UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SCALE-DEPENDENT INFLATION FOR MULTISCALE ENSEMBLE BASED DATA ASSIMILATION

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

Naicheng Xu
Norman, Oklahoma
2023

SCALE-DEPENDENT INFLATION FOR MULTISCALE ENSEMBLE BASED DATA ASSIMILATION

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Xuguang Wang (Chair)

Dr. Steven M. Cavallo

Dr. Yongming Wang

Acknowledgement

I would like to extend my deepest gratitude to my advisor, Dr. Xuguang Wang and my committee member, Dr. Yongming Wang, for their guidance and support during my research for this thesis. I would also like to thank my committee member, Dr. Steven Cavallo, for his constructive feedback and valuable suggestions to improve my study further. I'm immensely grateful to my friend, Yu-Shin Kim, for bringing my poster to the ISDA 2023 when I encountered a visa issue and could not travel internationally. In addition, I want to thank all Multiscale data Assimilation and Predictability (MAP) group members for the discussion and assistance that inspired me when I encountered trouble in research. This study was supported by NOAA (NA19OAR0220154). The computational resources in this study are provided by the University of Oklahoma Supercomputing Center for Education and Research.

TABLE OF CONTENTS

List of Tables

List of Figures

Abstract

The degree of the background ensemble deficiency, often manifested as ensemble underdispersion, can vary at different scales in the ensemble-based data assimilation. This study develops the new scale-dependent inflation (SDI) methods based on two scale-unaware inflation approaches, RTPS (Relaxation To Prior Spread) and SE (observation-dependent Sampling Error inflation). In the new scale-dependent RTPS inflation (RTPS-SDI), the posterior ensemble spread is relaxed toward the prior ensemble spread at each scale separately. In the Scale-dependent SE inflation (SE-SDI), mathematical derivation is performed so that posterior ensemble variance is individually adjusted toward the mean square error of ensemble analysis mean at each scale. The impact of RTPS-SDI and SE-SDI are examined and evaluated by implementing both approaches within the Multiscale Local Gain Form Ensemble Transform Kalman Filter (MLGETKF).

Four continuously cycled MLGETKF experiments are performed with the four inflation methods using a two-layer surface quasi-geostrophic turbulence model. During the DA cycling, RTPS-SDI and SE-SDI outperform RTPS and SE, respectively, in the reduction of analysis errors and the enhancement of ensemble spread nearly at all scales and all cycles. In addition, the improvements in RTPS-SDI over RTPS are greater than those of SE-SDI over SE. These improvements in both SDI methods are associated with their larger inflation at all scales, especially at larger scales, compared to their scale-unaware counterparts. In the subsequent forecast, both SDI methods show statistically significantly better forecast performance than their scale-unaware inflation experiments. RTPS-SDI is more accurate than RTPS for all scales in 1-4 days lead time. SE-SDI is more accurate than SE at all scales for 3-6 days lead time during the early cycles and shows a smaller forecast error with significance for 2-3 days.

1. Introduction

Ensemble-based data assimilation (DA) has been widely used in research and operational centers (Gasperoni et al. 2022; Wang et al. 2021; Johnson et al. 2015) for the ability of the use of flow-dependent background error covariances (BECs). In contrast to the pure variational DA approach with static BECs, many studies have documented the benefits of flow-dependent BECs in global and regional applications (Wang et al. 2007a, 2008a,b, 2013; Wang 2011; Buehner et al. 2013, 2015; Clayton et al. 2013; Gustafsson et al. 2014; Wang and Lei 2014; Lorenc et al. 2015; Kleist and Ide 2015a,b; Kutty and Wang 2015; Jones et al. 2023; Lu et al. 2021; Wang et al. 2021). However, the quality of flow-dependent BECs can be strongly degraded by the ensemble deficiency, often manifested as ensemble underdispersion. Such ensemble deficiency may be caused by the misrepresentation of model errors due to the poorly resolved physical processes and sampling errors due to the limited number of affordable ensemble members. Deficient representation of model errors and sampling errors can lead to the degradation of ensemble-based DA and, consequently, the reduction in forecast skill (Kelly et al., 2015; Gottwald and Majda. 2013).

Many studies have proposed a series of techniques to account for the ensemble deficiency. For instance, numerous studies adopted the spatial localization technique in ensemble-based DA to alleviate sampling errors by damping distant covariances to zero. The effectiveness of localization in removing noisy distant correlations has been demonstrated in many ensemble-based DA systems (e.g., Houtekamer and Mitchell 2001; Wang et al. 2008a, b; Bishop and Hodyss 2009). As current numerical weather prediction models can resolve the flows at a wide range of scales, the proper application of localization requires constraining sampling errors at all resolved scales. Thus, recent studies proposed multiscale localization methods to reduce sampling errors across multiple scales using ensemble-based DA systems. Zhang et al. (2009) designed a sequential approach by separately applying different localization radii in the assimilation of different groups of observations. Miyoshi and Kondo (2013) combined two different analysis increments by using different localization radii with the assimilation of the same observations. Recent studies started to use a simultaneous multiscale update by model space scale-dependent localization within the ensemble variational (EnVar,

Buehner 2012; Buehner and Shlyaeva 2015) and pure ensemble Kalman Filter (EnKF, Wang et al. 2021) frameworks. Particularly, Wang et al. (2021) developed a multiscale local gain ensemble Kalman filter (MLGETKF), which allows simultaneous multiscale updates for both the ensemble mean and perturbations. Compared to the scale unaware local gain ensemble Kalman filter (LGETKF, Bishop et al. 2017), MLGETKF shows a significant analysis improvement at large scale and improves the subsequent forecast performance. Techniques were also proposed to address the misrepresentation of model errors using stochastic physics schemes to increase ensemble spread. For example, Reynolds et al. (2008) utilized the stochastically perturbed parameterization tendencies (SPPT) scheme in their ensemble transform DA system to increase ensemble variances in the tropics. Fujita et al. (2007) and Meng and Zhang (2008) showed the effectiveness of multi-physics schemes in the improvement of ensemble spread and background estimates in their mesoscale DA experiments. Leutbecher et al. (2007) demonstrated that the stochastic kinetic energy backscatter (SKEB) scheme in their ensemble 4D-Var DA experiments increases ensemble spread and subsequently improves probabilistic forecasts skill in the tropics.

In addition, inflation methods have been used to treat ensemble deficiency due to sampling errors or/and model error mis-representation. Multiple inflation methods were employed to directly increase ensemble spread. The multiplicative inflation methods, such as relaxation-to-prior perturbations (RTPP, Zhang et al. 2004) and relaxation-to-prior spread (RTPS, Whitaker and Hamill. 2012), are commonly used in EnKF to increase the ensemble spread in the global (Terasaki et al., 2019; Otsuka et al., 2016; Sluka et al., 2016) and regional (Zeng et al., 2018; Kotsuki et al., 2017; Harnisch & Keil. 2015) applications. Hodyss et al. (2016) noted that the inflated EnKF ensemble posterior variance using RTPS and RTPP is inconsistent with a Bayesian posterior, in which posterior variance depends on prior variance and observations. Therefore, they proposed and examined an alternative multiplicative inflation (observation-dependent Sampling Error inflation, denoted as SE) method to account for sampling errors by inflating EnKF posterior variance toward the Bayesian posterior variance. It was found that SE outperforms RTPS in addressing the ensemble deficiency using a two-level primitive equation model.

Despite that Wang et al. (2021) proposed to account for sampling errors through scale-

dependent localization, their study still adopted the scale-unaware RTPS method to increase ensemble spread. This RTPS method as well as the aforementioned inflation methods commonly ignore the dependence of ensemble deficiency with respect to various scales and use the same inflation for all scales. Wang et al. (2021) discussed that the ensemble spread in MLGETKF is still underestimated at all scales and shows greater underestimation towards larger scales with the use of the single-scale inflation. Therefore, they suggested that the ensemble deficiency requires to be further addressed in a scale-dependent inflation fashion.

In this study, we propose a scale-dependent inflation (SDI) method to alleviate the ensemble deficiency and maintain the ensemble spread in a scale-dependent manner. Specifically, the SDI method separately increases ensemble spread at various scales based on each scale's ensemble deficiency. The SDI method is first implemented on top of RTPS, denoted as scale-dependent RTPS (RTPS-SDI). Compared to scale-unaware RTPS using a fixed inflation factor for all scales, scale-dependent RTPS separately relaxes posterior ensemble variance toward corresponding prior ensemble variance for each scale using the varied inflation factors. The SDI method is also derived and implemented for SE, denoted as scale-dependent SE (SE-SDI), aiming to separately adjust ensemble variance toward the corresponding analysis Mean Square Error (MSE) at each scale.

In this study, both SDI approaches are first introduced and described. These methods are implemented and examined within the MLGETKF framework (Wang et al. 2021) using a dynamical Surface Quasigeostrophic (SQG) model to address three scientific objectives. The first objective is to evaluate the SDI methods versus their scale-unaware counterparts in alleviating ensemble deficiency and improving analysis accuracy and forecast performance. The first objective is achieved by comparing scale-unaware inflation approaches and their corresponding scale-dependent counterparts. In the second objective, we assess the relative improvements of applying SDI to different single-scale inflation methods, i.e., RTPS and SE, in this study. The third objective is to explain the improvements of SDI approaches relative to scale-unaware approaches during different DA phases by understanding their temporal and spatial differences in inflation factors. For objective 3, we hypothesize that the SDI's inflation factors at different scales will demonstrate different behaviors and indicate scale-dependent ensemble deficiencies.

This paper is organized as follows. Section 2 briefly describes the MLGETKF algorithms and the formulas of the scale-unaware RTPS (SE) and their scale-dependent counterpart. Section 3 presents the model, DA configuration, and experiment design. Section 4 presents the results from the comparisons between the scale-unaware methods (RTPS and SE) and their scale-dependent counterparts in the analyses and forecasts. Section 5 concludes the paper and provides a discussion.

2 Methodology

This section first briefly introduces the MLGETKF algorithm. The following introduces the formulas of two scale-unaware inflation methods, RTPS and SE, and their scale-dependent counterparts. In this part, the variables in the regular font represent scalars, and the variables in the bold font represent vectors or matrixes.

2.1 Multiscale local gain form ensemble transform Kalman filter (MLGETKF)

In this study, the MLGETKF (Wang et al. 2021) is used as the DA framework to test the scale-dependent inflation approaches. MLGETKF is a multiscale data assimilation algorithm developed from the scale unaware LGETKF (Bishop et al. 2017). This MLGETKF method allows the simultaneous multiscale update by utilizing model space scale-dependent localization. The derivation details have been introduced by Wang et al. (2021). A brief introduction of MLGETKF is provided as follows.

The scale separation of ensemble covariances is accomplished by decomposing each background ensemble perturbation. The Fast Fourier transform (FFT) is used to decompose the kth ensemble background perturbation $\mathbf{X}_k^b$ into M scales as:

$$\mathbf{X}_k^b = \sum_{m=1}^{M} \mathbf{X}_{k,m}^b. \quad (1)$$

where $\mathbf{X}_{k,m}^b$ is the kth background ensemble perturbation containing the mth scale, $k = 1, 2, \ldots, K$, and $m = 1, 2, \ldots, M$. K is the ensemble size.

Let $\mathbf{X}_m^b = \frac{[\mathbf{X}_{1,m}^b, \mathbf{X}_{2,m}^b, \ldots, \mathbf{X}_{K,m}^b]}{\sqrt{K-1}}$ contain the normalized, decomposed background perturbations for the mth scale band confined in the local volume. $\mathbf{X}^b = [\mathbf{X}_1^b, \mathbf{X}_2^b, \ldots, \mathbf{X}_M^b]$ denotes the square root of the raw background ensemble covariance in a local volume. Then MLGETKF updates the mean and perturbation at the center of the local volume using

$$\begin{cases} \boldsymbol{L}\overline{\mathbf{x}^a} = \boldsymbol{L}\overline{\mathbf{x}^b} + \boldsymbol{L}\mathbf{Z}^{\mathrm{ML}}\mathbf{C}^{\mathrm{ML}}(\boldsymbol{\Gamma}^{\mathrm{ML}} + \boldsymbol{I})^{-1}(\mathbf{C}^{\mathrm{ML}})^T\left(\tilde{\mathbf{H}}\mathbf{Z}^{\mathrm{ML}}\right)^T\mathbf{R}^{-\frac{1}{2}}\left[\mathbf{y} - \boldsymbol{H}\left(\overline{\mathbf{x}^b}\right)\right] \\ \boldsymbol{L}\mathbf{X}^a = \boldsymbol{L}\mathbf{X}^b - \boldsymbol{L}\mathbf{Z}^{\mathrm{ML}}\mathbf{C}^{\mathrm{ML}}[\boldsymbol{I} - (\boldsymbol{\Gamma}^{\mathrm{ML}} + \boldsymbol{I})^{-\frac{1}{2}}](\boldsymbol{\Gamma}^{\mathrm{ML}} + \boldsymbol{I})^{-1}(\mathbf{C}^{\mathrm{ML}})^T\left(\tilde{\mathbf{H}}\mathbf{Z}^{\mathrm{ML}}\right)^T\tilde{\mathbf{H}}\mathbf{X}^b. \end{cases} \quad (2)$$

where $\boldsymbol{L}$ is the operator to select state variables at the center of the local volume, $\overline{\mathbf{x}^a}$ and $\overline{\mathbf{x}^b}$ are the ensemble analysis and background mean, respectively, $\mathbf{X}^a$ is the ensemble posterior

perturbation in a local volume, $\mathbf{y}$ represents observations, $\mathbf{R}$ is the observation error covariance, and $\mathbf{H}$ is the observation operator that converts a state variable from model space to observation space. The square root of the localized ensemble covariance $\mathbf{Z}^{\mathrm{ML}}$ can be obtained by modulating the normalized raw ensemble perturbations with the square roots of the localization matrix. The term $\widetilde{\mathbf{H}} = \mathbf{R}^{-\frac{\mathbf{1}}{\mathbf{2}}}\mathbf{H}$, and $\mathbf{C}^{\mathrm{ML}}$ and $\mathbf{\Gamma}^{\mathrm{ML}}$ are the eigenvectors and eigenvalues of matrix $\left(\widetilde{\mathbf{H}}\mathbf{Z}^{\mathrm{ML}}\right)^{\boldsymbol{T}}\left(\widetilde{\mathbf{H}}\mathbf{Z}^{\mathrm{ML}}\right)$, respectively. $\mathbf{I}$ is the identity matrix that has the same size as $\mathbf{\Gamma}^{\mathrm{ML}}$.

2.2 Scale-unaware RTPS and scale-dependent RTPS-SDI inflation methods

2.2.1 Scale-unaware RTPS method

RTPS (Whitaker and Hamill 2012) is one commonly used inflation method in ensemble-based DA, which relaxes the posterior spread $\sqrt{\mathbf{P}_a}$ toward the prior spread $\sqrt{\mathbf{P}_b}$ independently at each analysis grid point by

$$\sqrt{\mathbf{P}_a} \leftarrow (1-\alpha)\sqrt{\mathbf{P}_a} + \alpha\sqrt{\mathbf{P}_b}. \quad (3)$$

and thus, the *k*th inflated posterior ensemble perturbation $\mathbf{X}_k^a$ is,

$$\mathbf{X}_k^a \leftarrow \left(\alpha\frac{\sqrt{\mathbf{P}_b} - \sqrt{\mathbf{P}_a}}{\sqrt{\mathbf{P}_a}} + 1\right)\mathbf{X}_k^a. \quad (4)$$

In Eqs. (3) and (4), $\mathbf{P}_b$ and $\mathbf{P}_a$ represent the prior and posterior ensemble variances, respectively, and α is a tunable parameter. The corresponding inflation factor $\mathbf{g}$ in the scale unaware RTPS method is defined as,

$$\mathbf{g} = \alpha\frac{\sqrt{\mathbf{P}_b} - \sqrt{\mathbf{P}_a}}{\sqrt{\mathbf{P}_a}} + \mathbf{1}. \quad (5)$$

2.2.2 Scale-dependent RTPS method (RTPS-SDI)

SDI is proposed in this study to address the ensemble deficiency at different scales separately. Therefore, we develop RTPS-SDI to inflate the posterior ensemble spread at different scales toward the corresponding prior ensemble spread. The procedure of inflating posterior ensemble perturbation $\mathbf{X}_k^a, k, = 1, 2, \ldots, K$, in RTPS-SDI is introduced as follows.

Similar to Eq. (1), the scale separation of background and posterior ensemble spread is achieved by decomposing the corresponding ensemble perturbation. FFT is used to separately

decompose the kth prior and posterior ensemble perturbation $\mathbf{X}_k^b$ and $\mathbf{X}_k^a$ into M scales as:

$$\begin{cases} \mathbf{X}_k^b = \sum_{m=1}^{M} \mathbf{X}_{k,m}^b \\ \mathbf{X}_k^a = \sum_{m=1}^{M} \mathbf{X}_{k,m}^a , \end{cases} \quad (6)$$

where $\mathbf{X}_{k,m}^a$is the kth posterior ensemble perturbation containing the mth scale. Thus, the $\mathbf{P}_m^a$ and $\mathbf{P}_m^b$ can be calculated using,

$$\begin{cases} \mathbf{P}_m^b = \frac{1}{K-1} \sum_{k=1}^{K} \mathbf{X}_{k,m}^b (\mathbf{X}_{k,m}^b)^{\boldsymbol{T}} \\ \mathbf{P}_m^a = \frac{1}{K-1} \sum_{k=1}^{K} \mathbf{X}_{k,m}^a (\mathbf{X}_{k,m}^a)^{\boldsymbol{T}} . \end{cases} \quad (7)$$

Similar to the scale unaware RTPS in Eq (3), the posterior ensemble spread is relaxed toward the prior ensemble spread at each decomposed scale individually by

$$\sqrt{\mathbf{P}_m^a} \leftarrow (1-\alpha_m)\sqrt{\mathbf{P}_m^a} + \alpha_m \sqrt{\mathbf{P}_m^b} , \quad (8)$$

where $\sqrt{\mathbf{P}_m^a}$ and $\sqrt{\mathbf{P}_m^b}$ respectively represent the posterior and prior spread at the mth scale, and α_m is the corresponding tunable inflation parameter. The kth inflated posterior ensemble perturbation at the mth scale is obtained as,

$$\mathbf{X}_{k,m}^a \leftarrow \mathbf{g}_m \circ \mathbf{X}_{k,m}^a . \quad (9)$$

In Eq. (9), the sign ° denotes the Schur product, $\mathbf{g}_m$, m = 1, 2, …, M, represents the inflation factor at the mth scale, and is defined as

$$\mathbf{g}_m = \alpha_m \frac{\sqrt{\mathbf{P}_m^b} - \sqrt{\mathbf{P}_m^a}}{\sqrt{\mathbf{P}_m^a}} + \mathbf{1} . \quad (10)$$

The kth inflated full-scale posterior perturbation $\mathbf{X}_k^a$ is composed by the sum of all scale bands,

$$\mathbf{X}_k^a = \sum_{m=1}^{M} \mathbf{X}_{k,m}^a . \quad (11)$$

2.3 Scale-unaware SE and Scale-dependent SE-SDI

2.3.1 Scale-unaware SE method

The SE method was first proposed to inflate the ensemble posterior variance toward the

observation-dependent Bayesian posterior variance. Based on Eqs. (2.4) and (2.5) in Hodyss et al. (2016), the Bayesian posterior variance is equal to the posterior MSE. Therefore, the SE method aims to inflate the posterior ensemble variance toward the posterior MSE. The derivations of this SE method in this study are similar to those in Hodyss et al. (2016) and are introduced as follows.

In the scale-unaware DA, the analysis mean is defined by the classic Kalman filter equation,

$$\overline{\mathbf{x}^a} = \overline{\mathbf{x}^b} + \frac{\mathbf{P}^b\mathbf{H}^{\mathrm{T}}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R}}\left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}^b}\right). \quad (12)$$

The analysis MSE at grid point i by including Gaussian perturbations from sampling is written as,

$$\mathrm{S}_i = \mathrm{P}_i^t + \left(\frac{\mathrm{P}_i^a}{\mathrm{P}_i^f}\right)^2 \left(\frac{\mathrm{P}_i^f}{K} + \frac{2}{K-1}\left(\overline{\mathrm{x}_i^K} - \overline{\mathrm{x}_i^b}\right)^2\right)$$

$$\approx a * \mathrm{P}_i^a + \left(\frac{\mathrm{P}_i^a}{\mathrm{P}_i^b}\right)^2 \left(b * \frac{\mathrm{P}_i^b}{K} + c * \frac{2}{K-1}\left(\overline{\mathrm{x}_i^a} - \overline{\mathrm{x}_i^b}\right)^2\right), \quad (13)$$

where S_i is the analysis MSE at point i, P_i^t is the true error variance at point i, P_i^f is the true prior error variance at point i, $\overline{\mathrm{x}_i^K}$ is the true Kalman mean at given point. $\overline{\mathrm{X}_i^a}$ and $\overline{\mathrm{x}_i^b}$ are the ensemble analysis and background mean at point i, K represents the ensemble size, a, b and c are tunable parameters. The parameter a balances the difference by replacing P_i^t with P_i^a. The parameters b and c alleviate the difference in the sampling error terms caused by replacing P_i^f and $\overline{\mathrm{x}_i^K}$ with P_i^b and $\overline{\mathrm{x}_i^a}$. On the right-hand side of Eq. (13), the first term measures the true posterior variance. The second term is caused by the sampling error in the background estimate. The sampling error term is composed of two separate terms: the first term is the inflation caused by the sampling error in background mean, and the second term is the inflation caused by the sampling error in background variance.

To achieve the goal of this method, the corresponding inflation factor at grid point i is defined as

$$\mathrm{g}_i = \sqrt{\frac{\mathrm{S}_i}{\mathrm{P}_i^a}}. \quad (14)$$

As mentioned in Hodyss et al. (2016), this approximation is made to achieve a formula that may apply through simple vector manipulations without access to observations. Similar to Eq. (4), the *k*th inflated ensemble perturbation is obtained as

$$\mathbf{X}_k^a \leftarrow \mathbf{g} \circ \mathbf{X}_k^a , \quad (15)$$

where **g** is the matrix of the inflation factors in the entire domain.

2.3.2 Scale-dependent SE method (SE-SDI)

In this subsection, SDI is further derived and implemented in the SE inflation method to account for the sampling error at different scales and address the scale-dependent ensemble deficiency. The purpose of SE-SDI is to separately inflate the posterior ensemble variance at each scale toward the analysis MSE at the corresponding scale.

The analysis mean in each scale is defined by the multiscale DA analysis equation (Li et al. 2015)

$$\overline{\mathbf{x}_m^a} = \overline{\mathbf{x}_m^b} + \frac{\mathbf{P}_m^b \mathbf{H}^{\mathrm{T}}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R}} \left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}^b}\right), \quad (16)$$

where $m = 1,2, \ldots M$ represent the decomposed scales, $\overline{\mathbf{x}_m^b}$ and $\overline{\mathbf{x}_m^a}$ are the prior and posterior mean of the state variables in mth scale. Note that if we ignore the cross-scale covariances, the sum of Eq. (16) in all scale bands is equivalence to the Eq. (12).

Starting from Eq. (16), we further derive the analysis MSE at each scale in Appendix A. Specifically, the analysis MSE $\mathrm{S}_{i,m}$ at grid point i for the mth scale is calculated as,

$$\begin{aligned} \mathrm{S}_{i,m} &= \mathrm{P}_{i,m}^t + \left[\left(\frac{\mathrm{P}_{i,m}^a}{\mathrm{P}_{i,m}^f} \right)^2 \left(\frac{\mathrm{P}_{i,m}^f}{K} + \frac{2}{\mathrm{K}-1} \left(\overline{\mathrm{x}_{i,m}^K} - \overline{\mathrm{x}_{i,m}^b} \right)^2 \right) \right] \\ &\approx a_m * \mathrm{P}_m^a + \left[\left(\frac{\mathrm{P}_{i,m}^a}{\mathrm{P}_{i,m}^b} \right)^2 \left(b_m * \frac{\mathrm{P}_{i,m}^b}{K} + c_m * \frac{2}{\mathrm{K}-1} \left(\overline{\mathrm{x}_{i,m}^a} - \overline{\mathrm{x}_{i,m}^b} \right)^2 \right) \right]. \quad (17) \end{aligned}$$

Similar to the terms in Eq. (13), here $\mathrm{P}_{i,m}^t$, $\mathrm{P}_{i,m}^f$, and $\overline{\mathrm{x}_{i,m}^K}$ are the true error variance, true prior error variance and the true Kalman mean at point i for mth scale, respectively. a_m, b_m, and c_m are tunable parameters for the mth scale. The parameter a_m alleviates the difference by replacing $\mathrm{P}_{i,m}^t$ with $\mathrm{P}_{i,m}^a$.The parameters b_m and c_m are used to alleviate the differences in the sampling error terms caused by replacing $\mathrm{P}_{i,m}^f$ and $\overline{\mathrm{x}_{i,m}^K}$ with $\mathrm{P}_{i,m}^b$ and $\overline{\mathrm{x}_{i,m}^a}$.

The inflation factors $g_{i,m}$ at grid point i for mth scale are defined by inflating the posterior variance to the corresponding analysis MSE,

$$g_{i,m} = \sqrt{\frac{S_{i,m}}{P_{i,m}^{a}}}. \quad (18)$$

The kth inflated posterior ensemble perturbation at the mth scale is obtained as,

$$\mathbf{X}_{k,m}^{a} \leftarrow \mathbf{g}_m \circ \mathbf{X}_{k,m}^{a}, \quad (19)$$

where $\mathbf{g}_m$ is the inflation factor in the entire domain at the mth scale. Similar to RTPS-SDI, the kth inflated full-scale posterior perturbation $\mathbf{X}_k^a$ is composed by the sum of the inflated perturbation $\mathbf{X}_{k,m}^{a}$ in different scales:

$$\mathbf{X}_k^a = \sum_{m=1}^{M} \mathbf{X}_{k,m}^{a}. \quad (20)$$

3. Experimental design

This study evaluates the SDI method in the cycled MLGETKF DA system using a Surface Quasigeostrophic (SQG) model following Wang et al. (2021) and Tulloch and Smith (2009). The SQG model equation governs the evolution of potential temperature on the two bounding surfaces positioned 10 km apart. This model mimics the mesoscale of the atmosphere by simulating a range of scales with the KE spectrum satisfying the -5/3 slope (FIG. 1a). The model domain has 64 × 64 horizontal grid points with a grid spacing of ~300 km.

The SQG model is initialized by combining a random and sinusoidal scaled potential temperature field. The simulation runs for a total of 300 days, and the truth simulation begins at the 100th day onward since the first 100 days are removed to account for the spinup time. Figure 1b demonstrates that 100 days is sufficient for the natural variability of the SQG model to become stable. Potential temperature observations are simulated by adding Gaussian noise to the truth field with a standard deviation of 1 K, and their locations are randomly chosen from the model grid without replacement. The number of observations is 1024, occupying a quarter of the grid points in the entire domain. Figure 1c is an example of the simulated observations distributed over a potential temperature anomaly field from the truth. More details of the SQG model were introduced in Wang et al. (2021).

The DA configurations in this study mirror Wang et al. (2021). The initial ensemble consists of 20 members, randomly choosing from the climatology of the truth simulation. The multiscale DA algorithm, MLGETKF, is implemented to update the 20-member ensemble analyses. The assimilation is performed every 3 hours for 50 days, 400 cycles in total. The localization radii of 2000 and 4000 km are utilized to constrain the small- and large-scale ensemble covariances. A filtering function with a wavelength of 1818 km is used to decompose the raw ensemble perturbations into large and small scales ($m = 2$) for both MLGETKF update and SDI inflation. Ensemble mean forecasts with a 6-day lead time are initialized from the analysis ensemble mean from each of the last 350 cycles. The free forecasts from the first 50 DA cycles are omitted to allow a sufficient DA spinup time. Here, the spinup time is defined as the time required for the analysis error in DA system to reach a relatively stable status.

Four cycled DA experiments are designed (Table 1) to examine the four inflation

methods in Section 2. Comparisons in analysis accuracy and forecast performance between RTPS and RTPS-SDI and between SE and SE-SDI are conducted to evaluate the SDI method versus its scale-unaware counterparts in alleviating ensemble deficiency. We further compared the relative improvements achieved by applying SDI to RTPS and SE for the second objective introduced in Section 1. In addition, the temporal and spatial behaviors of inflation factors from four experiments are compared for the third objective. Sensitivity tests were performed to optimize the tunable parameters in the inflation methods. For RTPS and RTPS-SDI, initial tests were performed with α and α_m respectively increasing from 0.3 to 0.8 at an interval of 0.1. Tests were also conducted to optimize the tunable parameters in SE and SE-SDI. Notice that the values of a and a_m are chosen as 1 to avoid the ensemble spread deflation, SE and SE-SDI experiments only tune parameters b, c, b_m, and c_m. Specifically, b and b_m change from 1 to 7 at an interval of 1; c and c_m change from 1 to 7 at an interval of 1. The optimal α for RTPS is α = 0.6. The optimal parameters for RTPS-SDI are (α_L,α_S) = (0.4, 0.6). The best-performing SE corresponds to (b, c) = (3, 2), and the best-performing SE-SDI corresponds to (b_L,c_L,b_S,c_S) = (4, 6, 3, 2). Moreover, the computational costs for each DA cycle are comparable among the four experiments as the differences in the costs among the various inflation methods are negligible.

4. Results

4.1 Impacts of SDI on DA cycling

4.1.1 Time series of the RMS analysis error

This section verifies and diagnoses the analysis result in the DA cycling for the designed experiments in Section 3. We divide the 400 DA cycles into 3 phases to reveal the differences among experiments as follows (Fig. 2). The period of the first 50 cycles is determined as Phase 1, where all experiments are still spinning up. The 51-100 cycles are selected as Phase 2, where the analysis errors of both RTPS-SDI and SE-SDI experiments become stable. In the remaining 101-400 cycles (Phase 3), all experiments have relatively stable analysis errors.

In Fig. 2, the analysis RMSE in SE (solid red line) is lower than that in RTPS (solid black line) for nearly all cycles. Thus, SE inflation shows improved analysis accuracy compared to RTPS in all phases. Regarding the two SDI approaches, RTPS-SDI (dashed black line) outperforms SE-SDI (dashed red line) in Phase 1; SE-SDI (dashed red line) outperforms RTPS-SDI (dashed black line) in Phases 2 and 3. The outperformance of RTPS-SDI relative to SE-SDI in Phase 1 is attributed to the larger inflation values of RTPS-SDI in the earlier cycles. The larger inflation values will be explained in section 4.1.3. In addition, the two SDI approaches demonstrate an apparent analysis improvement compared to the scale-unaware inflations. Specifically, the analysis RMSE of RTPS-SDI (SE-SDI) is smaller than that of RTPS (SE) in Phases 1 and 2. While the differences among the four experiments in analysis RMSE are much smaller than the earlier DA cycles in Phase 3, RTPS-SDI and SE-SDI show smaller analysis RMSE than RTPS and SE, respectively. The spatiotemporally averaged RMS analysis RMSE of potential temperature in each phase for four experiments (Table 2) is generally consistent with Fig. 2. In Phases 1 and 2, RTPS-SDI outperforms RTPS with the reduced analysis RMSE of 0.4-0.9 K. Similarly, the analysis improvements of SE-SDI over SE are 0.3-0.5 K. In Phase 3, such improvements are smaller than in the early phases, and two SDI approaches have 0.007-0.03 K lower analysis RMSE than their scale-unaware counterparts. Table 2 also shows that the relative improvements of RTPS-SDI over RTPS are 37%, 30%, and 3% in Phases 1, 2, and 3, respectively. As for SE-SDI over SE, the relative improvements are 21%, 24%, and 1% in the corresponding phases. These results indicate that the relative

improvement for applying SDI to RTPS is much larger than that for SE.

4.1.2 Spectral space diagnostics

(i) RMS errors and ensemble spread for the analysis and the background

The kinetic energy (KE) for the ensemble mean background error and ensemble mean analysis error are examined in spectral space to further evaluate the differences between the SDI methods and their scale-unaware counterparts at each scale. Figures 3a-3c compare the ensemble mean error between RTPS-SDI and RTPS. RTPS-SDI reduces absolute analysis and background errors at all phases and nearly all scales relative to RTPS. The outperformance of RTPS-SDI over RTPS is more apparent in Phase 1 than in Phases 2 and 3. The error corrections are defined by the differences between background and analysis errors. In Phase 1, the error corrections at large scales are more marked than those at small scales for RTPS (black line vs. gray line in Fig. 3a) and RTPS-SDI (red line vs. yellow line in Fig. 3a). In Phases 2 and 3, these error corrections become smaller than that in Phase 1, and the large-scale error corrections are smaller than the small-scale error corrections. Differences in the error corrections between RTPS-SDI and RTPS are mostly exhibited in Phase 1, where RTPS-SDI produces larger error corrections than RTPS, especially at large scales. Figures 3d-3f show the comparisons between SE-SDI and SE. Similar to the comparisons between RTPS-SDI and RTPS, SE-SDI reduces both analysis and background errors relative to SE for all phases and nearly all scales, except at the very large scale in Phase 3. In Phase 1, the large-scale error corrections are greater than that at small scales for SE (black line vs. gray line in Fig.3d) and SE-SDI (red line vs. yellow line in Fig.3d). In Phases 2 and 3, the error corrections at large scales are also smaller than the small-scale error corrections. Different from RTPS-SDI, SE-SDI shows slightly larger background and analysis errors at the largest scale in Phase 3. To sum up, compared to the scale unaware RTPS and SE, the two SDI experiments produce analysis improvements in all phases and nearly at all scales, except SE-SDI at the largest scale in Phase 3.

A representative example to present the differences in analysis errors in physical space between RTPS-SDI and RTPS is shown in Fig. 4. Large-scale errors mostly agree with large-scale flow features. Small-scale errors are distributed over the entire domain with greater values over the high-temperature gradient regions. RTPS demonstrates larger analysis errors at both

large and small scales than RTPS-SDI. Compared to RTPS, the improvements by RTPS-SDI are more marked at small scales than at large scales, consistent with Fig. 3 (gray lines vs. yellow lines in Figs.3a-3c). Similarly, Fig. 5 compares SE and SE-SDI analysis errors in physical space. The scale-dependent analysis errors in SE and SE-SDI show features similar to those of RTPS-based experiments. SE-SDI decreases analysis errors compared to SE at both large and small scales. The decrease in analysis errors at small scales is more apparent than at large scales.

To further quantify the relative improvements of applying SDI on analysis, we measure the relative analysis error reduction in the spectral space averaged on all cycles in Fig. 6. The relative analysis error reduction is defined as the reduction of RTPS-SDI (SE-SDI) relative to scale-unaware RTPS (SE) divided by the RTPS (SE) analysis error. RTPS-SDI shows ~40% lower errors than RTPS averaged over all cycles. The error reductions show scale-varying features. Remarkably, the reductions gradually get greater from the smallest scale (600 km) to the largest scale (about 7000 km). Comparisons between SE and SE-SDI present that SE-SDI produces ~25% smaller errors than SE averaged over all cycles. Similar to the comparisons between RTPS-SDI and RTPS, larger scales also show more apparent reductions than smaller scales. These results suggest that both SDI experiments show apparent analysis improvements compared to their single-scale counterparts. The analysis improvements of RTPS-SDI over RTPS are greater than that of SE-SDI over SE. The differences in the relative improvements by applying SDI to SE and to RTPS may be explained as follows. RTPS, by design, adjusts analysis ensemble variance toward background ensemble variance without considering the scale-dependent differences in analysis ensemble variance. In contrast, SE partially addresses the scale-dependent sampling error. Therefore, applying SDI to RTPS obtains more apparent improvements than that to SE. Moreover, the greater error reductions at the large scale indicate that the SDI approach inflates the large scale more effectively. This more effective large-scale inflation will be explained in section 4.1.3.

(ii) Consistent ratio

Consistency ratio (CR) is commonly used to evaluate the quality of ensemble-based DA systems. It is defined as the ratio of the total ensemble spread to the RMSE of the analysis ensemble mean. Ideally, CR is 1.0, indicating the forecast error variance is properly estimated

by the total ensemble variance. Fig.7 shows CR to further evaluate the ensemble spread for the four experiments. Both SDI experiments have improved CRs relative to their single-scale counterparts, except that RTPS-SDI overly estimates CR at large scales in Phase 1 (Fig. 7a). The overestimation problem is caused by the excessively large background spread which is obtained from the very initial climatology ensemble. Details of the overestimation problem will be explained in section 4.1.3. In Phase 2, RTPS-SDI has substantial improvements compared to RTPS at all scales with the maximum improvements at smaller scales exceeding 0.2. The CR values of SE-SDI are generally 0.15–0.2 greater than SE. The larger improvements of SE-SDI over SE appear at the smaller scales (Fig. 7b). In Phase 3, the four experiments show much smaller CR differences than in earlier DA cycles. Despite that, both SDI experiments produce 0.05-0.07 larger CR values than their single-scale counterparts. In addition, the increased CR values of RTPS-SDI over RTPS are greater than that of SE-SDI over SE at all scales in Phase 1. On the contrary, the CR enhancement in SE-SDI over SE is higher than that in RTPS-SDI over RTPS at larger scales in Phases 2 and 3. These results imply that both SDI experiments have improved ensemble spread relative to their single-scale counterparts. SE-SDI has greater improvements over SE in ensemble spread at the majority of scales and DA cycles than RTPS-SDI over RTPS.

4.1.3 Temporal and spatial behaviors of the inflation

To explain and understand the analysis improvements by the SDI experiments over their scale-unaware counterparts, the temporal and spatial distribution of the inflation factors are diagnosed in this section. The value of the inflation factor shows a positive relationship with the ensemble spread during DA cycling. Figure 8 shows the temporal distribution of the inflation factors. In Phase 1, the RTPS and RTPS-SDI inflation factors start with excessive median values above 2.0 and decrease to a steadier value of around 1.1. As the RTPS-based approaches inflate the posterior variance toward the prior variance, these excessive values are caused by the large background spread in earlier DA cycles and explain the overestimated CR in RTPS and RTPS-SDI during Phase 1 (Fig. 7a). In contrast, the SE and SE-SDI inflation factors start with lower median values of 1.0 and increase to achieve stable values of 1.07~1.09. These relatively small inflation factors in SE and SE-SDI are associated with the small values

of the ratio term (P^a/P^b) in Eqs. 13 and 17. Along with the increase of DA cycles, the inflation factors get larger and become stable as the stable P^b. During the later DA cycles of Phase 1 and during Phase 2, the inflation factors in RTPS are similar to those in RTPS-SDI at the small scales with values of 1.06-1.08. In comparison, the inflation factors in RTPS-SDI at the large scales with values of 1.09-1.11 are greater than in RTPS and RTPS-SDI at the small scales. Similarly, the large-scale inflation factors in SE-SDI of ~1.09 show higher values than those in SE and the small-scale inflation factors of 1.07 in SE-SDI. These enhancements for the large scales cause a marked increase in ensemble spread and result in a reduced spinup time in the two SDI experiments. In Phase 3, the large-scale inflation factors in the two SDI experiments remain larger than that at the corresponding small scales and in the scale-unaware experiments. These results explain the improvements of the two SDI experiments in the analysis error correction and CR, especially at the large scales, compared to their scale-unaware counterparts. Moreover, RTPS and RTPS-SDI show greater temporal variability than SE and SE-SDI in Phase 3. The inflation median values of RTPS and RTPS-SDI at the small scales oscillate between 1.04-1.07. The inflation median values of RTPS-SDI at the large scales range from 1.07-1.13. The inflation factors of SE and SE-SDI at the small scales keep around 1.07. SE-SDI maintains the inflation median values of 1.09 at the large scales. The greater inflation temporal variabilities in RTPS and RTPS-SDI are caused by the oscillation of the background spread, which is related to the oscillation of natural variability in the SQG model. For instance, peak values of the inflation factors in RTPS and RTPS-SDI in Fig. 8c occur at 100-110 cycles and 270-300 cycles, consistent with the oscillation peak of natural variability at around 900 and 1100 cycles of SQG nature run (Fig. 1b).

We further choose the inflation factors at the 300th cycle as an example to reveal the spatial differences among experiments. Figure. 9 shows the inflation patterns in four experiments overlay on the background potential temperature fields at the corresponding scales. The inflation factors in RTPS and small-scale RTPS-SDI are located on the entire domain and reflect small-scale patterns (Figs. 9a, 9c). In comparison, the large-scale RTPS-SDI inflation demonstrates flow-dependent patterns related to the background. For example, significant inflation are shown in the center region with high potential temperature gradients (Fig. 9b). On the contrary, SE and SE-SDI inflation are account for sampling errors, thus don't demonstrate

clear relationship with the background flow. In the two SDI experiments, the inflation factors at large and small scales reflect the corresponding scale structures (Figs. 9b, c and 9e, f). Remarkably, the single-scale inflation factors primarily show a similar spatial pattern to the small-scale SDI inflation factors (Figs. 9a, c and 9d, f), indicating that the full-scale spread is mainly dominated by that at the small scales. Therefore, the corresponding single-scale inflation may be inappropriate for the variance at the large scales. The large-scale inflation factors in both SDI experiments are individually calculated based on the large-scale ensemble deficiency and thus are larger than in their scale-unaware counterparts. Therefore, both SDI experiments obtain greater ensemble spread at the large scale than their scale-unaware counterparts. Furthermore, each scale can impact other scales through the multiscale covariance matrix during the update process (Wang et al. 2021). Thus, the ensemble improvements at large scales by the SDI experiments can gradually deliver to small scales during DA cycling, resulting in the full-scale CR increase and analysis improvements in the subsequent stable cycles.

4.2 Impact of SDI on forecast

Figures 10 and 11 examine the ensemble mean forecast errors averaged over Phases 2 and 3 separately. Forecasts from Phase 1 are omitted as none of the experiments reach the spinup status in this period. In addition to the full-scale forecast error, we decompose the forecast error into large and small scales to examine the forecast performance at different scales. The separation wavelength is chosen as 1818 km, which is consistent with the scale separation used in the DA update and inflation.

Figures 10a, b show the forecast performance of RTPS and RTPS-SDI in Phases 2 and 3. In Phase 2, the RTPS-SDI deterministic forecast is statistically significantly more accurate than the RTPS for all scales up to 3-4 days lead time. However, there is a less accurate period for full-scale RTPS-SDI forecast in 4-5 days, which is mainly associated with the worse performance at the large scale. During Phase 3, the RTPS-SDI forecast is more accurate than RTPS for 2-3 days lead time at all scales with statistical significance. Figures 10c,d compare the deterministic forecast errors between SE-SDI and SE. In Phase 2, SE-SDI shows statistically significantly lower forecast errors up to 5-6 days lead time at full and large scales

and up to 3 days lead time at small scales. During Phase 3, the difference in the forecast error between SE and SE-SDI does not pass the 95% or 99.5% significance tests. At full and small scales, SE-SDI shows a smaller forecast error during 2-3 days with significance and passes the 90% significance test. Overall, the SDI experiments show a slight forecast improvement compared to their scale-unaware counterparts. Moreover, Fig. 10 indicates that the forecast errors show varied behaviors at different scales. In all experiments, the performance of early forecasts is dominated by small-scale forecast errors. After 3-4 days lead time, later forecasts are degraded due to the increasing large-scale forecast errors. Therefore, the improvements at large scales are critical to improve the forecast performance in 4-6 days or even longer lead time.

Note that applying SDI to RTPS (RTPS-SDI vs. RTPS) obtains more apparent forecast relative improvement than SE (SE-SDI vs. SE), which is consistent with the relative improvements during DA cycling. Figures 11a, b show that the forecast of SE is more accurate than that of RTPS with statistical significance. During Phase 2, SE shows statistically significantly lower forecast errors up to 3 days lead time at full and large scales and up to 2 days lead time at small scales. In Phase 3, SE produces more accurate forecasts than RTPS up to 2-3 days lead time at all scales. However, RTPS-SDI and SE-SDI obtain the forecast performance at nearly the same level without passing significant differences in most of the lead times (Figs. 11c, d). These results indicate that the application of SDI has more apparent improvement on RTPS than on SE. Therefore, while RTPS has higher forecast errors than SE, RTPS-SDI could achieve a similar forecast performance to SE-SDI.

5 Conclusion and discussion

Ensemble-based data assimilation (DA) often suffers from ensemble sampling deficiency with the varied extents of insufficient ensemble spread at different scales. This study develops the scale-dependent inflation (SDI) within the MLGETKF framework (Wang et al. 2021) to alleviate the scale-dependent ensemble deficiency. In this study, SDI is first implemented on top of the commonly used RTPS method by separately relaxing the posterior ensemble spread toward the corresponding prior ensemble spread at each scale (RTPS-SDI). In contrast, RTPS inflates the posterior ensemble spread using the same inflation factor at all scales. To explore the generic usage of SDI, this study also derives and implements SDI on top of the previously proposed SE method (Hodyss et al. 2016) by individually adjusting posterior ensemble variance toward the Bayesian posterior at each scale (SE-SDI), whereas SE similarly inflates posterior ensemble variance in a scale-unaware manner.

Four experiments are performed with the four inflation methods for 400 DA cycles using a two-layer surface quasigeostrophic (SQG) turbulence model. During the DA cycling, RTPS-SDI and SE-SDI outperform RTPS and SE, respectively, in reducing analysis errors at nearly all scales and all cycles. Specifically, both SDI experiments achieve a faster spinup in the earlier DA cycles and have an average 25%-40% relative analysis error reduction in the later DA cycles than their scale-unaware counterparts. The improvements in analysis errors of RTPS-SDI over RTPS are greater than those of SE-SDI over SE. In addition, RTPS-SDI obtains a 0.05-0.2 larger consistency ratio (CR) than RTPS, whereas SE-SDI has 0.07-0.2 greater CR than SE. These improvements in both SDI methods are associated with their greater inflation at all scales, especially at larger scales, compared to their scale-unaware counterparts. In the subsequent forecast, both SDI methods show statistically significantly better forecast performance than their scale-unaware inflation experiments. RTPS-SDI is more accurate than RTPS for all scales in 3-4 days lead time during the cycles of 50-100 and for 2-3 days lead time during the cycles of 100-400. SE-SDI is more accurate than SE at all scales for 5-6 days lead time during the cycles of 50-100 and shows full- and small-scale forecast improvement for 2-3 days during the cycles of 100-400. Similar to the relative improvement in analysis, RTPS achieves a more noticeable improvement in forecasting than SE.

As an initial effort to examine the newly proposed scale-dependent inflation method, the posterior ensemble variance is adjusted within two scale bands only to address the scale-dependent ensemble sampling deficiency. More scale bands may be required for the sophisticated atmospheric models. Moreover, efforts will be needed to determine scale separation lengths and the tunable parameters in Eqs. (8) and (17). In addition, this study examines the SDI methods using a simple SQG model with perfect model assumption. Future work will apply these methods to real-data models.

Appendix:

A. Derivation of Scale-Dependent SE inflation (SE-SDI)

The scale-dependent SE inflation is derived as follows. SE is further developed into SE-SDI using the multiscale DA analysis equation (Li et al., 2015):

$$\overline{\mathbf{x}_m^a} = \overline{\mathbf{x}_m^b} + \frac{\mathbf{P}_m^b \mathbf{H}^{\mathrm{T}}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R}}\left(\mathbf{y} - \mathbf{H}\overline{\mathbf{x}^b}\right), \quad (16)$$

where $m = 1,2, \ldots M$ represent the decomposed scales, $\overline{\mathbf{x}_m^b}$ and $\overline{\mathbf{x}_m^a}$ are the prior and posterior mean of the state variables in the mth scale, $\mathbf{P}_m^b$ is the corresponding sample background covariance matrix in the mth scale. The relationship between the full scale and decomposed scales satisfies:

$$\begin{cases} \overline{\mathbf{x}^b} = \sum_{m=1}^{M} \overline{\mathbf{x}_m^b} \\ \mathbf{P}^b = \sum_{m=1}^{M} \mathbf{P}_m^b \\ \overline{\mathbf{x}^a} = \sum_{m=1}^{M} \overline{\mathbf{x}_m^a} \end{cases}. \quad (A1)$$

Using Eq. (16), we first obtain the analysis covariance at different scales as follows. The Kalman gain at the mth scale is calculated by,

$$\mathbf{K}_m = \frac{\mathbf{P}_m^b \mathbf{H}^{\mathrm{T}}}{\mathbf{H}\mathbf{P}^b\mathbf{H}^{\mathrm{T}} + \mathbf{R}}. \quad (A2)$$

The mth scale's analysis error $\mathbf{e}_m^a$ is obtained as,

$$\mathbf{e}_m^a = \mathbf{x}_m^a - \mathbf{x}_m^t = \mathbf{x}_m^b - \mathbf{x}_m^t + \mathbf{K}_m(\mathbf{y} - \mathbf{H}\mathbf{x}^t + \mathbf{H}\mathbf{x}^t - \mathbf{H}\mathbf{x}^b) = \mathbf{e}_m^b + \mathbf{K}_m(\mathbf{e}^o - \mathbf{H}\mathbf{e}^b), \quad (A3)$$

where $\mathbf{x}_m^a, \mathbf{x}_m^b$ and $\mathbf{x}_m^t$ are the analysis, background, and truth of the state variables in the mth scale, respectively, $\mathbf{e}_m^a$ and $\mathbf{e}_m^b$ are the analysis and background error of the state variables, $\mathbf{e}^o$ is the observation error. Then, the analysis error covariance in the mth scale ($\mathbf{P}_m^a$) can be calculated by,

$$\mathbf{P}_m^a = \mathbf{e}_m^a(\mathbf{e}_m^a)^{\mathrm{T}} = [\mathbf{e}_m^b + \mathbf{K}_m(\mathbf{e}^o - \mathbf{H}\mathbf{e}^b)][\mathbf{e}_m^b + \mathbf{K}_m(\mathbf{e}^o - \mathbf{H}\mathbf{e}^b)]^{\mathrm{T}}. \quad (A4)$$

If we assume the observation error is uncorrelated with background error and the background errors at different scales are uncorrelated with each other, the Equation (A4) can be further derived as:

$$\mathbf{P}_m^a = \mathbf{P}_m^b - \mathbf{P}_m^b\mathbf{H}^{\mathrm{T}}\mathbf{K}_m^{\mathrm{T}} + \mathbf{K}_m(\mathbf{R})\mathbf{K}_m^{\mathrm{T}} - \mathbf{K}_m\mathbf{H}\mathbf{P}_m^b + \mathbf{K}_m\mathbf{H}(\mathbf{P}^b)\mathbf{H}^{\mathrm{T}}\mathbf{K}_m^{\mathrm{T}} = (\mathbf{I} - \mathbf{K}_m\mathbf{H})\mathbf{P}_m^b. \quad (A5)$$

Similar to Hodyss et al. (2016), our further derivation is taken pointwise. We include sampling errors in the ensemble Kalman filter estimate of the posterior mean and variance at given point i for each scale,

$$\begin{cases} \overline{\mathrm{x}_{i,m}^{f}} = \overline{\mathrm{x}_{i,m}^{b}} + \delta_{i,m}^{u} \\ \mathrm{P}_{i,m}^{f} = \mathrm{P}_{i,m}^{b} + \delta_{i,m}^{v} \end{cases} m = 1,2,\dots,M\ , (\mathrm{A6})$$

where $\overline{\mathrm{x}_{i,m}^{f}}$ and $\mathrm{P}_{i,m}^{f}$ are the true prior mean and true prior perturbation variance at grid point i for mth scale, $\delta_{i,m}^{u}$ and $\delta_{i,m}^{v}$ are the sampling errors in the estimate of the prior mean and variance at point i for mth scale, respectively.

To simplify the derivation and consistent with Hodyss et al. (2016), we assume the observation operator $H = I$ since SE-based inflations do not require access to the observations, where I is an identity matrix and has the same dimension as the operator H. Eq. (16) is therefore simplified to:

$$\overline{\mathrm{x}_{i,m}^{a}} = \overline{\mathrm{x}_{i,m}^{b}} + \frac{\mathrm{P}_{i,m}^{b}}{\mathrm{P}_{i}^{b} + \mathrm{R}_{i}}\left(\mathrm{y}_{i} - \overline{\mathrm{x}^{b}}\right), \quad (A7)$$

where subscript i in Eq. (B7) represents variables at the grid point i. Using a Taylor expansion in terms of $\delta_{i,m}^{u}$ and $\delta_{i,m}^{v}$, we obtain the posterior mean in this point as,

$$\overline{\mathrm{x}_{i,m}^{a}} \approx \overline{\mathrm{x}_{i,m}^{K}} + \frac{\mathrm{P}_{i,m}^{a}}{\mathrm{P}_{i,m}^{f}}\left(\delta_{i,m}^{u} + \frac{\overline{\mathrm{x}_{i,m}^{K}} - \overline{\mathrm{x}_{i,m}^{f}}}{\mathrm{P}_{i,m}^{f}}\delta_{i,m}^{v}\right), \quad (A8)$$

where $\overline{\mathrm{x}_{i,m}^{K}}$ is the true Kalman mean at point i for the mth scale. The sampling error between the true posterior mean $\overline{\mathrm{x}_{i,m}}$ and the sample posterior mean $\overline{\mathrm{x}_{i,m}^{a}}$ at given point is further calculated as,

$$E\left[\left(\overline{\mathrm{x}_{i,m}} - \overline{\mathrm{x}_{i,m}^{a}}\right)^{2}\right] = E\left[\left(\left(\overline{\mathrm{x}_{i,m}} - \overline{\mathrm{x}_{i,m}^{K}}\right) - \frac{\mathrm{P}_{i,m}^{a}}{\mathrm{P}_{i,m}^{f}}\left(\delta_{i,m}^{u} + \frac{\overline{\mathrm{x}_{i,m}^{K}} - \overline{\mathrm{x}_{i,m}^{f}}}{\mathrm{P}_{i,m}^{f}}\delta_{i,m}^{v}\right)\right)^{2}\right]. \quad (A9)$$

As the unbiased assumption of $\overline{\mathrm{x}^{b}}$ and P^{b} and the independent assumption of δ^{u} and δ^{v}, the cross terms $E\left[\delta_{i,m}^{u}\delta_{i,m}^{v}\right] = E\left[\delta_{i,m}^{v}\delta_{i,m}^{u}\right] = E\left[\delta_{i,m}^{v}\right] * E\left[\delta_{i,m}^{u}\right] = 0$. Therefore, Eq. (A9) is simplify as:

$$E\left[\left(\overline{\mathrm{x}_{i,m}}-\overline{\mathrm{x}_{i,m}^{a}}\right)^{2}\right]$$

$$=\mathrm{E}\left[\left(\overline{\mathrm{x}_{i,m}}-\overline{\mathrm{x}_{i,m}^{K}}\right)^{2}\right]+\left(\frac{\mathrm{P}_{i,m}^{a}}{\mathrm{P}_{i,m}^{f}}\right)^{2}*\left(E\left[{\delta_{i,m}^{u}}^{2}\right]+\left(\frac{\overline{\mathrm{x}_{i,m}^{K}}-\overline{\mathrm{x}_{i,m}^{f}}}{\mathrm{P}_{i,m}^{f}}\right)^{2}E\left[{\delta_{i,m}^{v}}^{2}\right]\right)-2$$

$$*E\left[\left(\overline{\mathrm{x}_{i,m}}-\overline{\mathrm{x}_{i,m}^{K}}\right)*\frac{\mathrm{P}_{i,m}^{a}}{\mathrm{P}_{i,m}^{f}}\left(\delta_{i,m}^{u}+\frac{\overline{\mathrm{x}_{i,m}^{K}}-\overline{\mathrm{x}_{i,m}^{f}}}{\mathrm{P}_{i,m}^{f}}\delta_{i,m}^{v}\right)\right].\quad(A10)$$

In Eq. (A10), $E\left[{\delta_{i,m}^{u}}^{2}\right]$ and $E\left[{\delta_{i,m}^{v}}^{2}\right]$ are regarded as the variance due to the sampling of the sample mean and the sampling of the sample variance, respectively. Under the Gaussian assumption, $\overline{\mathrm{x}_{m}}$ is equal to $\overline{\mathrm{x}_{m}^{K}}$, resulting in the first and cross terms in Eq. (A10) being zero. Using the Gaussian central limit theorem (CLT), $E\left[{\delta_{i,m}^{u}}^{2}\right]$ and $E\left[{\delta_{i,m}^{v}}^{2}\right]$ result in Eq. (A11).

$$\begin{cases}E\left[{\delta_{i,m}^{u}}^{2}\right]=E\left(\left(\overline{\mathrm{x}_{i,m}^{b}}-\overline{\mathrm{x}_{i,m}^{f}}\right)^{2}\right)=E\left(\left(\overline{\mathrm{x}_{i,m}^{b}}-E\left[\overline{\mathrm{x}_{i,m}^{b}}\right]\right)^{2}\right)=\dfrac{\mathrm{P}_{i,m}^{f}}{N_{e}}\\ E\left[{\delta_{i,m}^{v}}^{2}\right]=E\left(\mathrm{P}_{i,m}^{b}-\mathrm{P}_{i,m}^{f}\right)^{2}\right)=E\left(\left(\mathrm{P}_{i,m}^{b}-E\left[\mathrm{P}_{i,m}^{b}\right]\right)^{2}\right)=\dfrac{2}{N_{e}-1}\left(\mathrm{P}_{i,m}^{f}\right)^{2}.\end{cases}\quad(A11)$$

Substituting Eq. (A11) into Eq. (A10), we obtain,

$$E\left[\left(\overline{\mathrm{x}_{i,m}}-\overline{\mathrm{x}_{i,m}^{a}}\right)^{2}\right]=\left(\frac{\mathrm{P}_{i,m}^{a}}{\mathrm{P}_{i,m}^{f}}\right)^{2}\left(\frac{\mathrm{P}_{i,m}^{f}}{N_{e}}+\frac{2}{N_{e}-1}\left(\overline{\mathrm{x}_{i,m}^{K}}-\overline{\mathrm{x}_{i,m}^{f}}\right)^{2}\right).\quad(A12)$$

Furthermore, the analysis MSE is obtained as,

$$\mathrm{S}_{i,m}=\int_{-\infty}^{\infty}\left(\mathrm{x}_{i,m}^{t}-\overline{\mathrm{x}_{i,m}^{a}}\right)^{2}\rho\left(\mathrm{x}_{i,m}^{t}|y\right)d\mathrm{x}_{i,m}^{t}=E\left[\left(\mathrm{x}_{i,m}^{t}-\overline{\mathrm{x}_{i,m}^{a}}\right)^{2}\right]$$

$$=E\left[\left(\mathrm{x}_{i,m}^{t}-\overline{\mathrm{x}_{i,m}}+\overline{\mathrm{x}_{i,m}}-\overline{\mathrm{x}_{i,m}^{a}}\right)^{2}\right]=E\left[\left(\mathrm{x}_{i,m}^{t}-\overline{\mathrm{x}_{i,m}}\right)^{2}\right]+E\left[\left(\overline{\mathrm{x}_{i,m}}-\overline{\mathrm{x}_{i,m}^{a}}\right)^{2}\right]$$

$$=\mathrm{P}_{i,m}^{t}+\left[\left(\frac{\mathrm{P}_{i,m}^{a}}{\mathrm{P}_{i,m}^{f}}\right)^{2}\left(\frac{\mathrm{P}_{i,m}^{f}}{N_{e}}+\frac{2}{N_{e}-1}\left(\overline{\mathrm{x}_{i,m}^{K}}-\overline{\mathrm{x}_{i,m}^{f}}\right)^{2}\right)\right],\quad(A13)$$

where $\mathrm{S}_{i,m}$ is the analysis MSE, $\mathrm{x}_{i,m}^{t}$ is the true state, $\mathrm{P}_{i,m}^{t}$ is the true posterior variance, y is the observation, $E\left[\left(\overline{\mathrm{x}_{i,m}}-\overline{\mathrm{x}_{i,m}^{a}}\right)^{2}\right]$ represents the sampling error term, and $\rho\left(\mathrm{x}_{i,m}^{t}|y\right)$ is the conditional true state distribution based on observation. N_e is the limited ensemble size. The subscript i and m in these variables represent the given grid points and the decomposed scales.

Note that the true background and analysis estimates in Eqs. (A12) and (A13) are usually unknown in a specific DA system. In practice, we replace these unknown terms in Eq.

(A13) with their sample background or analysis and use tunable parameters (a_m, b_m, c_m) to balance the difference as,

$$\mathrm{S}_{i,m} \approx a_m * \mathrm{P}^{a}_{i,m} + \left[\left(\frac{\mathrm{P}^{a}_{i,m}}{\mathrm{P}^{b}_{i,m}}\right)^2 \left(b_m \frac{\mathrm{P}^{b}_{i,m}}{N_e} + c_m \frac{2}{N_e - 1}\left(\overline{\mathrm{x}^{a}_{i,m}} - \overline{\mathrm{x}^{b}_{i,m}}\right)^2\right)\right]. \quad (17)$$

Tunable parameter a_m is usually chosen as 1 to avoid the ensemble spread deflation.

REFERENCE

Anderson, J. L., & Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Monthly Weather Review, 127(12), 2741–2758. https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2

Bishop, C. H., J. S. Whitaker, and L. Lei, 2017: Gain Form of the Ensemble Transform Kalman Filter and Its Relevance to Satellite Data Assimilation with Model Space Ensemble Covariance Localization. Mon. Wea. Rev., 145, 4575–4592, https://doi.org/10.1175/MWR-D-17-0102.1.

Buehner, M., J. Morneau, and C. Charette, 2013: Four-dimensional ensemble-variational data assimilation for global deterministic weather prediction. Nonlinear Processes Geophys., 20, 669–682, https://doi.org/10.5194/npg-20-669-2013.

——, and Coauthors, 2015: Implementation of deterministic weather forecasting systems based on ensemble–variational data assimilation at Environment Canada. Part I: The global system. Mon. Wea. Rev., 143, 2532–2559, https://doi.org/10.1175/MWR-D-14-00354.1.

Clayton, A. M., A. C. Lorenc, and D. M. Barker, 2013: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office. Quart. J. Roy. Meteor. Soc., 139, 1445–1461, https://doi.org/10.1002/qj.2054.

Eyre, J. R. (2016). Observation bias correction schemes in data assimilation systems: A theoretical study of some of their properties. Quarterly Journal of the Royal Meteorological Society, 142(699), 2284-2291. https://doi.org/10.1002/qj.2819

Gasperoni, N. A., X. Wang, and Y. Wang, 2022: Using a cost-effective approach to increase background ensemble member size within the GSI-based EnVar system for improved radar analyses and forecasts of convective systems. Mon. Wea. Rev., 150, 667-689, https://doi.org/10.1175/MWR-D-21-0148.1.

Gottwald, G. A. and Majda, A. J.: A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks, Nonlin. Processes Geophys., 20, 705–712, https://doi.org/10.5194/npg-20-705-2013, 2013.

Gustafsson, N., J. Bojarova, and O. Vignes, 2014: A hybrid variational ensemble data assimilation for the High-Resolution Limited Area Model (HIRLAM). Nonlinear Processes Geophys., 21, 303–323, https://doi.org/10.5194/npg-21-303-2014.

Harnisch, F., & Keil, C. (2015). Initial conditions for convective-scale ensemble forecasting provided by ensemble data assimilation. Monthly Weather Review, 143(5), 1583-1600. https://doi.org/10.1175/MWR-D-14-00209.1

Harris, B. A., & Kelly, G. (2001). A satellite radiance-bias correction scheme for data assimilation. Quarterly Journal of the Royal Meteorological Society, 127(574), 1453-1468. https://doi.org/10.1002/qj.49712757418

Hodyss, D., W. F. Campbell, and J. S. Whitaker, 2016: Observation-Dependent Posterior Inflation for the Ensemble Kalman Filter. Mon. Wea. Rev., 144, 2667–2684, https://doi.org/10.1175/MWR-D-15-0329.1.

Houtekamer, P. L., Mitchell, H. L., & Deng, X. (2009). Model error representation in an operational ensemble Kalman filter. Monthly Weather Review, 137(7), 2126–2143. https://doi.org/10.1175/2008MWR2737.1

Johnson, A., Wang, X., Carley, J. R., Wicker, L. J., & Karstens, C. (2015). A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. Monthly Weather Review, 143(8), 3087–3108. https://doi.org/10.1175/mwr-d-14-00345.1

Jones, E. A., and X. Wang, 2023: A multi-resolution ensemble hybrid 4DEnVar with variable ensemble sizes to improve global and tropical cyclone track numerical prediction. Mon. Wea. Rev., in press, https://doi.org/10.1175/MWR-D-22-0186.1.

Kelly, D., Majda, A. J., & Tong, X. T. (2015). Concrete ensemble Kalman filters with rigorous catastrophic filter divergence. Proceedings of the National Academy of Sciences, 112(34), 10589-10594. https://doi.org/10.1073/pnas.1511063112

Kleist, D. T., and K. Ide, 2015a: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. Mon. Wea. Rev., 143, 452–470, https://doi.org/10.1175/MWR-D-13-00350.1.

——, and ——, 2015b: 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.

Kotsuki, S., Ota, Y., & Miyoshi, T. (2017). Adaptive covariance relaxation methods for ensemble data assimilation: Experiments in the real atmosphere. Quarterly Journal of the Royal Meteorological Society, 143(705), 2001-2015. https://doi.org/10.1002/qj.3060

Kumar, S. V., Reichle, R. H., Harrison, K. W., Peters-Lidard, C. D., Yatheendradas, S., & Santanello, J. A. (2012). A comparison of methods for a priori bias correction in soil moisture data assimilation. Water Resources Research, 48(3). https://doi.org/10.1029/2010WR010261

Kutty, G., and X. Wang, 2015: A comparison of the impacts of radiosonde and AMSU radiance observations in GSI based 3DEnsVar and 3DVar data assimilation systems for NCEP GFS. Adv. Meteor., 2015, 1–17, https://doi.org/10.1155/2015/280546.

Leutbecher, M., Buizza, R., & Isaksen, L. (2007). Ensemble forecasting and flow-dependent estimates of initial uncertainty. In Proc. ECMWF Workshop on Flow-Dependent Aspects of Data Assimilation (pp. 185–201).

Li, Z., J. C. McWilliams, K. Ide, and J. D. Farrara, 2015: A Multiscale 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., N. E. Bowler, A. M. Clayton, S. R. Pring, and D. Fairbairn, 2015: Comparison of hybrid-4DEnVar and hybrid-4DVar data assimilation methods for global NWP. Mon. Wea. Rev., 143, 212–229, https://doi.org/10.1175/MWR-D-14-00195.1.

Lu, X., and X. Wang, 2021: Improving the Four-Dimensional Incremental Analysis Update (4DIAU) with the HWRF 4DEnVar Data Assimilation System for Rapidly Evolving Hurricane Prediction. Mon. Wea. Rev., 149, 4027-4043, https://doi.org/10.1175/MWR-D-21-0068.1.

Nastrom, G. D., and K. S. Gage, 1985: A Climatology of Atmospheric Wavenumber Spectra of Wind and Temperature Observed by Commercial Aircraft. J. Atmos. Sci., 42, 950–960, https://doi.org/10.1175/1520-0469(1985)042<0950:ACOAWS>2.0.CO;2.

Otsuka, S., Kotsuki, S., & Miyoshi, T. (2016). Nowcasting with data assimilation: A case of global satellite mapping of precipitation. Weather and Forecasting, 31(5), 1409-1416. https://doi.org/10.1175/WAF-D-16-0039.1

Reynolds, C. A., Teiseira, J., & Mclay, J. G. (2008). Impact of stochastic convection on the ensemble transform. Monthly Weather Review, 136(11), 4517–4526. https://doi.org/10.1175/2008MWR2453.1

Sluka, T. C., Penny, S. G., Kalnay, E., & Miyoshi, T. (2016). Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. Geophysical Research Letters, 43(2), 752-759. https://doi.org/10.1002/2015GL067238

Terasaki, K., Kotsuki, S., & Miyoshi, T. (2019). Multi-year analysis using the NICAM-LETKF data assimilation system. SOLA, 15, 41-46. https://doi.org/10.2151/sola.2019-009

Tulloch, R., and K. S. Smith, 2009: A Note on the Numerical Representation of Surface Dynamics in Quasigeostrophic Turbulence: Application to the Nonlinear Eady Model. J. Atmos. Sci., 66, 1063–1068, https://doi.org/10.1175/2008JAS2921.1.

Wang, X., C. Snyder, and T. M. Hamill, 2007b: On the theoretical equivalence of differently proposed ensemble–3DVAR hybrid analysis schemes. Mon. Wea. Rev., 135, 222–227, https://doi.org/10.1175/MWR3282.1.

——, D. M. Barker, ——, and ——, 2008a: A hybrid ETKF–3DVAR data assimilation scheme for the WRF Model. Part I: Observation system simulation experiment. Mon. Wea. Rev., 136, 5116–5131, https://doi.org/10.1175/2008MWR2444.1.

——, ——, ——, and ——, 2008b: A hybrid ETKF–3DVAR data assimilation scheme for the WRF Model. Part II: Real observation experiments. Mon. Wea. Rev., 136, 5132–5147, https://doi.org/10.1175/2008MWR2445.1.

——., 2011: Application of the WRF hybrid ETKF–3DVAR data assimilation system for hurricane track forecasts. Wea. Forecasting, 26, 868–884, https://doi.org/10.1175/WAF-D-10-05058.1.

——, D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP global forecast system: Single-resolution experiments. Mon. Wea. Rev., 141, 4098–4117, https://doi.org/10.1175/MWR-D-12-00141.1.

——, H. G. Chipilski*, C. H. Bishop, E. Satterfield, N. Baker, J. Whittaker, 2020: A

Multiscale Local Gain Form Ensemble Transform Kalman Filter (MLGETKF). Mon. Wea. Rev., 149, 605-622, https://doi.org/10.1175/MWR-D-20-0290.1.

——, and T. Lei, 2014: GSI-based four-dimensional ensemble–variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP global forecast system. Mon. Wea. Rev., 142, 3303–3325, https://doi.org/10.1175/MWR-D-13-00303.1.

Wang, Y., and X. Wang, 2021: Rapid Update with EnVar Direct Radar Reflectivity Data Assimilation for the NOAA Regional Convection-Allowing NMMB Model over the CONUS: System Description and Initial Experiment Results. Atmosphere, 12, 1286, https://doi.org/10.3390/atmos12101286.

——, and ——, 2021: Development of Convective-Scale Static Background Error Covariance within GSI-Based Hybrid EnVar System for Direct Radar Reflectivity Data Assimilation. Mon. Wea. Rev., 149, 2713-2736, https://doi.org/10.1175/MWR-D-20-0215.1.

Whitaker, J. S., & Hamill, T. M. (2012). Evaluating methods to account for system errors in ensemble data assimilation. Monthly Weather Review, 140(9), 3078–3089. https://doi.org/10.1175/MWR-D-11-00276.1

Zeng, Y., Janjić, T., de Lozar, A., Blahak, U., Reich, H., Keil, C., & Seifert, A. (2018). Representation of model error in convective-scale data assimilation: additive noise, relaxation methods, and combinations. Journal of Advances in Modeling Earth Systems, 10(11), 2889-2911. https://doi.org/10.1029/2018MS001375

Zhang, F., Snyder, C., & Sun, J. (2004). Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. Monthly Weather Review, 132(5), 1238 – 1253. https://doi.org/10.1175/15200493(2004)132<1238:IOIEAO>2.0.CO;2

TABLES

Table 1: The name and brief introduction of four experiments

Exp. Name	Description
RTPS	Use RTPS for MLGETKF posterior inflation (Wang X. et al. 2021)
RTPS-SDI	Use scale dependent RTPS for MLGETKF posterior inflation
SE	Use SE inflation for MLGETKF posterior inflation
SE-SDI	Use scale dependent SE inflation for MLGETKF posterior inflation

Table 2: The spatiotemporally averaged RMS analysis error of potential temperature in the three phases for four experiments

Phase	RTPS [K]	RTPS-SDI [K]	SE [K]	SE-SDI [K]	RTPS-SDI vs RTPS	SE-SDI vs SE
Phase 1	2.312	1.449	2.118	1.666	37%	21%
Phase 2	1.381	0.968	1.201	0.912	30%	24%
Phase 3	0.850	0.821	0.807	0.800	3%	1%

The first column shows all phases. The second to fifth columns show the averaged RMS analysis error in each experiment. The sixth and seventh columns demonstrate the relative improvement for applying SDI in RTPS and SE, respectively.

FIGURES

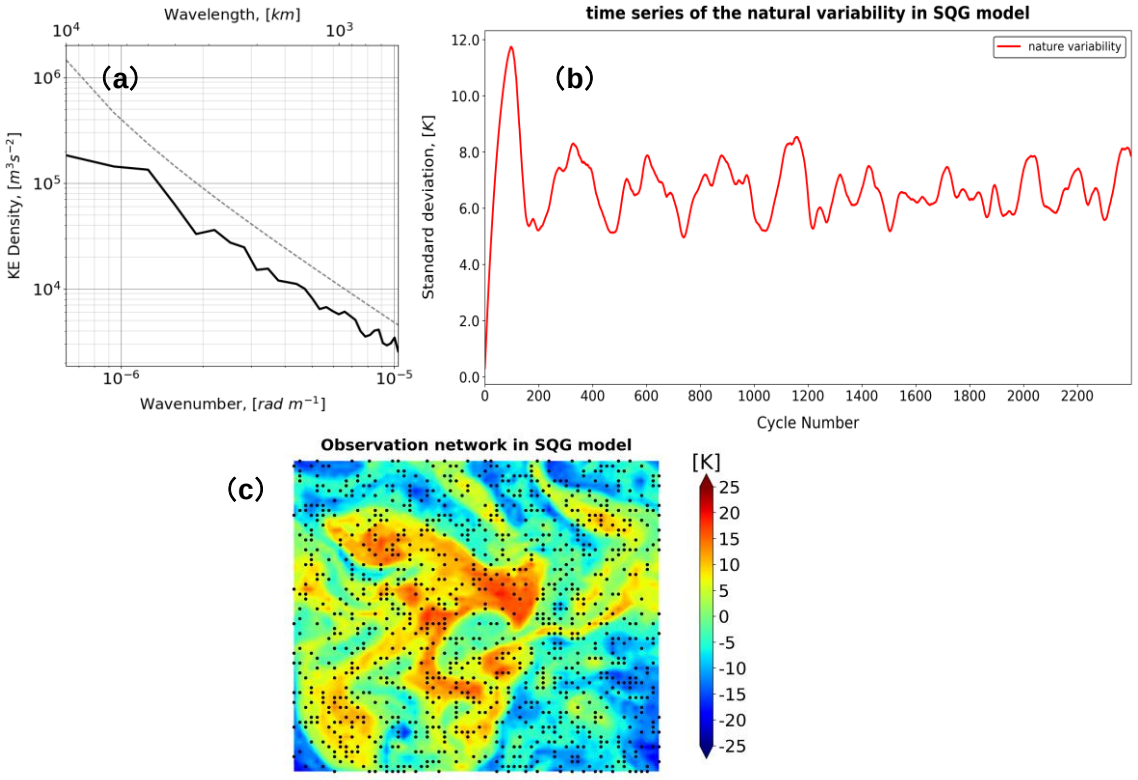

FIG. 1. Illustration of the SQG model. (a) Average KE density (solid) compared to a reference 5/3 spectral slope (Dashed). (b) Time series of the nature variability in the SQG model. (c) Spatial distribution of the simulated observations (black points) covered on a potential temperature anomaly field from the truth.

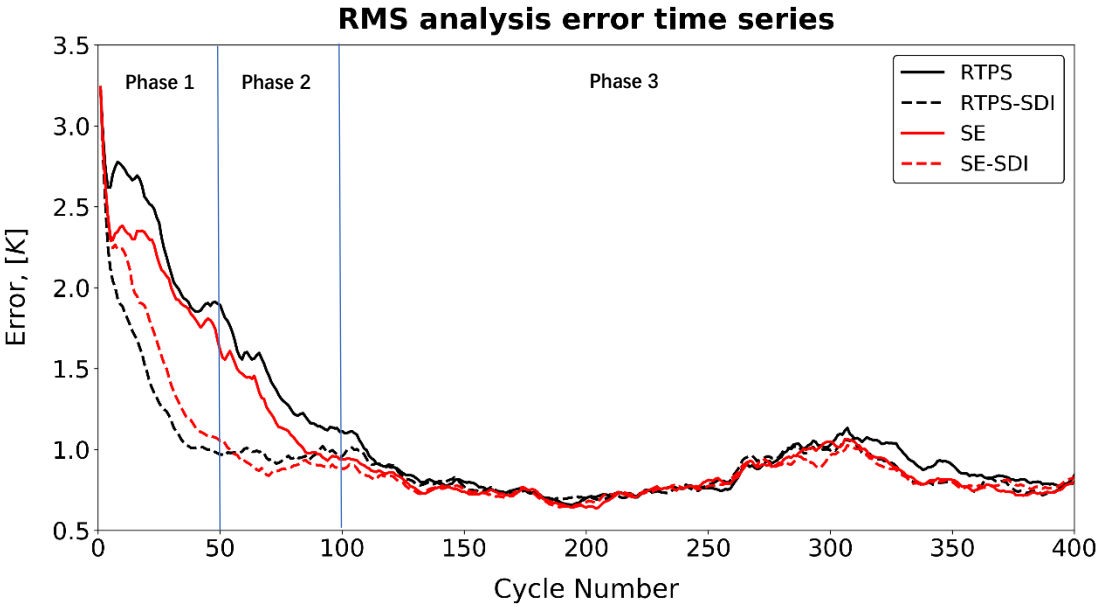

FIG. 2. The time series of the full-scale RMS potential temperature analysis errors. The black solid line, black dashed line, red solid line, and red dashed line represent RTPS, RTPS-SDI, SE, and SE-SDI, respectively.

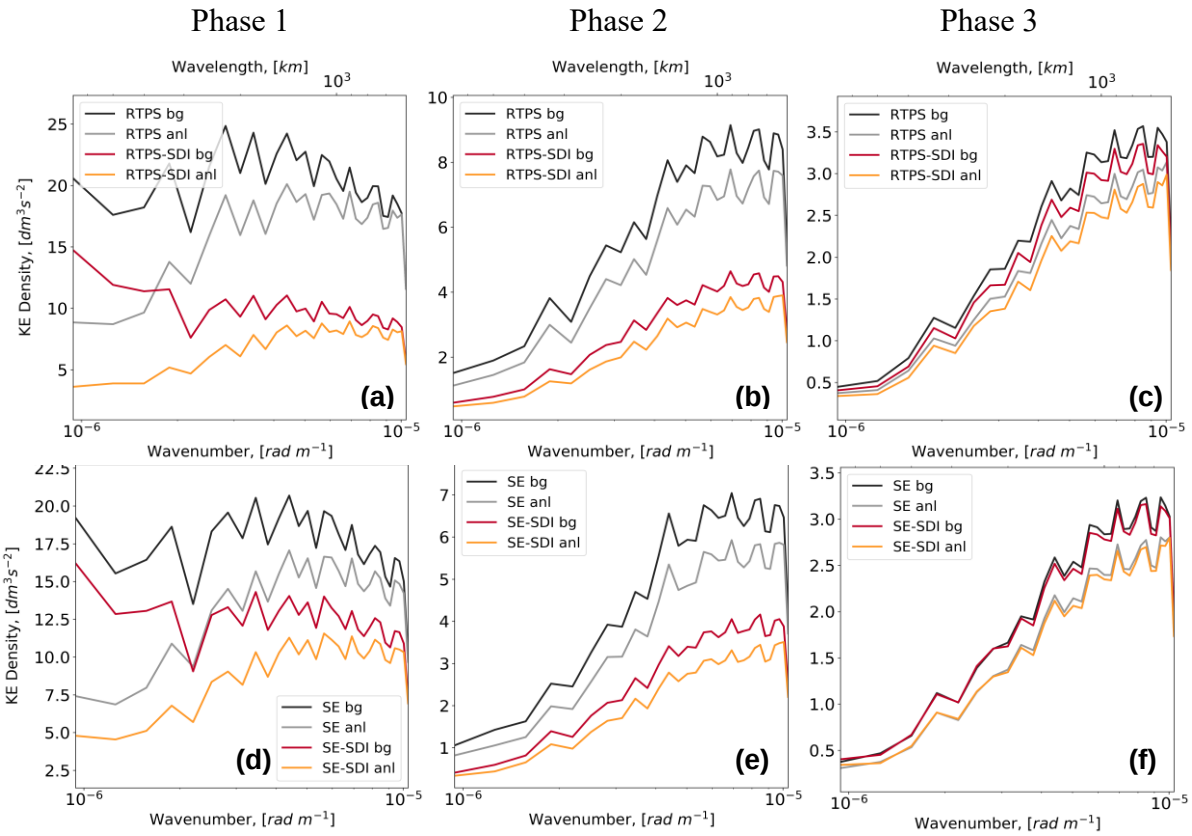

FIG. 3. Average kinetic energy (KE) spectra of the ensemble mean errors associated with different experiments. (a)-(c) comparison of RTPS and RTPS-SDI in phases 1, 2 and 3, respectively. RTPS is represented by black (background) and gray (analysis), RTPS-SDI is represented by red (background) and yellow (analysis). (d)-(f) comparison of SE and SE-SDI in different phases. Errors in SE use black (background) and gray (analysis) lines, and errors in SE-SDI use red (background) and yellow (analysis) lines.

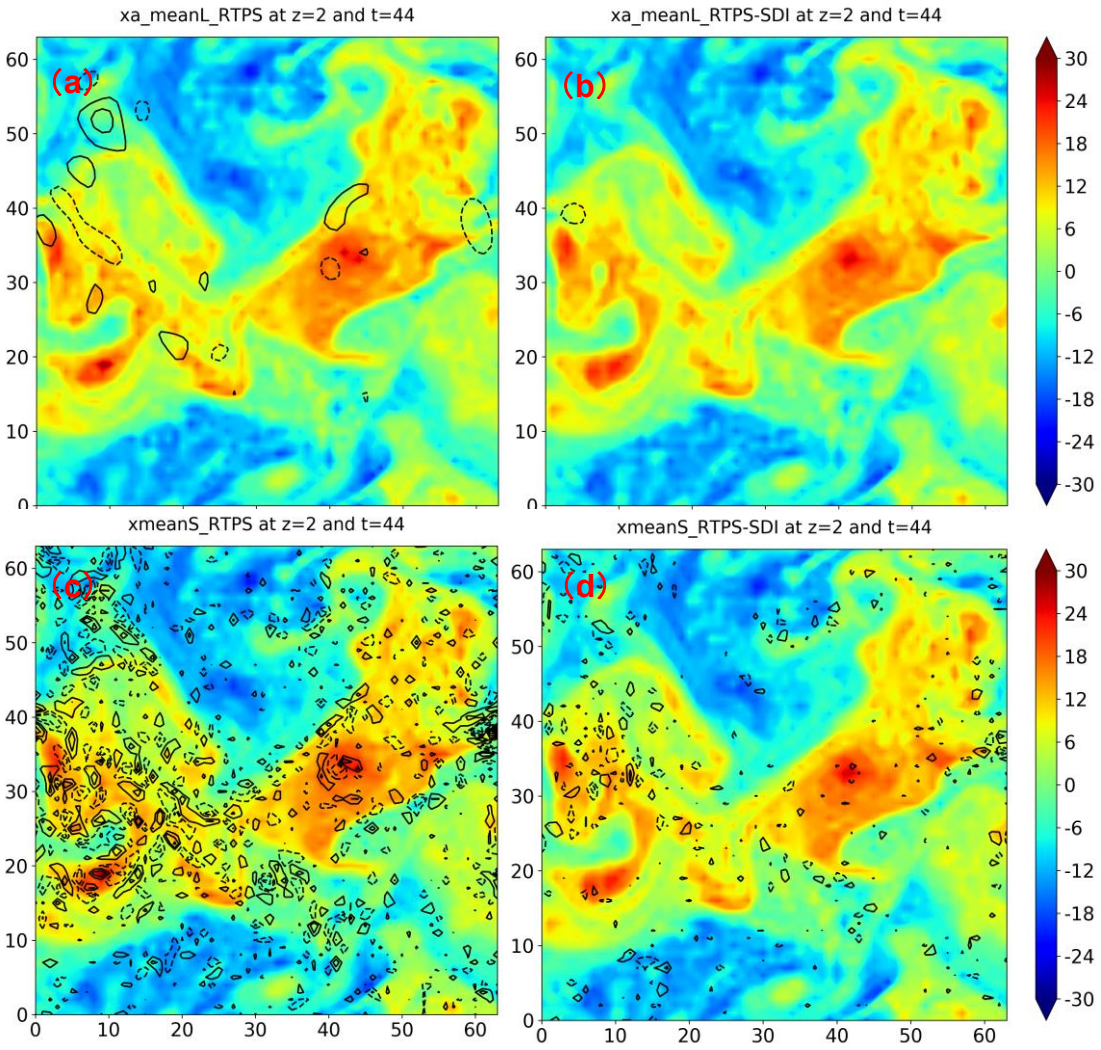

FIG. 4. Different scale's potential temperature ensemble mean analysis error at the 44th cycle (upper layer) for RTPS and RTPS-SDI. (a) RTPS (large scale), (b) RTPS-SDI (large scale), (c) RTPS (small scale), and (d) RTPS-SDI (small scale). The ensemble mean analysis errors are demonstrated as black contours with ranks of $\pm 1.5, \pm 3.5, \text{and} \pm 7K$. Solid (dashed) contours represent the analysis errors above (below) the ensemble mean values. The background shadow is the upper-level analysis mean potential temperature anomaly.

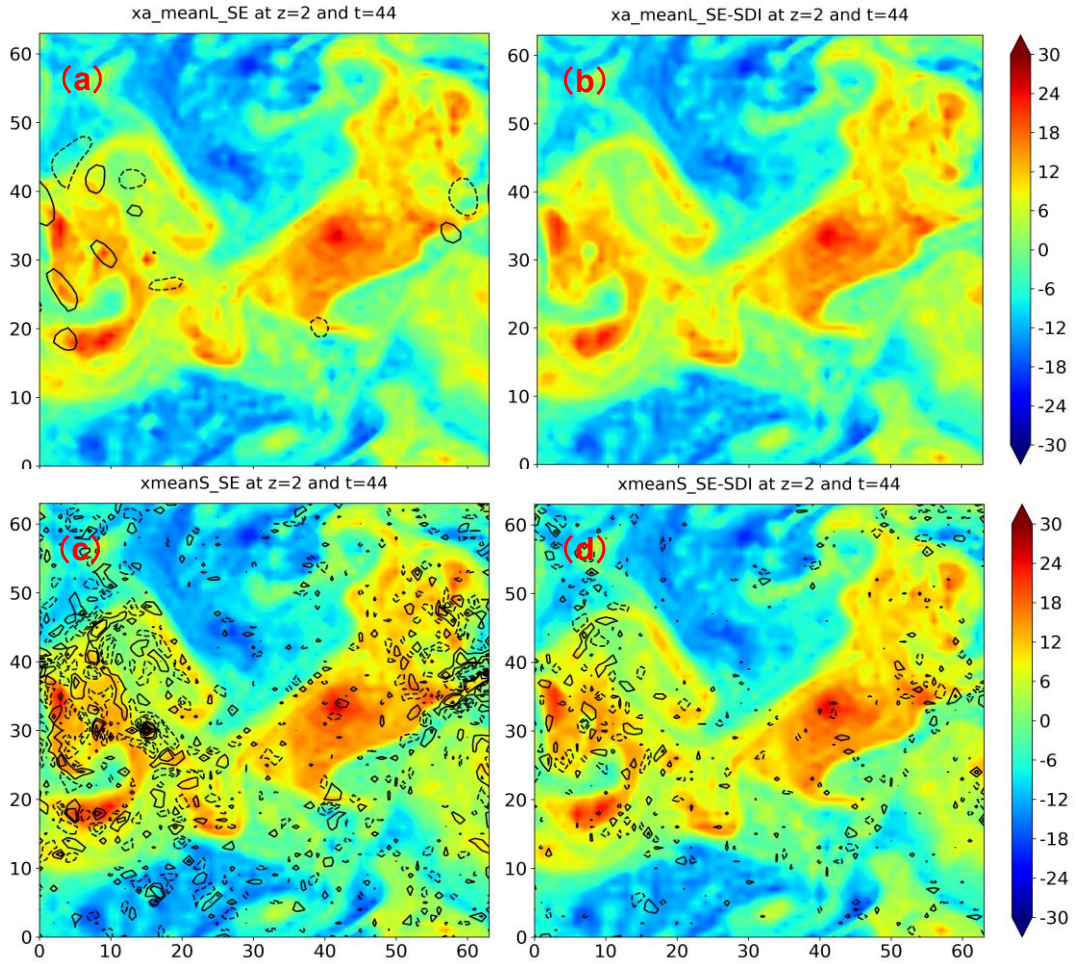

FIG. 5. Different scale's potential temperature ensemble mean analysis error at the 44th cycle (upper level) for SE and SE-SDI. (a) SE (large scale), (b) SE-SDI (large scale), (c) SE (small scale), and (d) SE-SDI (small scale). The analysis errors are demonstrated as black contours.

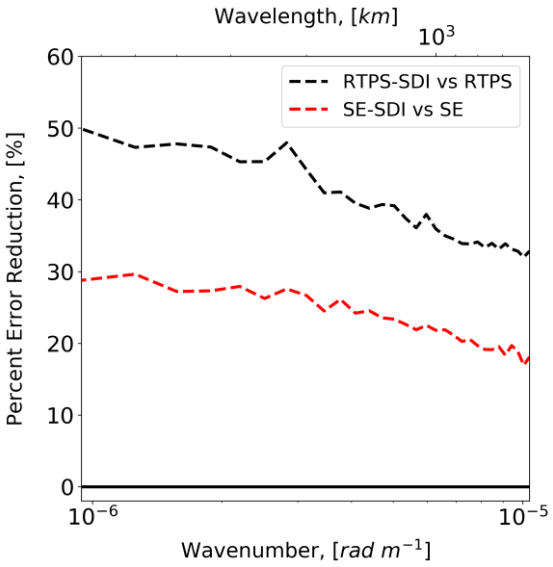

FIG. 6. Relative analysis error reduction in spectral space. Relative average KE analysis error reduction in SDI approaches relative to single-scale inflation in spectral space. The black dashed line is RTPS-SDI vs. RTPS, and the red dashed line is SE-SDI vs. SE

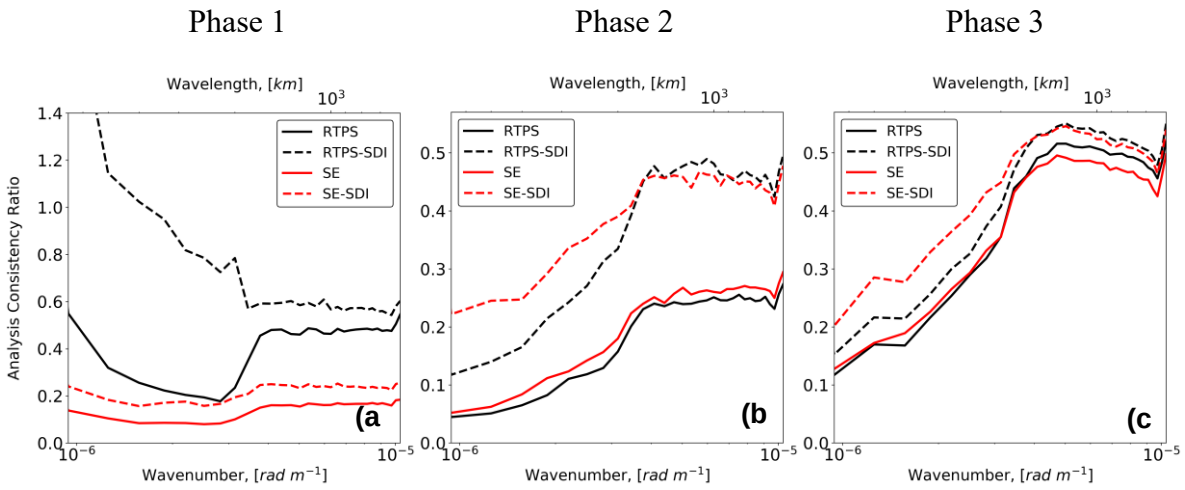

FIG. 7. Scale-dependent analysis consistency ratio in terms of KE norm for RTPS (black solid lines), RTPS-SDI (black dashed lines), SE (red solid lines), and SE-SDI (red dashed lines). (a)-(c) represents the consistency ratio in Phases 1, 2, and 3, respectively.

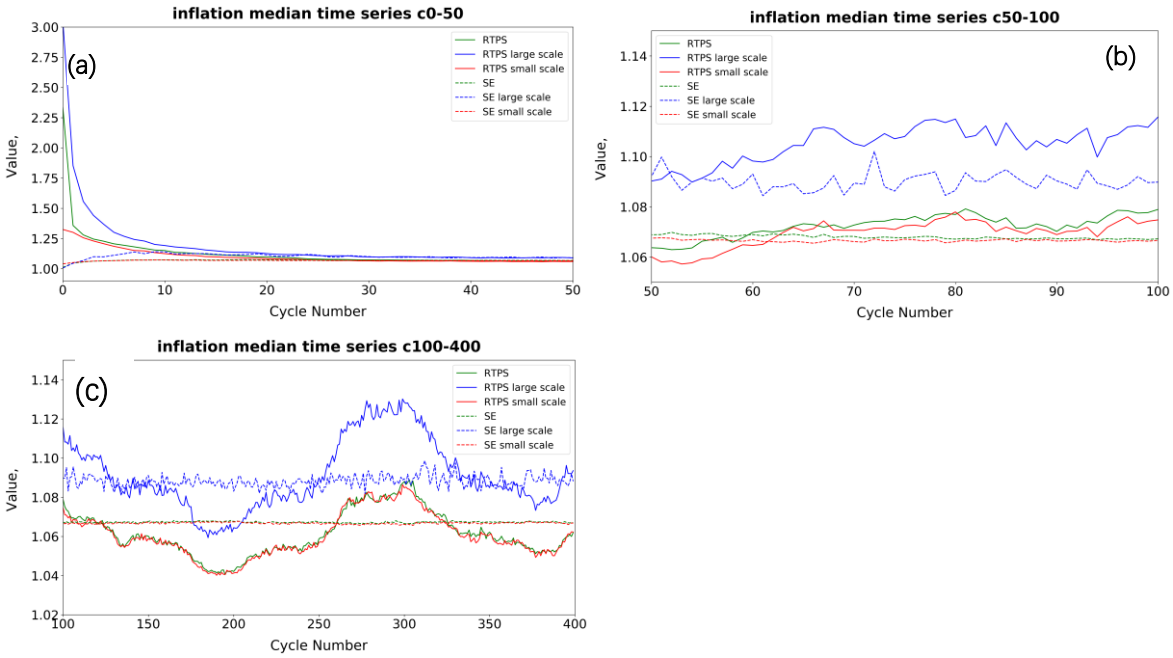

FIG. 8. The temporal distribution of the inflation factors. (a)-(c) inflation median in Phases 1,2, and 3, respectively. The green lines represent two single scale inflation: RTPS (solid) and SE (dashed). The blue and red lines represent large- and small-scale inflation in RTPS-SDI (solid) and SE-SDI (dashed).

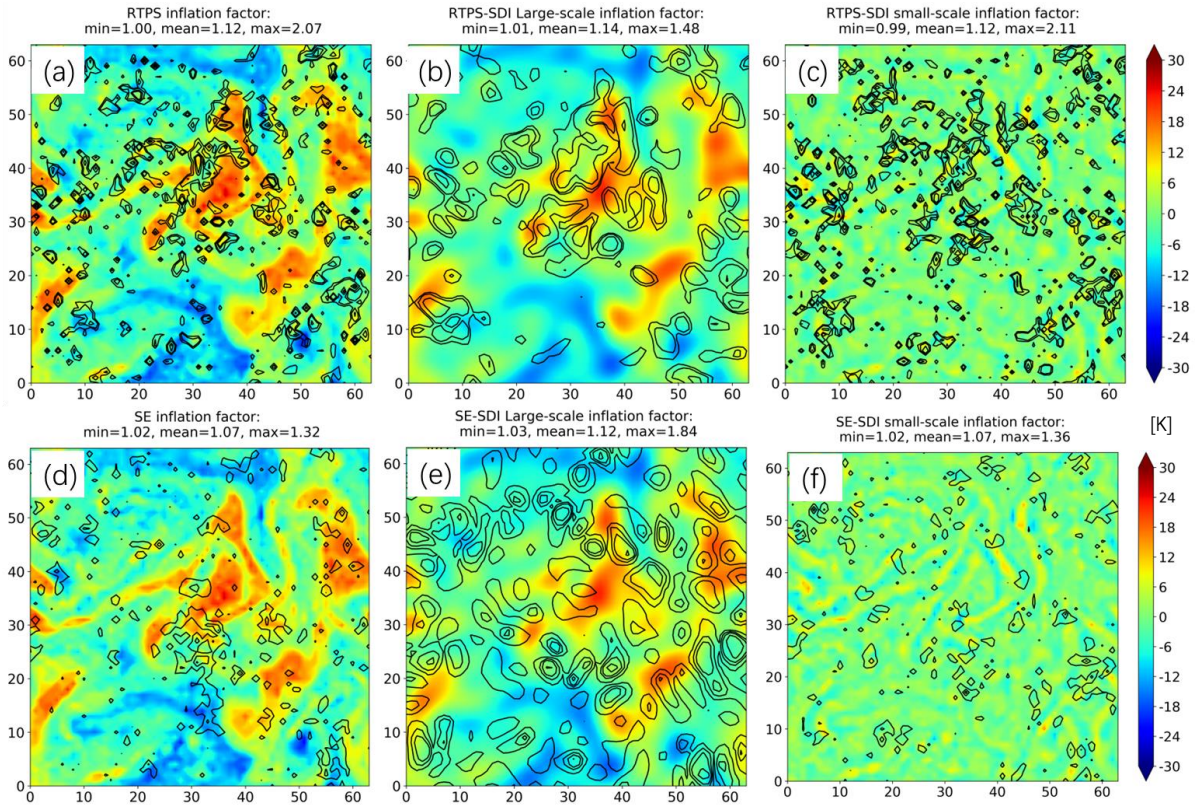

FIG. 9. The spatial distribution of the inflation factors overlay on the background at cycle 300, layer 2. The superscripts represent different inflation factors: (a) RTPS, (b) RTPS-SDI at large scale, (c) RTPS-SDI at small scale, (d) SE, (e) SE-SDI at large scale, and (f) SE-SDI at small scale. The inflation factors are demonstrated as black contours with ranks of [1.16,1.2,1.3,1.4] for (a)(b)(c) and with rank of [1.1,1.2,1.3,1.4] for (d)(e)(f). The color bar is the background field at the corresponding scales: (a)(d): Full-scale, (b)(e): Large-scale, (c)(f): Small-scale.

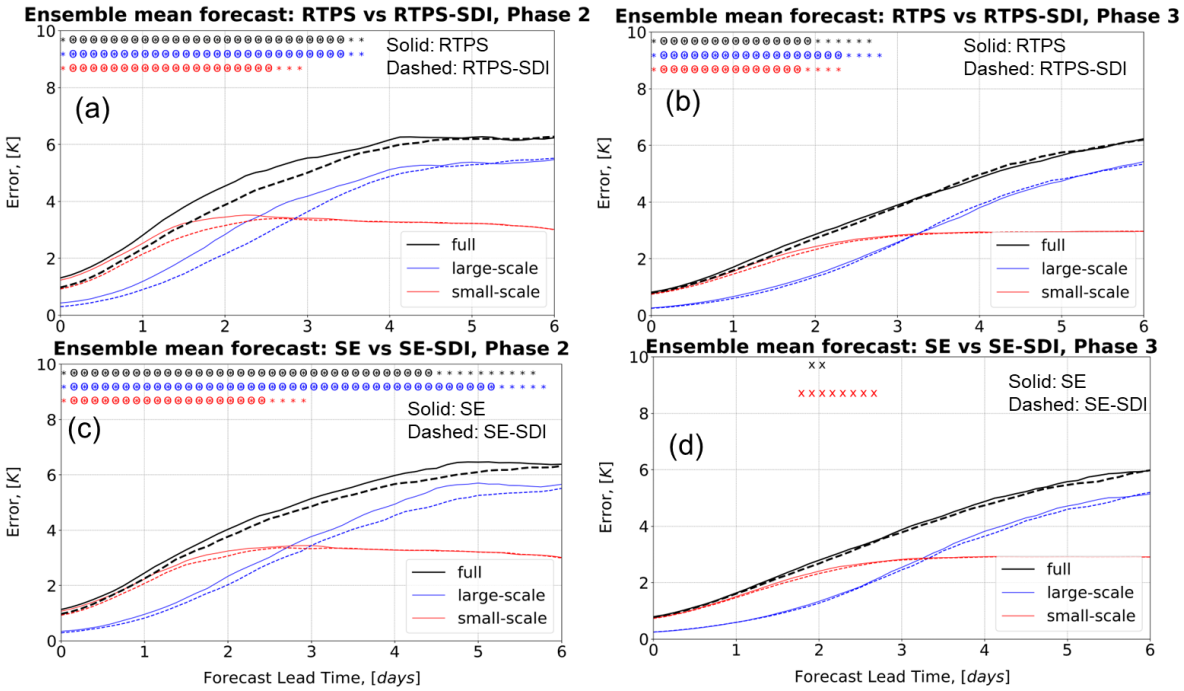

FIG. 10. Forecast performance in terms of average potential temperature ensemble mean forecast errors. Full-scale forecast error (black) is plotted with the decomposed large (blue)- and small (red)- scale forecast error. (a)(b) compares RTPS (solid) to RTPS-SDI (dash) in Phases 2 and 3, respectively. (c)(d) compares SE (solid) to SE-SDI (dash) in Phases 2 and 3, respectively. The annotations are the results of the significance test (based on a two-sides Mann-Whitney U-test). '*' represents the forecast improvement of SDI methods compared to their scale-unaware counterparts passing the 95% confidence level in that time, and '⁂' represents the improvement passing the 99.5% confidence level. 'x' represents the improvement passing the 90% confidence level. Different colors of annotations represent the significance test taken in full (black)-, large (blue)-, and small (red)-scales.

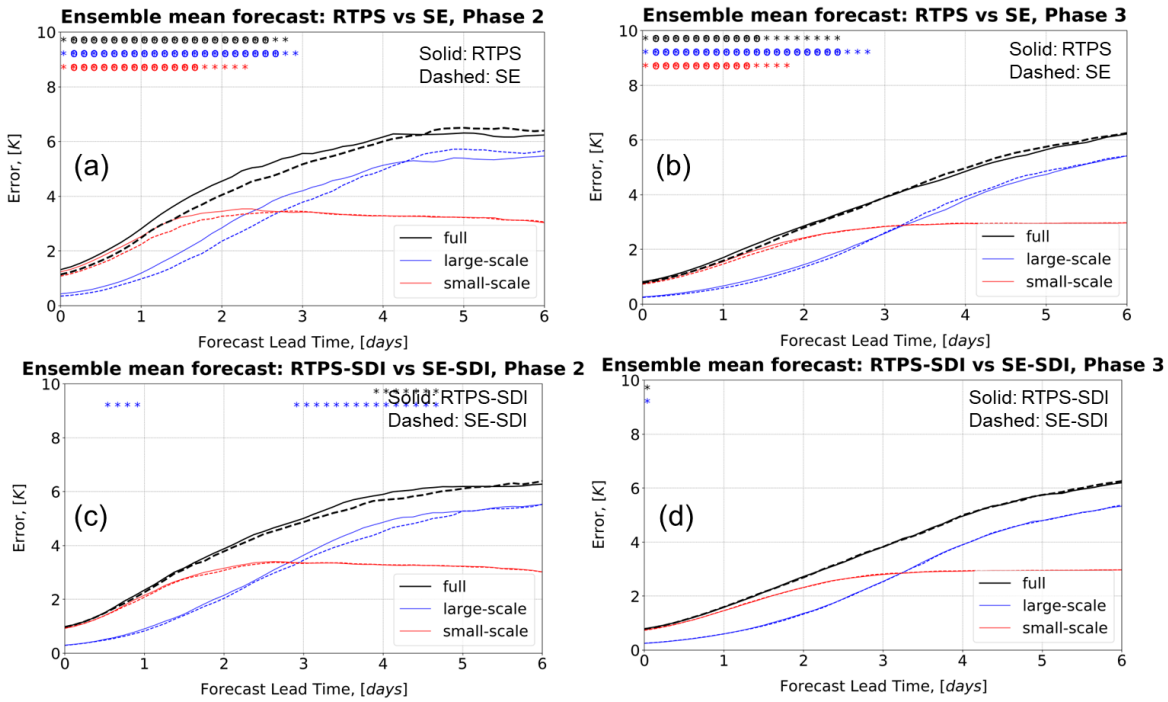

FIG. 11. Comparing Forecast performance in RTPS-based and SE-based methods. (a)(b) compares RTPS (solid) to SE (dash) in Phases 2 and 3, respectively. (c)(d) compares RTPS-SDI (solid) to SE-SDI (dash) in Phases 2 and 3, respectively. '*' represents the forecast errors in SE(SE-SDI) is significantly smaller than RTPS(RTPS-SDI) passing the 95% confidence level, and '⊛' represents the SE(SE-SDI) forecast improvement passing the 99.5% confidence level compared to RTPS(RTPS-SDI).