Kondo 2013; Buehner and Shlyaeva 2015). The MSL method is implemented in this study by decomposing the model state into several scale components and applying a separate localization ROI when updating each component. The model state is decomposed into N_t scale components $\mathbf{x}_t = \tilde{\mathbf{U}}[\tilde{\mathbf{f}}_t \circ (\tilde{\mathbf{U}}^T \mathbf{x})]$, where t indexes the scale components. Let $\mathbf{x}_{n,t}^b$ denote the tth scale component of the nth prior ensemble member. Starting from the prior states $\mathbf{x}_n^{(1)} = \mathbf{x}_n^b$, the following steps are applied for $t = 1, \ldots, N_t$ to update the states.

1) The observation priors corresponding to the intermediate updated ensemble states are computed as $\mathbf{y}_n^{(t)} = H[\mathbf{x}_n^{(t)}]$.
2) Use $\mathbf{x}_{n,t}$ as prior ensemble, $\mathbf{y}^o$ as observations, and $\mathbf{y}_n^{(t)}$ as observation priors; use ROI_t as localization radius and σ^2 as observation error variance; apply the serial EnSRF update step described in (7)–(10), and the resulting analysis increments are $\delta \mathbf{x}_n$. Add these increments to update the ensemble as $\mathbf{x}_n^{(t+1)} = \mathbf{x}_n^{(t)} + \delta \mathbf{x}_n$.
3) If $t < N_t$, let $t = t + 1$ and go to step 1, otherwise the final posterior ensemble is $\mathbf{x}_n^a = \mathbf{x}_n^{(N_t+1)}$.

The MSL method can be combined with the MSO method for applications involving both multiple correlation length scales and spatially correlated observation errors. The combined method will be referred to as MSOL, which will go through each pair of sth observation component and tth state component, and compute serial EnSRF analysis increments to iteratively update the ensemble states to reach the final posterior. Note that the multiscale approach can also be applied to other variants of ensemble filters.

3. Lorenz-96 model experiments

In this section, the Lorenz (1996) model is employed as the test model to demonstrate the suboptimality due to mischaracterized observation error statistics in EnSRF and evaluate the MSO method. The governing equations for the Lorenz-96 model state $\mathbf{x}$ are

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F, \qquad (17)$$

$i = 1, 2, \ldots, N_x$, where $N_x = 40$, and F is a forcing parameter. The true model forcing parameter is set to

$F* = 8$, for which chaotic behavior occurs. Cycling DA experiments are conducted over 100 000 cycles with a cycling period of 0.2 time units, and performance metrics (RMSE and CR) are averaged over these cycles. The identity matrix is used as forward operator (i.e., observations are the same as state variables, $\mathbf{y}_n^b = \mathbf{x}_n^b$, and all state variables are observed). The true observation error statistics are set to $\sigma* = 1$ and $L* = 5$.

a. Impact from mischaracterization of observation error correlation

This section evaluates the suboptimality in DA performance due to mis-specified L. In DA practice, several sources of suboptimality may exist simultaneously, including mis-specified observation error statistics, model errors, sampling errors due to limited ensemble size, and the deviation from linear and Gaussian approximation due to nonlinear model dynamics and forward operators. Here, the impact of changing L will be evaluated in several scenarios with or without sampling errors (small N) and model parameter errors (wrong F). In each scenario, the localization ROI and prior inflation factor λ are manually tuned by trying ROI = 5, 10, 15, . . . , 60 grid points, and $\lambda = 1, 1.02, 1.04, . . . , 1.2$ (or up to 1.5 if model errors exist) to find the best solution. These different scenarios provide a more comprehensive evaluation of the suboptimality due to mis-specified L.

A benchmark case is chosen with correctly specified $\sigma = 1$ and $L = 5$, the perfect model ($F = 8$), and ensemble size of $N = 40$. Note that the benchmark solution is still suboptimal because the EnSRF makes linear and Gaussian assumptions and the Lorenz-96 model has nonlinear dynamics. Nonetheless, it is still considered the benchmark case since this study only focuses on mischaracterized observation error statistics. Also note that using very large ensemble sizes in this case does not further improve performance due to the nonlinear model dynamics (Anderson 2010). For the benchmark case, the best ROI = 60 grid points and $\lambda = 1.04$. With correctly specified parameters, only a little localization and inflation is needed, almost all prior error covariances are retained with those at greater distances given less weight.

To understand the impact from mis-specified L, a set of experiments are performed with L that deviates from the benchmark case. The changes in DA performance are evaluated by investigating the error spectra averaged over all analysis cycles. Figure 2 shows the spectra of true observation error (Λ^{o*}), specified observation error (Λ^o), prior error (Λ^{b*}) and ensemble spread (Λ^b), and posterior error (Λ^{a*}) and ensemble spread (Λ^a) for $L = 0, 0.5, 1, 3, 5$ (benchmark), and 8. The benchmark case (Fig. 2a) has minimum spectral error variances and

CR close to 1 for all wavenumbers. When $L > L*$ (Fig. 2b), observation error spectral variances are underspecified at smaller scales compared to the truth, resulting in underdispersive posterior ensemble. Note that observation error spectral variance at $k = 0$ is actually overspecified (not shown on plot) and the overall observation error variance is still matching with the true since $\sigma = \sigma*$. But this overspecification only happens at the largest scale and the posterior solution seems to be dominated by the underspecification at small scales. On the other hand, when $L < L*$ (Fig. 2c), the observation error spectral variances are overspecified at smaller scales, causing an overdispersive posterior ensemble. For the $L = 3$ and $L = 8$ cases, the mismatches between true and specified observation error spectral variances are not large and do not considerably increase the posterior error variances. As L deviates more from $L*$ (Figs. 2d–f), the posterior error variances become much larger than the benchmark case. While posterior error variances keep increasing as L decreases to 0, the posterior ensemble consistency seems to be worst for $L = 1$. For $L = 0$ (Fig. 2f), CR is close to 1 again for most of the spectrum except for wavenumbers 0–2. In this case, observation error spectral variances are underspecified at large scales and overspecified at small scales, and the effects on ensemble spread from such under and overspecification seem to cancel out. However, since observation error spectral variances are not correctly specified during assimilation, the posterior error variances do increase. The solution from $L = 0$ is similar to the case of running the serial EnSRF where observation error correlations are not accounted for. The solutions from the serial EnSRF and the EnSRF with $L = 0$ differ only due to their different ways of applying localization.

In practice, the specified σ is often inflated during DA to compensate for correlated observation errors (Courtier et al. 1998; Bormann et al. 2016). Another set of experiments are performed by deviating both σ and L from the benchmark case to see if adjusting σ can help compensate the misspecification of L. Figure 3 plots the overall posterior RMSE and CR (averaged over all scales) as a function of σ ranging from 0.4 to 1.6 and L ranging from 0 to 10 grid points. The "+" sign marks the benchmark case which has the lowest RMSE and CR close to 1. Results show that small deviations from $L*$ ($L = 3$–10) can be well compensated by adjusting σ. For example, $\sigma = 1.2$ for $L = 7$ and $\sigma = 0.8$ for $L = 3$ both yield DA performance that is comparable to the benchmark. However, as L deviates more from the truth ($L < 2$), adjusting σ becomes no longer effective and overall posterior RMSE increases. As seen from Figs. 2d–f, the shape of the specified observation error spectrum
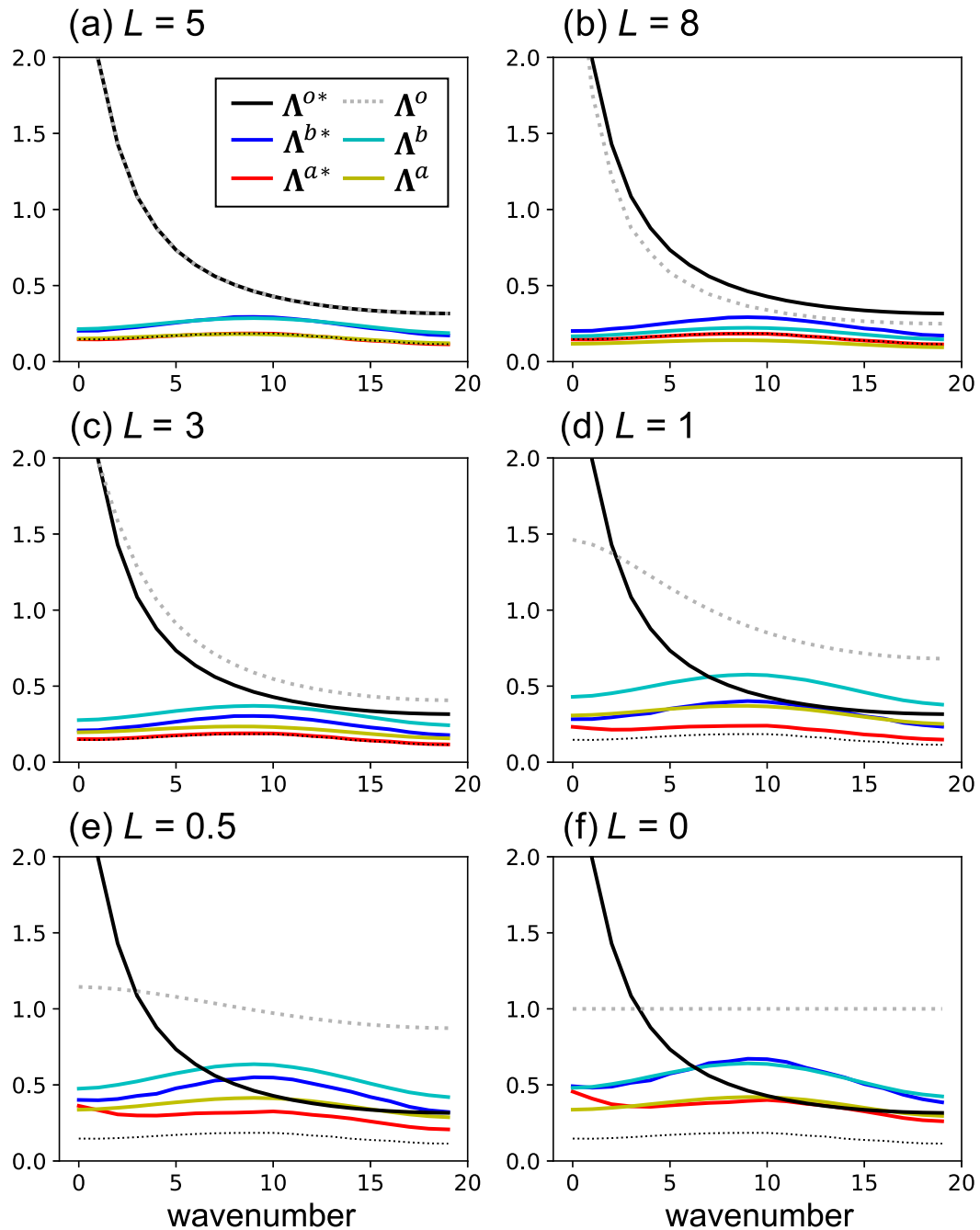
FIG. 2. Error spectrum averaged over the first 5000 analysis cycles for the cases using EnSRF with $N = 40$, $F = 8$, ROI $= 60$, $\lambda = 1.04$, $\sigma = 1$, and (a)–(f) $L = 5, 8, 3, 1, 0.5$, and 0. Each panel shows the true observation error ($\mathbf{\Lambda}^{o*}$), specified observation error ($\mathbf{\Lambda}^o$), prior error ($\mathbf{\Lambda}^{b*}$) and spread ($\mathbf{\Lambda}^b$), and posterior error ($\mathbf{\Lambda}^{a*}$) and spread ($\mathbf{\Lambda}^a$). The spectrum plotted here corresponds to the square root of the diagonals of $\mathbf{\Lambda}$ matrices. The posterior error spectrum of the benchmark case ($L = L^*$) is shown as black dotted lines on each panel for reference.

differs too much from the truth and cannot be matched by simply adjusting the overall error level. For $L < L^*$, inflating σ can better match its large-scale error variances, but at a price of sacrificing the small-scale components where the mismatch gets even worse. Similarly,

trying to match small-scale error variances will sacrifice the large-scale components.

The best DA performance for each specified L is found by manually tuning σ, ROI, and λ, and results are summarized in the first section of Table 1. The benchmark
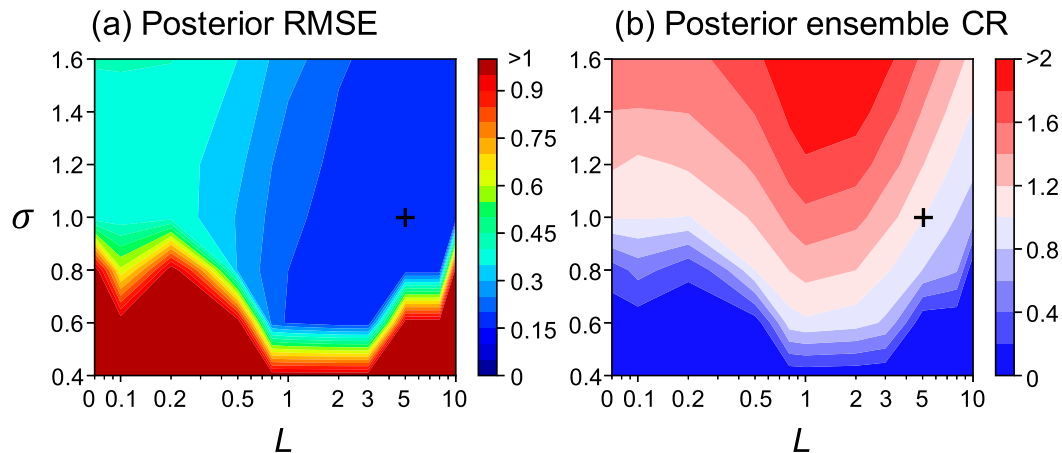
FIG. 3. (a) RMSE and (b) CR of the posterior ensemble averaged over the first 5000 analysis cycles plotted as a function of specified observation error standard deviation σ and correlation length scale L. The cases using EnSRF with $N = 40$ and $F = 8$ are shown. The plus sign indicates the true values, $\sigma^* = 1$ and $L^* = 5$, used in the benchmark case. The L values are plotted in log scale.

case does not require too much localization and inflation because of the near-optimal ensemble size and the perfect model. As a result, cases with $L = 3$–10 have low RMSE and CR close to 1 because L only deviates slightly from the truth ($L = 5$). Note that the shape of the observation error spectrum does not change linearly with L, therefore doubling L from 5 to 10 does not result in big mismatches in error spectra. The suboptimality becomes clear as L drops below 2, where the posterior RMSE increases considerably. The same experiments are repeated for two other scenarios: more sampling errors ($N = 20$, $F = 8$), and imperfect model ($N = 40$, $F = 7.5$), whose results are also summarized in Table 1. In these scenarios, extra sources of suboptimality are introduced to the system, causing DA performance to degrade. The manually tuned ROI and λ counter these suboptimality as much as possible (larger inflation factor and tighter localization distances are used). In these scenarios, the suboptimality due to mis-specified L still appear in a similar fashion as in the first scenario, indicating that the behavior described above is robust.

b. Test of multiscale observation approach

In this section the MSO method is tested with the serial EnSRF using the Lorenz-96 model. Since the model does not feature multiscale dynamics (errors grow and instantaneously project to all wavenumbers), multiscale localization is not applied. The wavenumber space is evenly divided into N_s bands for each scale component, and the cases using $N_s = 1, 2, 3, 4, 5,$ and 7 are tested. Note that $N_s = 1$ is the original single-scale serial EnSRF. Figure 4 shows the spectra of prior and posterior error and ensemble spread for cases with increasing N_s. Note that results from $N_s = 4$ and 5 do not visually differ much from $N_s = 7$, thus not shown. For

$N_s = 2$, the observations are decomposed into two scale components, the larger-scale observation errors are inflated to $\sigma \approx 1.3$ and the smaller-scale errors are deflated to $\sigma \approx 0.4$. Results suggest that separating observations into just two scale components already reduces most of the suboptimality due to mis-specified L. As N_s

TABLE 1. Best performance using the EnSRF found for the Lorenz-96 model experiments under three scenarios: benchmark ($N = 40$, $F = 8$), more sampling errors ($N = 20$, $F = 8$), and imperfect model ($N = 40$, $F = 7.5$). For each scenario, the specified L ranging from 0 to 10 grid points are tested. For each case, the posterior RMSE and CR are shown along with the manually tuned observation error standard deviation σ, localization ROI, and prior inflation factor λ.

| Scenarios	RMSE	CR	σ	ROI	λ
$N = 40, F = 8$:					
$L = 0$	0.360	1.07	1	50	1.06
$L = 1$	0.197	1.03	0.6	55	1.08
$L = 3$	0.159	1.02	0.8	55	1.04
$L = 5$ (benchmark)	0.158	1.01	1	55	1.04
$L = 7$	0.158	1.02	1.2	60	1.04
$L = 10$	0.159	0.99	1.4	60	1.04
$N = 20, F = 8$:					
$L = 0$	0.370	1.09	1	35	1.12
$L = 1$	0.208	1.02	0.6	40	1.14
$L = 3$	0.164	1.04	0.8	40	1.08
$L = 5$	0.162	0.99	1	40	1.06
$L = 7$	0.161	1.01	1.2	40	1.06
$L = 10$	0.162	0.98	1.4	40	1.06
$N = 40, F = 7.5$:					
$L = 0$	0.448	1.05	1	40	1.20
$L = 1$	0.320	0.92	0.6	45	1.40
$L = 3$	0.280	0.94	0.8	45	1.40
$L = 5$	0.280	0.92	1	45	1.40
$L = 7$	0.280	0.94	1.2	45	1.40
$L = 10$	0.281	0.91	1.4	45	1.40
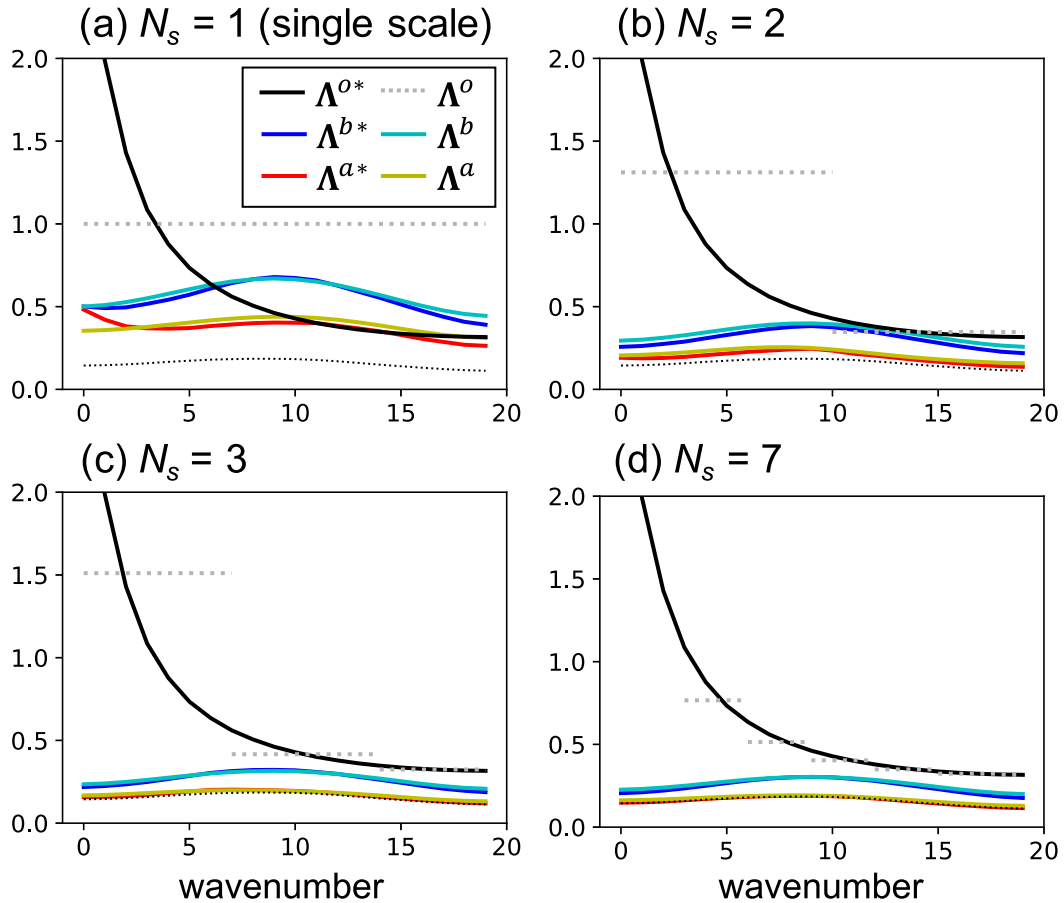
FIG. 4. Error spectrum averaged over all the analysis cycles for the test cases using serial EnSRF in the $N = 40$ and $F = 8$ scenario. The cases using the MSO method are shown where observations are decomposed into (a)–(d) $N_s = 1$ (single scale), 2, 3, and 7 scale components. Each panel shows the true observation error ($\mathbf{\Lambda}^{o*}$), specified observation error ($\mathbf{\Lambda}^{o}$), prior error ($\mathbf{\Lambda}^{b*}$) and spread ($\mathbf{\Lambda}^{b}$), and posterior error ($\mathbf{\Lambda}^{a*}$) and spread ($\mathbf{\Lambda}^{a}$). The spectrum plotted here corresponds to the square root of the diagonals of $\mathbf{\Lambda}$ matrices. The posterior error spectrum from running EnSRF with correctly specified L is shown as black dotted lines on each panel for reference.

increases, the filter solution approaches the benchmark EnSRF solution with correctly specified L. More quantitatively, the overall posterior RMSE and CR are summarized in the first section of Table 2. The MSO method is also tested in the other two scenarios with additional suboptimalities, and results are also summarized in Table 2. As N_s increases, the overall posterior RMSE also decreases for these scenarios despite the fact that performance is worse than the first scenario due to additional error sources. One exception is for the second scenario with more sampling errors, where the improvement in performance stops at $N_s = 5$ and posterior RMSE increases back up for $N_s = 7$. The accumulation of sampling noises may be the culprit here. When the serial EnSRF assimilates the sth observation scale component, error covariances between the observation at scale s and the model state need to be estimated by the ensemble. Such cross-scale error covariance is inevitably

contaminated with sampling noises not fully removed by a distance-based localization function. This may also explain why the $N_s = 7$ case still have larger posterior RMSE compared to the benchmark EnSRF case, since no cross-scale covariance is needed in the EnSRF.

4. Two-layer quasigeostrophic model experiments

The multiscale approach is further tested in a higher-dimensional setting using the two-layer quasigeostrophic (QG) model (Smith et al. 2002). The model describes two-dimensional vorticity dynamics in geophysical flows with baroclinic instabilities. Although much simpler than actual weather prediction models, the QG model describes flows spanning many scales, and therefore is useful for this study. Horizontal model grid dimension is 128×128 with doubly periodic boundary conditions. Let x and y be the coordinates in the zonal and meridional

TABLE 2. Test of serial EnSRF with the MSO method for the Lorenz-96 model experiments under different scenarios. For each scenario, the EnSRF solution with correctly specified $L = 5$ provides the "best performance" for reference, and the serial EnSRF is run with the observations decomposed into $N_s = 1, 2, \ldots, 7$ scale components. Note that $N_s = 1$ is the single scale serial EnSRF without MSO. For each case, the posterior RMSE and CR are shown along the manually tuned localization ROI and prior inflation factor λ.

Scenarios	RMSE	CR	ROI	λ
$N = 40, F = 8$:				
EnSRF ($L = 5$)	0.158	1.01	55	1.04
Serial EnSRF ($L = 0$):				
$N_s = 1$ (single scale)	0.370	1.05	50	1.06
$N_s = 2$	0.200	1.09	50	1.10
$N_s = 3$	0.171	1.03	55	1.06
$N_s = 4$	0.165	1.03	55	1.06
$N_s = 5$	0.163	1.04	55	1.06
$N_s = 7$	0.162	1.05	55	1.06
$N = 20, F = 8$:				
EnSRF ($L = 5$)	0.162	0.99	40	1.06
Serial EnSRF ($L = 0$):				
$N_s = 1$ (single scale)	0.372	1.09	35	1.12
$N_s = 2$	0.217	1.13	35	1.22
$N_s = 3$	0.180	1.06	35	1.12
$N_s = 4$	0.174	1.07	35	1.12
$N_s = 5$	0.173	1.07	35	1.12
$N_s = 7$	0.176	1.10	35	1.14
$N = 40, F = 7.5$:				
EnSRF ($L = 5$)	0.280	0.92	45	1.40
Serial EnSRF ($L = 0$):				
$N_s = 1$ (single scale)	0.449	1.05	40	1.20
$N_s = 2$	0.319	1.01	45	1.40
$N_s = 3$	0.294	0.95	45	1.40
$N_s = 4$	0.286	0.94	45	1.40
$N_s = 5$	0.284	0.94	45	1.40
$N_s = 7$	0.283	0.94	45	1.40

directions, and subscripts 1 and 2 indicate the top and bottom layer, respectively. Background streamfunction is $\overline{\psi}_1 = -Uy$ and $\overline{\psi}_2 = Uy$, where U is the mean flow. The governing equations are

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

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

$$(18)$$

where $q_i = \nabla^2\psi_i - (-1)^i(k_d^2/2)(\psi_2 - \psi_1)$, $i = 1, 2$, are the perturbation potential vorticity, ψ is the perturbation streamfunction, and J denotes the Jacobian term. Model parameters include the Rossby deformation wavenumber (k_d), the meridional gradient of the Coriolis force (β), and the Ekman bottom drag strength (b). The scale defined by k_d^{-1} features energy injection due to baroclinic instability, while β and b define the Rhines scale at which the inverse

cascade of kinetic energy halts. This model configuration follows the control experiment in Ying et al. (2018): $k_d = 20$, $\beta = 16$, $U = 0.2$, and $b = 0.5$. The potential temperature θ is chosen as the model state variable for assimilation in this study, because its spectrum resembles the real atmospheric temperature and kinetic energy spectra. Let $k = \sqrt{k_x^2 + k_y^2}$ be the horizontal total wavenumber, so the maximum wavenumber for the model grid is $k_{max} = 63$. The θ field is converted from streamfunction in spectral space,

$$\hat{\theta}(k) = -k\,\hat{\psi}(k). \tag{19}$$

A nature run is first obtained to provide the verifying truth. Figure 5a shows a snapshot of the true θ field. Synthetic observations are then generated by adding randomly drawn errors to the truth field. A uniform observation grid with horizontal spacing of 3 model grid points is used. The θ variable is directly observed, but only available at the top model layer. Observation errors are drawn from a first-order autoregressive spatial random process. Figures 5b and 5c show snapshots of observation error fields with standard deviation $\sigma^* = 3$ and spatial correlation $r^* = 0$ and 0.6, respectively.

Only the serial EnSRF is implemented for the QG model (the EnSRF is not computationally feasible) and tested with the MSO and MSL methods. To determine the ROIs specified in the MSL method, a set of sensitivity experiments are performed to determine the best ROIs for different scales. The QG model states are arbitrarily decomposed into three scale components: large ($k = 0$–5), medium ($k = 6$–12), and small ($k = 13$–63) scales. Figure 6 compares the spectra of posterior error and ensemble spread using ROIs ranging from 8 to 24 grid points for the $r^* = 0$ and 0.6 cases. The reference spectrum has a -3 slope, which corresponds to the forward enstrophy cascade in QG dynamics. The uncorrelated observation error spectrum has a $+1$ slope (white noise), and the correlated one has increased spectral error variances at large scale and decreased at small scale. Note that the spectra are plotted on a log-log scale. For $r^* = 0$ (Fig. 6a), the posterior error spectral variances are concentrated at the medium scale, since the large scale is well constrained by the observations and the small scale has low reference error variances to begin with. Since the observing network is at lower spatial resolution, the small scale lacks observation information and the posterior errors are mostly saturated (close to reference level). The best ROI is 24 grid points for the large scale and 12 grid points for the small scale, consistent with findings from Ying et al. (2018). For $r^* = 0.6$ (Fig. 6b), the posterior error spectral variances are more concentrated in the large scale. Underspecification of
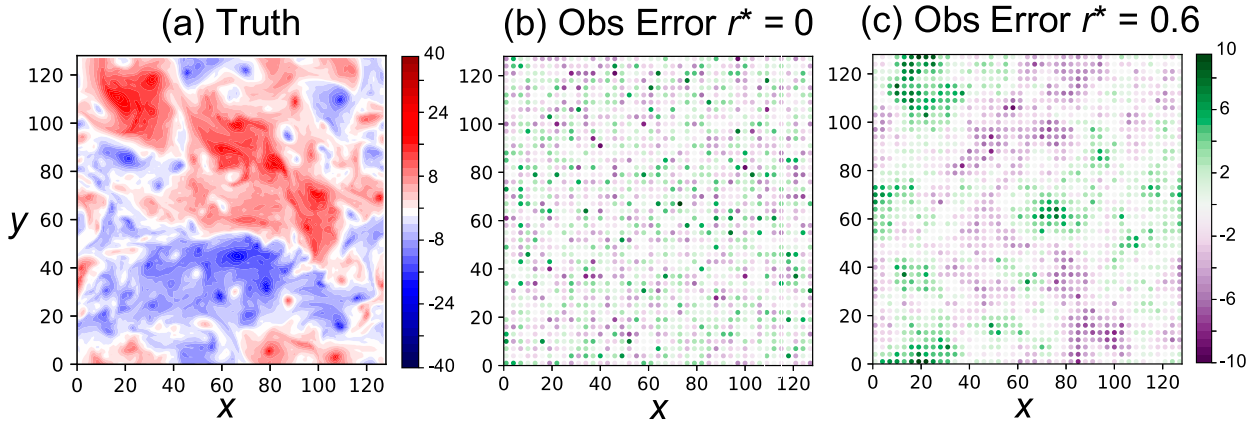
FIG. 5. Snapshots of (a) the QG model θ field from the truth simulation, and its corresponding observation error fields for the (b) $r* = 0$ (uncorrelated errors) and (c) $r* = 0.6$ cases. The observation locations are indicated by the dots.

observation error spectral variances at the large scale causes the posterior ensemble to be underdispersive. This case seems to display a stronger scale dependency in best ROIs. The favorable ROIs actually exceed the range shown in Fig. 6b. The best ROI is 32 grid points for the large scale and 4 grid points for the small scale.

The three scale components defined above are used for both the MSO and MSL methods. The ROIs are set to 24, 16, and 8 grid points for the three components in the MSL method. For the MSO method, the spectral observation error variance adjustment factors λ_s^o are set to 2.4, 1.5, and 0.8 for the three scale components according to the mismatches between the specified and true observation error spectral variances. An adaptive prior inflation algorithm is implemented based on innovation statistics (Hollingsworth and Lönnberg 1986;

Desroziers et al. 2005) to avoid manual tuning. At each analysis cycle, a maximum likelihood estimate gives the prior inflation factor:

$$\lambda = \sqrt{\frac{\sum_{j=1}^{N_y} \text{var}(y_j^b)}{\sqrt{\sum_{j=1}^{N_y} (y_j^o - y_j^b)^2 - \sum_{j=1}^{N_y} \sigma_{o,j}^2}}}. \quad (20)$$

A relaxation to prior perturbation method (Zhang et al. 2004) is also applied with a fixed coefficient $\alpha = 0.5$ (chosen by trial and error) to provide some posterior inflation, since prior and posterior inflation may ameliorate different kinds of suboptimalities during DA (El Gharamti et al. 2019).
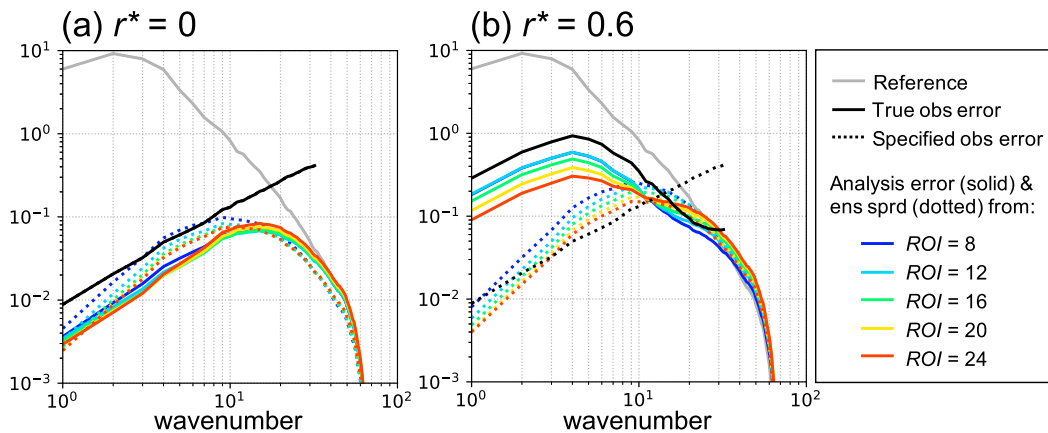

FIG. 6. Error spectra averaged over 200 analysis cycles compared for the (a) $r* = 0$ and (b) $r* = 0.6$ cases. The reference spectrum (gray line) is error from a free ensemble forecast without data assimilation. The true and specified observation error spectra are shown as black solid and dotted lines, respectively. The posterior error spectra from serial EnSRF using ROI = 8–24 grid points are shown as solid lines color coded from blue to red, their corresponding ensemble spread spectra are shown as dotted lines with the same colors.
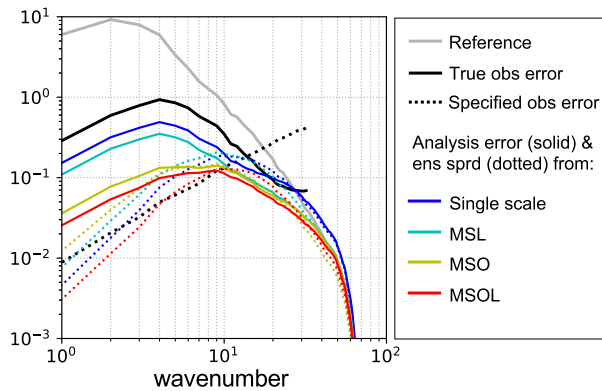
FIG. 7. Posterior error spectra averaged over 200 analysis cycles compared for cases with $r* = 0.6$ and using the original single scale serial EnSRF (blue), and serial EnSRF with the MSL (cyan), MSO (yellow), and MSOL (red) methods. Their corresponding ensemble spread spectra are shown as dotted lines with the same colors. The reference spectrum (gray line) is error from a free ensemble forecast without data assimilation. The true and specified observation error spectra are shown as black solid and dotted lines, respectively.

TABLE 3. Test of the multiscale approach to assimilate observations with $r* = 0.6$ using the QG model. Two scenarios using perfect and imperfect models are tested. In each scenario, results from the single scale serial EnSRF with ROI = 16, and the serial EnSRF with the MSL, MSO, and MSOL methods are compared. The posterior θ field is decomposed into large, medium, and small scales, and averaged posterior RMSE and CR are computed for each scale.

Scenarios	Large scale		Medium scale		Small scale	
	RMSE	CR	RMSE	CR	RMSE	CR
Perfect model:						
Single scale	1.853	0.13	1.564	0.63	2.547	1.11
MSL	1.331	0.27	1.140	0.97	1.694	1.40
MSO	0.498	0.75	0.789	1.03	1.617	0.92
MSOL	0.370	0.43	0.678	0.99	1.427	1.24
Imperfect model:						
Single scale	2.522	0.17	1.912	0.70	2.479	1.03
MSL	1.467	0.35	1.181	1.09	1.720	1.21
MSO	0.646	0.90	0.930	1.12	1.774	0.78
MSOL	0.517	0.45	0.819	1.07	1.557	1.17

Cycling DA experiments are conducted using $N = 20$ ensemble members and assimilating the observations with correlated errors ($r* = 0.6$) every 0.05 time units for 200 cycles. The cycling period is long enough so that the posterior errors have time-lagged correlations of no more than 0.2, and thus can be treated as quasi-independent samples for performance metrics. Tests are first conducted in a perfect model scenario. Four experiments are compared: the original single-scale serial EnSRF with ROI = 16, and the serial EnSRF with the MSO, MSL, and MSOL methods.

Figure 7 compares the averaged spectra of posterior error and ensemble spread from each method. More quantitatively, the first section of Table 3 summarizes the averaged posterior RMSE and CR for the three scale components. All comparisons of RMSE between different methods have passed the student t test at $p < 0.01$ significance level. The underspecified observation error spectral variances at the large scale causes the single-scale filter solution to be severely underdispersive (CR = 0.13) at the large scale. The MSL method lowers the posterior error spectral variances for all wavenumbers, but the large-scale component is still underdispersive (CR = 0.27). Spatially correlated observation errors can also be considered as locally biased large-scale observation errors (Wilks 1995, section 5.2.3). Assimilating these observations assuming uncorrelated errors will cause the posterior ensemble to overfit the observations and also become locally biased (hence the low CR). The MSO method significantly improves the large-scale component of the posterior ensemble, reducing RMSE and bringing the CR much closer to 1.

The posterior ensemble no longer overfits the observations because the large-scale observation error variances are inflated. The MSOL method produces the lowest posterior RMSE for all three scale components. The MSL method seems to cause the large-scale posterior ensemble to be underdispersive and small-scale ensemble to be overdispersive, which is adjusted during the model forecast steps to produce an overall consistent prior ensemble (not shown). The interplay between scale-dependent error growth during forecast and error reduction during DA calls for further investigation in more complex models.

The experiments are repeated under another scenario where model parameter error exists in the forecast steps. The forecast model has $k_d = 10$ such that the peak baroclinic instability generation is shifted toward larger scales, causing smaller scales ($k > 10$) to be less energetic than the true model ($k_d = 20$). The second section in Table 3 shows that the presence of model error increases the overall posterior RMSEs. The time averaged adaptive prior inflation factor also increases from 0.98 to 1.07. However, the MSOL method still improves the performance over the single-scale method, and the impact from the MSO and MSL methods are similar to those found in the perfect model scenario. The only difference is that the MSL method improves the small-scale component more than the MSO method for the imperfect model scenario.

5. Discussion

The current study uses an idealized observing network to demonstrate the suboptimality in ensemble filtering

performance due to mis-specified observation error statistics and test the proposed MSO method as a remedy. The idealized observing network is uniform in space and time, its observation error is modeled as a first-order autoregressive process with constant variance σ^2 and correlation length scale L in space and time. Such idealized observations are hard to find in real application scenarios. Typically, observations are not distributed uniformly in space and time, and the statistics of representation errors also vary in space and time due to the flow dependency in model forecast errors (Janjić and Cohn 2006; Waller et al. 2014). The true observation error statistics are ultimately unknown, which is even more challenging. In this section, potential approaches for extending the MSO method to more realistic scenarios will be discussed.

For a nonuniform observing network, scale decomposition cannot be performed with a pair of fast Fourier transforms. A spatial averaging approach similar to superobbing can be used instead. Observations are averaged within each grid box of a uniform grid at a certain resolution. The smoothed observation defined on the low-resolution grid represents its large-scale component. Due to the spatial inhomogeneity of observations, some grid boxes will have too few observations to provide robust statistics for the large scale, therefore shall be marked as missing values and discarded by the filter. To find the smaller-scale component, the large-scale component can be subtracted from the original observations and spatial averaging can be applied again with higher-resolution grid corresponding to the smaller scale. This procedure will be repeated for the total number of scale components, resulting in a multigrid representation of the original observations, which will then be used as observation scale components in the MSO method.

The MSO method uses an adjustment factor λ_s^o to change the observation error variance for the sth scale component. In this study, the correct factors are given by

$$\lambda_s^o = \sqrt{\Lambda_s^{o*}/\Lambda_s^o}, \qquad (21)$$

where Λ_s^{o*} is the true observation error spectral variance and Λ_s^o is the originally specified one. In real application scenarios, Λ_s^{o*} is unknown and has to be estimated through observation-space innovation statistics (e.g., Hollingsworth and Lönnberg 1986; Desroziers et al. 2005; Bowler 2017; Howes et al. 2017). For example, the innovation statistics for the sth scale component can be written as

$$(\lambda_s^o)^2 \Lambda_s^o + (\lambda_s^b)^2 \Lambda_s^b = E[(y_s^o - \overline{y}_s^b)^2], \qquad (22)$$

where y_s^o is the observation and $\overline{y}_s^b$ is the observation prior ensemble mean, λ_s^o and λ_s^b are adjustment factors for the specified observation error and prior error spectral variances. This study only considers the case where $\Lambda_s^o \neq \Lambda_s^{o*}$ due to mischaracterization of observation error correlation length scale L. Therefore, assume $\Lambda_s^b = \Lambda_s^{b*}$, $\lambda_s^b = 1$, and (22) will provide a maximum likelihood estimate for λ_s^o. In a more general case, the prior error variances may also be mis-specified due to other suboptimalities (e.g., sampling noises), then λ_s^b can be estimated as an adaptive inflation factor to adjust the prior error variances. The method described in Li et al. (2009) can provide simultaneous estimate for both adjustment factors λ_s^o and λ_s^b using the innovation statistics. Since spatial correlation in representation errors may be actually caused by the prior errors, adjusting the prior inflation factors λ_s^b to account for the effect of these spatial correlations may be an alternative method. Adaptive inflation methods using Bayesian formulations (Anderson 2009; Miyoshi 2011; El Gharamti 2018) can be borrowed to estimate a spatially and temporally varying λ_s^o, so that the MSO method can be applied to nonuniform observing networks.

Another concern is the increase in computational cost for the MSO method. As described in this paper, the MSO method increases the number of observations by a factor of N_s, which is the number of scale components. Using a multigrid representation, however, would only increase the number of observations by a factor of no more than 2, since larger-scale components would have fewer number of observations. Different model variables may have different spectral error variance distributions. For example, wind and temperature tend to have errors in the larger scales compared to precipitation and vertical motion. Complex scale interactions among state variables may also occur. Determining the cutoff wavenumbers for each scale component may be a nontrivial task for real models. Adaptive algorithms will be needed to turn on the MSO method when mismatches in spectral error variances are detected and scale decomposition is performed to optimally adjust the spectral error variances. The cost-benefit analysis of spending computational resources on ensemble size and model resolution has been performed in previous studies (e.g., Lei and Whitaker 2017). Similar analysis can also be performed for the number of scale components for real application.

6. Conclusions

In this study, the suboptimality in ensemble square root filters (EnSRF) due to mischaracterizing the observation error spatial correlations is evaluated using the

Lorenz-96 model. The true observation errors have spatial correlations that are neglected by the serial EnSRF, resulting in suboptimal filter performance. The mismatches between specified and true observation error spectral variances cause the analysis ensemble to be overdispersive at small scales and underdispersive at large scales. As the spectral distribution is vastly different between correlated and uncorrelated observation errors, a scale-dependent adjustment of observation error variances is needed to reduce such suboptimality, which motivates the multiscale approach. The proposed multiscale observation (MSO) method decomposes observations into multiple scale components using a spectral bandpass filter and assimilates each component with adjusted observation error variances to update the model state iteratively. Lorenz-96 model experiments indicate that the solution from the serial EnSRF with the MSO method approaches the solution from the EnSRF with correctly specified observation error correlations as the number of scales increases. The MSO method is further tested with a more complicated two-layer quasigeostrophic (QG) model. The serial EnSRF is implemented with the MSO method to assimilate observations with spatially correlated errors. The MSO method is capable of reducing the suboptimality due to mischaracterization of observation error correlations in both perfect and imperfect model scenarios. The multiscale localization (MSL) method is also applied to allow the use of different localization radii when updating the model state at different scales. The QG model state spans a range of spatial scales, therefore its large-scale component favors a longer localization radius, and small-scale favors a shorter radius. The MSL method improves filter performance compared to the single-scale serial EnSRF with a fixed localization radius. Combining the MSO and MSL, the MSOL multiscale approach gives the best filter performance in terms of analysis error and consistency between error and ensemble spread. Although the MSO method is only tested in very idealized settings in this study (uniform observing network, simple observation error model, and known true statistics), the underlying concept of adjusting the observation error spectral variances is potentially applicable to realistic scenarios. Using innovation statistics to estimate a spatially and temporally varying observation error adjustment factor for each scale component, the MSO method can be potentially applied to nonuniform observing networks with inhomogeneous error statistics. Future studies will further investigate the cost and benefit of the MSO method in realistic prediction systems.

REFERENCES

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

——, 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642, https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2.

——, 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus*, **61A**, 72–83, https://doi.org/10.1111/j.1600-0870.2007.00361.x.

——, 2010: A non-Gaussian ensemble filter update for data assimilation. *Mon. Wea. Rev.*, **138**, 4186–4198, https://doi.org/10.1175/2010MWR3253.1.

——, 2016: Assimilating observations with spatially and temporally correlated errors in a global atmospheric model. *10th CAWCR Workshop*, Melbourne, Australia, CSIRO–Australian Bureau of Meteorology, Geophysical Research Abstracts, Vol. 18, EGU2016-3104.

——, and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758, https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.

Bedard, J., and M. Buehner, 2020: A practical assimilation approach to extract smaller-scale information from observations with spatially correlated errors: An idealized study. *Quart. J. Roy. Meteor. Soc.*, **146**, 468–482, https://doi.org/10.1002/qj.3687.

Berger, H., and M. Forsythe, 2004: Satellite wind superobbing. Met Office Forecasting Research Tech. Rep. 451, 33 pp.

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

Bormann, N., and P. Bauer, 2010: Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I: Methods and application to ATOVS data. *Quart. J. Roy. Meteor. Soc.*, **136**, 1036–1050, https://doi.org/10.1002/qj.616.

——, M. Bonavita, R. Dragani, R. Eresmaa, M. Matricardi, and A. McNally, 2016: Enhancing the impact of IASI observations through an updated observation-error covariance matrix. *Quart. J. Roy. Meteor. Soc.*, **142**, 1767–1780, https://doi.org/10.1002/qj.2774.

Bowler, N. E., 2017: On the diagnosis of model error statistics using weak constraint data assimilation. *Quart. J. Roy. Meteor. Soc.*, **143**, 1916–1928, https://doi.org/10.1002/qj.3051.

Buehner, M., and A. Shlyaeva, 2015: Scale-dependent background error covariance localization. *Tellus*, **67A**, 28027, https://doi.org/10.3402/tellusa.v67.28027.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126<1719: ASITEK>2.0.CO;2.

Campbell, W. F., E. A. Satterfield, B. Ruston, and N. L. Baker, 2017: Accounting for correlated observation error in a dual-formulation 4D variational data assimilation system. *Mon. Wea. Rev.*, **145**, 1019–1032, https://doi.org/10.1175/MWR-D-16-0240.1.

Collard, A. D., 2004: On the choice of observation errors for the assimilation of AIRS brightness temperatures: A theoretical study. ECMWF Tech. Memo. AC/90.

Courtier, P., and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1783–1807, https://doi.org/10.1002/qj.49712455002.

Dando, M. L., A. J. Thorpe, and J. R. Eyre, 2007: The optimal density of atmospheric sounder observations in the Met Office NWP system. *Quart. J. Roy. Meteor. Soc.*, **133**, 1933–1943, https://doi.org/10.1002/qj.175.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Quart. J. Roy. Meteor. Soc.*, **131**, 3385–3396, https://doi.org/10.1256/qj.05.108.

El Gharamti, M., 2018: Enhanced adaptive inflation algorithm for ensemble filters. *Mon. Wea. Rev.*, **146**, 623–640, https://doi.org/10.1175/MWR-D-17-0187.1.

——, K. Raeder, J. Anderson, and X. Wang, 2019: Comparing adaptive prior and posterior inflation for ensemble filters using an atmospheric general circulation model. *Mon. Wea. Rev.*, **147**, 2535–2553, https://doi.org/10.1175/MWR-D-18-0389.1.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, https://doi.org/10.1029/94JC00572.

Fowler, A. M., S. L. Dance, and J. A. Waller, 2018: On the interaction of observation and prior error correlations in data assimilation. *Quart. J. Roy. Meteor. Soc.*, **144**, 48–62, https://doi.org/10.1002/qj.3183.

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.

Geer, A. J., and Coauthors, 2017: The growing impact of satellite observations sensitive to humidity, cloud and precipitation. *Quart. J. Roy. Meteor. Soc.*, **143**, 3189–3206, https://doi.org/10.1002/qj.3172.

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.

Hilton, F., A. Collard, V. Guidard, R. Randriamampianina, and M. Schwaerz, 2009: Assimilation of IASI radiances at European NWP centres. *Proc. Workshop on the Assimilation of IASI Data in NWP*, Reading, United Kingdom, ECMWF, 73 pp., https://www.ecmwf.int/node/15331.

Hodyss, D., and E. Satterfield, 2017: The treatment, estimation, and issues with representation error modelling. *Data Assimilation for Atmospheric, Oceanic, and Hydrologic Applications*, S. K. Park and L. Xu, Eds., Vol. III, Springer, 177–194.

Hoffman, R. N., 2018: The effect of thinning and superobservations in a simple one-dimensional data analysis with mischaracterized error. *Mon. Wea. Rev.*, **146**, 1181–1195, https://doi.org/10.1175/MWR-D-17-0363.1.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111–136, https://doi.org/10.1111/j.1600-0870.1986.tb00460.x.

Houtekamer, P. L., and H. L. Mitchell, 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.

Howes, K. E., A. M. Fowler, and A. S. Lawless, 2017: Accounting for model error in strong-constraint 4DVar data assimilation. *Quart. J. Roy. Meteor. Soc.*, **143**, 1227–1240, https://doi.org/10.1002/qj.2996.

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.

——, and Coauthors, 2018: On the representation error in data assimilation. *Quart. J. Roy. Meteor. Soc.*, **144**, 1257–1278, https://doi.org/10.1002/qj.3130.

Lei, L., and J. S. Whitaker, 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.

Li, H., E. Kalnay, and T. Miyoshi, 2009: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 523–533, https://doi.org/10.1002/qj.371.

Liu, Z. Q., and F. Rabier, 2002: The interaction between model resolution observation resolution and observation density in data assimilation: A one-dimensional study. *Quart. J. Roy. Meteor. Soc.*, **128**, 1367–1386, https://doi.org/10.1256/003590002320373337.

——, and ——, 2003: The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Quart. J. Roy. Meteor. Soc.*, **129**, 3013–3035, https://doi.org/10.1256/qj.02.170.

Lorenz, E. N., 1996: Predictability: A problem partly solved. *Proc. Seminar on Predictability*, Reading, United Kingdom, ECMWF, 1–18, https://www.ecmwf.int/node/10829.

Migliorini, S., and B. Candy, 2019: All-sky satellite data assimilation of microwave temperature sounding channels at the Met Office. *Quart. J. Roy. Meteor. Soc.*, **145**, 867–883, https://doi.org/10.1002/qj.3470.

Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535, https://doi.org/10.1175/2010MWR3570.1.

——, and K. Kondo, 2013: A multi-scale localization approach to an ensemble Kalan filter. *SOLA*, **9**, 170–173, https://doi.org/10.2151/SOLA.2013-038.

——, E. Kalnay, and H. Li, 2013: Estimating and including observation-error correlations in data assimilation. *Inverse Probl. Sci. Eng.*, **21**, 387–398, https://doi.org/10.1080/17415977.2012.712527.

Rainwater, S., C. Bishop, and W. F. Campbell, 2015: The benefits of correlated observation errors for small scales. *Quart. J. Roy. Meteor. Soc.*, **141**, 3439–3445, https://doi.org/10.1002/qj.2582.

Ruggiero, G. A., E. Cosme, J. Brankart, J. Le Sommer, and C. Ubelmann, 2016: An efficient way to account for observation error correlations in the assimilation of data from the future SWOT high-resolution altimeter mission. *J. Atmos. Oceanic Technol.*, **33**, 2755–2768, https://doi.org/10.1175/JTECH-D-16-0048.1.

Simonin, D., J. A. Waller, S. P. Ballard, S. L. Dance, and N. K. Nichols, 2019: A pragmatic strategy for implementing spatially

correlated observation errors in an operational system: An application to Doppler radial winds. *Quart. J. Roy. Meteor. Soc.*, **145**, 2772–2790, https://doi.org/10.1002/qj.3592.

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.

Stewart, L. M., 2010: Correlated observation errors in data assimilation. Ph.D. thesis, University of Reading, 195 pp., https://www.reading.ac.uk/web/files/maths/StewartPhD2010.pdf.

——, S. L. Dance, and N. K. Nichols, 2008: Correlated observation errors in data assimilation. *Int. J. Numer. Methods Fluids*, **56**, 1521–1527, https://doi.org/10.1002/fld.1636.

——, ——, and ——, 2013: Data assimilation with correlated observation errors: Experiments with a 1-D shallow water model. *Tellus*, **65A**, 19546, https://doi.org/10.3402/tellusa.v65i0.19546.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490, https://doi.org/10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.

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.

Waller, J. A., S. L. Dance, A. S. Lawless, N. K. Nichols, and J. R. Eyre, 2014: Representativity error for temperature and humidity using the Met Office high-resolution model. *Quart. J. Roy. Meteor. Soc.*, **140**, 1189–1197, https://doi.org/10.1002/qj.2207.

——, D. Simonin, S. L. Dance, N. K. Nichols, and S. P. Ballard, 2016: Diagnosing observation error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics.

Mon. Wea. Rev., **144**, 3533–3551, https://doi.org/10.1175/MWR-D-15-0340.1.

Weston, P. P., W. Bell, and J. R. Eyre, 2014: Accounting for correlated error in the assimilation of high resolution sounder data. *Quart. J. Roy. Meteor. Soc.*, **140**, 2420–2429, https://doi.org/10.1002/qj.2306.

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.

——, ——, X. Wei, Y. Song, and Z. Toth, 2008: Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Wea. Rev.*, **136**, 463–482, https://doi.org/10.1175/2007MWR2018.1.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. International Geophysics Series, Vol. 59, Elsevier, 467 pp.

Ying, Y., F. Zhang, and J. L. Anderson, 2018: On the selection of localization radius in ensemble filtering for multiscale quasigeostrophic dynamics. *Mon. Wea. Rev.*, **146**, 543–560, https://doi.org/10.1175/MWR-D-17-0336.1.

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

——, 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.