Variable-dependent and selective multivariate localization for ensemble-variational data assimilation in the tropics

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1175/MWR-D-23-0201.1

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Variable-Dependent and Selective Multivariate Localization for Ensemble–Variational Data Assimilation in the Tropics

JOSHUA CHUN KWANG LEE [a,b], JAVIER AMEZCUA [c,d,e], AND ROSS NOEL BANNISTER [c,d]

[a] *Department of Meteorology, University of Reading, Reading, United Kingdom*
[b] *Centre for Climate Research Singapore, Meteorological Service Singapore, Singapore*
[c] *University of Reading, Reading, United Kingdom*
[d] *National Centre for Earth Observation, Reading, United Kingdom*
[e] *Tecnologico de Monterrey, Campus Ciudad de Mexico, Ciudad de Mexico, Mexico*

ABSTRACT: Two aspects of ensemble localization for data assimilation are explored using the simplified nonhydrostatic ABC model in a tropical setting. The first aspect (i) is the ability to prescribe different localization length scales for different variables (variable-dependent localization). The second aspect (ii) is the ability to control (i.e., to knock out by localization) multivariate error covariances (selective multivariate localization). These aspects are explored in order to shed light on the cross-covariances that are important in the tropics and to help determine the most appropriate localization configuration for a tropical ensemble–variational (EnVar) data assimilation system. Two localization schemes are implemented within the EnVar framework to achieve (i) and (ii). One is called the isolated variable-dependent localization (IVDL) scheme and the other is called the symmetric variable-dependent localization (SVDL) scheme. Multicycle observation system simulation experiments are conducted using IVDL or SVDL mainly with a 100-member ensemble, although other ensemble sizes are studied (between 10 and 1000 members). The results reveal that selective multivariate localization can reduce the cycle-averaged root-mean-square error (RMSE) in the experiments when cross-covariances associated with hydrostatic balance are retained and when zonal wind/mass error cross-covariances are knocked out. When variable-dependent horizontal and vertical localization are incrementally introduced, the cycle-averaged RMSE is further reduced. Overall, the best performing experiment using both variable-dependent and selective multivariate localization leads to a 3%–4% reduction in cycle-averaged RMSE compared to the traditional EnVar experiment. These results may inform the possible improvements to existing tropical numerical weather prediction systems that use EnVar data assimilation.

KEYWORDS: Tropics; Data assimilation; Ensembles

1. Introduction

Ensemble–variational (EnVar) data assimilation methods have recently gained traction and have been widely tested in several operational global and regional numerical weather prediction (NWP) systems (Buehner et al. 2013; Clayton et al. 2013; Wang et al. 2013; Gustafsson et al. 2014; Hu et al. 2017; Montmerle et al. 2018; Singh and Prasad 2019; Bédard et al. 2020; Kadowaki et al. 2020) and in research case studies focusing on extreme weather events (Schwartz et al. 2013; Shen et al. 2016; Lu et al. 2017; Gao et al. 2019; Kutty et al. 2020). The main idea relies on using ensemble-derived background error statistics to replace the climatological error statistics used in the variational approach. Often, a hybrid EnVar approach is adopted by weighting the ensemble-derived and climatological error statistics with respective weights that depend on the ensemble size (Hamill and Snyder 2000). Where the ensemble is sufficiently large, one might solely rely on ensemble-derived error statistics by placing full weight on it in the variational algorithm.

Most studies reported a benefit from using EnVar data assimilation, either in the hybrid or pure EnVar form, as opposed to traditional three-dimensional or four-dimensional variational (3D-Var or 4D-Var) approaches. They attributed the benefit broadly to the flow dependency introduced by the ensemble-derived error statistics. This flow dependency generically encompasses time appropriateness of error variances and flow consistency of spatial covariances, as well as flow consistency of multivariate error relationships (i.e., cross-covariances between different variables in the model). The benefit stemming from each component has yet to be clearly distinguished. Shen et al. (2016) and Lu et al. (2017) highlighted through tropical cyclone case studies that using ensemble-derived error statistics yielded more realistic multivariate error cross-covariances, particularly in the vicinity of the cyclone vortex. Johnson et al. (2015) and Gao et al. (2019) found, through case studies over the United States and China, that using ensemble-derived error statistics resulted in more dynamically coherent analyses of mesoscale convective systems. In these case studies, the flow consistency of multivariate error relationships from the ensemble-derived error statistics was found to be a key contributor leading to the improvements in the quality of the analysis and subsequent forecasts.

Corresponding author: Joshua Chun Kwang Lee, joshua.lee@pgr.reading.ac.uk

While capturing appropriate multivariate error relationships may help to retrieve a dynamically consistent and balanced analysis, sampling noise may also contaminate the error covariances. This is because the ensemble size is usually far smaller than the degrees of freedom of the state, and therefore, the estimated error covariance matrix will be rank-deficient. Houtekamer and Mitchell (2001) suggested using a Schur product of a correlation matrix (referred to as a localization matrix) with the ensemble-derived background error covariance matrix to apply spatial covariance localization and mitigate spurious long-range correlations. This is widely adopted in traditional EnVar implementations (e.g., in Wang et al. 2008), especially in most operational weather prediction centers adopting EnVar techniques. However, there are two potential limitations with current traditional approaches. First, the same spatial localization is usually applied to all variables, irrespective of their characteristic length scales associated with the system dynamics. For example, Huang et al. (2021) and Caron and Buehner (2022) specified variable-independent localization scales. This assumption of the same spatial localization was shown to be rather unrealistic (Lei et al. 2015), especially at convective scales (Destouches et al. 2021; Necker et al. 2023). The error cross-covariance localization should ideally also reflect a mix of the characteristic length scales of the variables involved, but again, this is not usually done in traditional EnVar implementations. Second, in traditional EnVar schemes, the ability to do multivariate localization (the knocking out of correlations between variables) is not currently implemented, so multivariate error relationships between all variables are determined by the spatially localized ensemble. In the absence of a reasonable physically based constraint, some of the ensemble-derived relationships may be useful, as seen in tropical cyclone case studies (Shen et al. 2016; Lu et al. 2017) and over Southeast Asia (Lee and Barker 2023). However, other cross-covariances and their characteristic length scales may not be well represented by a limited-size ensemble and are likely to be dominated by sampling noise. It is possible that these cross-covariances are noninformative and may predominantly be introducing noise to the analysis yet cannot be removed using traditional localization frameworks.

In the tropics, the disadvantages of the abovementioned two potential limitations may become more apparent given the nature of convective weather and less balanced flow in the region. One would desire, for instance, the flexibility to prescribe smaller localization length scales for convection-related variables, which may typically involve vertical wind and hydrometeors, following Destouches et al. (2021). Additionally, appropriate multivariate localization may also be required since it is not trivial to specify multivariate error relationships for the tropics, particularly between mass (e.g., temperature and pressure) and wind variables. Some operational systems prescribe geostrophic balance or linear balance in their background error covariance model (e.g., Lorenc et al. 2000), but in the tropics, this procedure effectively treats the mass and wind variables univariately since the Coriolis parameter is small there. It remains to be seen if all multivariate error relationships estimated by an ensemble are physically meaningful or even required in the tropics. In this light, an improved EnVar implementation for the tropics should allow for *variable-dependent localization* (a concept suggested by Necker et al. 2020) and a way to constrain the multivariate error relationships (keeping some cross-covariances and not others), which we term as *selective multivariate localization*. Notwithstanding this, neither have been explored in the tropics yet.

To this end, one possible modification to traditional EnVar is to use scale-dependent localization (Buehner and Shlyaeva 2015; Huang et al. 2021; Caron and Buehner 2022), which allows the localization length scales at different scales to be specified independently. Therefore, the large-scale and small-scale errors are allowed to have different error characteristics. However, in these studies, for each scale, all variables still share the same localization length scales. Wang and Wang (2023) further extended this to include both variable-dependent localization and scale-dependent localization, for a few tornadic supercell case studies over the United States. They introduced an approach to modify traditional EnVar and found that further applying variable-dependent localization was beneficial to see storm maintenance. Another possible modification was proposed by Stanley et al. (2021) to construct separate localization functions for the multivariate error cross-covariances (within a bivariate Lorenz-96 system with coupled data assimilation), but this has yet to be applied to the EnVar framework. Additionally, the approach does not allow one to select specific cross-covariances to be retained.

In this study, we implement and explore two approaches to grant the ability to apply variable-dependent and selective multivariate localization within the EnVar data assimilation framework. The first approach is termed as the *isolated variable-dependent localization (IVDL) scheme*. The second approach is termed as the *symmetric variable-dependent localization (SVDL) scheme*. Details are given in section 2, along with their similarities or novelties vis-à-vis existing schemes. Both schemes allow for variable-dependent localization; IVDL is not only more computationally efficient but also is less flexible than SVDL. By design, the IVDL scheme implicitly determines the multivariate localization, while the SVDL scheme explicitly prescribes the multivariate localization on top of spatial localization. This study aims to answer the following questions:

1) How many ensemble members are sufficient to show a significant degree of "signal" in the covariances, but still benefit from localization of sampling noise?
2) Which multivariate error relationships in the ensemble-derived error covariances are important/beneficial for EnVar data assimilation in the tropics?
3) Is variable-dependent spatial localization beneficial for EnVar data assimilation in the tropics?

Section 2 describes the design and implementation of IVDL and SVDL schemes. A simplified nonhydrostatic convective-scale model, the ABC model (Petrie et al. 2017), is used for this study. A tropical configuration of the ABC model (a longitude–height dry model, without diabatic processes) with data assimilation is set up to demonstrate the two schemes. Section 3 provides further details on the model and data assimilation framework. Section 4 describes the experiments and gives guidance on the ensemble size (question 1). When it comes to deciding on a suitable ensemble size, we consider the linear independence of the

ensemble members and sampling error. Section 5 evaluates the two schemes through empirical experiments. Due to the cheaper computational cost, we use only the IVDL scheme to explore selective multivariate localization by controlling which multivariate error relationships are retained (question 2), but we use both schemes to explore variable-dependent spatial localization (question 3). Section 6 discusses and concludes the results of this study.

2. Variable-dependent and selective multivariate localization applied with IVDL and SVDL

This work builds on the implementation in Lee et al. (2022, hereafter L22), who introduced hybrid EnVar data assimilation via the alpha control variable approach (Lorenc 2003) for the ABC model (section 3). We shall follow their notation for consistency. For the pure EnVar approach, a given alpha control variable transform $\mathbf{U}^\alpha$ acts on an alpha control vector $\boldsymbol{\chi}^{\alpha k}$ to give an alpha field (i.e., $\boldsymbol{\alpha}^k = \mathbf{U}^\alpha \boldsymbol{\chi}^{\alpha k}$), which controls the linear combination of ensemble perturbations $\mathbf{x}'^k_t$. There is one alpha control variable (and hence one alpha field) per ensemble perturbation member. The analysis increment $\delta\mathbf{x}$ at time t is then

$$\delta\mathbf{x} = \sum_{k=1}^{N} \mathbf{x}'^k_t \circ \boldsymbol{\alpha}^k, \tag{1}$$

where N is the number of ensemble members, k is the ensemble member index, and $\circ$ is the Schur product. The implied localization matrix $\mathbf{L}$ in the variational algorithm is $\mathbf{L} = \mathbf{U}^\alpha \mathbf{U}^{\alpha T}$, so changing the design of $\mathbf{U}^\alpha$ is key to controlling the application of variable-dependent and selective multivariate localization. Following Eq. (15) from L22 for the pure EnVar approach, the implied background error covariance matrix is therefore

$$\mathbf{B}_e = (\mathbf{U}^\alpha \mathbf{U}^{\alpha T}) \circ (\mathbf{X}^f_t \mathbf{X}^{fT}_t), \tag{2}$$

where $\mathbf{X}^f_t$ is the matrix whose columns contain the ensemble perturbations $\mathbf{x}'^k_t$ divided by $\sqrt{N-1}$.

a. The IVDL scheme

We first describe the implementation of IVDL. In appendix B of L22, a proof of equivalence between their approach in designing $\mathbf{U}^\alpha$ and the traditional EnVar approach of Wang et al. (2008)—who presented it slightly differently—is provided. L22 further showed that their choice of $\mathbf{U}^\alpha$ does full intervariable localization (where no multivariate error cross-covariances are retained). In Wang et al. (2008), the length of $\boldsymbol{\chi}^{\alpha k}$ for each ensemble member k is given by the number of horizontal grid points N_g, whereas in L22, the length of $\boldsymbol{\chi}^{\alpha k}$ is further multiplied by the number of prognostic variables N_{var}. This obviously influences the dimensions of $\mathbf{U}^\alpha$ [the number of columns of $\mathbf{U}^\alpha$ must match the length of $\boldsymbol{\chi}^{\alpha k}$ and the number of rows must match the length of the state perturbations $\mathbf{x}'^k_t$ for the Schur product in Eq. (1)]. Here, we reproduce the parts of the L22 proof to illustrate how variable-dependent and selective multivariate localization can be implemented. For an arbitrary state with three one-dimensional (horizontal) physical variables (e.g., p, q, and r) per grid point ($\mathbf{x} \in \mathbb{R}^{3N_g}$),

the alpha control variable transform implied by Wang et al. (2008) can be mathematically represented by the choice $\mathbf{U}^\alpha = \tilde{\mathbf{U}}^\alpha$:

$$\tilde{\mathbf{U}}^\alpha = \begin{bmatrix} \mathbf{U}^\alpha_p \\ \mathbf{U}^\alpha_q \\ \mathbf{U}^\alpha_r \end{bmatrix}, \tag{3}$$

in contrast to the approach in L22, which takes the choice $\mathbf{U}^\alpha = \hat{\mathbf{U}}^\alpha$:

$$\hat{\mathbf{U}}^\alpha = \begin{bmatrix} \mathbf{U}^\alpha_p & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}^\alpha_q & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}^\alpha_r \end{bmatrix}. \tag{4}$$

Here, $\mathbf{0}$ is an $N_g \times N_g$ null matrix and $\mathbf{U}^\alpha_p$, $\mathbf{U}^\alpha_q$, and $\mathbf{U}^\alpha_r$ can each have the form of an eigenvector matrix scaled by the square root of the eigenvalue matrix associated with the eigendecomposition of a spatial correlation (localization) matrix for a specified length scale h^α. For example, $\mathbf{U}^\alpha_r = \mathbf{F}_r \mathbf{\Lambda}_r^{1/2}$, where $\mathbf{F}_r$ contains the eigenvectors and $\mathbf{\Lambda}_r$ contains the eigenvalues for variable r. For this study, we have used a Gaspari–Cohn localization function (Gaspari and Cohn 1999; see L22 for details) to prescribe the correlation matrix.

Note that Eq. (3) presents a mathematically consistent interpretation of the approach in Wang et al. (2008). In their implementation, they use recursive filters instead of the eigen-approach mentioned above and apply the same transform to each variable, i.e., $\mathbf{U}^\alpha_p = \mathbf{U}^\alpha_q = \mathbf{U}^\alpha_r$. This precludes the possibility of using a different length scale for each variable (although this can be relaxed if required). In the L22 approach, however, a different spatial correlation matrix for each variable is used to achieve variable-dependent localization but forces full multivariate localization. Note that extra memory cost is required to store the eigenvectors and eigenvalues for each variable. In our implementation, this is the case even if, e.g., two variables share the same h^α; the same eigenvectors and eigenvalues are stored twice—once for each variable.

Next, we show how selective multivariate localization can be achieved in the IVDL scheme. Equation (4) shows the most primitive form of $\mathbf{U}^\alpha$. This can be considered one limiting/extreme case where full intervariable localization is implied because of its design. In L22, they further highlighted that it was possible to extend Eq. (4) to include selective multivariate localization by introducing a mapping matrix $\hat{\mathbf{I}}$ comprising scaled blocks of either null or identity matrices ($\hat{\mathbf{U}}^\alpha \hat{\mathbf{I}}$; see below). We refer to this extension using the variants of $\hat{\mathbf{I}}$ as selective multivariate localization because we can select which block matrices of $\hat{\mathbf{I}}$ are identity matrices and which are null matrices. If all block matrices in $\hat{\mathbf{I}}$ are identity matrices, we get the other limiting/extreme case where all error cross-covariances are retained (no intervariable localization). Here, $\hat{\mathbf{I}}$ is given in this case by

$$\hat{\mathbf{I}} = \frac{1}{\sqrt{3}} \begin{bmatrix} \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} \\ \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} \\ \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} \end{bmatrix}, \tag{5}$$

where $\mathbf{I}_{N_g}$ is the $N_g \times N_g$ identity matrix and 3 is the number of variables whose cross-covariances are retained (all $N_{\text{var}} = 3$ variables in this case). We can then choose $\mathbf{U}^\alpha$ to be given by

$$\mathbf{U}^\alpha = \mathbf{U}^\alpha\hat{\mathbf{I}} = \frac{1}{\sqrt{3}}\begin{bmatrix} \mathbf{U}_p^\alpha & \mathbf{U}_p^\alpha & \mathbf{U}_p^\alpha \\ \mathbf{U}_q^\alpha & \mathbf{U}_q^\alpha & \mathbf{U}_q^\alpha \\ \mathbf{U}_r^\alpha & \mathbf{U}_r^\alpha & \mathbf{U}_r^\alpha \end{bmatrix}. \tag{6}$$

One can compute the implied localization matrices $\mathbf{L}$ using $\tilde{\mathbf{U}}^\alpha$ and $\hat{\mathbf{U}}^\alpha\hat{\mathbf{I}}$, from Wang et al. (2008) and L22 ($\tilde{\mathbf{U}}^\alpha\tilde{\mathbf{U}}^{\alpha\text{T}}$ and $\hat{\mathbf{U}}^\alpha\hat{\mathbf{I}}\hat{\mathbf{U}}^{\alpha\text{T}}$, respectively), to see that they are equivalent (proven elementwise in appendix B of L22).

Now, consider a variant of $\hat{\mathbf{I}}$ where only p and q cross-covariances are retained in the localization scheme, and $\hat{\mathbf{I}}$ is then given by

$$\hat{\mathbf{I}} = \begin{bmatrix} \frac{1}{\sqrt{2}}\mathbf{I}_{N_g} & \frac{1}{\sqrt{2}}\mathbf{I}_{N_g} & \mathbf{0} \\ \frac{1}{\sqrt{2}}\mathbf{I}_{N_g} & \frac{1}{\sqrt{2}}\mathbf{I}_{N_g} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{\sqrt{1}}\mathbf{I}_{N_g} \end{bmatrix}. \tag{7}$$

Here, we have partitioned the variables into two groups with two parameters and one parameter, respectively, where variables p and q are allowed to be correlated in the assimilation, but each is uncorrelated with r. With this setup, many permutations of selective multivariate localization are possible, depending on how the sets are determined. Each set of variables is treated independently in $\boldsymbol{\chi}^{\alpha k}$ (i.e., as a partition) by knocking out selected multivariate error cross-covariances. We can verify that the implied localization matrix $\mathbf{L}$ of Eq. (7) is given by

$$\mathbf{L} = \hat{\mathbf{U}}^\alpha\hat{\mathbf{I}}\hat{\mathbf{U}}^{\alpha\text{T}} = \begin{bmatrix} \mathbf{U}_p^\alpha\mathbf{U}_p^{\alpha\text{T}} & \mathbf{U}_p^\alpha\mathbf{U}_q^{\alpha\text{T}} & \mathbf{0} \\ \mathbf{U}_q^\alpha\mathbf{U}_p^{\alpha\text{T}} & \mathbf{U}_q^\alpha\mathbf{U}_q^{\alpha\text{T}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}_r^\alpha\mathbf{U}_r^{\alpha\text{T}} \end{bmatrix}, \tag{8}$$

demonstrating how selective multivariate localization can be achieved. This approach can be extended in an obvious way to grouping/isolating any number of variables.

One should also note that with the current approach—using eigenvectors decomposed from correlation matrices—special care must be taken when dealing with periodic domains since the correlation matrix is circulant and may not be positive semidefinite. The implication is that if the localization length scales for p and q are different and any eigenvectors associated with negative eigenvalues are truncated, $\mathbf{U}_p^\alpha\mathbf{U}_q^{\alpha\text{T}}$ and $\mathbf{U}_q^\alpha\mathbf{U}_p^{\alpha\text{T}}$ (the associated off-diagonal blocks) will not strictly be cross-correlation matrices (not shown). The off-diagonal block matrices with this extra symmetry are explored with SVDL in the next section.

Here, we have described IVDL in detail to provide clarity on how one might technically implement it in an NWP system. This approach is outlined in Wang and Wang (2023) and is referred to as basic scale-dependent localization variable-dependent localization Basic-SDLVDL (with scale-dependent aspects in their case), but they implemented and tested a variant minor extension scale-dependent localization variable-dependent localization (MinorE-SDLVDL) instead. Mathematically, Basic-

SDLVDL and MinorE-SDLVDL are equivalent (Wang and Wang 2023). Prior to this study, L22 had already discussed the technical implementation of IVDL (although not named IVDL then) and the approach to apply alpha fields to one or all variables, along with the proof of equivalence. Another more computationally efficient approach has since been proposed by Menetrier (https://doi.org/10.5281/zenodo.7547230).

b. The SVDL scheme

We note that while the full localization matrix is always symmetric, the off-diagonal blocks of $\mathbf{L}$ are not themselves symmetric if the localization length scales (and hence transforms) for p and q are different (i.e., $\mathbf{U}_p^\alpha\mathbf{U}_q^{\alpha\text{T}} \neq \mathbf{U}_q^\alpha\mathbf{U}_p^{\alpha\text{T}}$). Buehner and Shlyaeva (2015) also found the asymmetry in their between-scale cross-covariances when applying scale-dependent localization. This is not necessarily a criticism of the existing approaches but is rather a prompt to pose the question if imposing extra symmetry in $\mathbf{U}^\alpha$ could improve the performance of a multivariate localization scheme.

In SVDL, the full localization matrix is specified explicitly:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{p,p} & \mathbf{L}_{p,q} & \mathbf{L}_{p,r} \\ \mathbf{L}_{q,p} & \mathbf{L}_{q,q} & \mathbf{L}_{q,r} \\ \mathbf{L}_{r,p} & \mathbf{L}_{r,q} & \mathbf{L}_{r,r} \end{bmatrix}, \tag{9}$$

and this is made, by construction, to have all block correlation matrices symmetric, e.g., $\mathbf{L}_{p,q} = \mathbf{L}_{p,q}^{\text{T}}$. SVDL can be considered a "brute force" approach—all block correlation matrices (including off-diagonal ones) within Eq. (9) are fully prescribed. The eigendecomposition is then performed on the full localization matrix (instead of blocks of it like in IVDL) to retrieve $\mathbf{U}^\alpha = \mathbf{L}^{1/2}$ to use in the variational algorithm. Due to the application of the eigendecomposition on the full localization matrix, the computational cost of the SVDL approach is estimated to be $\mathcal{O}[(N_g N_{\text{var}})^3]$ compared to $N_{\text{var}}\mathcal{O}(N_g^3)$ for IVDL if based solely on the computational complexity of eigendecomposition. For a small N_{var}, this may still be acceptable even for a full NWP system, although this needs further testing.

To prescribe the off-diagonal correlation matrices (e.g., $\mathbf{L}_{p,q}$), the average correlation length scale $\overline{h}^\alpha$ of two associated variables (p and q in this example) is computed, which is then used as the length scale in the Gaspari–Cohn localization function to construct $\mathbf{L}_{p,q}$. Other approaches may also be considered instead of using $\overline{h}^\alpha$, e.g., computing localization functions separately and taking their average. As the off-diagonal matrices are constructed like autocorrelation matrices, they are symmetric, unlike in IVDL. Additionally, each off-diagonal matrix pair (e.g., $\mathbf{L}_{p,q}$ and $\mathbf{L}_{q,p}$) uses exactly the same $\overline{h}^\alpha$ to construct the localization function, so the full localization matrix is automatically symmetric. Furthermore, to apply selective multivariate localization, one could set selective off-diagonal correlation matrices of $\mathbf{L}$ to $\mathbf{0}$, similar to Eq. (8).

It is also important to note that in SVDL, $\mathbf{L}$ is constructed with block correlation matrices (or with $\mathbf{0}$ in selected off-diagonal blocks) but may not be a correlation matrix as a whole. The implication is that without further adjustment, it is not possible to guarantee positive semidefiniteness. Stanley et al. (2021)
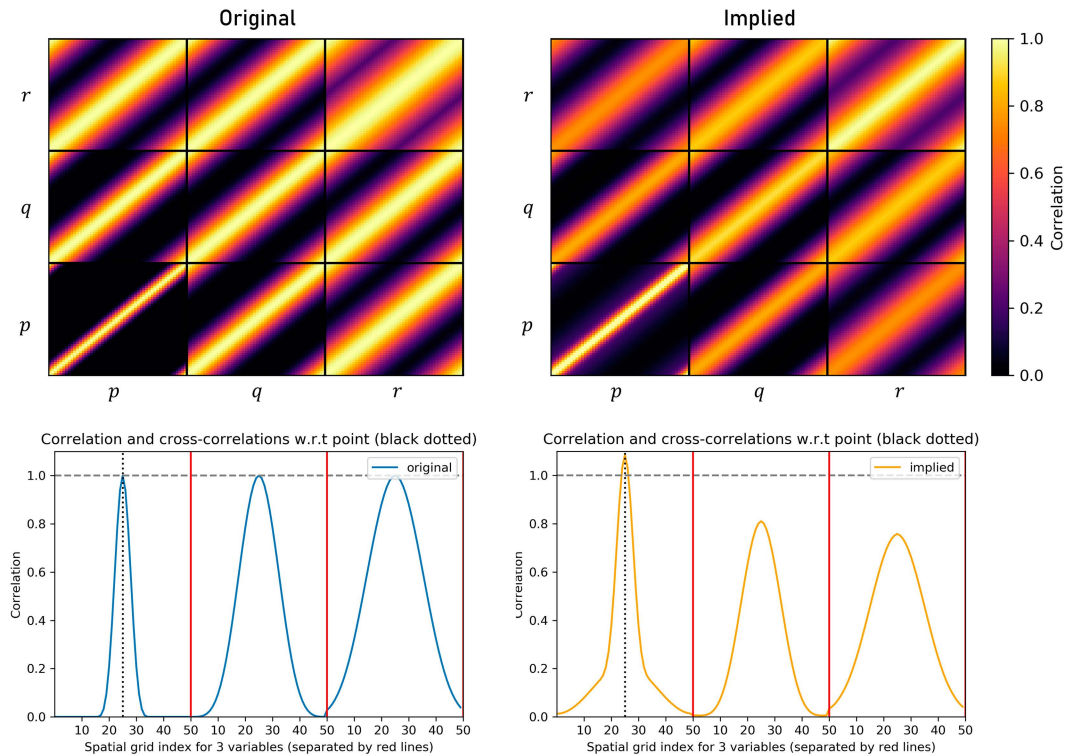
FIG. 1. Full localization matrix [Eq. (9)] with variable-dependent localization using SVDL for a one-dimensional periodic domain of 50 points and three variables (p, q, and r; localization length scale of 5, 10, and 15 points, respectively). (top left) The original matrix is prescribed explicitly, whereas (top right) the implied matrix is reconstructed from eigenvectors after truncating negative eigenvalues and rescaling. Autocorrelations and cross correlations with respect to the midpoint (index 25) of variable p are shown for the (bottom left) original matrix and (bottom right) implied matrix. The black dotted lines in the bottom panels are at value 1, which is the desired value of the peak correlations.

proposed how one might prescribe the off-diagonal correlation matrices such that the implied **L** is positive semidefinite. SVDL does not use their approach; this is to maintain flexibility to prescribe some of the off-diagonal blocks to **0**. This also means that an extra step is required to guarantee positive semidefiniteness. Following L22, any negative eigenvalues are truncated and the remaining eigenvalues are rescaled (e.g., by a uniform factor given by the ratio of the original sum of eigenvalues to the sum of eigenvalues after truncation) to restore the original total variance. Figure 1 shows how variable-dependent localization can be explicitly prescribed in SVDL, ensuring that the off-diagonal block matrices are symmetric. After decomposing the original localization matrix and reconstructing, the implied localization matrix is not identical to the original. Even with rescaling, the truncation of negative eigenvalues has the effect of damping, particularly on the cross correlations. The kurtosis of the correlations is also slightly altered. This effect is more severe for periodic domains where circulant matrices may be involved. There may be alternatives to using a uniform factor to rescale, but they are not investigated here.

SVDL allows full control of the localization, including multivariate error cross-covariances, but is computationally expensive. As

mentioned, it does enforce symmetry in the off-diagonal correlation matrices unlike previous approaches (IVDL or in Wang and Wang 2023) and allows the specification of null matrices on certain cross-correlation components unlike in Stanley et al. (2021). Nevertheless, it remains to be seen if symmetry is beneficial as there may not be a physical justification (see section 3b for figures illustrating selective multivariate localization with IVDL and the differences between IVDL and SVDL). Due to the expensive—but flexible—formulation of SVDL, one could also easily apply a cross-localization weight factor to diminish the error cross-covariances, similar to that discussed in Stanley et al. (2021), but this is not investigated here.

3. Model and data assimilation framework

a. Development of the ABC-DA system

To evaluate variable-dependent and selective multivariate localization for the tropics, we use the ABC model (Petrie et al. 2017), which solves a modified set of the compressible Euler equations. This model uses a vertical slice formulation (a two-dimensional longitude–height plane) and contains only dry dynamics. It is named after its key parameters: the pure gravity wave frequency A, the controller of acoustic wave speed B, and

the constant of proportionality between pressure and density perturbations C. Additionally, a Coriolis parameter f can be set based on the desired latitudinal position of the chosen longitude–height plane. This allows for a deep tropical environment to be mimicked by selecting a very small value for f. In this configuration, a value of $f = 10^{-5}$ s^{-1} is used. This corresponds approximately to a value of f at a latitude of 4°N. The other model parameters are also set as $A = 0.02$ s^{-1}, $B = 0.01$, and $C = 10^4$ m^2 s^{-2}. There are five prognostic variables: zonal wind u, meridional wind v, vertical wind w, scaled density perturbation $\tilde{\rho}'$ (a pressure-like variable), and buoyancy perturbation b' (a potential temperature–like variable), which govern the model dynamics. The ABC model is thus sufficiently complex as a multivariate dynamical system, while retaining simplicity to expedite research and development.

The associated data assimilation was introduced in Bannister (2020), supporting incremental 3D-Var and 3D-Var first guess at appropriate time (3D-Var-FGAT). The system is solely based on variational data assimilation. Initial implementation had an arguably crude form of generating an ensemble by considering multiple latitudinal slices from a three-dimensional operational model's output file (a version of the Unified Model). This was used for calibrating the background error covariance matrix, and no ensemble-based methods (e.g., ensemble Kalman filter or square root filters) were implemented. Further development of the ABC-DA system by L22 introduced hybrid ensemble–variational data assimilation via the alpha control variable transform approach (Lorenc 2003). Concurrently, L22 also introduced other ensemble generation and propagation approaches. The random field perturbations method (Magnusson et al. 2009) was used to cold start an ensemble, and the ensemble bred vectors (EBV) method (Balci et al. 2012) was introduced to propagate the ensemble at each data assimilation cycle—this was computationally cheaper than the traditional ensemble Kalman filter or square root filters and did not suffer from filter collapse (see L22 for details). This parallel-run ensemble was necessary to support hybrid 3D-Var and hybrid 3D-Var-FGAT in the ABC-DA system. From the ensemble forecasts, the error modes $\mathbf{x}'^k_t$ can be computed and used with $\mathbf{U}^\alpha$ as in Eq. (1). Using these newly implemented features in the ABC-DA system, L22 showed that hybrid 3D-Var outperformed 3D-Var and pure EnVar methods in the ABC-DA configured for the tropical environment. However, for the purpose of this study, we will focus on the pure EnVar framework.

b. Illustration of IVDL and SVDL

Before exploring the research questions using variable-dependent and selective multivariate localization in assimilation experiments, we illustrate how IVDL and SVDL can control the localization with ABC model variables. First, selective multivariate localization is illustrated using IVDL. For demonstration, the state variables have been grouped into two sets: (i) u and v; (ii) w, $\tilde{\rho}'$, and b'. This means that variables in the same set retain their cross correlations (and thus cross-covariances after the Schur product with the ensemble-derived error covariances), but variables in different sets have cross correlations knocked out by localization. Figure 2

shows the implied localization functions with respect to u and $\tilde{\rho}'$ points. It is clear that only u and v cross correlations are retained in the first set, and w, $\tilde{\rho}'$, and b' cross correlations are retained in the second set. Between variables of different sets, no cross correlations are retained by the design of $\mathbf{U}^\alpha$. Other grouping options have been implemented in the ABC-DA system, which we use to isolate important multivariate error relationships (see section 5a). This will enable us to explore question 2 on which multivariate error relationships are beneficial for EnVar data assimilation in the tropics.

Next, variable-dependent localization is illustrated using both IVDL and SVDL. For demonstration, only the vertical localization length scale is changed between variables. Figure 3 shows the comparison of IVDL- and SVDL-implied localization functions with respect to u and b' points. There are subtle differences in Figs. 3a and 3b due to the truncation of negative eigenvalues in SVDL. Additionally, note how there are differences in the u–b' and b'–u cross correlations using IVDL (Figs. 3c,d; left), which is due to the asymmetry in the off-diagonal block matrices. Using SVDL on the other hand (Figs. 3c,d; right), the u–b' and b'–u cross correlations are identical. As discussed in section 2b, it is not known a priori whether the extra symmetry imposed by SVDL is beneficial to data assimilation, but SVDL is certainly more flexible than IVDL. This will enable us to explore question 3 on whether variable-dependent spatial localization is beneficial for EnVar data assimilation in the tropics.

4. Description of the experiments

a. Setup for the ABC-DA system

To evaluate the performance of IVDL and SVDL in data assimilation to learn about tropical multivariate covariances and the best localization settings, we conduct a series of hourly cycling observation system simulation experiments (OSSEs) similar to those in L22. To represent the incompleteness of the observation network in an NWP system, only u, v, and $\tilde{\rho}'$ are observed at a set of points in a subdomain (longitudinal distance between 50 and 500 km; height between 9 and 14 km, i.e., the upper portion of vertical slice of 546-km length by 16-km height). The observation operator used is bilinear interpolation. This setup is more akin to an NWP system with only satellite-related point observations (e.g., satellite-derived wind) that are available in the upper troposphere and stratosphere. This setup may also accentuate the impact of selective multivariate localization because unobserved variable fields are updated solely on localized multivariate error cross-covariances. At each cycle, 100 observations of each of the abovementioned variables are assimilated. The observations are sampled from a "truth" run (using a time step of 4 s) with the same values of A, B, and C, but with added Gaussian noise based on the observation error standard deviation. For this setup, the observation error standard deviations are 0.1, 0.1, and 1.5×10^{-4} m s^{-1}, respectively. All observations are valid at the analysis time of each cycle and all experiments use the same observations.
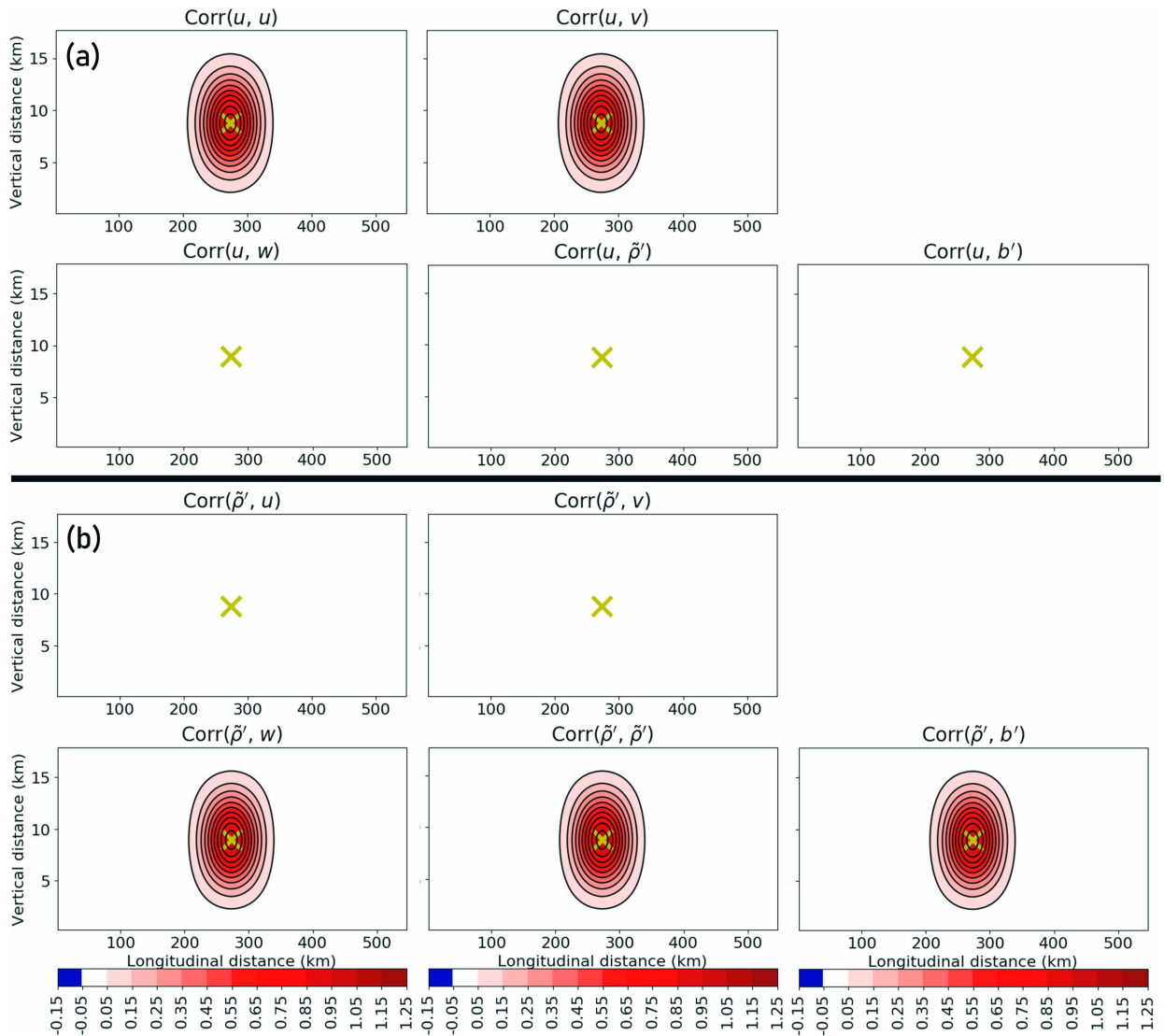
FIG. 2. Implied localization functions with respect to a point (yellow cross) using IVDL for cross correlations of all variables with respect to (a) u and (b) $\tilde{\rho}'$. The state variables have been grouped into two sets: (i) u and v; (ii) w, $\tilde{\rho}'$, and b' to illustrate selective multivariate localization.

The random field perturbations method (section 3.1 of L22) is first used to generate a 1000-member ensemble set (i.e., the pairs of states are sampled from a 150-day truth run with hourly outputs; $24 \times 150 = 3600$ available states). The number of available states has to be sufficiently large to avoid repetition in the pairs of states drawn (with replacement back into pool) by the random field perturbations method. The new 1000-member ensemble is centered around one randomly chosen ensemble analysis from the last cycle of the experiment labeled EBVd in L22 (cold start) and allowed to spin up for 75 cycles (75 h) so that the system would have lost memory of the cold start initialization using random field perturbations. At each cycle, the ensemble perturbations are updated using the EBV method, similar to L22. No inflation is used, but by definition of the EBV method, a scaling based on a fixed factor divided by the max norm of the

perturbations from the previous cycle (see Balci et al. 2012 or L22 for details) is used to get the updated perturbations. After spinning up for 75 cycles, the ensemble perturbations from the last spinup cycle are used for the start of the experiments, computed according to the number of ensemble members chosen for a particular experiment (these are subsets of the 1000-member ensemble; see section 4b). For example, for a 1000-member ensemble, 999 perturbations are computed using the member-minus-mean approach; for a 100-member ensemble, the first 100 members are retained and 99 perturbations are computed. Each experiment in section 5 is run for 100 cycles.

b. Guidance on ensemble size for experiments

Since the pure EnVar approach is used for the experiments (i.e., no climatological error statistics are used), we
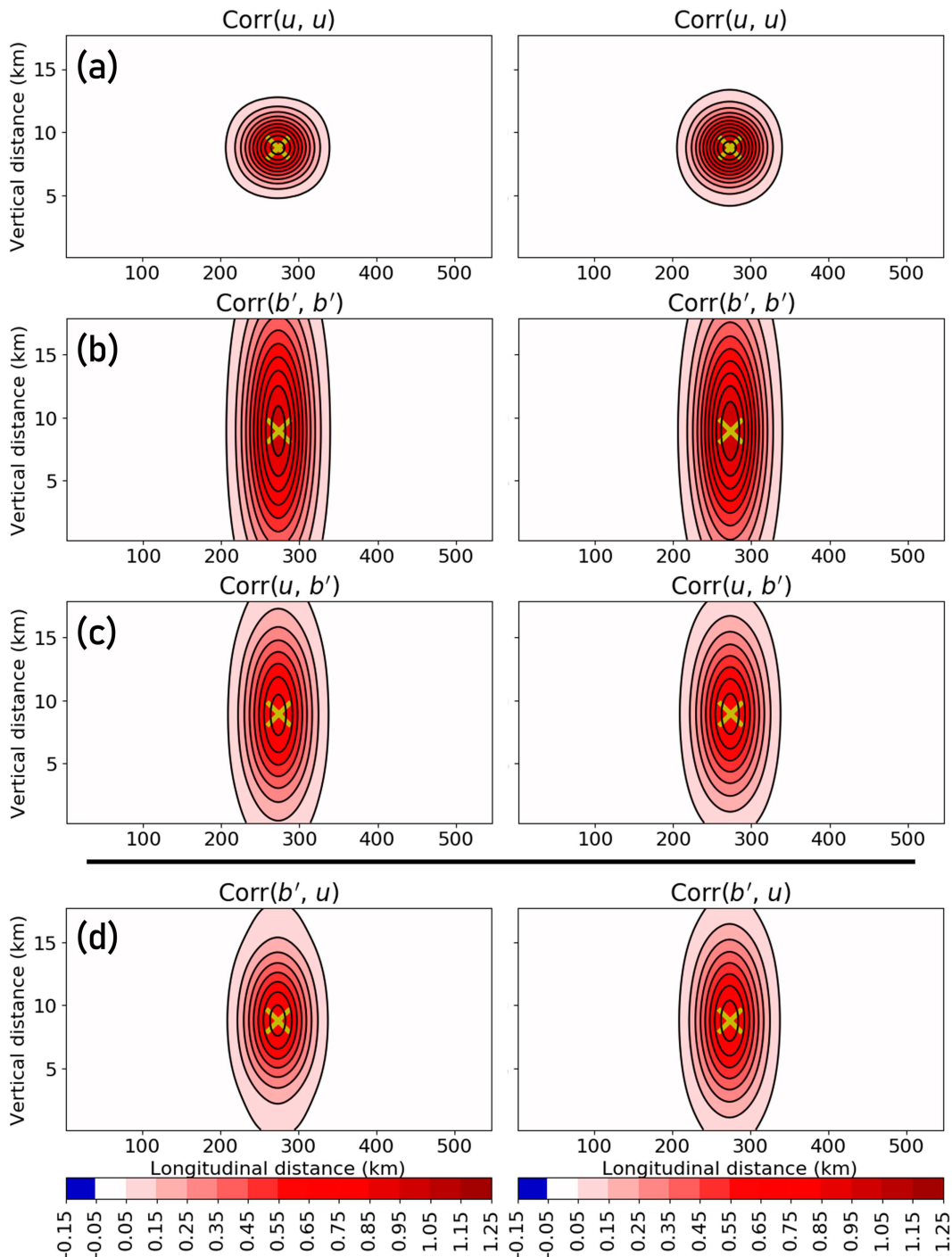
FIG. 3. Implied localization functions with respect to a point (yellow cross) using (left) IVDL and (right) SVDL for autocorrelations of (a) u and (b) b' and cross correlations of (c) b' with respect to u and (d) u with respect to b'. The vertical localization length scales are larger in b' to illustrate variable dependence.

conduct further analysis to provide guidance on the ensemble size that is suitable for the tropical ABC-DA system. First, we assess the degree of orthogonality of the ensemble, by computing the linear independence of each successive ensemble perturbation with respect to previous perturbations at the start of the

experiment (i.e., only the perturbations from the first cycle, even though they get updated throughout the experiment run), following Bannister et al. (2017, manuscript submitted to *Geosci. Model Dev.*). We do this separately for each variable, and so, the results can differ for each
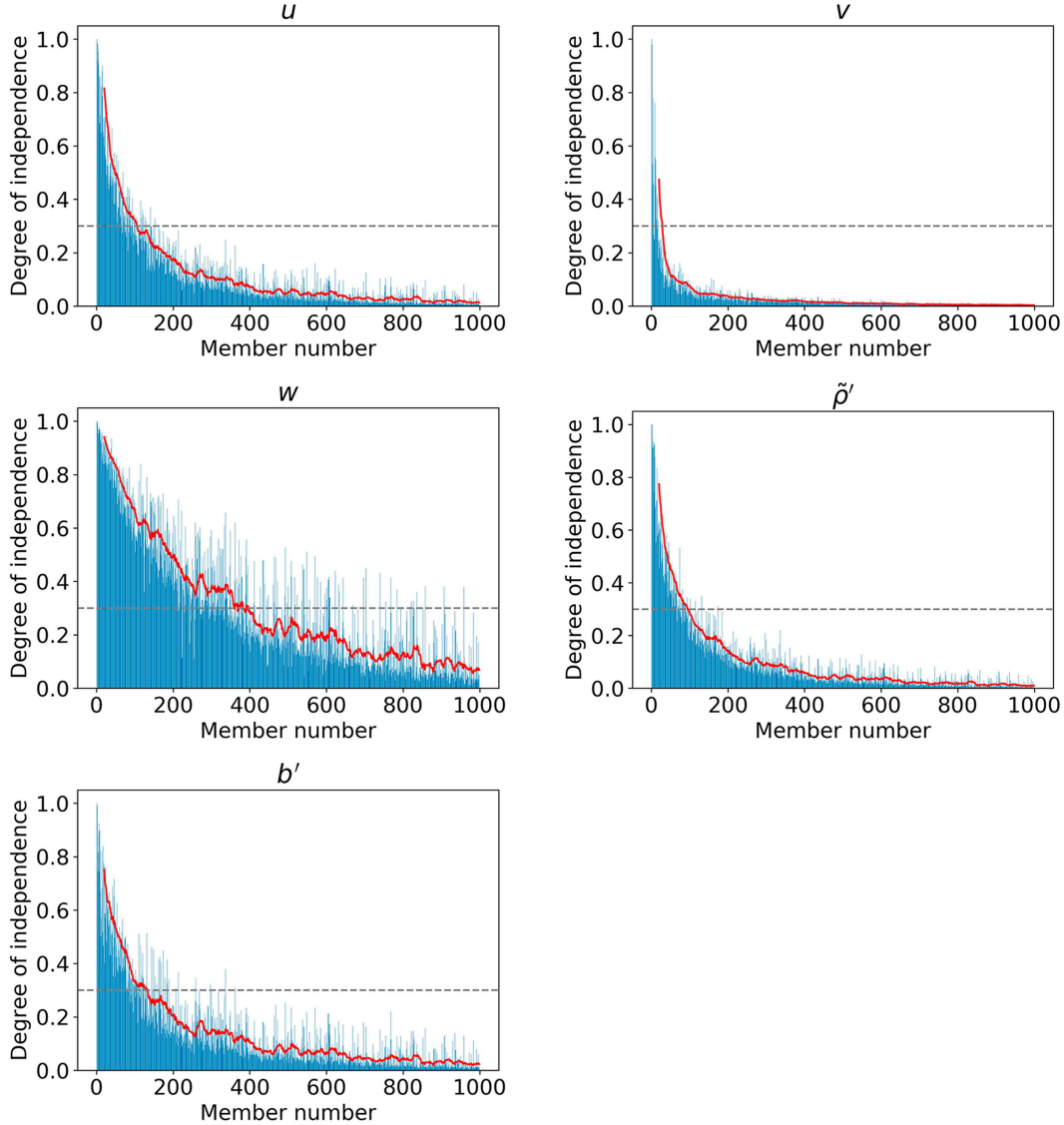
FIG. 4. Degree of linear independence $\widehat{N}_k$ of each successive ensemble perturbation for all five prognostic variables. Perturbations are valid at the start of the experiments. An arbitrary threshold of 0.3 is indicated by a gray dotted line. The 20-member rolling averages are indicated in red.

one. Given the full set of 999 perturbations and ignoring time indices for brevity, the Gram–Schmidt procedure is used to successively compute a set of orthonormalized vectors $\hat{\mathbf{x}}'^k$,

$$\hat{\mathbf{x}}'^k = \frac{1}{\widehat{N}_k}\left(\frac{\mathbf{x}'^k}{|\mathbf{x}'^k|} - \sum_{j=1}^{k-1}\left\langle\frac{\mathbf{x}'^k}{|\mathbf{x}'^k|}, \hat{\mathbf{x}}'^j\right\rangle\hat{\mathbf{x}}'^j\right), \qquad (10)$$

where $|\cdot|$ denotes the inner product, $\langle\mathbf{a}, \mathbf{b}\rangle = \mathbf{a}^\tau\mathbf{b}$, and $\widehat{N}_k$ is chosen to ensure that each successive perturbation $\hat{\mathbf{x}}'^k$ has unit length. If $\widehat{N}_k$ is small, then the newly orthogonalized vector $\hat{\mathbf{x}}'^k$ has only a small component that is linearly independent from the previously considered vectors of index $1, \ldots, k-1$. The magnitude of $\widehat{N}_k$ is thus a measure of the degree of linear independence of the ensemble perturbation $\hat{\mathbf{x}}'^k$.

Figure 4 shows the degree of linear independence for each successive ensemble perturbation for the five prognostic variables at the start of the experiment. It reflects the capability of each new random field perturbation to sufficiently explore the additional direction in the subspace. This is more likely if the system dynamics develop with strong nonlinearities within the first hour (since we are using 1-h forecast ensemble perturbations from a randomly generated analysis ensemble). Based on Fig. 4, and using the 20-member rolling average in red, we note that the degree of independence of each successive ensemble perturbation varies for each variable. We use the $\widehat{N}_k = 0.3$ level to help decide on whether a sufficient level of linear independence is reached. For w, the degree of independence remains above 0.3 threshold until after the 400th
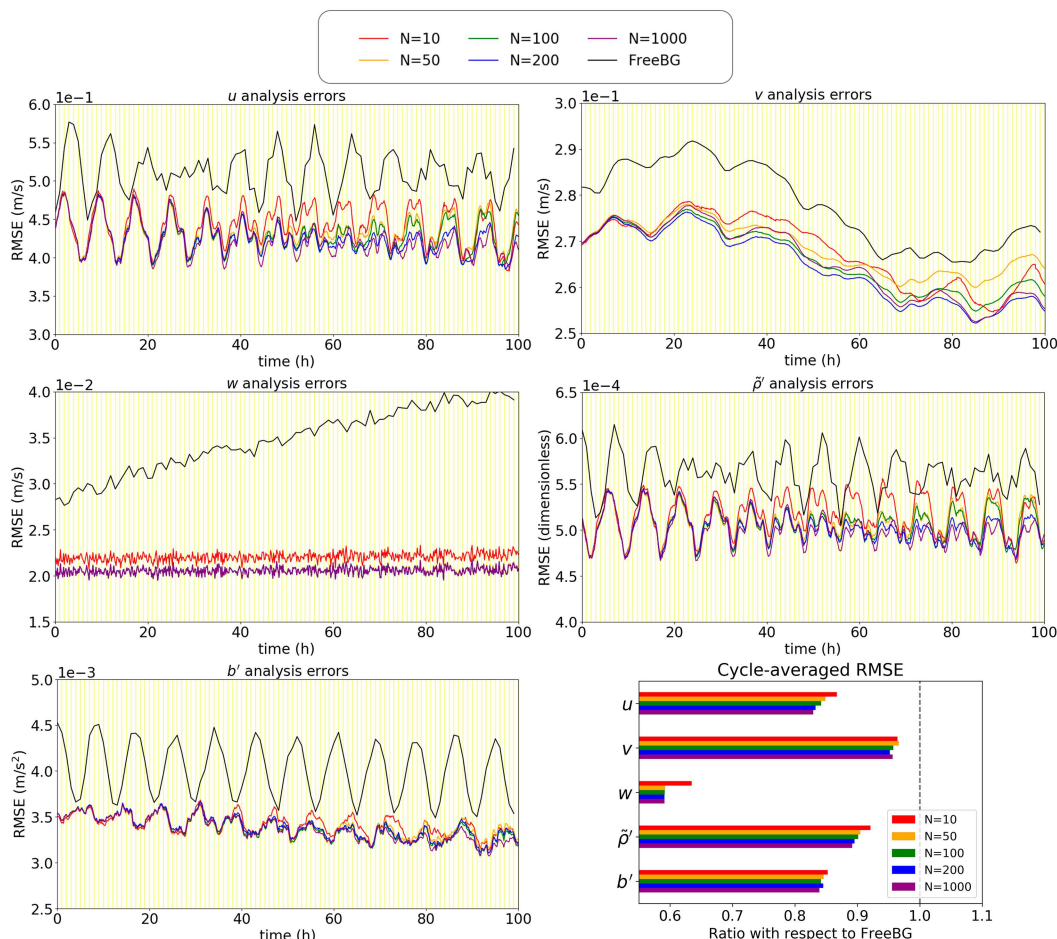
FIG. 5. All panels except bottom right show time series of root-mean-square analysis errors for the ensemble sensitivity experiments (10, 50, 100, 200, and 1000 members) and FreeBG. No localization is used in these experiments. The vertical yellow lines are the analysis times. Analysis errors are defined with respect to the "truth" run, computed every 10 min within the respective assimilation windows for experiments and every hour for the FreeBG. The bottom right panel shows the ratio of the cycle-averaged RMSE for each experiment with respect to the FreeBG for the five ABC model variables.

member. On the other hand, for v, this occurs at about the 40th member. For u and $\tilde{\rho}'$, the threshold is met after the 100th member, and for b', the threshold is met at the 120th member. This suggests that nonlinearities largely develop in the w field when computing the "truth" run, which are captured by the random field perturbations method, and/or are evolved within the first hour. The nonlinearities captured by the random field perturbations method in other variables are weaker. The results suggest that if the ensemble size is larger than about 100, most of the ensemble perturbations would virtually be linear combinations of others and thus have limited impacts on EnVar data assimilation [since the analysis increment is a linear combination of perturbations, Eq. (1)]. This method does not, however, give any indication that the sampling error is sufficiently small for localization to be unnecessary. The results do highlight though how each field has its own characteristics based on the system dynamics, so variable-dependent localization in particular should be worth exploring.

We further conducted a sensitivity test to the number of ensemble members using the OSSE framework described above. We tried experiments with 10, 50, 100, 200, and 1000 members. Horizontal and vertical localization are not applied in this sensitivity test to reveal the impacts of sampling noise on the (raw) ensemble-derived error covariances. The runs are evaluated using the RMSE with respect to the truth run. This was the approach taken in Bannister (2020, 2021) and L22 to assess the performance of their experiments. It is also a straightforward metric to measure the deviation of the forecasts from the truth run.

Figure 5 contains the time series of RMSE for each variable over the 100 cycles and a summary of the cycle-averaged RMSE for each variable compared to the free background run (FreeBG, which is the reference forecast without any data assimilation starting from the cold start background state). It shows that in general, the cycle-averaged RMSE decreases as the number of ensemble members increases,
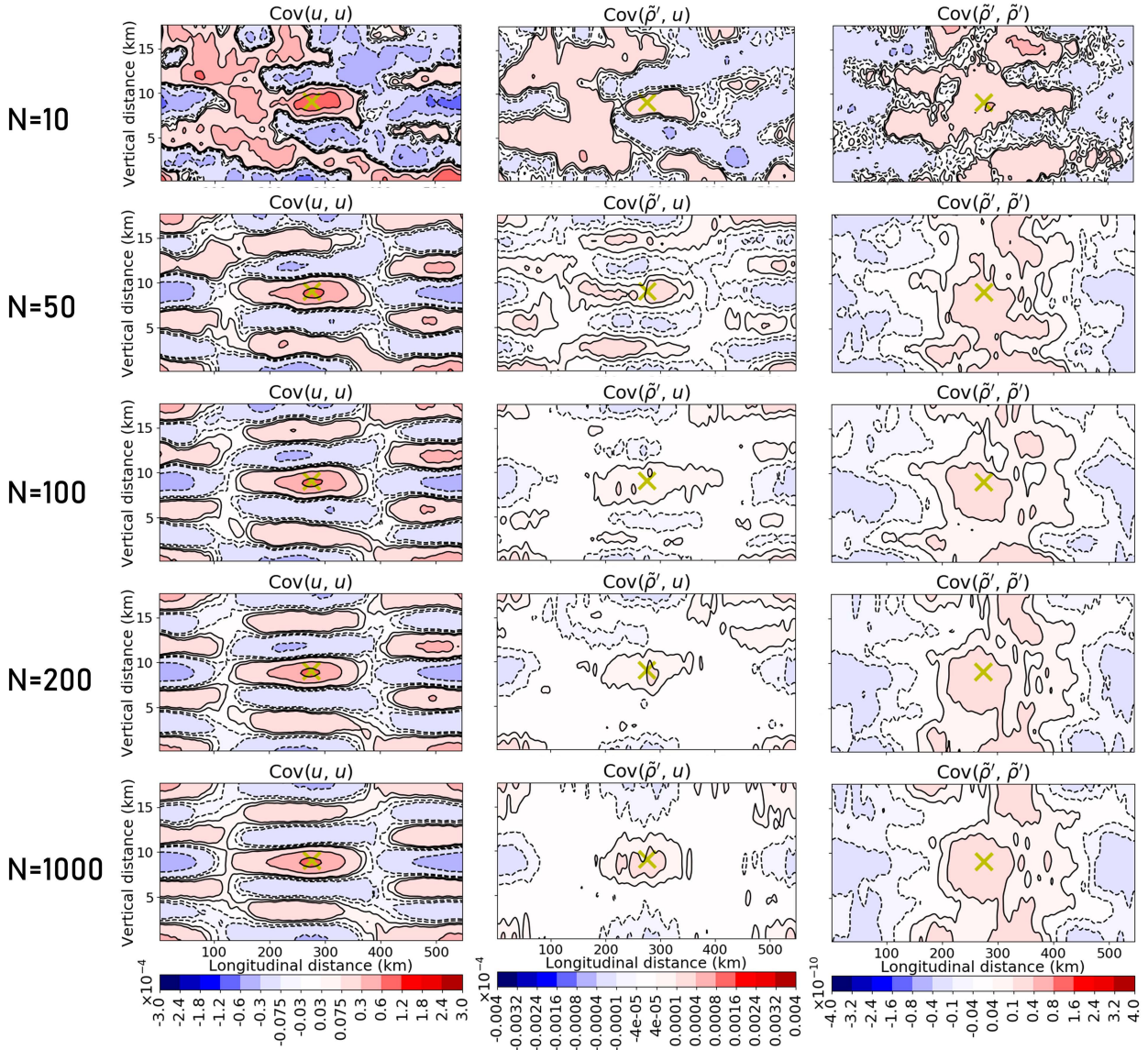
FIG. 6. Raw ensemble-derived error autocovariances of (left) u, (center) cross-covariances of u with respect to $\tilde{\rho}'$, and (right) autocovariances of $\tilde{\rho}'$ as a function of the number of ensemble members N (increasing from top to bottom). Negative values have contours that are dashed, and contour intervals are nonuniform to elucidate any features. The covariances are computed with respect to a point (yellow cross) near the center of the domain. The ensemble perturbations are drawn from the first cycle of the sensitivity test.

particularly for u and $\tilde{\rho}'$. The error reduction is about 2%–4% depending on the variable when the ensemble size is increased from 10 to 1000. The highest cycle-averaged RMSE is seen in the runs with 10 and 50 members, respectively, as expected due to the lack of localization to address sampling error. Also, the reduction in cycle-averaged RMSE between the runs with 200 and 1000 members is small compared to that between the runs with 10 and 50 members, for almost all variables. For w, the decrease in RMSE with an increasing number of ensemble members is less pronounced beyond 50 members. As w is highly nonlinear, it is unsurprising that the FreeBG deviates from the truth

run more rapidly than the other variables and so the error increases over time.

Also, note the different cycle-averaged RMSE patterns for u and v. This difference may be due to the two-dimensional nature of the ABC model (no latitude dependence). In this setup, vorticity ($\partial v/\partial x$) is associated with v and divergence ($\partial u/\partial x$) is associated with u. We would expect vorticity and divergence to have different time scales, which are indeed observed in the different cycle-averaged RMSE patterns for u and v.

Next, we examine the raw ensemble-derived error covariances between u and $\tilde{\rho}'$ using the ensemble perturbations drawn from the first cycle of the sensitivity test. Figure 6 shows how a

TABLE 1. Horizontal and vertical localization length scales $h_{\text{horiz}}^{\alpha}$ and h_{vert}^{α} for the experiments to evaluate variable-dependent localization. "NIL" means no localization is used in the relevant direction (horizontal, vertical). IVDL refers to the isolated variable-dependent localization scheme (section 2a) and SVDL refers to the symmetric variable-dependent localization scheme (section 2b).

Experiment	Multivariate localization	Scheme	u	v	w	$\tilde{\rho}'$	b'
1a–1g	See text	IVDL	20 km, NIL	20 km, NIL	20 km, NIL	20 km, NIL	20 km, NIL
2a	One set: (i) u, v, w, $\tilde{\rho}'$, b'	SVDL	50 km, NIL	50 km, NIL	20 km, NIL	20 km, NIL	20 km, NIL
2b	One set: (i) u, v, w, $\tilde{\rho}'$, b'	SVDL	20, 5 km	10, 2 km	10, 2 km	20 km, NIL	20 km, NIL
2c	Two sets: (i) u, v and (ii) w, $\tilde{\rho}'$, b'	IVDL	50 km, NIL	50 km, NIL	20 km, NIL	20 km, NIL	20 km, NIL
2d	Two sets: (i) u, v and (ii) w, $\tilde{\rho}'$, b'	IVDL	50, 5 km	50, 5 km	20 km, NIL	20 km, NIL	20 km, NIL

selection of covariance structures changes as the number of ensemble members is increased. It shows that the ensemble-derived error covariances become less noisy, particularly in the u–u and $\tilde{\rho}'$–$\tilde{\rho}'$ autocovariances. Notably, the ensemble size threshold at which the covariance structures start to appear consistent is 100 members. Above the threshold (i.e., 200 and 1000 members), the structures do not differ substantially. We also note that with 1000 members, the u–u autocovariances contain wavelike patterns which result in nonnegligible longer-range spatial covariances. These patterns have previously been seen in Bannister (2020) and L22, albeit for $\tilde{\rho}'$ autocovariances instead. The wavelike patterns suggest that the main error modes in the ABC-DA system could be strongly influenced by periodic waves resonating in the domain, as also noted by L22 (i.e., they are real features rather than artifacts of sampling error).

Given the results in Figs. 5 and 6, and weighing the computational costs of running very large (1000-member) ensembles, a suitable ensemble size of 100 is used in further experiments to explore the possibilities of variable-dependent and selective multivariate localization. This number of members is large enough to show coherent structures in the unlocalized covariances, but still shows evidence of sampling errors.

5. Results from data assimilation experiments using localization

a. Exploring the importance of multivariate error relationships

In this section, we run EnVar data assimilation experiments to test different multivariate localization options. All data assimilation experiments start with the same initial background and ensemble perturbations as the 100-member experiment in section 4b. The observations are also the same. The details of the selective multivariate localization experiment variants are listed here. As a reminder, variables that are in different groups do not retain cross-covariances with variables in other groups. For example, in experiment 1c below, localization is used to ensure that u and v errors are made to be completely uncorrelated with w, $\tilde{\rho}'$, and b' errors.

1a) One set: limiting case where all multivariate error cross-covariances are retained, as in traditional EnVar implementations.
1b) Five sets: limiting case where no multivariate error cross-covariances are retained (full intervariable localization).
1c) Two sets: (i) u and v and (ii) w, $\tilde{\rho}'$, and b'.

1d) Two sets: (i) v, w, $\tilde{\rho}'$, and b' and (ii) u.
1e) Two sets: (i) u, v, w, and b' and (ii) $\tilde{\rho}'$.
1f) Two sets: (i) u, w, $\tilde{\rho}'$, and b' and (ii) v.
1g) Three sets: (i) u, $\tilde{\rho}'$, and b', (ii) v, and (iii) w.

For this subsection, all variables use the same horizontal localization length scales of 20 km and use the IVDL scheme (see first row of Table 1 and section 5b for justification of choice). No vertical localization is used, following appendix C of L22 which showed that the vertical localization of $\tilde{\rho}'$ and b' could result in hydrostatic imbalances. Here, IVDL is first used to enable selective multivariate localization; variable dependence is incorporated in the next subsection.

To demonstrate the impact of selective multivariate localization on the analysis, the analysis increments for the first cycle of experiments 1a, 1b, 1c, 1d, and 1g are plotted in Fig. 7. These experiments represent the diverse possibilities and the implications of choosing a different number of sets and/or different number of variables in each set (experiments 1e and 1f are not shown as they are similar to 1d; partitioning into two sets: four variables and one variable). A comparison of successive experiment pairs allows for the impact of specific multivariate error relationships to be disentangled. For example, comparing experiments 1a and 1d reveals the impact of isolating u. Note that since EBV is used to propagate the ensemble (without data assimilation), the analysis increments shown are for the control member which assimilates the observations. Here, the impact of u observations on all other variables through the error cross-covariances is substantial; the analysis increments in experiment 1d are less widespread than in 1a, especially over unobserved regions. For the full intervariable localization limiting case (experiment 1b), there are analysis increments for observed variables only, as expected. The $\tilde{\rho}'$ analysis increment patterns are broadly similar to those from experiments where $\tilde{\rho}'$ is not influenced by u observations [i.e., experiments 1c, 1d, and 1e (latter not shown)]. Similarly, the v increment patterns are broadly similar to those from experiments 1f (not shown) and 1g. When v is isolated in experiment 1g, the impact on the analysis increments of other variables is subtle due to relatively small error cross-covariances associated with v. Note that the widespread analysis increments in experiment 1a do not yet reveal if the multivariate error relationships associated with u are meaningful or undesirable. Unlike the midlatitudes, where the multivariate error covariances may be explained largely by geostrophic and hydrostatic theories, the essential multivariate error covariances in the tropics are not well known.
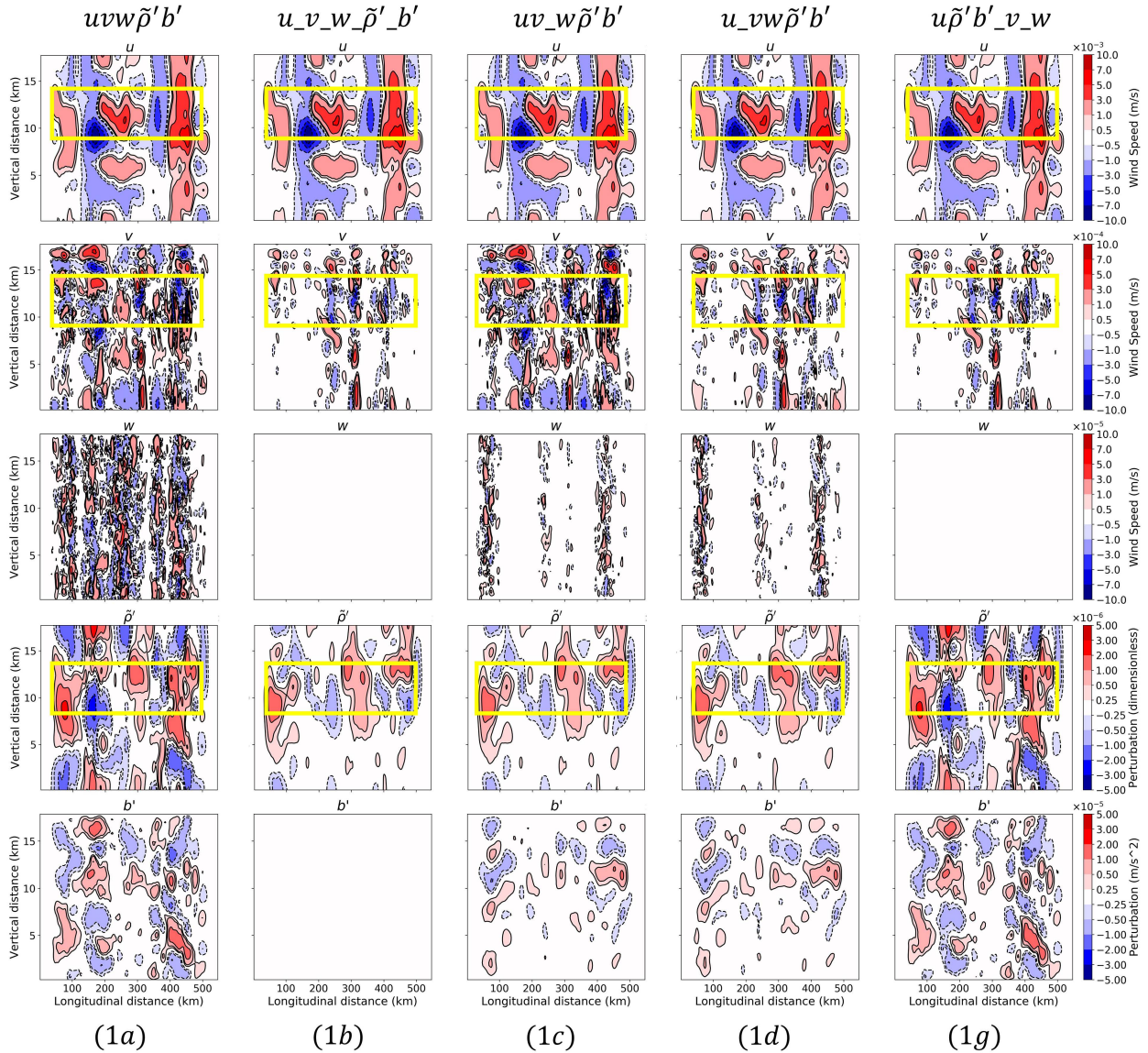
FIG. 7. Analysis increments from the first cycle of experiments 1a, 1b, 1c, 1d, and 1g (shown from left to right; see row 1 of Table 1), for the five prognostic variables (shown from top to bottom). All experiments start with the same background ensemble, assimilate the same observations (of u, v, and $\tilde{\rho}'$, which are equally spaced within the yellow box), and use the same spatial localization. Selective multivariate localization is applied with IVDL; variables are partitioned into sets (see text for description), demarcated by underscores (e.g., $u\_wv\tilde{\rho}'b'$ refers to experiment 1d).

To identify important/beneficial multivariate error relationships in the ensemble-derived error covariances, we analyze the performance of each experiment over the course of 100 cycles, benchmarked against two limiting cases: (i) where all multivariate error cross-covariances are retained (experiment 1a) and (ii) where no multivariate error cross-covariances are retained (experiment 1b). Experiments that have smaller cycle-averaged RMSE than benchmark (i) suggest that certain multivariate error cross-covariances are not important and may be introducing more noise into the analysis than signal. Likewise, experiments that perform better than benchmark (ii) suggest that certain multivariate

error cross-covariances are important/beneficial for EnVar data assimilation in the tropics. The experiments which isolate individual variables (1d, 1e, and 1f) are plotted in Fig. 8. The remaining experiments (1c and 1g) are plotted in Fig. 9.

The most salient feature from both figures is that the experiments have similar cycle-averaged RMSE to either experiment 1a or 1b (the two limiting cases), for all variables except w. This may be unsurprising based on the comparison of analysis increments in Fig. 7. Experiments 1c, 1d, and 1e have similar values as experiment 1b (except w), while experiments 1f and 1g have similar values as 1a. Note that experiment 1b has
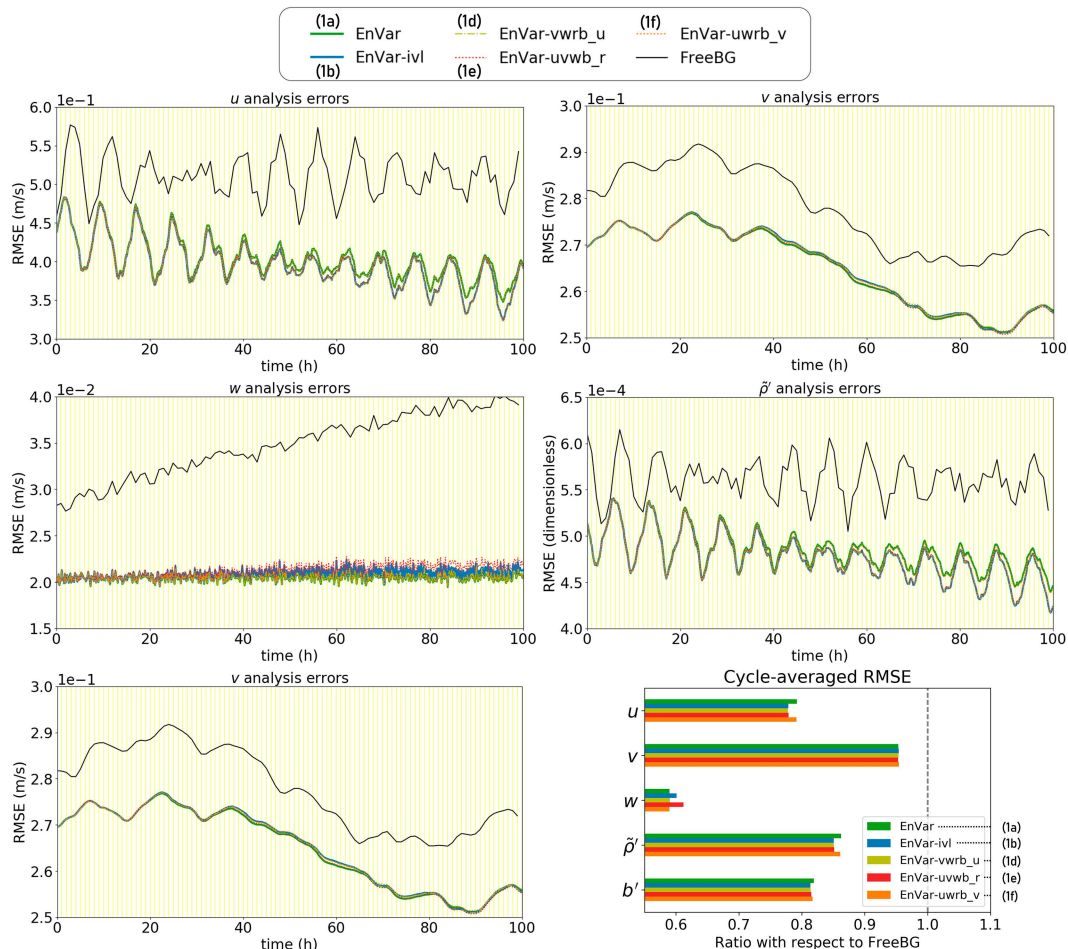
FIG. 8. As in Fig. 5, but for experiments 1a (EnVar; limiting case), 1b (EnVar-ivl; limiting case), 1d (EnVar-vwrb_u), 1e (EnVar-uvwb_r), and 1f (EnVar-uwrb_v) compared to the FreeBG.

smaller cycle-averaged RMSE than 1a for all variables except w. The key points from the seven experiments are as follows.

- Experiments that remove the $\tilde{\rho}'$–b' error cross-covariances (1b and 1e) lead to a substantially larger RMSE for w.
- Experiments that remove the u–$\tilde{\rho}'$ error cross-covariances (1b, 1c, 1d, and 1e) lead to a notably smaller RMSE for two observed variables u and $\tilde{\rho}'$ and one unobserved variable b'.

The first point relates to hydrostatic imbalances in the analysis. From the prognostic equations for w, there are source and sink terms which relate to b' and the vertical gradient of $\tilde{\rho}'$. Any imbalance between the two will result in changes to w as the system evolves. Clearly, this imbalance is undesirable, as seen from the experiments. Lorenc (2003) previously discussed this issue in the context of mass–wind balance. Here, we find that maintaining signals of hydrostatic balance where the ensemble forecasts are hydrostatically balanced—by retaining the $\tilde{\rho}'$–b' error cross-covariances—is important/beneficial for the tropical ABC-DA system. Vetra-Carvalho et al. (2012)

previously showed that for regions (in the United Kingdom) where convection was weak, hydrostatic balance holds very well. However, in regions where moist convection was involved, hydrostatic balance was not preserved. To generalize our results for the tropics, moist processes would need to be considered in the ABC model (Zhu and Bannister 2023), but that version does not yet have data assimilation incorporated. Notwithstanding the limitations, the tropical dry dynamics representing vertical wind (dry convection) and the mass–wind interactions are still relevant to explore within the ABC-DA system.

The second point relates to mass–wind sampling errors in the analysis. As discussed previously, there is no clear mass–wind balance relationship for the tropics and many centers implicitly treat mass and wind variables univariately in their climatological background error covariance matrices. Here, we find that explicitly treating the mass and wind variables univariately in EnVar data assimilation (by knocking out covariances between u and $\tilde{\rho}'$) is beneficial for the tropical ABC-DA system. From Fig. 6, it appears that even with 100 members, the mass–wind error relationship still contains
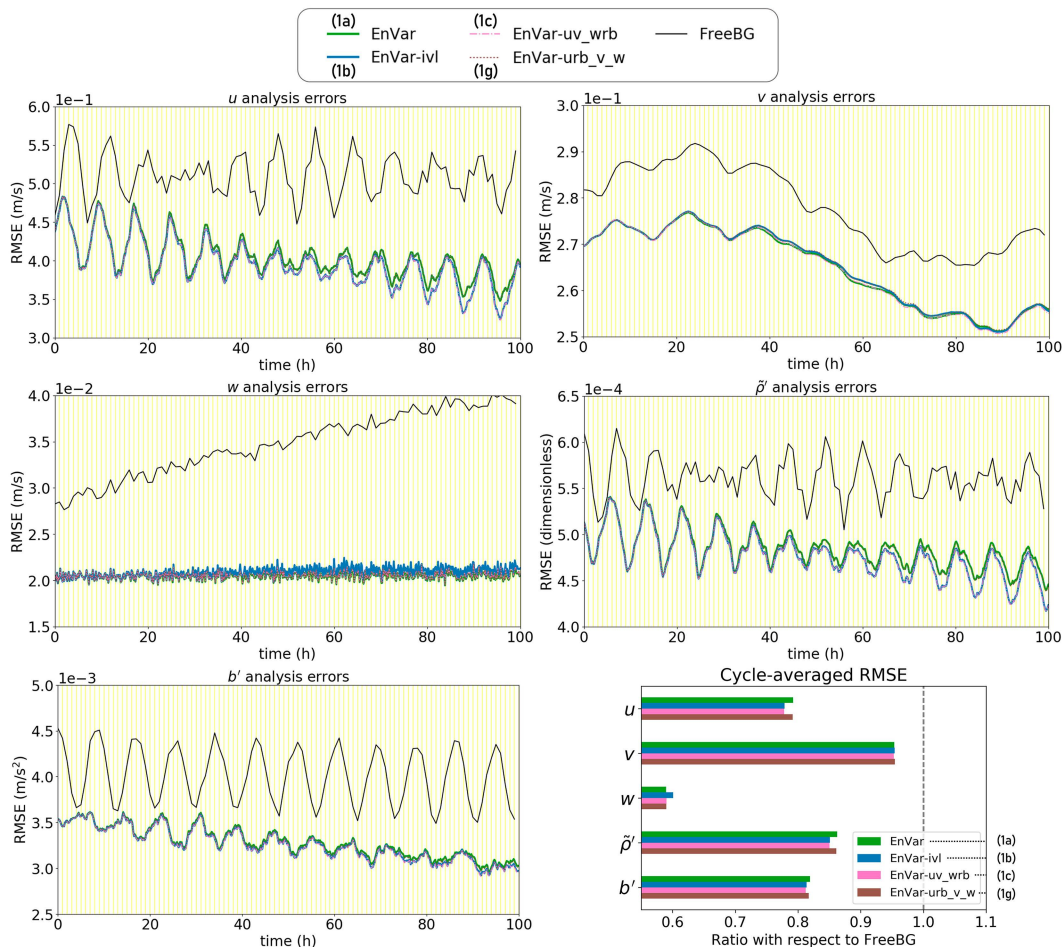
FIG. 9. As in Fig. 5, but for experiments 1a (EnVar; limiting case), 1b (EnVar-ivl; limiting case), 1c (EnVar-uv_wrb), and 1g (EnVar-urb_v_w) compared to the FreeBG.

some sampling noise. Referencing earlier results from Fig. 7, the impact of u observations on other mass variables ($\tilde{\rho}'$ and b') is likely negative when u-related error cross-covariances are retained. The results suggest that the mass–wind error relationship prescribed directly from a 100-member ensemble is not beneficial for the tropical ABC-DA system. It remains a challenge to find a scale-dependent balance between mass and wind errors for the tropics (e.g., handling the larger scales based on large-scale balances, but handling the smaller scales based on convective-scale balances, if they exist at all). Until then, our results suggest that they should be treated entirely univariately.

We further examined if other multivariate error relationships are important/beneficial in experiments 1f and 1g, but isolating v and w does not have a substantial impact on the analysis increments nor cycle-averaged RMSE. We further repeated experiments 1a, 1c, 1d, and 1e two more times with different random seeds for the observations (not shown), arriving at the same conclusions. There is limited additional benefit of repeating the other experiments, since they would yield similar results (as seen above) which would lead to the

same conclusions. The abovementioned two points therefore suggest that (i) where dry dynamics are concerned, hydrostatic balance is important for EnVar data assimilation in the tropics and (ii) treating mass and wind errors univariately is also beneficial for EnVar data assimilation in the tropics.

b. Exploring the benefits from variable-dependent spatial localization

Next, we explore the benefits of variable-dependent localization for EnVar data assimilation in the tropics by assimilating with different localization length scale values for different model variables. Thus far, all experiments have included only horizontal localization of uniform length scales. For some of the subsequent experiments, we use both horizontal and vertical localization of length scales $h_{\text{horiz}}^{\alpha}$ and h_{vert}^{α}, respectively (see section 2 for localization function details, but applied separately for each spatial dimension). As before, the details of the variable-dependent localization experiment variants are listed here:
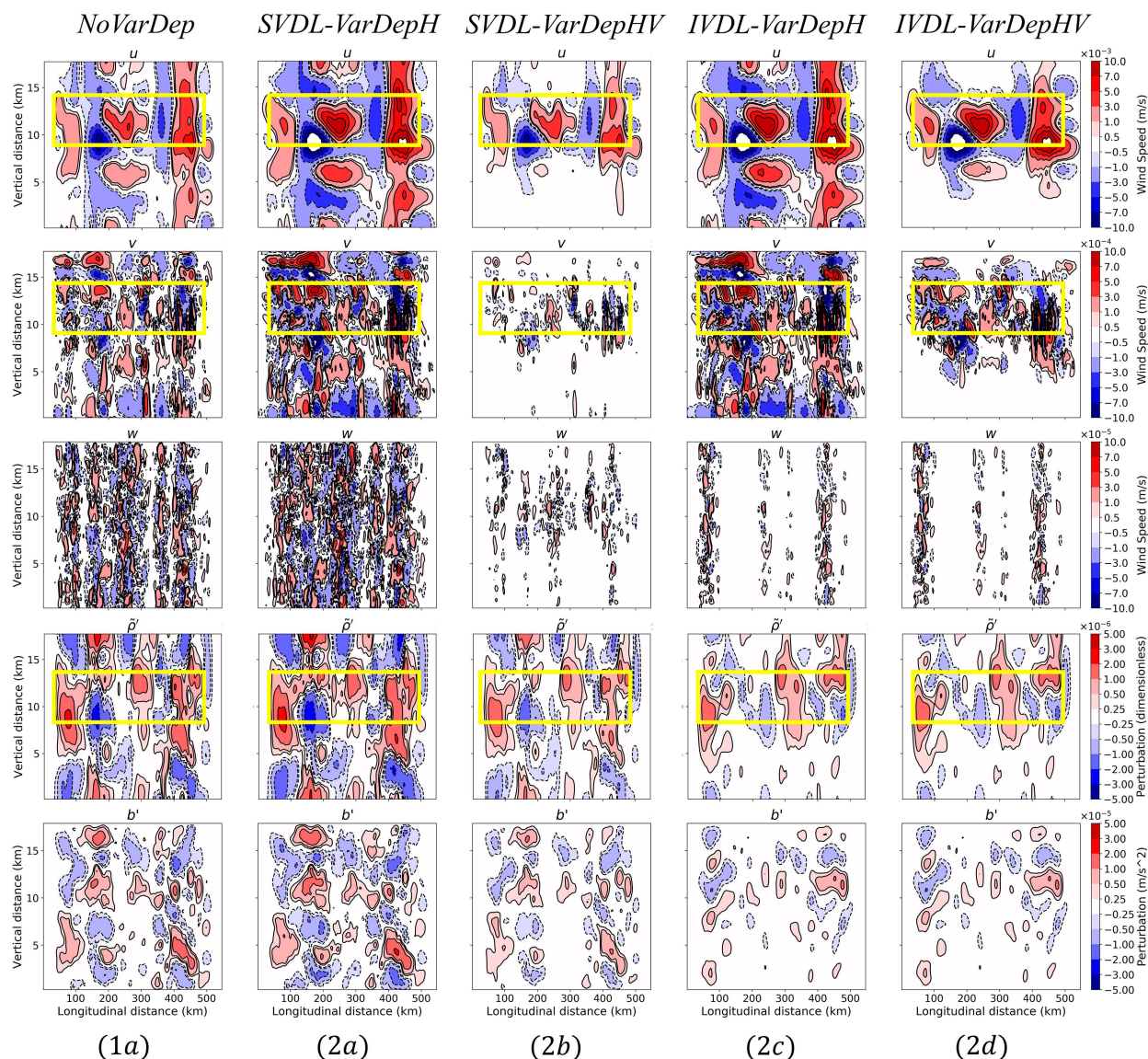
FIG. 10. As in Fig. 7, but for experiments 1a (NoVarDep), 2a (SVDL-VarDepH), 2b (SVDL-VarDepHV), 2c (IVDL-VarDepH), and 2d (IVDL-VarDepHV).

2a) As in experiment 1a, but applying SVDL and changing horizontal localization for horizontal wind variables only.

2b) As in experiment 1a, but applying SVDL and varying horizontal and vertical localization for wind variables only.

2c) As in experiment 1c, applying IVDL but changing horizontal localization for horizontal wind variables only.

2d) As in experiment 1c, applying IVDL but changing horizontal and vertical localization for horizontal wind variables only.

For reference, Table 1 shows a summary of h_{horiz}^{α} and h_{vert}^{α} used for the experiments. The default length scales for experiments 1a–1g were determined based on the horizontal distance between adjacent observations (≈23 km), following L22. Experiments 2a and 2c implement variable-dependent horizontal localization only. This is to assess if the flexibility granted by horizontal localization alone is beneficial. Experiments 2b and 2d further implement variable-dependent vertical localization to assess if it is also beneficial. Experiments using SVDL (2a and 2b) retain all multivariate error relationships, so they are compared to experiment 1a as the benchmark. Experiments using IVDL (2c and 2d) require that variables in the same set use the same h_{horiz}^{α} (see issues with periodic domains highlighted in section 2a), and they are compared to experiment 1c as the benchmark. Experiment 1c is also suitable to be the benchmark since it was the best performing out of the selective multivariate localization experiments.

Figure 10 shows how variable-dependent localization changes the analysis increments for the first cycle for experiments 1a, 2a,
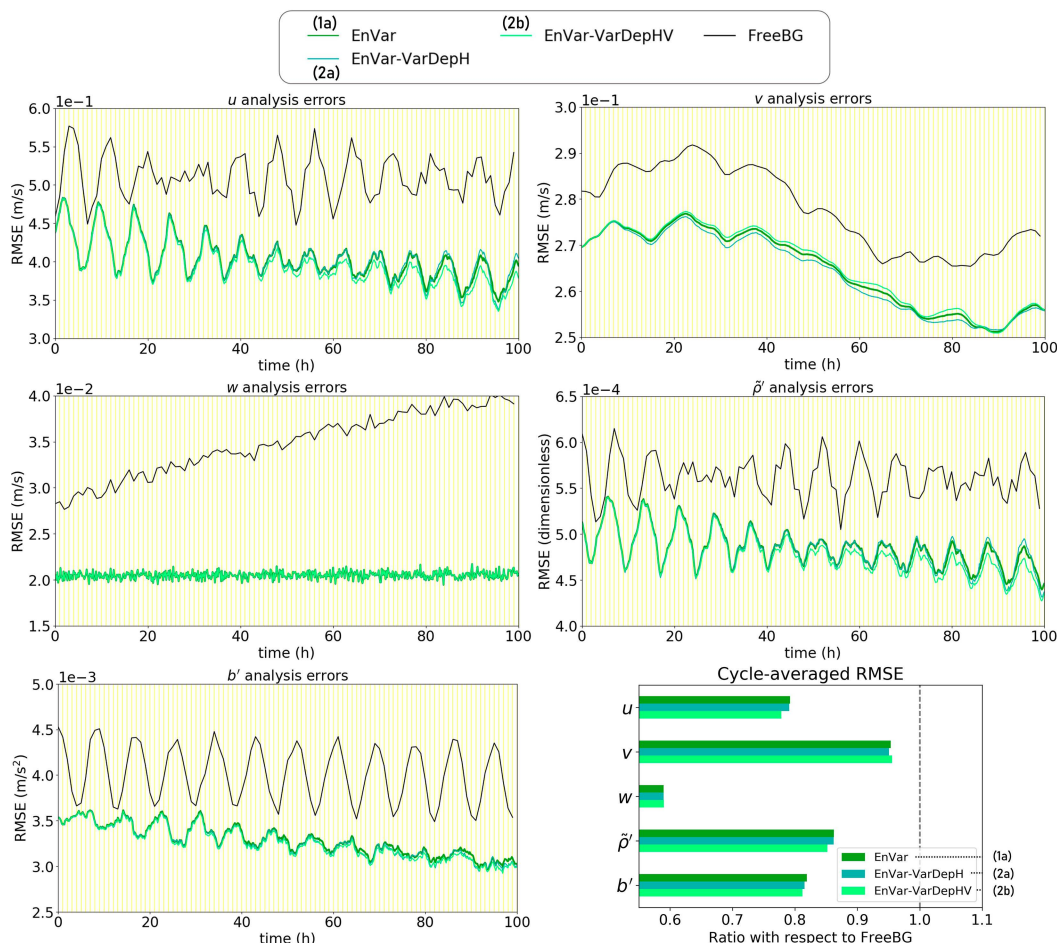
FIG. 11. As in Fig. 5, but for experiments 1a (EnVar; limiting case), 2a (EnVar-VarDepH), and 2b (EnVar-VarDepHV) compared to the FreeBG.

2b, 2c, and 2d. As expected, the u and v analysis increments in experiments 2a and 2c have broader structures than before since $h_{\mathrm{horiz}}^{\alpha}$ is larger. Similarly, when vertical localization is further introduced in experiments 2b and 2d, the u and v analysis increments are confined to the observed regions. Note how the w, $\tilde{\rho}'$, and b' analysis increments are similar between experiments 2c and 2d despite changes to the vertical localization length scales of the wind variables. This highlights how both variable-dependent and selective multivariate localization can be simultaneously applied to the ensemble-derived error covariances (using IVDL) to constrain the observation impact on specific variables.

Next, we examine the impacts of variable-dependent localization on the cycle-averaged RMSE. Figure 11 shows that the cycle-averaged RMSE for experiment 2a (when u and v use different $h_{\mathrm{horiz}}^{\alpha}$ from the rest of the variables) is marginally smaller than 1a for most variables. The oscillations in the u and $\tilde{\rho}'$ RMSE evolution are also marginally more pronounced, likely related to the gravity waves (and their associated frequencies) within the domain. Similar results are seen comparing experiment 2c with 1c (Fig. 12); using different $h_{\mathrm{horiz}}^{\alpha}$ for both u and v

leads to marginally improved forecasts, despite both experiments omitting error cross-covariances between u and $\tilde{\rho}'$ (unlike in experiment 2a, which uses SVDL). When variable-dependent vertical localization is further introduced in experiments 2b and 2d, the cycle-averaged RMSEs in general are further reduced, even for b' which is unobserved. Across all experiments, we find that 2d leads to the smallest cycle-averaged RMSE, about 3%–4% smaller (for u, $\tilde{\rho}'$, and b') compared to experiment 1a (the benchmark) which uses the widely adopted traditional localization approach. We have further repeated experiments 1a and 2d two more times with different random seeds for the observations. We have also repeated the experiments using a 50-member ensemble instead of a 100-member ensemble to ensure the conclusions are not heavily dependent on ensemble size. Figure 13 shows that the cycle-averaged RMSE differences of u, $\tilde{\rho}'$, and b' within groups of experiment 1a or 2d configurations are much smaller than the differences between runs of experiment 1a and their respective 2d configurations.

We use a paired t test and a Kolmogorov–Smirnov test to quantify the statistical significance. The two tests compare the
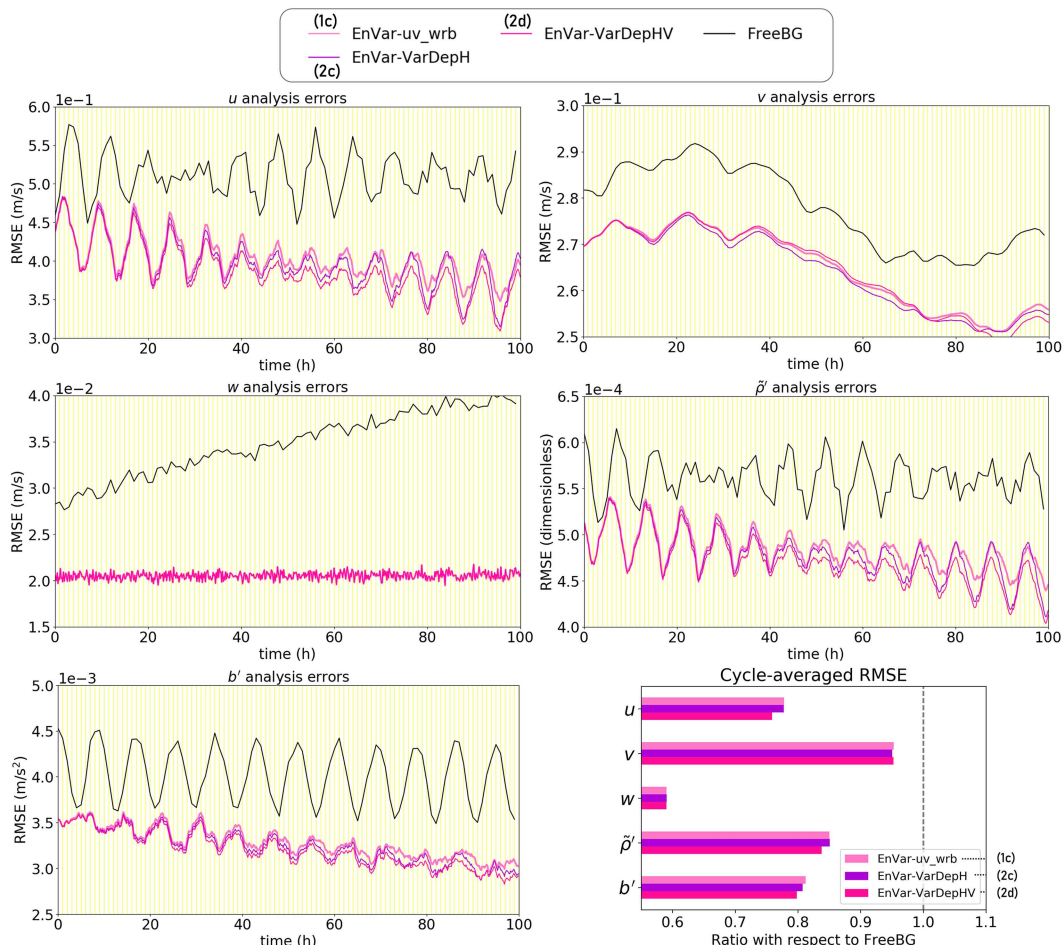
FIG. 12. As in Fig. 5, but for experiments 1c (EnVar-uv_wrb), 2c (EnVar-VarDepH), and 2d (EnVar-VarDepHV) compared to the FreeBG.

distributions of the pairs of experiments 1a and 2d using a 100-member ensemble (three samples). The paired t test assumes that the cycle-averaged RMSE follows a normal distribution, while the Kolmogorov–Smirnov test does not. For a small sample size, it is useful to use both tests. Using the paired t test, the p values for u, $\tilde{\rho}'$, and b' are much less than 0.01, while the p values for v and w are 0.301 and 0.0163, respectively. Using the Kolmogorov–Smirnov test, the Kolmogorov–Smirnov statistics for u, $\tilde{\rho}'$, and b' are each 1, with a p value of 0.1. For v and w, the Kolmogorov–Smirnov statistic is each 0.667, with a p value of 0.6. A p value of less than 0.1 indicates that the result is statistically significant at the 90% confidence level. Note that experiments with 50-member ensemble are omitted from the statistical significance tests as they are repeats of the 100-member ensemble experiments (same corresponding observations) and are therefore not independent samples from the 100-member experiments. Overall, the results show that experiment 2d produces a smaller cycle-averaged RMSE than 1a, which is statistically significant at the 90% confidence level (at least) for u, $\tilde{\rho}'$, and b' but not statistically significant for v and w.

6. Conclusions

a. Summary and key results

In this study, the benefits of retaining or rejecting multivariate error relationships and controlling the localization length scales separately for each variable in ensemble–variational (EnVar) data assimilation in the tropics are explored. This is conducted using a simplified nonhydrostatic model, the ABC model. Two approaches are implemented within the EnVar framework of the ABC model, which we refer to as the isolated variable-dependent localization (IVDL) and symmetric variable-dependent localization (SVDL) schemes. These grant the ability to (i) prescribe different spatial localization length scales for different variables and (ii) control (i.e., knockout by localization) multivariate error cross-covariances. The IVDL determines the multivariate localization in an implicit way, while the SVDL prescribes the multivariate localization in an explicit way.

Using IVDL and SVDL in multicycle observation system simulation experiments (OSSEs) with the ABC-DA system, we explore which multivariate relationships in the ensemble-
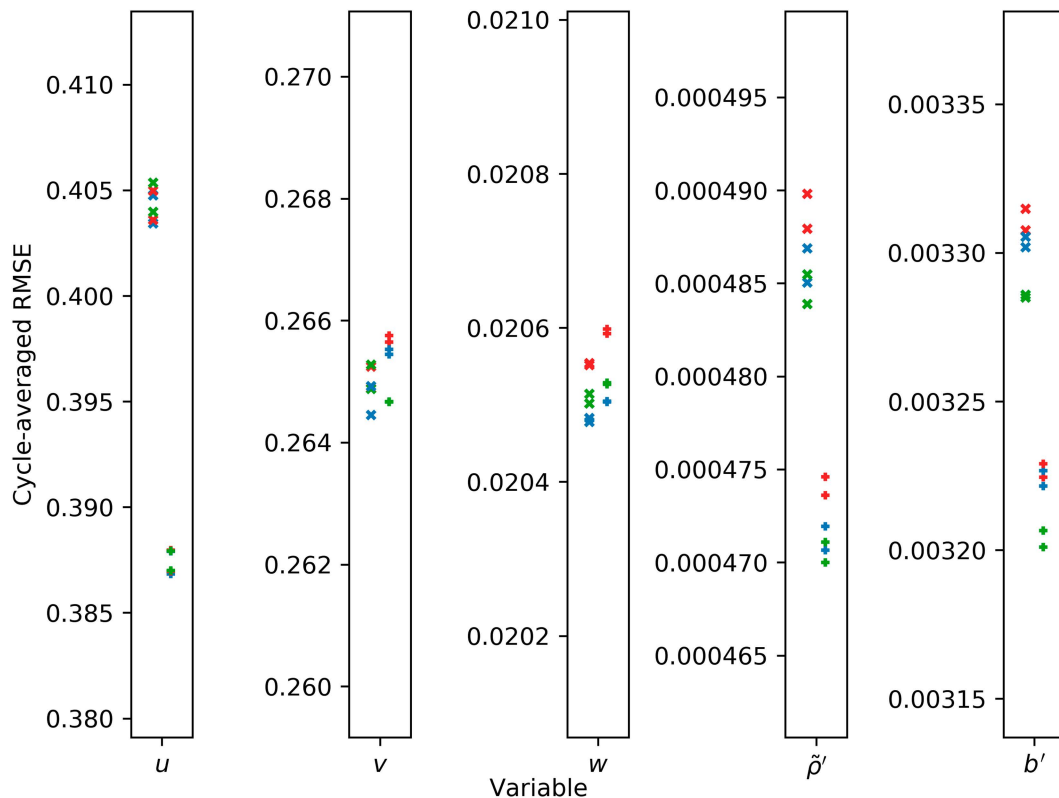
FIG. 13. Comparison of the cycle-averaged RMSE for experiments 1a (crosses) and 2d (pluses) for all five prognostic variables. Each experiment is run three times with different random seeds for the observations (blue, red, and green) and using a 50- or 100-member ensemble (a total of six runs).

derived error covariances are important/beneficial and the benefits of variable-dependent spatial localization for EnVar data assimilation in the tropics.

Using up to 1000 ensemble members (generated at each cycle as ensemble bred vectors), we first decide on a suitable ensemble size to test the localization methods (question 1 posed in section 1). We compute the degree of linear independence of each successive ensemble perturbation. This provides an indication of whether the ensemble bred vectors suitably span the subspace. This analysis shows that each variable has different characteristics due to the system dynamics. The fact that the degree of independence is different for each variable suggests that variable-dependent spatial localization is justifiable. We further show that for the ABC-DA system, increasing the number of ensemble members results in a decrease in the cycle-averaged root-mean-square error (RMSE) in the OSSEs. We find that the threshold at which covariance structures start to appear consistent (but still have appreciable sampling error) is about 100 members for the ABC model. This ensemble size is used for most of the remaining experiments.

Further experiments reveal the following:

- Using selective multivariate localization is beneficial, particularly when covariances associated with hydrostatic balance are retained and when the zonal wind errors are decoupled from the mass (scaled density perturbation) errors in the background error covariances (question 2). These lead to

reduced cycle-averaged RMSE in corresponding experiments. The results suggest that when tropical dry dynamics is concerned (as with the ABC model), hydrostatic balance can still be important even at convective scales. There is also little basis for retaining the mass–wind error covariances that are presented by the ensemble, which we believe could introduce more sampling noise than useful signal in the error cross-covariances, even with as many as 100 ensemble members.

- Using variable-dependent localization is beneficial for EnVar data assimilation in the tropics, albeit to a smaller extent compared to selective multivariate localization (question 3). For this particular setup, we show that using different horizontal length scales for wind and mass variables (longer length scales for wind) reduces the cycle-averaged RMSE, perhaps because they are more optimal for this system.

- The best-performing experiment uses both variable-dependent localization and selective multivariate localization (retaining covariances associated with hydrostatic balance and omitting covariances associated with mass–wind error relationships). It leads to a 3%–4% smaller cycle-averaged RMSE than the experiment using a traditional EnVar setup, where there is often little control over multivariate localization nor over separate localization length scale for each variable. This is of the same order of magnitude improvement that is typical of upgrades to, for instance, the EnVar system of Environment and Climate Change Canada (e.g., Caron and Buehner 2022).

b. Discussion and future work

In section 2b, we highlighted how SVDL allows for extra symmetry to be enforced in the off-diagonal block matrices. Although there is no mathematical need for such extra symmetry in **L** (only that **L** as a whole is symmetric), we thought that this could be nonetheless investigated as a plausible option. Comparing experiments 2a with 1a and 2c with 1c shows that variable-dependent spatial localization using SVDL and IVDL, respectively, reduces the RMSE of respective benchmarks by about the same percentages. From other further experiments (not shown), there also does not appear to be additional benefits of having symmetric off-diagonal block matrices using SVDL compared to IVDL. For computational efficiency, it may be more prudent to implement IVDL over SVDL, especially over nonperiodic domains (see section 2a for this caveat). This is because the potential lack of positive semidefiniteness in circulant block matrices only arises in IVDL because of periodic boundary conditions, but they frequently arise in SVDL regardless of boundary conditions.

A more complete study of the optimal horizontal and vertical length scales, and the correlation shapes for variable-dependent localization, is beyond the scope of this paper. In the above experiments, the specific length scale and correlation function choices for wind variables led to improved forecasts, but this could be because the initial choices for experiments 1a and 1c were severely suboptimal. To briefly explore this further, we apply the empirical optimal localization approach of Necker et al. (2023) to find the optimal tapering factors—the factors for populating the optimal localization matrix for a given ensemble size. The optimal tapering factors are computed using the mean of middle five vertical levels, 28–32, and with respect to the middle horizontal grid point as a simple test. The 1000-member background ensemble and 100-member subsamples from the first cycle of the 1000-member experiment (section 4b) are used. The cutoff length of the optimal tapering factor curve with respect to the middle horizontal grid point is about 150 km for u, 80 km for $\tilde{\rho}'$, and 40 km for b' (not shown), corresponding to horizontal localization length scales of 75, 40, and 20 km, respectively (half of cutoff length), based on the commonly used Gaspari–Cohn localization function (Necker et al. 2023). The optimal tapering factor curves for v and w were too noisy, owing to the limited number of vertical levels used in this simple test. This simple test suggests that the initial choices were moderately suboptimal for some variables. Nonetheless, the key point from these experiments is that variable-dependent localization can be beneficial—the impacts are sizeable even in a simplified multivariate system. These results support the findings by Necker et al. (2023), Lei et al. (2015), and Wang and Wang (2023), who found that the localization length scales depend on the variable. Future work may focus on identifying suitable correlation functions and length scales, as above, to implement with variable-dependent localization to achieve further improvements in the forecasts.

There are still many avenues to explore with regard to localization. It has been shown that scale-dependent localization is useful in general (Caron and Buehner 2022; Wang and Wang 2023), although its performance in the tropics is still questionable (Caron and Buehner 2022), and so requires further study. The localization function is not necessarily best specified as a fixed function of distance. Anderson (2012) for instance showed that the localization function is a function of the sample correlation itself in addition to the ensemble size. This issue may be dealt with using adaptive localization schemes such as those proposed by Bishop and Hodyss (2007, 2009). These schemes define the localization function from the ensemble itself, but these schemes are quite expensive to apply.

One limitation of this study is that the ABC model is a dry dynamics model and so does not represent moist processes, even though moist processes are obviously important for convective-scale data assimilation. However, as mentioned in section 5a, the tropical dry dynamics representing vertical wind (dry convection) and the mass–wind interactions are still interesting and relevant to explore as these are captured by the ABC model. Another limitation relates to the lower dimensionality of the ABC model (two-dimensional plane), so divergence and vorticity are only related to either zonal or meridional wind. Despite the limitations, this study could still pave the way for further work in testing IVDL and SVDL in full NWP systems, especially in the tropics, to assess their feasibility and benefits in applying variable-dependent or selective multivariate localization to improve EnVar data assimilation. Should computational costs permit, one might even combine this work with that of Buehner and Shlyaeva (2015) to have scale-dependent, variable-dependent, and selective multivariate localization altogether.

Data availability statement. The source code is freely available on a GitHub repository at https://doi.org/10.5281/zenodo.8334645. Data and scripts for running the experiments are available on request.

REFERENCES

Anderson, J. L., 2012: Localization and sampling error correction in ensemble Kalman filter data assimilation. *Mon. Wea. Rev.*, **140**, 2359–2371, https://doi.org/10.1175/MWR-D-11-00013.1.

Balci, N., A. L. Mazzucato, J. M. Restrepo, and G. R. Sell, 2012: Ensemble dynamics and bred vectors. *Mon. Wea. Rev.*, **140**, 2308–2334, https://doi.org/10.1175/MWR-D-10-05054.1.

Bannister, R. N., 2020: The ABC-DA system (v1.4): A variational data assimilation system for convective-scale assimilation research with a study of the impact of a balance constraint. *Geosci. Model Dev.*, **13**, 3789–3816, https://doi.org/10.5194/gmd-13-3789-2020.

——, 2021: Balance conditions in variational data assimilation for a high-resolution forecast model. *Quart. J. Roy. Meteor. Soc.*, **147**, 2917–2934, https://doi.org/10.1002/qj.4106.

——, S. Migliorini, A. C. Rudd, and L. H. Baker, 2017: Methods of investigating forecast error sensitivity to ensemble size in a limited-area convection-permitting ensemble. *Geosci. Model Dev.*, https://doi.org/10.5194/gmd-2017-260, preprint.

Bédard, J., J.-F. Caron, M. Buehner, S.-J. Baek, and L. Fillion, 2020: Hybrid background error covariances for a limited-area deterministic weather prediction system. *Wea. Forecasting*, **35**, 1051–1066, https://doi.org/10.1175/WAF-D-19-0069.1.

Bishop, C. H., and D. Hodyss, 2007: Flow-adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. *Quart. J. Roy. Meteor. Soc.*, **133**, 2029–2044, https://doi.org/10.1002/qj.169.

——, and ——, 2009: Ensemble covariances adaptively localized with ECO-RAP. Part I: Tests on simple error models. *Tellus*, **61A**, 84–96, https://doi.org/10.1111/j.1600-0870.2008.00371.x.

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

——, 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.

Caron, J.-F., and M. Buehner, 2022: Implementation of scale-dependent background-error covariance localization in the Canadian global deterministic prediction system. *Wea. Forecasting*, **37**, 1567–1580, https://doi.org/10.1175/WAF-D-22-0055.1.

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.

Destouches, M., T. Montmerle, Y. Michel, and B. Ménétrier, 2021: Estimating optimal localization for sampled background-error covariances of hydrometeor variables. *Quart. J. Roy. Meteor. Soc.*, **147**, 74–93, https://doi.org/10.1002/qj.3906.

Gao, S., J. Min, L. Liu, and C. Ren, 2019: The development of a hybrid EnSRF-En3DVar system for convective-scale data assimilation. *Atmos. Res.*, **229**, 208–223, https://doi.org/10.1016/j.atmosres.2019.06.024.

Gaspari, G., and S. E. 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.

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.

Hamill, T. M., and C. Snyder, 2000: A hybrid ensemble Kalman filter–3D variational analysis scheme. *Mon. Wea. Rev.*, **128**, 2905–2919, https://doi.org/10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2.

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.

Hu, M., S. G. Benjamin, T. T. Ladwig, D. C. Dowell, S. S. Weygandt, C. R. Alexander, and J. S. Whitaker, 2017: GSI three-dimensional ensemble–variational hybrid data assimilation using a global ensemble for the regional rapid refresh model. *Mon. Wea. Rev.*, **145**, 4205–4225, https://doi.org/10.1175/MWR-D-16-0418.1.

Huang, B., X. Wang, D. T. Kleist, and T. Lei, 2021: A simultaneous multiscale data assimilation using scale-dependent localization in GSI-based hybrid 4DEnVar for NCEP FV3-based

GFS. *Mon. Wea. Rev.*, **149**, 479–501, https://doi.org/10.1175/MWR-D-20-0166.1.

Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A comparison of multiscale GSI-based EnKF and 3DVar data assimilation using radar and conventional observations for midlatitude convective-scale precipitation forecasts. *Mon. Wea. Rev.*, **143**, 3087–3108, https://doi.org/10.1175/MWR-D-14-00345.1.

Kadowaki, T., Y. Ota, and S. Yokota, 2020: Introduction of a new hybrid data assimilation system for the JMA global spectral model. Research activities in Earth system modelling. Working Group on Numerical Experimentation Tech. Review 22, 4 pp., https://www.jma.go.jp/jma//jma-eng/jma-center/rsmc-hp-pub-eg/techrev/text22-2.pdf.

Kutty, G., R. Gogoi, V. Rakesh, and M. Pateria, 2020: Comparison of the performance of hybrid ETKF-3DVar and 3DVar data assimilation scheme on the forecast of tropical cyclones formed over the Bay of Bengal. *J. Earth Syst. Sci.*, **129**, 233, https://doi.org/10.1007/s12040-020-01497-8.

Lee, J. C. K., and D. M. Barker, 2023: Development of a hybrid ensemble–variational data assimilation system over the western Maritime Continent. *Wea. Forecasting*, **38**, 425–444, https://doi.org/10.1175/WAF-D-22-0113.1.

——, J. Amezcua, and R. N. Bannister, 2022: Hybrid ensemble-variational data assimilation in ABC-DA within a tropical framework. *Geosci. Model Dev.*, **15**, 6197–6219, https://doi.org/10.5194/gmd-15-6197-2022.

Lei, L., J. L. Anderson, and G. S. Romine, 2015: Empirical localization functions for ensemble Kalman filter data assimilation in regions with and without precipitation. *Mon. Wea. Rev.*, **143**, 3664–3679, https://doi.org/10.1175/MWR-D-14-00415.1.

Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203, https://doi.org/10.1256/qj.02.132.

——, and Coauthors, 2000: The Met. Office global three-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **126**, 2991–3012, https://doi.org/10.1002/qj.49712657002.

Lu, X., X. Wang, Y. Li, M. Tong, and X. Ma, 2017: GSI-based ensemble-variational hybrid data assimilation for HWRF for hurricane initialization and prediction: Impact of various error covariances for airborne radar observation assimilation. *Quart. J. Roy. Meteor. Soc.*, **143**, 223–239, https://doi.org/10.1002/qj.2914.

Magnusson, L., J. Nycander, and E. Källén, 2009: Flow-dependent versus flow-independent initial perturbations for ensemble prediction. *Tellus*, **61A**, 194–209, https://doi.org/10.1111/j.1600-0870.2008.00385.x.

Montmerle, T., Y. Michel, E. Arbogast, B. Ménétrier, and P. Brousseau, 2018: A 3D ensemble variational data assimilation scheme for the limited-area AROME model: Formulation and preliminary results. *Quart. J. Roy. Meteor. Soc.*, **144**, 2196–2215, https://doi.org/10.1002/qj.3334.

Necker, T., S. Geiss, M. Weissmann, J. Ruiz, T. Miyoshi, and G.-Y. Lien, 2020: A convective-scale 1000-member ensemble simulation and potential applications. *Quart. J. Roy. Meteor. Soc.*, **146**, 1423–1442, https://doi.org/10.1002/qj.3744.

——, D. Hinger, P. J. Griewank, T. Miyoshi, and M. Weissmann, 2023: Guidance on how to improve vertical covariance localization based on a 1000-member ensemble. *Nonlinear Processes Geophys.*, **30**, 13–29, https://doi.org/10.5194/npg-30-13-2023.

Petrie, R. E., R. N. Bannister, and M. J. P. Cullen, 2017: The ABC model: A non-hydrostatic toy model for use in

convective-scale data assimilation investigations. *Geosci. Model Dev.*, **10**, 4419–4441, https://doi.org/10.5194/gmd-10-4419-2017.

Schwartz, C. S., Z. Liu, X.-Y. Huang, Y.-H. Kuo, and C.-T. Fong, 2013: Comparing limited-area 3DVar and hybrid variational-ensemble data assimilation methods for typhoon track forecasts: Sensitivity to outer loops and vortex relocation. *Mon. Wea. Rev.*, **141**, 4350–4372, https://doi.org/10.1175/MWR-D-13-00028.1.

Shen, F., J. Min, and D. Xu, 2016: Assimilation of radar radial velocity data with the WRF hybrid ETKF–3DVar system for the prediction of Hurricane Ike (2008). *Atmos. Res.*, **169**, 127–138, https://doi.org/10.1016/j.atmosres.2015.09.019.

Singh, S. K., and V. S. Prasad, 2019: Evaluation of precipitation forecasts from 3D-Var and hybrid GSI-based system during Indian summer monsoon 2015. *Meteor. Atmos. Phys.*, **131**, 455–465, https://doi.org/10.1007/s00703-018-0580-y.

Stanley, Z., I. Grooms, and W. Kleiber, 2021: Multivariate localization functions for strongly coupled data assimilation in the bivariate Lorenz 96 system. *Nonlinear Processes Geophys.*, **28**, 565–583, https://doi.org/10.5194/npg-28-565-2021.

Vetra-Carvalho, S., M. Dixon, S. Migliorini, N. K. Nichols, and S. P. Ballard, 2012: Breakdown of hydrostatic balance at convective scales in the forecast errors in the Met Office Unified Model. *Quart. J. Roy. Meteor. Soc.*, **138**, 1709–1720, https://doi.org/10.1002/qj.1925.

Wang, X., D. M. Barker, C. Snyder, and T. M. Hamill, 2008: A hybrid ETKF–3DVar data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Mon. Wea. Rev.*, **136**, 5116–5131, https://doi.org/10.1175/2008MWR2444.1.

——, 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.

Wang, Y., and X. Wang, 2023: Simultaneous multiscale data assimilation using scale-and variable-dependent localization in EnVar for convection allowing analyses and forecasts: Methodology and experiments for a tornadic supercell. *J. Adv. Model. Earth Syst.*, **15**, e2022MS003430, https://doi.org/10.1029/2022MS003430.

Zhu, J., and R. N. Bannister, 2023: The Hydro-ABC model (version 2.0): A simplified convective-scale model with moist dynamics. *Geosci. Model Dev.*, **16**, 6067–6085, https://doi.org/10.5194/gmd-16-6067-2023.