Hybrid ensemble-variational data assimilation in ABC-DA within a tropical framework

Joshua Chun Kwang Lee[a], Javier Amezcua[b], and Ross Noel Bannister[b]

[a]Centre for Climate Research Singapore, Meteorological Service Singapore, Singapore
[b]University of Reading and UK National Centre for Earth Observation, Reading, UK

Correspondence: Joshua Chun Kwang Lee (joshua_lee@nea.gov.sg)

Abstract. Hybrid ensemble-variational data assimilation (DA) methods have gained significant traction in recent years. These methods aim to alleviate the limitations and maximise the advantages offered by ensemble or variational methods. Most existing hybrid applications focus on the mid-latitudinal context; almost none have explored its benefits in the tropical context. In this article, hybrid ensemble-variational DA is introduced to a tropical configuration of a simplified non-hydrostatic convective-scale fluid dynamics model and its existing variational framework, the ABC-DA system.

The hybrid ensemble-variational DA algorithm is developed based on the alpha control variable approach, often used in numerical weather prediction. Aspects of the algorithm such as localisation (used to mitigate sampling error caused by finite ensemble sizes) and weighting parameters (used to weight the ensemble and climatological contributions to the background error covariance matrix) are implemented. To produce the flow-dependent error modes (ensemble perturbations) for the ensemble-variational DA algorithm, an ensemble system is also designed for the ABC model, which is run alongside the hybrid DA system. A random field perturbations method is used to generate an initial ensemble, which is then propagated using the ensemble bred vectors method. This setup allows the ensemble to be centred on the hybrid control analysis. Visualisation software has been developed to focus on the diagnosis of the ensemble system.

To demonstrate the hybrid ensemble-variational DA in the ABC-DA system, sensitivity tests using observing system simulation experiments are conducted within a tropical framework. A 30-member ensemble was used to generate the error modes for the experiments. In general, the best performing configuration (with respect to the "truth") for the hybrid ensemble-variational DA system used an 80%/20% weighting on the ensemble-derived/climatological background error covariance matrix contributions. For the horizontal wind variables though, full weight on the ensemble-derived background error covariance matrix (100%/0%) resulted in the smallest cycle-averaged analysis root-mean-square errors, mainly due to large errors in the meridional wind field when contributions from the climatological background error covariance matrix were involved, possibly related to a sub-optimal background error covariance model.

The ensemble bred vectors method propagated a healthy-looking DA-centred ensemble without bimodalities or evidence of filter collapse. The ensemble was under-dispersive for some variables, but for others, the ensemble spread approximately

matched the corresponding root-mean-square errors. Reducing the number of ensemble members led to slightly larger errors across all variables, due to the introduction of larger sampling errors into the system.

Copyright statement. To be filled by journal.

1 Introduction

Data assimilation (DA) methods can traditionally be classified into three categories: variational methods, which look for a maximum-a-posteriori (MAP) estimator, Kalman-based methods, which produce a minimum variance estimator (often in an ensemble implementation), and methods which attempt to estimate full probability density functions (PDFs) without making any parametric assumptions (e.g. Markov Chain Monte Carlo and particle filters). For an introductory discussion the reader is referred to e.g. Asch et al. (2016). Each traditional DA method is subject to its own advantages and limitations, which determine the applicability in operational numerical weather prediction (NWP) systems. A wide spectrum of modern DA methods have been proposed in recent years, including hybrid ensemble-variational (hybrid-EnVar) methods which have gained significant traction. Within the category of hybrid-EnVar methods, there exist different flavours due to subtleties in the derivation and permutations arising from the usage of different variational or ensemble methods. One can modify, for instance, the elements of the problem or the solution algorithm, yielding different varieties of hybrid variants. Bannister (2017) provides a comprehensive review of the latest hybrid-EnVar methods used in modern DA.

In this article, we focus on the hybrid covariance ensemble-variational approach (Hamill and Snyder, 2000). This differs from the hybrid gain ensemble-variational DA approach (Penny, 2014), which is also commonly used. Most existing hybrid applications focus on the mid-latitudinal context and highlight the advantage of introducing flow-dependency in the error statistics. However, almost none have explored the hybrid application in the tropical context, where the characteristics of the error statistics are still poorly understood. Here, we introduce the hybrid-EnVar method to an existing convective-scale DA framework (Bannister, 2020) for a simplified non-hydrostatic fluid dynamics model (ABC model; Petrie et al., 2017), with the hope that the upgraded system can provide insights on the benefits and highlight potential issues that may arise using hybrid-EnVar methods in the tropical context. We note that this study is also the first to use a tropical configuration of the ABC-DA system.

The aims of this study are as follows:

(a) to document and test a hybrid-EnVar DA system for the ABC model, and

(b) to test generating an ensemble suitable for hybrid-EnVar DA to function.

Section 2 contains details of the existing system used in this study. Section 3 documents the development of an ABC ensemble system, necessary to generate a meaningful ensemble of ABC states, which feed into the hybrid-EnVar DA system along with

the implementation of hybrid-EnVar DA system itself. Section 4 demonstrates the use of the hybrid-EnVar DA system within a tropical framework. Three appendices provide details that may be of interest to readers familiar with ensemble initialisation, and inter-variable localisation in ensemble DA.

2 The ABC-DA system

2.1 Model equations

The ABC model used in this study was originally developed by Petrie et al. (2017) and was designed as a simplified non-hydrostatic fluid dynamics model for use in convective-scale DA experiments. It comprises solving a set of simplified partial differential equations derived from the Euler equations. A vertical slice formulation containing only dry dynamics is used (two-dimensional *x*-*z* spatial grid). This section summarises the model equations and their properties. These are:

$$\frac{\partial u}{\partial t} + B\boldsymbol{u}\cdot\nabla u + C\frac{\partial \tilde{\rho}'}{\partial x} - fv = 0, \tag{1a}$$

$$\frac{\partial v}{\partial t} + B\boldsymbol{u}\cdot\nabla v + fu = 0, \tag{1b}$$

$$\frac{\partial w}{\partial t} + B\boldsymbol{u}\cdot\nabla w + C\frac{\partial \tilde{\rho}'}{\partial z} - b' = 0, \tag{1c}$$

$$\frac{\partial \tilde{\rho}'}{\partial t} + B\nabla\cdot(\tilde{\rho}\boldsymbol{u}) = 0, \tag{1d}$$

$$\frac{\partial b'}{\partial t} + B\boldsymbol{u}\cdot\nabla b' + A^2 w = 0 \tag{1e}$$

where $\boldsymbol{u} = (u, v, w)$ is the three-dimensional wind vector of zonal, meridional and vertical wind; $\tilde{\rho}'$ and b' are perturbation quantities from a reference state of scaled density ($\tilde{\rho}$) and buoyancy respectively (see Petrie et al., 2017). The coefficients A, B and C are tunable parameters which control the pure gravity wave frequency, the modulation of the advective and divergent terms, and the relationship between the pressure and density perturbations in the equation of state, respectively. The small-scale acoustic wave speed is given by $\sqrt{BC}$. Additionally, the Coriolis parameter f can be chosen depending on the desired latitudinal position of the vertical slice. Collectively the variables u, v, w, $\tilde{\rho}'$ and b' at every grid position in the domain are referred to as the state vector $\mathbf{x}$.

2.2 Variational data assimilation

Variational DA was subsequently implemented in the ABC model by Bannister (2020), termed as the ABC-DA system. As of version 1.4 (https://doi.org/10.5281/zenodo.3531926), 3DVar and 3DVar-FGAT (First Guess at Appropriate Time) are available in the ABC-DA system. The reader is directed to Bannister (2020) for the full details of this implementation, but here we summarise the key equations in the context of 3DVar-FGAT. The 3DVar-FGAT scheme is later in this article adapted into a hybrid scheme.

2.2.1 Incremental formulation

The objective of variational DA is to find an optimal state $\mathbf{x}^a$ which minimises a cost function $J(\mathbf{x})$ (e.g. Kalnay, 2003). This cost function usually comprises two terms: one for the departure of the state with respect to the background state $\mathbf{x}^b$, and one for the departure of the state (transformed to observation space) with respect to observations $\mathbf{y}$. A third term related to any model errors can be added in the so-called weak-constraint formulation, which is not needed in this work as we do not consider model errors in our set-up. Even though the terms in J are based on Mahalanobis distances, J can be non-quadratic (with respect to the state variable) due to the non-linearities of the (often) non-linear forecast model ($\mathcal{M}_{t-1\rightarrow t}$, used in the case of 4DVar) and observation operator ($\mathcal{H}_t$). Most variational systems implement an incremental formulation of the cost function (Courtier et al., 1994) which involves iteratively linearising $\mathcal{M}_{t-1\rightarrow t}$ and $\mathcal{H}_t$ around a reference state ($\mathbf{x}^r$) and framing the problem in terms of increments to $\mathbf{x}^r$ in a series of outer loops. This allows one to find an approximate solution of a complicated non-quadratic optimisation problem by tackling a series of easier quadratic ones. To illustrate, for a DA cycle with a window from $t = 0$ to T, the incremental form of the 3DVar-FGAT cost function is:

$$J(\delta\mathbf{x}) = \frac{1}{2}(\delta\mathbf{x} - \delta\mathbf{x}^b)^\top \mathbf{B}_c^{-1}(\delta\mathbf{x} - \delta\mathbf{x}^b) + \frac{1}{2}\sum_{t=0}^{T}(\mathbf{H}_t\delta\mathbf{x} - \mathbf{d}[t])^\top \mathbf{R}_t^{-1}(\mathbf{H}_t\delta\mathbf{x} - \mathbf{d}[t]) \tag{2}$$

where $\mathbf{x} = \mathbf{x}^r + \delta\mathbf{x}$, $\mathbf{x}^r$ is a reference state, $\delta\mathbf{x}$ is the state increment, and $\delta\mathbf{x}^b$ is the difference between the background and the reference. $\mathbf{B}_c$ is the background error covariance matrix, $\mathbf{R}_t$ is the observation error covariance matrix at time t, and $\mathbf{H}_t$ is the linearised observation operator at time t. We define the innovation:

$$\mathbf{d}[t] = \mathbf{y}[t] - \mathcal{H}_t[\mathcal{M}_{0\rightarrow t}[\mathbf{x}^r]]. \tag{3}$$

Note that Eq. (2) is the same as Eq. (7) of Bannister (2020), except that the linearised forecast model $\mathbf{M}_{t-1\rightarrow t}$ has been replaced here by the identity $\mathbf{I}$ (this replacement is what distinguishes 3DVar-FGAT from 4DVar). For the first outer loop, $\mathbf{x}^r$ is set as $\mathbf{x}^b$ (i.e. $\delta\mathbf{x}^b = 0$).

2.2.2 Estimation and modelling of $\mathbf{B}_c$

A vital component in variational DA is $\mathbf{B}_c$. It is the averaged (climatological) second moment of the PDF of forecast errors of the system (Bannister, 2008a). It determines the weighting between the use of observational and background information, and it allows for the spreading of observational information spatially and between variables. We can disentangle the construction of $\mathbf{B}_c$ by considering how the background errors are first estimated, and then used in the modelling of $\mathbf{B}_c$.

In the original implementation by Bannister (2020), the estimation of the background error statistics was performed by the extraction of multiple longitude-height slices from one or more Met Office Unified Model outputs (since these were conveneiently available), which were processed to create an "ensemble" of ABC states (and subsequently ABC forecasts). This set of forecast perturbations serve as proxies for the background errors used as training data for $\mathbf{B}_c$. The validity of this prescribed source of background error statistics has not been investigated, but the approach is convenient and practical. Another way to estimate

the training data is to compute forecast differences (with different lead times and valid at the same time) over a climatological period (the National Meteorological Center method; Parrish and Derber, 1992), but as of version 1.4, this is not coded in the ABC-DA system. Instead, we introduce a different method to compute the ensemble forecasts for the training data (Section 3.1).

In many systems, $\mathbf{B}_c$ is too large to explicitly be computed using the training data. For instance, in operational models the size of the state variable can be $\mathcal{O}\left(10^9\right)$. Instead, $\mathbf{B}_c$ is often modelled through the use of a so-called control variable transform $\mathbf{U}$. Even though the ABC model is small enough for the explicit computation of $\mathbf{B}$ to be feasible, it is still far more practical to use a control variable. We introduce a control vector $\delta\boldsymbol{\chi}$ which is related to a state vector $\delta\mathbf{x}$ by:

$$\delta\mathbf{x} = \mathbf{U}\delta\boldsymbol{\chi}. \tag{4}$$

The choice of the control vector, $\delta\boldsymbol{\chi}$, and control variable transform $\mathbf{U}$ is flexible, but they dictate the eventual cross-covariances between model variables of $\delta\mathbf{x}$. In order to improve the conditioning of the incremental cost function (for more efficient minimisation), the control variables are chosen to be uncorrelated and have unit variance. Substituting Eq. (4) into Eq. (2) yields a new pre-conditioned incremental cost function:

$$J(\delta\boldsymbol{\chi}) = \frac{1}{2}(\delta\boldsymbol{\chi} - \delta\boldsymbol{\chi}^b)^\top(\delta\boldsymbol{\chi} - \delta\boldsymbol{\chi}^b) + \frac{1}{2}\sum_{t=0}^{T}(\mathbf{H}_t\mathbf{U}\delta\boldsymbol{\chi} - \mathbf{d}[t])^\top\mathbf{R}_t^{-1}(\mathbf{H}_t\mathbf{U}\delta\boldsymbol{\chi} - \mathbf{d}[t]) \tag{5}$$

where $\delta\mathbf{x}^b = \mathbf{U}\delta\boldsymbol{\chi}^b$. Since $\mathbf{B}_c$ is a symmetric and positive matrix, $\mathbf{U}$ may be chosen to be a lower triangular matrix (using a Cholesky decomposition). The implied $\mathbf{B}_c$ is given by minimising Eq. (5) with the transform in Eq. (4); $\mathbf{B}_c = \mathbf{U}\mathbf{U}^\top$ (Bannister, 2008b). It is evident that the use of a carefully designed $\mathbf{U}$ removes the need to compute $\mathbf{B}_c^{-1}$ in order to minimise the cost function. By contrast, since observation errors are assumed to be uncorrelated in the ABC-DA system, $\mathbf{R}_t$ is diagonal. Hence, there is no requirement for a separate transform since $\mathbf{R}_t^{-1}$ can be easily computed.

The calibration of $\mathbf{U}$ (and thus the implied $\mathbf{B}_c$) is usually only performed once at the start of any cycling experiment using "climatological" background error statistics, and then used for every DA cycle.

2.2.3 ABC-DA system minimisation algorithm

In the ABC-DA system, a conjugate gradient algorithm is used to find the minimiser of the cost function. Differentiating Eq. (5) with respect to $\delta\boldsymbol{\chi}$ yields the gradient $\nabla_{\delta\boldsymbol{\chi}}J$, given by:

$$\nabla_{\delta\boldsymbol{\chi}}J = \delta\boldsymbol{\chi} - \delta\boldsymbol{\chi}^b + \mathbf{U}^\top\sum_{t=0}^{T}\mathbf{H}_t^\top\mathbf{R}_t^{-1}(\mathbf{H}_t\mathbf{U}\delta\boldsymbol{\chi} - \mathbf{d}[t]) \tag{6}$$

where $\mathbf{U}^\top$ and $\mathbf{H}_t^\top$ are the adjoints of $\mathbf{U}$ and $\mathbf{H}_t$ respectively. The reader is directed to Bannister (2020) for more details. The modifications required for specific steps in order to enable hybrid-EnVar DA are highlighted later in Section 3.2.2.

3 Technical implementation of the data assimilation and forecast framework

Hybrid-EnVar schemes stem from a combination of two approaches: ensemble methods and variational methods. For the former, the archetypical example is the Kalman filter (EnKF) in its different formulations. The reader is referred to e.g. Evensen (2006) for an introduction. The variational approach has been discussed in Section 2.2. Instead of a one-off retrieval of the background error statistics from a climatological source (Section 2.2.2), the purpose of having an ensemble is to estimate time-dependent background error statistics from the ensemble forecasts valid at each cycle. As such, the background error statistics vary as the system evolves.

Accordingly, a parallel ensemble system that runs alongside the hybrid (single-trajectory) analysis is required in order to provide the background error statistics at each cycle. In this study, we explore the ensemble bred vectors (a variant of the bred vector method; EBV) method to evolve the ensemble system, which will be described in Section 3.3.1. The following sections cover the step-by-step implementation of a cycling hybrid-EnVar DA system in the ABC model, in accordance with the schematic diagram (Fig. 1) which shows the coupling between the deterministic components and the parallel-run ensemble system using the two different ensemble propagation methods. Figure 1 is explained in the remainder of Section 3.

3.1 Generation of initial ensemble of states for ABC ensemble system

This section discusses the generation of an initial ensemble, the first step in Fig. 1 (red segments), which is needed in the case of a cold start. Subsequent propagation of the ensemble can then proceed after this problem is addressed.

In the ABC model, an initial two-dimensional state can be computed from a longitude-height slice of a Unified Model output, which is a convenient approach adopted by Petrie et al. (2017). In this light, the simplest method to generate an initial ensemble is to extract different longitude-height slices from the same Unified Model output, similar to how a population of training data is generated in Bannister (2020) for the calibration of the static background error covariances as mentioned in Section 2.2.2. Another method is to simply add statistical noise to the initial ABC model state, although it is not straightforward to determine the distributions for the noise sampling (which could vary for different variables and include multi-variate correlations), so that the solutions are consistent with the underlying dynamics. The model evolution in the first few cycles may spuriously dampen or amplify the added statistical noise if it is drawn from an incorrectly-chosen distribution.

For this study, we adopt the random field perturbation method proposed by Magnusson et al. (2009) to generate the initial ensemble. The main idea relies on choosing two (assumed independent) ABC states and calculating their differences. The differences are treated as perturbations and can then be scaled to maintain a fixed amplitude between ensemble members and/or cycles, and are subsequently added to the initial ABC state computed above to generate an initial (arbitrary-sized) ensemble of states. Linear balances are approximately preserved in the resulting ensemble as only linear operations are performed on the

fields (Magnusson et al., 2009).

Unlike in an operational NWP system where archived past analyses are available, here a long "truth" simulation needs to be performed using the ABC model starting from a chosen initial state (the "truth run" in Fig. 1). To generate each ensemble member, two states from the same "truth" simulation are chosen at random. These need to be sufficiently separated in time for the assumption of independence to be valid. Following the above steps, the initial ensemble of states is given by:

$$\mathbf{x}_{cs}^{k} = \mathbf{x}_{cs}^{c} + \frac{1}{\sqrt{2}} r^{k} (\mathbf{x}_{kt1}^{tr} - \mathbf{x}_{kt2}^{tr}), \tag{7a}$$

$$r^{k} = \frac{\epsilon^{rf}}{|\mathbf{x}_{kt1}^{tr} - \mathbf{x}_{kt2}^{tr}|_{E_{tot}}^{2}}, \quad \epsilon^{rf} = \overline{|\mathbf{x}_{kt1}^{tr} - \mathbf{x}_{kt2}^{tr}|_{E_{tot}}^{2}} \tag{7b}$$

where $\mathbf{x}_{cs}^{k}$ represents the kth initial state of N ensemble members, $\mathbf{x}_{cs}^{c}$ is the initial unperturbed (hereafter referred to as control) state; the superscript "cs" refers to cold start, $\mathbf{x}_{kt1}^{tr}$ and $\mathbf{x}_{kt2}^{tr}$ are the two random states drawn from the same "truth" simulation at different times ($kt1$ and $kt2$), and r^{k} depends on the scaling factor ϵ^{rf} defined according to the total energy norm ($|\bullet|_{E_{tot}}^{2} = E_{tot}$; see Eq. (A1)) of the perturbations, to maintain a fixed amplitude (the ensemble mean $\overline{|\bullet|_{E_{tot}}^{2}}$) between ensemble members. The reason for the $\frac{1}{\sqrt{2}}$ factor is included because we are considering differences between two states so the variance of the difference is a reflection of the sum of their error variances, rather than considering differences between a state and a mean (see Appendix of Berre et al., 2006). More details and justification of the method are covered in Appendix A.

After generating the initial ensemble, the cold start members are propagated to the analysis time of the first cycle at $T+0$ (from $\mathbf{x}_{cs}^{k}$ to $\mathbf{x}_{0}^{fk}$ and from $\mathbf{x}_{cs}^{c}$ to $\mathbf{x}_{0}^{fc}$ by $\mathcal{M}_{cs\rightarrow 0}$; Fig. 1, lower brown segments). Note that for subsequent cycles, the analysis ensemble (produced from Section 3.3) and control analysis are propagated instead of cold start members (i.e. from $\mathbf{x}_{t}^{ak}$ to $\mathbf{x}_{t+1}^{fk}$ and from $\mathbf{x}_{t}^{ac}$ to $\mathbf{x}_{t+1}^{fc}$ by $\mathcal{M}_{t\rightarrow t+1}$; Fig. 1, upper right brown segments). These ensemble forecasts are then used in the DA step, described in the next section.

3.2 Hybrid ensemble-variational data assimilation

The hybrid-EnVar approach seeks to implement a hybrid background error covariance $\mathbf{B}_h$ which is a linear combination of a climatological and an ensemble-derived background error covariance matrix ($\mathbf{B}_c$ and $\mathbf{B}_e$), in the form following Hamill and Snyder (2000):

$$\mathbf{B}_h = \beta_c^2 \mathbf{B}_c + \beta_e^2 \mathbf{B}_e \tag{8}$$

where β_c^2 and β_e^2 are (positive) scalar weights often determined empirically for the algorithm. These weights are often chosen to add to unity, but this need not be the case. This approach computes $\mathbf{B}_h$ explicitly, but is not practical in an NWP system. For the ABC-DA system, the alpha control variable approach of Lorenc (2003) is instead implemented, which constructs an implied version of Eq. (8) using an alteration of the standard variational cost function and control variables (see Section 3.2.2). Wang et al. (2007) demonstrates the mathematical equivalence of both approaches.

Given the control background, which is a short-range forecast from the previous cycle ($\mathbf{x}^b = \mathbf{x}_t^{fc}$), the hybrid-EnVar DA yields the hybrid control analysis $\mathbf{x}_t^{ac}$ (Fig. 1, blue segments), which needs the ensemble members to implicitly construct the $\mathbf{B}_e$ part (recall that $\mathbf{B}_c$ in Eq. (8) is derived from the $\mathbf{U}$ transform, and $\mathbf{B}_e$ is derived from the ensemble). The steps to retrieve the hybrid control analysis are described in Section 3.2.2, but we first explain how $\mathbf{B}_e$ can be computed from the ensemble.

3.2.1 Computation of the ensemble-derived background error covariance matrix for the control analysis

At each cycle, one may compute a rectangular matrix $\mathbf{X}_t^f$ whose columns contain the scaled differences between the ensemble forecasts (i.e. $\mathbf{x}_t^{fk}$ for the kth member forecast valid at time t) and the ensemble mean ($\bar{\mathbf{x}}_t^f$):

$$\mathbf{X}_t^f = \frac{1}{\sqrt{N-1}}(\mathbf{x}_t^{f1} - \bar{\mathbf{x}}_t^f, \mathbf{x}_t^{f2} - \bar{\mathbf{x}}_t^f, \ldots, \mathbf{x}_t^{fN} - \bar{\mathbf{x}}_t^f) = (\mathbf{x}_t'^1, \mathbf{x}_t'^2, \ldots, \mathbf{x}_t'^N) \tag{9}$$

where $\mathbf{x}_t'^k$ are the scaled error modes valid at time t. The ensemble-derived background error covariance matrix (at time t) $\mathbf{P}_e^f[t]$ is explicitly given by the outer product:

$$\mathbf{P}_e^f[t] = \mathbf{X}_t^f \mathbf{X}_t^{f\top}. \tag{10}$$

As we shall see in Section 3.2.2, this matrix is not computed explicitly, although parts of it are computed explicitly for visualisation purposes later in this article.

In the limit where N tends to infinity, or where N is far greater than the degrees of freedom of the state n ($N \gg n$), $\mathbf{P}_e^f[t]$ may be full rank. In practice, however, a small number of ensemble members ($N \ll n$) will inevitably lead to sampling error and a rank-deficient matrix. Houtekamer and Mitchell (2001) proposed mitigating this problem by performing a Schur product of $\mathbf{P}_e^f$ with a correlation matrix (or localisation matrix) $\mathbf{L}$:

$$\mathbf{B}_e = \mathbf{L} \circ \mathbf{P}_e^f[t]. \tag{11}$$

This seeks to address the sampling error by damping the long-range background error covariances, as well as effectively increasing the rank of $\mathbf{P}_e^f[t]$. The spatial and multi-variate aspects of the localisation matrix are further discussed in Section 3.2.3, including how this can be performed without constructing explicit matrices.

3.2.2 Alpha control variable transform

Following the approach of Lorenc (2003), we introduce an ensemble-related penalty in the variational cost function. This requires constructing so-called alpha fields $\boldsymbol{\alpha}^k$ (part of a new set of mutually uncorrelated control variables) associated with each ensemble member k, and constrained to have covariance $\mathbf{L}$ (the localisation matrix, as used in Eq. (11)). The number of elements in $\boldsymbol{\alpha}^k$ must be the same as the state vector of $\mathbf{x}_t'^k$ (number of model gridpoints N_g $\times$ number of model variables

N_{var}). The modified cost function is:

$$J(\delta\boldsymbol{\chi},\boldsymbol{\alpha}^1,\boldsymbol{\alpha}^2,\ldots,\boldsymbol{\alpha}^N) = \overbrace{\frac{1}{2}(\delta\boldsymbol{\chi}-\delta\boldsymbol{\chi}^b)^\top(\delta\boldsymbol{\chi}-\delta\boldsymbol{\chi}^b)}^{J_b} + \overbrace{\frac{1}{2}\sum_{t=0}^{T}(\mathbf{H}_t\delta\mathbf{x}-\mathbf{d}[t])^\top\mathbf{R}_t^{-1}(\mathbf{H}_t\delta\mathbf{x}-\mathbf{d}[t])}^{J_o} + \overbrace{\frac{1}{2}\sum_{k=1}^{N}\boldsymbol{\alpha}^{k\top}\mathbf{L}^{-1}\boldsymbol{\alpha}^k}^{J_e} \tag{12a}$$

$$\text{with } \delta\mathbf{x} = \beta_c\mathbf{U}\delta\boldsymbol{\chi} + \beta_e\sum_{k=1}^{N}\mathbf{x}_t'^k \circ \boldsymbol{\alpha}^k, \tag{12b}$$

where J_b, J_o and J_e are the background, observation and ensemble penalties respectively. Equation (12a) is an extension of Eq. (5), and Eq. (12b) is an extension of Eq. (4), the hybrid control variable transform. Together these equations make up the hybrid scheme.

Similar to the way that $\mathbf{B}_c$ can be decomposed as $\mathbf{B}_c = \mathbf{U}\mathbf{U}^\top$, $\mathbf{L}$ can be decomposed in terms of the alpha control variable transform, $\mathbf{U}^\alpha$, i.e. $\mathbf{L} = \mathbf{U}^\alpha\mathbf{U}^{\alpha\top}$. Consider an alpha control vector $\boldsymbol{\chi}^{\alpha k}$ (again associated with ensemble member k) which is related to the alpha field $\boldsymbol{\alpha}^k$ via:

$$\boldsymbol{\alpha}^k = \mathbf{U}^\alpha\boldsymbol{\chi}^{\alpha k}. \tag{13}$$

Substituting Eq. (13) into Eq. (12a) yields:

$$J(\delta\boldsymbol{\chi},\boldsymbol{\chi}^{\alpha 1},\boldsymbol{\chi}^{\alpha 2},\ldots,\boldsymbol{\chi}^{\alpha N}) = J_b + J_o + \frac{1}{2}\sum_{k=1}^{N}\boldsymbol{\chi}^{\alpha k\top}\boldsymbol{\chi}^{\alpha k} \tag{14a}$$

$$\text{with } \delta\mathbf{x} = \beta_c\mathbf{U}\delta\boldsymbol{\chi} + \beta_e\sum_{k=1}^{N}\mathbf{x}_t'^k \circ (\mathbf{U}^\alpha\boldsymbol{\chi}^{\alpha k}), \tag{14b}$$

$$\mathbf{x}_t^{ac} = \mathbf{x}^r + \delta\mathbf{x}^a. \tag{14c}$$

The variational problem (Eq. (14a)) is minimised with respect to the collective set of control vectors, comprising a part that is associated with $\mathbf{B}_c$ ($\delta\boldsymbol{\chi}$), and parts that are associated with $\mathbf{B}_e$ ($\boldsymbol{\chi}^{\alpha 1},\boldsymbol{\chi}^{\alpha 2},\ldots,\boldsymbol{\chi}^{\alpha N}$). Together, these are combined using the hybrid transform (Eq. (14b)) to give the particular $\delta\mathbf{x}$ that minimises Eq. (14a), namely $\delta\mathbf{x}^a$. This gives the analysis $\mathbf{x}_t^{ac}$ in Eq. (14c).

The total implied covariance matrix (that is effectively seen by the DA) is formally given by:

$$\mathbf{B}_h = \beta_c^2\mathbf{U}\mathbf{U}^\top + \beta_e^2(\mathbf{U}^\alpha\mathbf{U}^{\alpha\top}) \circ (\mathbf{X}_t^f\mathbf{X}_t^{f\top}), \tag{15}$$

which is a linear combination of the implied $\mathbf{B}_c$ and $\mathbf{B}_e$ (without explicitly constructing either), and is element-wise equivalent to the explicit hybrid covariance in Eq. (8).

Next, we reproduce the minimisation algorithm steps Section 3.5 of Bannister (2020), and highlight (in blue) the modifications required when the hybrid-EnVar scheme is enabled:

1. Set the reference state at $t = 0$ to the background state $\mathbf{x}^r = \mathbf{x}^b$. Decide values for N, β_c, and β_e.

2. Do the outer loop.

   (a) For the first outer loop, $\delta\boldsymbol{\chi}^b = 0$; otherwise, compute $\delta\boldsymbol{\chi}^b = \mathbf{U}^{-1}(\mathbf{x}^b - \mathbf{x}^r)$.

   (b) Compute $\mathbf{x}^r[t]$ over the time window, $1 \leq t \leq T$, with the non-linear model $\mathbf{x}^r[t] = \mathcal{M}_{t-1\to t}(\mathbf{x}^r[t-1])$.

   (c) Compute the reference state's observations: $\mathbf{y}^{mr}[t] = \mathcal{H}_t(\mathbf{x}^r[t])$.

   (d) Compute the differences: $\mathbf{d}[t] = \mathbf{y}[t] - \mathbf{y}^{mr}[t]$.

   (e) Set $\delta\boldsymbol{\chi} = 0$, $\delta\mathbf{x} = 0$, and $\boldsymbol{\chi}^{\alpha k} = 0$, $1 \leq k \leq N$.

   (f) Do the inner loop.

      i. Integrate the perturbation trajectory over the time window, $1 \leq t \leq T$, with the linear forecast model: $\delta\mathbf{x}[t] = \mathbf{M}_{t-1\to t}\delta\mathbf{x}[t-1]$.

      ii. Compute the perturbations to the model observations: $\delta\mathbf{y}^m[t] = \mathbf{H}_t\delta\mathbf{x}[t]$.

      iii. Compute $\boldsymbol{\Delta}[t]$ vectors as defined as $\boldsymbol{\Delta}[t] = \mathbf{H}_t^\top\mathbf{R}_t^{-1}(\delta\mathbf{y}^m[t] - \mathbf{d}[t])$.

      iv. Set the adjoint state $\boldsymbol{\lambda}[T+1] = 0$.

      v. Integrate the following adjoint equation backwards in time, $T \geq t \geq 0$: $\boldsymbol{\lambda}[t] = \boldsymbol{\Delta}[t] + \mathbf{M}_{t\to t+1}^\top\boldsymbol{\lambda}[t+1]$.

      vi. Compute the gradient as follows: $\nabla_{\delta\boldsymbol{\chi}}J = \delta\boldsymbol{\chi} - \delta\boldsymbol{\chi}^b + \beta_c\mathbf{U}^\top\boldsymbol{\lambda}[0]$, and $\nabla_{\boldsymbol{\chi}^{\alpha k}}J = \boldsymbol{\chi}^{\alpha k} + \beta_e\mathbf{U}^{\alpha\top}(\mathbf{x}_t'^k \circ \boldsymbol{\lambda}[0])$. These are the gradients with respect to each control vector segment, $1 \leq k \leq N$.

      vii. Use the conjugate gradient algorithm to adjust $\delta\boldsymbol{\chi}$ and $\boldsymbol{\chi}^{\alpha k}$ to reduce the value of J. Note that the cost function is $J = J_b + J_o + J_e$ (Eq. (14a)).

      viii. Compute the new increment in model space using the control variable transform and alpha control variable transform: $\delta\mathbf{x} = \beta_c\mathbf{U}\delta\boldsymbol{\chi} + \beta_e\sum_{k=1}^N \mathbf{x}_t'^k \circ (\mathbf{U}^\alpha\boldsymbol{\chi}^{\alpha k})$ (Eq. (14b)).

      ix. Go to step 2fi until the inner-loop convergence criterion is satisfied.

   (g) Update the reference state: $\mathbf{x}^r \to \mathbf{x}^r + \delta\mathbf{x}$.

   (h) Go to step 2a until the outer-loop convergence criterion is satisfied. At convergence, set the hybrid control analysis $\mathbf{x}_t^{ac} = \mathbf{x}^r$.

3. Run a non-linear forecast from $\mathbf{x}_t^{ac}$ for the background of the next cycle and longer forecasts if required.

3.2.3 Inter-variable and spatial localisation

Localisation of the ensemble-derived background error covariance matrix, as in Eq. (11), is required to mitigate sampling error, which can dominate the computed covariance between distant points (Hamill et al., 2001). Localisation opens up a range of options and raises some pertinent questions: Should we localise only in space (and should these spatial localisation matrices depend on the variable), or should we additionally include localisation between different model variables? This depends on the

design of $\boldsymbol{\chi}^{\alpha k}$ and $\mathbf{U}^{\alpha}$ in Eq. (13), and the implied $\mathbf{L}$. If $\boldsymbol{\chi}^{\alpha k}$ only depends on gridpoint location (i.e. it need only be of length N_g), then $\mathbf{U}^{\alpha}$ must be rectangular ($N_g N_{var} \times N_g$) so that $\boldsymbol{\alpha}^k$ has length $N_g N_{var}$ required for the Schur product in Eq. (12b). This approach was adopted by Wang et al. (2008a), except that $\boldsymbol{\chi}^{\alpha k}$ was only dependent on *horizontal* gridpoint locations. By design, $\mathbf{U}^{\alpha}$ functions to "use the same $\boldsymbol{\chi}^{\alpha k}$ for each model variable and model level" (i.e. repeated rows in $\mathbf{U}^{\alpha}$) so the Schur product in Eq. (14b) can be computed. This point was not highlighted in the description of Eq. (1) of Wang et al. (2008a).

If $\boldsymbol{\chi}^{\alpha k}$ only has N_g elements, the implied $\mathbf{L} = \mathbf{U}^{\alpha}\mathbf{U}^{\alpha\top}$ can only involve spatial localisation, so full inter-variable covariances (as found from the raw ensemble) are retained (see Appendix B). Alternatively, if $\boldsymbol{\chi}^{\alpha k}$ is full length (i.e. of length $N_g N_{var}$) with independent fields for each model variable, $\mathbf{U}^{\alpha}$ is square ($N_g N_{var} \times N_g N_{var}$) so it is possible to use this transform to damp the ensemble-derived covariances between different variables and spatial locations. Nonetheless, there is flexibility to still retain the full inter-variable covariances in $\mathbf{L}$ depending on the design of $\mathbf{U}^{\alpha}$. For the record, a proof of the equivalence between this approach ($\boldsymbol{\chi}^{\alpha k}$ with length $N_g N_{var}$) with full inter-variable covariances retained, and the approach where $\boldsymbol{\chi}^{\alpha k}$ is of length N_g is included in Appendix B. More complex designs of $\mathbf{U}^{\alpha}$ which allow the retention of inter-variable covariances only between certain model variables, or using different spatial localisation length-scales for different model variables are also possible. In practice, these may be useful in convective-scale data assimilation, particularly when hydrometeor variables are involved (Xuguang Wang, personal communication).

In the ABC-DA system, $\mathbf{U}^{\alpha}$ is further decomposed into the horizontal $\mathbf{U}^{\alpha}_{horiz}$ and vertical $\mathbf{U}^{\alpha}_{vert}$ localisation transforms, similar to the decomposition of $\mathbf{U}$ in Bannister (2020). The series of transforms is given by:

$$\mathbf{U}^{\alpha} = \mathbf{U}^{\alpha}_{vert}\mathbf{U}^{\alpha}_{horiz} \tag{16}$$

thus treating the vertical and horizontal localisation separately.

Initial tests constructed $\mathbf{U}^{\alpha}_{horiz}$ using a Fourier decomposition (as is done in the standard horizontal transform of $\mathbf{U}$ – see Bannister (2020), but this yielded undesirable small negative correlations at longer localisation distances (presumably related to the Gibbs phenomenon, not shown). Thus a different approach was adopted as the basis for populating $\mathbf{U}^{\alpha}_{horiz}$, using the eigen-decomposition of a target horizontal localisation matrix $\mathbf{L}_{horiz} = \mathbf{U}^{\alpha}_{horiz}\mathbf{U}^{\alpha\top}_{horiz}$ (where $\mathbf{U}^{\alpha}_{horiz} = \mathbf{F}^{\alpha}_{horiz}(\boldsymbol{\Lambda}^{\alpha}_{horiz})^{1/2}$, and $\mathbf{F}^{\alpha}_{horiz}$ and $\boldsymbol{\Lambda}^{\alpha}_{horiz}$ are the eigenvectors and eigenvalues respectively of the imposed horizontal localisation matrix). We start by constructing $\mathbf{L}_{horiz}$ using the fifth-order piecewise Gaspari-Cohn function with a horizontal localisation length-scale h^{α} (equivalent to c, with $a = 1/2$, in Eq. (4.10) of Gaspari and Cohn, 1999). This function is approximately Gaussian over a compact support. As the ABC model uses periodic boundary conditions, $\mathbf{L}_{horiz}$ must be designed to be circulant and account for "overlapping tails" of the Gaspari-Cohn function when h^{α} is larger than half the domain size. In the "overlapping tails" regime, the correlation function does not satisfies the "space-limited" requirement described in (Gaspari and Cohn, 1999). Thus, the resulting $\mathbf{L}_{horiz}$ is found to not be positive semi-definite when h^{α} is large and tends to infinity, so the horizontal eigenvectors associated with the negative eigenvalues need to be truncated. Offline testing in idealised setups and within the

ABC-DA system showed that with the remaining eigenvectors, $\mathbf{U}^{\alpha}_{horiz}\mathbf{U}^{\alpha\top}_{horiz}$ is a good approximation for $\mathbf{L}_{horiz}$. It is also possible to scale the remaining eigenvalues to restore the initial total variances for a better approximation. Figure 2 illustrates the implied correlation function ($h^{\alpha} = 250\ \mathrm{km}$) with respect to longitudinal gridpoint 50, for an ABC-DA system with 364 longitudinal gridpoints and a 1.5 km horizontal grid, retrieved using the above steps. The original Gaspari-Cohn function with "overlapping tails" is compared with the implied correlation function reconstructed from the eigenvectors and eigenvalues of the eigen-decomposition of $\mathbf{L}_{horiz}$ (Fig. 2a), the resulting implied correlation function when negative eigenvectors/eigenvalues are truncated (Fig. 2b), and the resulting implied correlation function after further restoration of the initial total variances by scaling (Fig. 2c). Note that in this example, the threshold for which negative eigenvalues appear is $h^{\alpha} \approx 138.33\ \mathrm{km}$, found empirically. In the current version of the ABC-DA system, the scaling to restore initial total variances is not implemented yet.

To populate $\mathbf{U}^{\alpha}_{vert}$, a similar approach is adopted; a Gaspari-Cohn function is used with vertical localisation length-scale v^{α}. Note that the target vertical localisation matrix $\mathbf{L}_{vert}$ is a correlation matrix and so must be positive semi-definite, so truncation of eigenvectors is not required. Since $\mathbf{U}^{\alpha}_{horiz}$ and $\mathbf{U}^{\alpha}_{vert}$ are separate, it is possible to have a different $\mathbf{L}_{vert}$ for each horizontal eigenvector. However, for simplicity, the default setup in the ABC-DA system uses the same $\mathbf{L}_{vert}$ for each horizontal eigenvector. As for $\mathbf{L}_{horiz}$ the vertical eigenvectors are retrieved through the eigen-decomposition of $\mathbf{L}_{vert}$ and used to populate $\mathbf{U}^{\alpha}_{vert}$ such that $\mathbf{L}_{vert} = \mathbf{U}^{\alpha}_{vert}\mathbf{U}^{\alpha\top}_{vert}$ (where $\mathbf{U}^{\alpha}_{vert} = \mathbf{F}^{\alpha}_{vert}(\mathbf{\Lambda}^{\alpha}_{vert})^{1/2}$, and $\mathbf{F}^{\alpha}_{vert}$ and $\mathbf{\Lambda}^{\alpha}_{vert}$ are the eigenvectors and eigenvalues respectively of the imposed vertical localisation matrix).

3.3 Generation of ABC analysis ensemble

After the initial ensemble has been generated (Section 3.1) using the method of Magnusson et al. (2009), and the initial hybrid control analysis has been retrieved (Section 3.2), the next step is to generate analysis ensembles (Fig. 1, green segments). The ensemble then proceed, via the forecast model (Fig. 1, upper right brown segments), as a forecast ensemble which is used in the next hybrid DA step. Various methods have been used in previous studies, such as singular vectors (Buizza et al., 1993), bred vectors (Toth and Kalnay, 1993; 1997), perturbed observations (Houtekamer and Derome, 1995), Ensemble Kalman filter (EnKF; Evensen, 1994), Ensemble Transform Kalman Filter (ETKF; Bishop et al., 2001), and other square-root filters. Here, we focus mainly on the EBV method. This method has useful information about the nature of dynamical error growth about the analysis state at each cycle, but is uninformed about the observation network.

The ensemble, which is run in parallel to the hybrid DA are important components in hybrid-EnVar since they provide the means to compute $\mathbf{X}^{f}_{t}$ in Eq. (9). The success of the scheme depends on the extent to which the ensemble forecasts can appropriately represent the background error statistics for the ABC-DA system, so proper design of the ensemble system is critical. We construct the analysis ensemble around the hybrid control analysis (i.e. adding ensemble perturbations to the hybrid control analysis; see below), so the ensemble is "DA-centred".

3.3.1 Ensemble bred vectors

In this approach, we consider a variant of the bred vectors method – the EBV method (Balci et al., 2012). The basic bred vectors method (Toth and Kalnay, 1993) is generally simple to implement and has a cheap computational cost. The idea relies on breeding perturbations by running the non-linear forecast model for a fixed period for pairs of forecast ensemble members, taking the difference between the two forecasts, and then scaling the difference to have a specified and fixed amplitude. This process "breeds" the fastest growing error modes. This is repeated to retrieve the required number of error modes, and the resulting perturbations are respectively added to the hybrid control analysis to generate an analysis ensemble. The intention is that these perturbations should adequately sample the space of possible analysis errors.

The main difference between the bred vectors and EBV methods lies in the scaling of the perturbations at each cycle. In the bred vectors method, the perturbations are scaled to maintain a fixed amplitude across cycles for each ensemble member. The scaling is independent for each ensemble member and there is therefore no mechanism to compare the dynamics with perturbations of the other members. The EBV method (Balci et al., 2012) on the other hand involves a global scaling factor, which depends on the amplitude of the largest perturbation, and offers better insights into the relative behaviour of nearby ensemble trajectories. Perturbations that have an amplitude smaller than the largest perturbation of the ensemble then play a smaller role after scaling; in other words, ensemble trajectories that are clustered around the control member trajectory are less important for identifying dominant directions of error growth. An in-depth comparison of the bred vectors and EBV methods is provided in Balci et al. (2012).

To generate the analysis ensemble, a target maximum amplitude ϵ_0 is required, but this opens the question on what to choose for ϵ_0. Here, we use $\epsilon_0 = \epsilon^{rf}$ (the mean total energy norm of the initial ensemble of states), although other choices are possible, such as running a series of experiments and finding the average analysis error to estimate the analysis uncertainty. This scaling factor is fixed across perturbations, so at each cycle the perturbations are scaled by the same ratio r_t^{ebv}, which is used and defined as follows:

$$\delta \mathbf{x}_t^{fk} = \frac{1}{\sqrt{2}} r_t^{ebv} (\mathbf{x}_t^{fk} - \mathbf{x}_t^{fc}), \quad r_t^{ebv} = \frac{\epsilon_0}{\max[|\mathbf{x}_t^{fk} - \mathbf{x}_t^{fc}|^2_{E_{tot}}]}, \tag{17a}$$

$$\mathbf{x}_t^{ak} = \mathbf{x}_t^{ac} + \delta \mathbf{x}_t^{fk} \tag{17b}$$

where $\mathbf{x}_t^{fk}$ and $\mathbf{x}_t^{fc}$ are the kth ensemble and control forecast from the previous cycle respectively, $\delta \mathbf{x}_t^{fk}$ is the kth scaled ensemble perturbation at time t. The kth member of the analysis ensemble $\mathbf{x}_t^{ak}$ is centred on the hybrid control analysis $\mathbf{x}_t^{ac}$ produced by the DA step (Section 3.2). The $\frac{1}{\sqrt{2}}$ factor is not necessarily required because we are computing differences between individual ensemble member forecasts and the same control member forecast, but we have included it as a deflation factor with our choice of ϵ_0. It is worth noting that the EBV method is not formally consistent with Kalman filter theory, but will not suffer from filter collapse as long as the ϵ_0 chosen is well-tuned.

It is not uncommon to use such a set-up (i.e. separate hybrid deterministic and ensemble systems for, respectively, the first and second moments of the posterior). While the hybrid control analysis involves both the ensemble and climatological contributions to the background error covariance matrix Eq. (15), the computation of analysis perturbations involves only the forecast ensemble and neglects the climatological contributions. While this is a formal discrepancy, we assume that this setup is an adequate from a practical perspective.

4 Data assimilation experiments using the hybrid-EnVar scheme in a tropical setting

For this study, the Unified Model output is retrieved from a tropical convective-scale NWP system over the Maritime Continent (SINGV-DA; Heng et al., 2020). SINGV-DA operates on a 1.5 km core horizontal grid, with a model top height of 38.5 km. Longitude-height slices of fields u and v around 2°N are extracted from the SINGV-DA output by placing these fields onto the 1.5 km ABC model grid for the lowest 60 levels (up to around 18 km height), resulting in a 364 $\times$ 60 ABC model grid. These initial u and v fields are then modified to make them compatible with the ABC model's periodic boundary conditions, and the remaining fields, w, $\tilde{\rho}'$, and b' are derived following the procedure in Section 4.1 of Bannister (2020). For the ensemble system, 30 initial ensemble members are generated following Section 3.1, excluding the control state reconfigured from the longitude-height slice. This is a typical ensemble size used in operational NWP systems.

To represent a tropical setting of the ABC model, a value of $f = 10^{-5}\ \mathrm{s}^{-1}$ is used. This corresponds approximately to a value of f at a latitude of 4°N in an NWP system. The other model parameters are set as follows: $A = 0.02\ \mathrm{s}^{-1}$, $B = 0.01$, $C = 10^4\ \mathrm{m}^2\mathrm{s}^{-2}$. A series of hourly-cycling multi-cycle DA observation system simulation experiments are conducted to demonstrate the incorporation of ensemble-derived background error covariances in hybrid-EnVar DA. The hybrid extension of 3DVar-FGAT may be termed hybrid-En3DVar-FGAT.

We run four experiments with the following configurations:

(a) 100% $\mathbf{B}_c$ (i.e. no flow-dependency, equivalent to 3DVar-FGAT),

(b) 50% $\mathbf{B}_c$, 50% $\mathbf{B}_e$; hybrid-En3DVar-FGAT (i.e. flow-dependency with an equal contribution from $\mathbf{B}_c$ and $\mathbf{B}_e$),

(c) 20% $\mathbf{B}_c$, 80% $\mathbf{B}_e$; hybrid-En3DVar-FGAT (i.e. flow-dependency with most contribution from $\mathbf{B}_e$),

(d) 100% $\mathbf{B}_e$; pure En3DVar-FGAT (i.e. no contribution from $\mathbf{B}_c$).

These experiments are referred to as EBV(a) to EBV(d) accordingly. Note that configuration (a) does not use ensemble information, but the experiment is named as EBV(a) for ease of reference.

4.1 Implied background error covariances

To show the workings of Eq. (8) (or equivalently Eq. (15)) and the localisation, we compute a selection of implied background error covariances, with the various weights assigned to $\mathbf{B}_c$ and $\mathbf{B}_e$ (as the above configurations (a) to (d)). This is similar to performing single observation experiments and retrieving the analysis increments. For $\mathbf{B}_e$, spatial localisation lengthscales of $h^\alpha = 100\ \mathrm{km}$ and $v^\alpha = 5\ \mathrm{km}$ are set, and with no inter-variable localisation. The implied background error covariances are valid for the time of the first cycle after a cold start (i.e. at T+0) and use one-hour forecast ensemble perturbations. For $\mathbf{B}_c$, the same ensemble is used as training data to calibrate $\mathbf{U}$.

Figure 3 shows the implied background error covariances of $\tilde{\rho}'$, v and b' with respect to a fixed $\tilde{\rho}'$ point in the middle of the domain for the four configurations (four rows). Configuration (a) (top row) shows the implied background error covariances that are modelled purely by $\mathbf{U}$, and configuration (d) (bottom row) shows the purely ensemble-derived covariances with spatial localisation (implied by $\mathbf{U}^\alpha$). Configurations (b) and (c) are linear combinations with different weights, as demonstrated by Eq. (8).

For $\tilde{\rho}'$-$\tilde{\rho}'$ covariances in configuration (a) for pure 3DVar-FGAT, the central region of auto-correlation has horizontal and vertical lengthscales of 100 and $2\ \mathrm{km}$ respectively, and is surrounded by oscillations, possibly reflecting the dominant gravity wave propagation. The vertical lengthscale here is smaller than that found in the mid-latitude study of Bannister (2020), and such a contrast between low and higher latitudes is seen in other studies, e.g. Ingleby (2001). This can be compared with configuration (d) for pure (and localised) En3DVar-FGAT, which shows a narrower but taller region of auto-correlation. Most of the oscillations are beyond the localisation region so are not visible apart from small negative values to the west of the auto-correlation.

The $\tilde{\rho}'$-v covariances in configuration (a) follow from the use of geostrophic balance, which is manifested in $\mathbf{U}$ (Eq. (2a) in Bannister (2020)). The v pattern is consistent with an anti-cyclonic field around the source point (i.e. positive and negative v covariances west and east of the positive $\tilde{\rho}'$ source point respectively). Since f is small, these covariances are also small. Contrasting this with configuration (d), it appears that the ensemble-derived covariances are more substantial, and are of opposite sign, suggesting that there exists some (other) mass-wind relationship manifested in $\mathbf{B}_e$ (e.g. related to equatorial gravity wave processes not represented in $\mathbf{U}$).

The $\tilde{\rho}'$-b' covariances in configuration (a) follow from the use of hydrostatic balance, again manifested in $\mathbf{U}$ (Eq. (3) in Bannister (2020)). Hydrostatic balance relates b' increments with the vertical gradient of $\tilde{\rho}'$ increments, and the top-left and top-right panels are confirmed to be consistent in this way. In configuration (d), the $\tilde{\rho}'$-b' covariances have similar vertical patterns within the localisation region, although are weaker. As noted in Bannister (2020), the $\mathbf{B}_c$ covariances tend to be larger than their $\mathbf{B}_e$

counterparts even though both use the same training data for calibrating the transforms, but the broad structures are similar.

Note that the implied background error covariances between $\tilde{\rho}'$ and u and w are each zero in configuration (a) by definition of $\mathbf{U}$ (Bannister, 2020). By contrast, in configurations (b), (c) and (d), the implied background error covariances are prescribed directly between the associated model variables, implied by $\mathbf{U}^{\alpha}$ (not shown).

Even though the multi-variate background error relationships relevant to the tropics are likely to be different from those at mid-latitudes, the same balance conditions designed for mid-latitudes are often used in $\mathbf{U}$ for tropical settings (as they are here). By exploring ensemble-derived multi-variate background error relationships, we may be able to identify alternative balances inherent in the dynamical fields. This will be explored in a separate study.

4.2 Details of observation system simulation experiments

In all experiments, 200 observations of each variable (u, v, w, $\tilde{\rho}'$ and b'), which are equally spaced throughout the domain, are assimilated at every hourly cycle. The observations are sampled from a "truth" run, with added observation noise following a Gaussian distribution. The observation error standard deviations are chosen to be approximately 10% of the variable's root-mean-square value as seen in the "truth" run. These are 0.2 ms^{-1}, 0.2 ms^{-1}, 0.01 ms^{-1}, 1.5×10^{-4} and 1.5×10^{-3} ms^{-2} for u, v, w, $\tilde{\rho}'$ and b' respectively. All generated observations are valid at the background/analysis time of each cycle, so there is no difference in the analysis between 3DVar and 3DVar-FGAT (and indeed 4DVar if it were implemented). The number of observations are $\approx 1\%$ of the degrees of freedom of the state (both spatial and multi-variate), to mimic how observations are sparse in the tropical setting.

The initial background of the deterministic system is determined from the initial "truth" plus a small background noise perturbation $\delta\mathbf{x} = \mathbf{U}\delta\boldsymbol{\chi}$, where $\delta\boldsymbol{\chi}$ is drawn randomly from $\mathcal{N}(0, \mathbf{I})$. In order to reduce the effect of random noise on the experiments, the ABC-DA system is first spun-up for 50 one-hour cycles, with the expectation that the DA-centred ensemble system and deterministic system will have lost memory of the particular way that the system was initialised from a cold start. The information from the 50th cycle of spin-up is then used in the first cycle of all the actual experiments.

During the spin-up configuration testing, we noticed that the inclusion of vertical localisation in $\mathbf{B}_e$ was particularly detrimental to the evolution of the w field. Investigation revealed that this was due to the introduction of hydrostatic imbalance in the analysis increments (not shown). A similar well-known issue to do with horizontal localisation introducing geostrophic imbalance was discussed in Section 3c of Lorenc (2003). We include more comments on the hydrostatic imbalance issue in Appendix C. For this reason, vertical localisation was excluded in $\mathbf{B}_e$ in the spin-up process and in all experiments. After inspecting the other fields during spin-up configuration testing, we found that in most configurations, the hybrid control analysis gradually converged around the "truth" run as the observations were assimilated over the 50 spin-up cycles, which is logically expected. Particularly using the EBV(d) configuration, the evolution of the fields were reasonably in line with the "truth" run,

so this was the chosen configuration which was run for 50 spin-up cycles, referred to as the spin-up run.

To ensure a fair comparison in the results, the spin-up run provides the same starting background (50th cycle forecast), empirically tuned EBV ensemble and ensemble-derived error modes (if required) for the first cycle of all the actual experiments. Each experiment is run for 50 cycles and only differ in the DA algorithm configurations after spin-up. Where $\mathbf{B}_e$ is required, we use $h^{\alpha} = 20\ \mathrm{km}$ for the horizontal localisation, while not performing any inter-variable localisation (see Appendix B). The horizontal localisation length-scale was determined by comparing with the horizontal distance between adjacent observations ($\approx$23 km). For the minimisation, a total of 75 inner loops within a single outer loop is used. This was determined after testing to ensure that sufficient convergence was attained for all cycles in the experiments. This is demonstrated in Fig. 4, which shows the minimisation of the cost function for the first cycle of the EBV experiments. For this cycle, the cost function was minimised the fastest and slowest in EBV(d) and (a) respectively. The analysis misfit to assimilated observations was also the largest in EBV(d) with $J_o \approx 1500$ after minimisation. However, it is important to note that this metric is not a particularly useful indicator of analysis quality, but rather how each scheme draws the analysis towards the observations (Wang et al., 2008b). We would expect that the minimum of the cost function would approximate half the number of observations (i.e. $J_{min} \approx \frac{N_{obs}}{2} = 500$, the expected value of a chi-squared PDF), so EBV(c) neatly matches our expectations.

In addition to the experiments, a free background run, hereafter referred to as FreeBG, is performed starting from the same 50th cycle forecast of the spin-up run. This is used as the control run to assess if the DA in the experiments is adding value by bringing the deterministic run trajectories closer to the "truth" or if the trajectories are simply following the natural evolution of the system and neglecting the observational information.

4.3 Sensitivity to weighting of $\mathbf{B}_c$ and $\mathbf{B}_e$

Typically for the hybrid-EnVar scheme, tuning of the weights (β_c^2 and β_e^2) for $\mathbf{B}_c$ and $\mathbf{B}_e$ is performed empirically to to assess the best configuration which combines the benefits from both sources of background error statistics. Figure 5 shows the comparison of domain-averaged analysis errors (root-mean-square errors; RMSE) with respect to the "truth" for the EBV and FreeBG experiments.

The cycle-averaged analysis errors (Fig. 5, bottom right panel) for all prognostic variables except v are generally smaller for the EBV experiments compared to FreeBG, with an RMSE ratio less than 1. During the simulation, the w, $\tilde{\rho}'$ and b' errors were decreasing, suggesting that the deterministic run trajectories of the EBV experiments were converging around the "truth" because of the availability of observational information. The u analysis errors were decreasing in EBV(c) and EBV(d), but were increasing in EBV(a) and EBV(b). Throughout the 50 cycles, the v analysis errors were generally increasing in the EBV experiments. This peculiar issue was exacerbated when the weighting towards $\mathbf{B}_c$ was increased, suggesting that the issue originates from $\mathbf{B}_c$. A feature of the u, $\tilde{\rho}'$, and b' RMSE time sequences is the eight-hour periodicity, which is also apparent in the basic dynamical root-mean-square fields. A normal mode analysis (not shown) and inspection of the basic dynamical

fields reveals that there is a 16-hour period (local maxima to local maxima), which is within the period range of low-zonal-wavenumber gravity waves, suggesting that this feature is due to the dominant gravity waves in this system.

To test if the issue with the v analysis errors was due to the choice of training data, we repeated EBV(a) but with $\mathbf{B}_c$ calibrated using other training data (e.g. the ensemble perturbations from the 50th cycle of the spin-up run instead of those from the initial forecast ensemble). Even with more time-appropriate training data (but same variances), the issue was only partially resolved (smaller increase in RMSE; not shown). Also, this issue does not appear to occur in mid-latitudinal experimental setups in Bannister (2021). From the implied background error covariances in Section 4.1, there exists some mass-wind relationship in $\mathbf{B}_e$ that is not well-represented in $\mathbf{B}_c$ through the geostrophic balance relationship since f is small in the tropical setting. Repeating EBV(a), but omitting the geostrophic balance constraint entirely in the calibration of $\mathbf{B}_c$ (i.e. treating $\tilde{\rho}'$ and v background errors univariately) also did not resolve the issue (not shown). We speculate that the issue could be due to the absence of a suitable balance constraint for prescribing the mass-wind relationship for $\mathbf{B}_c$, or a likely lack of tuning of the variances for $\mathbf{B}_c$. Early results with a tuned $\mathbf{B}_c$ showed that the issue with the v analysis errors could be resolved by reducing the variances of all variables substantially. The warrants further investigation in a separate study, to tune the variances for the system or assess the possibility of deriving a balance relationship between v and $\tilde{\rho}'$ for the tropical setting.

Comparing between the EBV experiments, the u analysis errors and v analysis errors are generally the smallest in EBV(d), indicating that allocating full weight to $\mathbf{B}_e$ in this setup is ideal for minimising the horizontal wind-related analysis errors. The w, $\tilde{\rho}'$ and b' analysis errors are arguably the smallest in EBV(c), with the smallest cycle-averaged RMSE. The results presented here are not unsurprising given that previous studies evaluating hybrid-EnVar DA in simplified models (e.g. Hamill and Snyder, 2000) and NWP systems (e.g. Montmerle et al., 2018; Bédard et al., 2020) also show that the best configuration appears to rely on a combination of both $\mathbf{B}_e$ and $\mathbf{B}_c$, and not solely one or the other.

4.4 Ensemble trajectories and spread-error relationship

We can better appreciate the robustness of the ensemble by plotting the trajectories of the ensemble, its mean, the FreeBG, and the "truth" (Fig. 6). To avoid over-smoothing the local spatial variations in the fields, the trajectories are computed by taking a gridpoint-averaged value of the fields for a subset of the full domain; a box located at the centre of the domain (model levels 25 to 35, longitudinal gridpoints 127 to 237). We have also investigated the trajectories using other subsets (boxes) distributed around the domain, but the main ideas are the same so we have excluded discussion on them.

In Fig. 6, the spread of the EBV(d) ensemble is centred around the ensemble mean throughout the 50 cycles. The "truth" trajectory is also generally contained within the spread of the ensemble, particularly for u, $\tilde{\rho}'$ and b'. There were no evidence of filter collapse, nor bimodalities, which indicate that the DA-centred ensemble generated using the EBV method is healthy.

It is common practice to also compare the ensemble spread with the RMSE, which for a perfectly reliable large ensemble where observation density and errors are accounted for, the two quantities should be approximately the same (Leutbecher, 2009). In the EBV method, the ensemble spread is largely dependent on the choice of ϵ_0 since it does not account for the observation network, but this method is still worth considering as a "control method" and comparing with future methods consistent with Kalman filter theory. In the computation of the ensemble spread, Fortin et al. (2014) also cautions against using the wrong metric. Following their recommended approach, we define the gridpoint-averaged ensemble spread $\bar{S}_t$ using the square-root gridpoint-averaged ensemble variance:

$$\bar{S}_t = \sqrt{\frac{1}{N_g}\sum_{i=1}^{N_g} S^2[i,t]} \tag{18}$$

where the ensemble spread S is computed using Eq. (4) of Whitaker and Loughe (1998). N_g in this case refers to the number of gridpoints over which the average is taken (i.e. the points within the same box used for Fig. 6), and i is the gridpoint index which represents points in the box. The RMSE is computed as before, except now over points within this box.

Figure 7 shows $\bar{S}_t$ for each model quantity in the EBV(d) ensemble. These are benchmarked against the RMSE and the (time-stationary) implied background error standard deviations at model level 30 of $\mathbf{B}_c$, which are also plotted. For u and $\tilde{\rho}'$, the ensemble spread approximately matches the RMSE, particularly for later cycles as the hybrid control analysis converges around the "truth". For v, w and b', the ensemble is clearly under-dispersive. For all variables, the ensemble spread is also much smaller than the corresponding implied background error standard deviation at model level 30 of $\mathbf{B}_c$. This strongly suggests that the issue with v analysis errors highlighted in the previous section is due to lack of tuning of the variances of $\mathbf{B}_c$, which depends on the specific data assimilation setup. Note that the ensemble spread is computed with respect to the ensemble mean (Whitaker and Loughe, 1998), but the RMSE is computed between the hybrid control analysis (a surrogate to the ensemble mean) and the "truth". The spread-error relationship from this setup suggests that the DA-centering did not result in major statistical inconsistencies with the EBV ensemble. While the spread-error relationship is a useful diagnostic, it is not so straightforward to directly relate the ensemble spread to the eventual skill of the hybrid-EnVar DA system. Hence, it is not easy to determine whether to further inflate or deflate the EBV analysis perturbations by considering other choices of ϵ_0.

4.5 Sensitivity to number of ensemble members

As mentioned in Section 3.2.1, having a finite number of ensemble members will lead to sampling error in $\mathbf{P}_e^f[t]$. Logically, decreasing the number of ensemble members N used to compute $\mathbf{P}_e^f[t]$ should result in larger sampling errors. For a fixed $\mathbf{L}$ (as in Section 4.2), we demonstrate the sensitivity of the the skill of the hybrid-EnVar DA system to N in the ABC-DA system. We perform two additional experiments as variants of EBV(d) to maximise the impact of the ensemble size changes. The experiments follow the same configuration as EBV(d), but with only 20 and 10 ensemble members in the ensemble instead, referred to as EBV(d20) and EBV(d10) respectively, instead of the 30 members used until now.

Figure 8 shows the comparison of cycle-averaged analysis errors as N is varied. The RMSE is smallest in EBV(d) for almost all prognostic variables. Reducing the ensemble size from 30 to 20 in EBV(d20) leads to an increase in the RMSE, indicating poorer performance of the ABC-DA system. A further reduction of the ensemble size to 10 in EBV(d10) leads to the poorest performance overall. In this simple setup, these results are expected following the above argument that larger sampling errors are introduced into the system when N is smaller. For $\tilde{\rho}'$ and b', the RMSE ratio in EBV(d10) is even larger than in EBV(a), indicating that the pure EnVar setup may perform poorer than its 3DVar-FGAT counterpart when the ensemble size is too small (Fig. 8; bottom right panel).

It is important to highlight that the results are specific to this ABC-DA setup where the localisation length-scales are kept fixed (and are arguably quite tight) across the experiments. For other setups where the localisation length-scales are broader, the optimal ensemble size would be expected to be larger. It would also be worth exploring if a further increase in the ensemble size by orders of magnitude would yield a "saturation point" where there is little additional benefit to the system. However, one should also be aware of ensemble clustering (Amezcua et al., 2012) in very large ensembles. This issue has been shown to negatively impact hybrid-EnVar DA in simpler models such as the three variable Lorenz-63 model (Goodliff et al., 2015). In the case of the much larger ABC-DA system though, it is unlikely that N could practically be made large enough relative to n to be exposed to this handicap.

5 Summary

In this article, we document the development of the hybrid ensemble-variational data assimilation system for the ABC model (Petrie et al., 2017), built on the existing variational ABC-DA system (Bannister, 2020). The hybrid ensemble-variational algorithm that is introduced is based on the alpha control variable approach of Lorenc (2003). Key details related to the spatial and inter-variable localisation are discussed; the approach coded in the ABC-DA system allows flexibility in the localisation, for use in future exploratory studies. The hybrid ensemble-variational algorithm requires an ensemble system that is run parallel to the deterministic components to provide the flow-dependent error modes. To achieve this, the random field perturbations method is introduced in the ABC model for generating an initial ensemble. The ensemble bred vectors (EBV) method is also introduced in the ABC-DA system to propagate the ensemble, which is centred on the hybrid control analysis at each cycle.

Using a tropical setting of the ABC model, we test both ensemble propagation methods (30-member ensemble) in a series of hourly-cycling multi-cycle data assimilation observation system simulation experiments with hybrid ensemble-variational data assimilation. In the experiments, 3DVar-FGAT (First Guess at Appropriate Time) is employed together with EBV using different weightings assigned to the implied climatological (or static) background error covariance matrix ($\mathbf{B}_c$) and the implied ensemble-derived background error covariance matrix ($\mathbf{B}_e$); (a) 100% $\mathbf{B}_c$ (i.e. no flow-dependency, equivalent to 3DVar-FGAT), (b) 50% $\mathbf{B}_c$, 50% $\mathbf{B}_e$; hybrid-En3DVar-FGAT (i.e. flow-dependency with an equal contribution from $\mathbf{B}_c$ and $\mathbf{B}_e$), (c) 20% $\mathbf{B}_c$, 80% $\mathbf{B}_e$; hybrid-En3DVar-FGAT (i.e. flow-dependency with most contribution from $\mathbf{B}_e$), and (d) 100% $\mathbf{B}_e$;

pure En3DVar-FGAT (i.e. no contribution from $\mathbf{B}_c$).

The cycle-averaged analysis root-mean-square errors with respect to the "truth" for all prognostic variables except v were generally smaller for the EBV experiments compared to the free background. All experiments that involved the ensemble outperformed pure $\mathbf{B}_c$ for all variables. EBV(c) was the best performing configuration for w, $\tilde{\rho}'$ and b', while EBV(d) was the best performing configuration for u and v. We also noted that the v field gradually diverged from the "truth" during the simulations for experiments involving $\mathbf{B}_c$, even though fields of other variables were converging around the "truth" as logically expected. Through further assessment of the implied background error covariances and sensitivity tests, it was found that for the tropical setting of the ABC model, there exists some mass-wind relationship that is captured in $\mathbf{B}_e$ which is not well-represented by the (weak) geostrophic balance constraint in $\mathbf{B}_c$. We speculate that the issue with v for configurations that involve $\mathbf{B}_c$ could be due to the absence of a suitable balance constraint for prescribing the mass-wind relationship which may exist in the tropical setting of the ABC model, warranting further investigation in a separate study since it is not trivial to derive one. The results demonstrate the advantages of employing hybrid ensemble-variational data assimilation in the ABC-DA system over traditional variational data assimilation.

An inspection of the EBV(d) ensemble trajectories showed that the ensemble was centred around the ensemble mean throughout the experiment, with the "truth" trajectory generally contained within the spread of the ensemble. For v, w and b', the EBV ensemble was under-dispersive, but for u and $\tilde{\rho}'$, the ensemble spread approximately matched the corresponding RMSE. The EBV ensemble did not exhibit bimodalities or evidence of filter collapse, indicating that the DA-centred ensemble generated was healthy.

To illustrate the sensitivity to ensemble members, we performed two additional experiments as variants of EBV(d); EBV(d20) with 20 ensemble members and EBV(d10) with 10 ensemble members. The cycle-averaged analysis errors for almost all prognostic variables were smallest in EBV(d). Reducing the ensemble size from 30 to 20, and subsequently to 10 led to an increase in the RMSE, indicating poorer performance of the ABC-DA system. The results in this simple setup are consistent with the expectation that larger sampling errors are introduced into the system with a smaller ensemble, thus resulting in larger RMSE.

During the testing and development of the hybrid ensemble-variational method, localisation-related issues like hydrostatic imbalance in the analysis increments also became apparent. Similar issues have been documented in previous studies, but we have included additional comments in this article. Given the rapid adoption and broad shift towards hybrid ensemble-variational methods in convective-scale numerical weather prediction, we hope that the ABC-DA system can prove useful in providing further insights and highlight other potential issues that may arise in such methods. Particularly for the tropics, further work is required to better understand the characteristics of the ensemble-derived background errors, such as disentangling its flow-dependency or designing the localisation to isolate or identify important multi-variate relationships.

Code availability. The model and data assimilation system are written in Fortran 90/95, and the plotting code is written in Python. The upgraded system branches from ABC-DA v1.5. The source code, experiment and plotting scripts are open source and freely available on a Github repository at https://doi.org/10.5281/zenodo.6646951.

Appendix A: Details on the random field perturbations method

From Section 3.1, the random field perturbations method is used to generate the initial ensemble states for the ABC ensemble system. Equation (7a) describes the implementation where pairs of states are randomly chosen from a long "truth" run.

In Magnusson et al. (2009), there are additional constraints placed on the choice of random fields. The dates must be from different years and must be from the same season in order to eliminate inter-annual correlations in the perturbations yet preserve the seasonal characteristics of the variability. In the ABC model, we have attempted to capture the essence of these constraints even though there are no seasons in the ABC model.

For the experiments, the long "truth" run is generated with the initial control state $\mathbf{x}_{cs}^{c}$ as the initial condition, and is run for 50 days. ABC model dumps are produced every hour, resulting in a total of 1200 state dumps. A minimum threshold of 100 hours is set between the validity time of each random pair of states, for which they are assumed to be uncorrelated. In other words, pairs of states are selected randomly and are retained only if they are valid at least 100 hours apart. Additionally, Magnusson et al. (2009) did not indicate if the dates can be repeatedly selected (i.e. selection from a pool with replacement), so we have not imposed the additional constraint of selection from a pool without replacement. For the experiments, a total of 1200 state dumps is sufficiently large compared to the number of pairs required (number of ensemble members, 30 in most of this work).

One aspect that was highlighted in the implementation was the choice of fixed perturbation amplitude to scale the random field perturbations. It is not possible to follow the exact approach of Magnusson et al. (2009), using the average analysis error statistics, in the ABC model. However, we use the same metric (total energy norm) to gauge the initial fixed perturbation amplitude. As described in Section 3.1, the random field perturbations are scaled towards their mean total energy norm. This approach ensures that the random fields perturbations have the same fixed perturbation amplitude, but differ in directions of error growth. Further testing with the ABC model showed that reducing the fixed perturbation amplitude yielded smaller errors in the experiments, so a deflation factor of 5 was eventually adopted.

The total energy norm (E_{tot}) for the random field perturbations are computed using:

$$E_{tot} = E_k + E_b + E_e \tag{A1a}$$

$$E_k = \int \frac{\tilde{\rho}(u^2 + v^2 + w^2)}{2} \rho_0 \, \mathrm{d}V \tag{A1b}$$

$$E_b = \int \frac{\tilde{\rho} b'^2}{2A^2} \rho_0 \, \mathrm{d}V \tag{A1c}$$

$$E_e = \int \frac{C\tilde{\rho}'^2}{2B} \rho_0 \, \mathrm{d}V \tag{A1d}$$

where E_k, E_b, E_e are the kinetic, buoyant and elastic energy respectively, $\rho_0 = 1.225$ kg m^{-3} is a reference air density, and $\mathrm{d}V$ is the volume of a gridbox in the ABC model. Note that as mentioned in Section 3.3.1, we also use E_{tot} in the ensemble bred vectors method to scale the ensemble perturbations for subsequent cycles. Prior to the experiments, we performed some initial testing using the inner product norm instead of E_{tot}, which yielded similar results between the two norm choices. Since E_{tot} is a metric that is physically meaningful, it was eventually used for the scaling of the random field perturbations and ensemble perturbations from the ensemble bred vectors method for the experiments.

Appendix B: Accounting for inter-variable covariances – proof of equivalence of two approaches

As highlighted in Section 3.2.2, $\mathbf{L}$ (the localisation matrix) can be partitioned into a matrix $\mathbf{U}^\alpha$:

$$\mathbf{L} = \mathbf{U}^\alpha \mathbf{U}^{\alpha\top} \tag{B1}$$

We seek to prove that two approaches used to code $\boldsymbol{\chi}^{\alpha k}$ and $\mathbf{U}^\alpha$ (described below) give the same result when $\mathbf{L}$ is applied on the same model space vector $\mathbf{v}$ (of length $N_g N_{var}$). Note that $\mathbf{L}$ is $N_g N_{var} \times N_g N_{var}$, which in principle means that the inter-variable localisation matrix can be set to have any correlation structure, including the limiting cases of full localisation between different variables (where the corresponding matrix elements are 0), and no localisation (matrix elements are 1). Recall that $\mathbf{L}$ is used in the DA via a Schur product with $\mathbf{P}_e^f[t]$ (Eq. (11)). While the number of rows in $\mathbf{U}^\alpha$ is constrained to be $N_g N_{var}$, the number of columns can be chosen. The fewer the columns, the smaller the corresponding size of the $\boldsymbol{\chi}^{\alpha k}$ vectors (Eq. (14b)), but the less flexible the implied localisation matrix. The first approach considered is based on Wang et al. (2008a) (N_g columns) and the second approach is coded in the ABC-DA system ($N_g N_{var}$ columns), inspired by Bannister (2017). The first approach requires less memory and computation, but has less flexibility than the second approach in terms of multi-variate localisation choices.

For simplicity, the proof is demonstrated using pure EnVar with one non-zero element in $\mathbf{v}$. This is a similar procedure to computing a column of the implied $\mathbf{B}_c$ or $\mathbf{B}_e$, but now the implied $\mathbf{L}$ is being probed. It is easier to visualise the interactions of the matrix elements by partitioning $\mathbf{v}$ into segments of size $N_g \times 1$ based on the ABC prognostic variables, i.e. $\mathbf{v} = (\mathbf{v}_u, \mathbf{v}_v, \mathbf{v}_w, \mathbf{v}_{\tilde{\rho}'}, \mathbf{v}_{b'})^\top$. Similarly, we can consider blocks, each of size $N_g \times N_g$, used to construct $\mathbf{U}^\alpha$ (i.e. $\mathbf{U}_u^\alpha, \mathbf{U}_v^\alpha, \mathbf{U}_w^\alpha, \mathbf{U}_{\tilde{\rho}'}^\alpha$,

and $\mathbf{U}^{\alpha}_{b'}$), which will determine the spatial localisations (horizontal and vertical) for each variable.

The main difference between the two approaches is in the design of $\mathbf{U}^{\alpha}$. In the first approach, based on Wang et al. (2008a), $\mathbf{U}^{\alpha}$ (denoted $\tilde{\mathbf{U}}^{\alpha}$) is rectangular ($N_g N_{var} \times N_g$, and $\boldsymbol{\chi}^{\alpha k}$ has N_g elements), given by:

$$\tilde{\mathbf{U}}^{\alpha} = \begin{bmatrix} \mathbf{U}^{\alpha}_{u} \\ \mathbf{U}^{\alpha}_{v} \\ \mathbf{U}^{\alpha}_{w} \\ \mathbf{U}^{\alpha}_{\tilde{\rho}'} \\ \mathbf{U}^{\alpha}_{b'} \end{bmatrix}. \tag{B2}$$

Applying $\mathbf{L}$ (denoted $\tilde{\mathbf{L}}$ for first approach) to $\mathbf{v}$ yields:

$$\tilde{\mathbf{L}}\mathbf{v} = \tilde{\mathbf{U}}^{\alpha}\tilde{\mathbf{U}}^{\alpha\top}\mathbf{v} = \begin{bmatrix} \mathbf{U}^{\alpha}_{u} \\ \mathbf{U}^{\alpha}_{v} \\ \mathbf{U}^{\alpha}_{w} \\ \mathbf{U}^{\alpha}_{\tilde{\rho}'} \\ \mathbf{U}^{\alpha}_{b'} \end{bmatrix} \begin{bmatrix} \mathbf{U}^{\alpha\top}_{u} & \mathbf{U}^{\alpha\top}_{v} & \mathbf{U}^{\alpha\top}_{w} & \mathbf{U}^{\alpha\top}_{\tilde{\rho}'} & \mathbf{U}^{\alpha\top}_{b'} \end{bmatrix} \mathbf{v} \tag{B3a}$$

with the elements given by:

$$(\tilde{\mathbf{L}}\mathbf{v})_i = \sum_{j=1}^{N_g} \left\{ (\tilde{\mathbf{U}}^{\alpha})_{i,j} \sum_{i'=1}^{N_g N_{var}} (\tilde{\mathbf{U}}^{\alpha\top})_{j,i'} \mathbf{v}_{i'} \right\} \tag{B3b}$$

If there is only one non-zero element (the qth element of $\mathbf{v}$), this simplifies to:

$$(\tilde{\mathbf{L}}\mathbf{v})_i = \sum_{j=1}^{N_g} (\tilde{\mathbf{U}}^{\alpha})_{i,j} (\tilde{\mathbf{U}}^{\alpha\top})_{j,q} \mathbf{v}_q = \sum_{j=1}^{N_g} (\tilde{\mathbf{U}}^{\alpha})_{i,j} (\tilde{\mathbf{U}}^{\alpha})_{q,j} \mathbf{v}_q \tag{B3c}$$

In the second approach, $\mathbf{U}^{\alpha}$ (denoted $\hat{\mathbf{U}}^{\alpha}$) is square ($N_g N_{var} \times N_g N_{var}$ and $\boldsymbol{\chi}^{\alpha k}$ has $N_g N_{var}$ elements), given by:

$$\hat{\mathbf{U}}^{\alpha} = \begin{bmatrix} \mathbf{U}^{\alpha}_{u} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}^{\alpha}_{v} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}^{\alpha}_{w} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}^{\alpha}_{\tilde{\rho}'} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}^{\alpha}_{b'} \end{bmatrix} \tag{B4}$$

where $\mathbf{0}$ is a $N_g \times N_g$ block containing zeroes. This is the default configuration that is coded in the ABC-DA system, which gives an implied $\mathbf{L}$ (denoted $\hat{\mathbf{L}}$ for the second approach):

$$\hat{\mathbf{L}} = \hat{\mathbf{U}}^{\alpha}\hat{\mathbf{U}}^{\alpha\top} = \begin{bmatrix} \mathbf{U}_u^{\alpha}\mathbf{U}_u^{\alpha\top} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U}_v^{\alpha}\mathbf{U}_v^{\alpha\top} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U}_w^{\alpha}\mathbf{U}_w^{\alpha\top} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_{\tilde{\rho}'}^{\alpha}\mathbf{U}_{\tilde{\rho}'}^{\alpha\top} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U}_{b'}^{\alpha}\mathbf{U}_{b'}^{\alpha\top} \end{bmatrix} \tag{B5}$$

Notice that here, $\hat{\mathbf{L}}$ does a full inter-variable localisation, so that the Schur product of $\hat{\mathbf{L}}$ with $\mathbf{P}_e^f[t]$ will not retain any inter-variable covariances. This may be useful if N is small and sampling noise is problematic in $\mathbf{P}_e^f[t]$.

Next, we introduce a mapping matrix $\hat{\mathbf{I}}$, which consists of $N_p \times N_p$ blocks of identity matrices ($\mathbf{I}_{N_g}$, each of size $N_g \times N_g$):

$$\hat{\mathbf{I}} = \frac{1}{\sqrt{N_p}} \begin{bmatrix} \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} \\ \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} \\ \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} \\ \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} \\ \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} & \mathbf{I}_{N_g} \end{bmatrix}, \tag{B6}$$

where N_p is the number of model variables whose inter-variable covariances are retained by the mapping matrix (i.e. $N_p = N_{var} = 5$ in the above). Note that other designs of $\hat{\mathbf{I}}$ (e.g. replacing some blocks with $\mathbf{0}$) will allow only the desired retention of specific covariances between certain model variables.

Using the second approach of coding $\mathbf{U}^{\alpha}$ and $\boldsymbol{\chi}^{\alpha k}$, it is possible to retain the full inter-variable covariances and achieve the exact same outcome as the first approach by defining $\mathbf{U}^{\alpha} = \hat{\mathbf{U}}^{\alpha}\hat{\mathbf{I}}$. The implied localisation matrix is thus $\mathbf{L} = \frac{1}{N_p}\hat{\mathbf{U}}^{\alpha}\hat{\mathbf{I}}\hat{\mathbf{I}}\hat{\mathbf{U}}^{\alpha\top}$. As before, applying $\mathbf{L}$ to $\mathbf{v}$ yields:

$$\mathbf{L}\mathbf{v} = \frac{1}{N_p}\hat{\mathbf{U}}^{\alpha}\hat{\mathbf{I}}\hat{\mathbf{I}}\hat{\mathbf{U}}^{\alpha\top}\mathbf{v} = \frac{1}{N_p} \begin{bmatrix} \mathbf{U}_u^{\alpha} & \mathbf{U}_u^{\alpha} & \mathbf{U}_u^{\alpha} & \mathbf{U}_u^{\alpha} & \mathbf{U}_u^{\alpha} \\ \mathbf{U}_v^{\alpha} & \mathbf{U}_v^{\alpha} & \mathbf{U}_v^{\alpha} & \mathbf{U}_v^{\alpha} & \mathbf{U}_v^{\alpha} \\ \mathbf{U}_w^{\alpha} & \mathbf{U}_w^{\alpha} & \mathbf{U}_w^{\alpha} & \mathbf{U}_w^{\alpha} & \mathbf{U}_w^{\alpha} \\ \mathbf{U}_{\tilde{\rho}'}^{\alpha} & \mathbf{U}_{\tilde{\rho}'}^{\alpha} & \mathbf{U}_{\tilde{\rho}'}^{\alpha} & \mathbf{U}_{\tilde{\rho}'}^{\alpha} & \mathbf{U}_{\tilde{\rho}'}^{\alpha} \\ \mathbf{U}_{b'}^{\alpha} & \mathbf{U}_{b'}^{\alpha} & \mathbf{U}_{b'}^{\alpha} & \mathbf{U}_{b'}^{\alpha} & \mathbf{U}_{b'}^{\alpha} \end{bmatrix} \begin{bmatrix} \mathbf{U}_u^{\alpha\top} & \mathbf{U}_v^{\alpha\top} & \mathbf{U}_w^{\alpha\top} & \mathbf{U}_{\tilde{\rho}'}^{\alpha\top} & \mathbf{U}_{b'}^{\alpha\top} \\ \mathbf{U}_u^{\alpha\top} & \mathbf{U}_v^{\alpha\top} & \mathbf{U}_w^{\alpha\top} & \mathbf{U}_{\tilde{\rho}'}^{\alpha\top} & \mathbf{U}_{b'}^{\alpha\top} \\ \mathbf{U}_u^{\alpha\top} & \mathbf{U}_v^{\alpha\top} & \mathbf{U}_w^{\alpha\top} & \mathbf{U}_{\tilde{\rho}'}^{\alpha\top} & \mathbf{U}_{b'}^{\alpha\top} \\ \mathbf{U}_u^{\alpha\top} & \mathbf{U}_v^{\alpha\top} & \mathbf{U}_w^{\alpha\top} & \mathbf{U}_{\tilde{\rho}'}^{\alpha\top} & \mathbf{U}_{b'}^{\alpha\top} \\ \mathbf{U}_u^{\alpha\top} & \mathbf{U}_v^{\alpha\top} & \mathbf{U}_w^{\alpha\top} & \mathbf{U}_{\tilde{\rho}'}^{\alpha\top} & \mathbf{U}_{b'}^{\alpha\top} \end{bmatrix} \mathbf{v}, \tag{B7a}$$

with the elements given by:

$$(\mathbf{L}\mathbf{v})_i = \frac{1}{N_p} \sum_{j=1}^{N_g N_p} \left\{ (\mathbf{U}^{\alpha})_{i,j} \sum_{i'=1}^{N_g N_p} (\mathbf{U}^{\alpha\top})_{j,i'} \mathbf{v}_{i'} \right\}. \tag{B7b}$$

Note how in this case, the rows of $\mathbf{U}^{\alpha\top}$ are the same as in $\tilde{\mathbf{U}}^{\alpha\top}$ from the first approach (Eq. (B2)), but repeated N_p times. If there is only one non-zero element (the qth element of $\mathbf{v}$), then the computation simplifies to:

$$(\mathbf{L}\mathbf{v})_i = \frac{1}{N_p}\sum_{j=1}^{N_g N_p}(\mathbf{U}^{\alpha})_{i,j}(\mathbf{U}^{\alpha\top})_{j,q}\mathbf{v}_q = \frac{1}{N_p}\sum_{j=1}^{N_g} N_p(\tilde{\mathbf{U}}^{\alpha})_{i,j}(\tilde{\mathbf{U}}^{\alpha})_{q,j}\mathbf{v}_q = \sum_{j=1}^{N_g}(\tilde{\mathbf{U}}^{\alpha})_{i,j}(\tilde{\mathbf{U}}^{\alpha})_{q,j}\mathbf{v}_q, \tag{B7c}$$

noting that when $N_p = N_{var}$, full inter-variable covariances are retained. In the computation of any inner products of $\boldsymbol{\chi}^{\alpha k}$ in the variational algorithm, such as for the minimisation, or in the computation of J_e, these also have to be scaled by $\frac{1}{N_p}$ accordingly.

The key thing to note here is that when using both approaches with $\tilde{\mathbf{U}}^{\alpha}$ and $\mathbf{U}^{\alpha}$ respectively, the implied localisation matrices are the same ($\tilde{\mathbf{L}} = \mathbf{L}$), as demonstrated by Eq. (B3c) and Eq. (B7c) being the same. Given the greater flexibility, the second approach was coded in the ABC-DA system.

Appendix C: Hydrostatic imbalance due to vertical localisation

According to Eq. (3) of Bannister (2020), the hydrostatic balance relation in the ABC model (also used in the control variable transform) is given by:

$$C\frac{\partial\tilde{\rho}'}{\partial z} = b' \tag{C1}$$

From Eq. (1), the prognostic w equation indicates that the change in w following an air parcel (i.e. a Lagrangian frame of reference) is given by the source/sink terms $C\frac{\partial\tilde{\rho}'}{\partial z}$ and b'. This neatly corresponds to hydrostatic balance. In other words, hydrostatic imbalance will lead to sources/sinks in w as the system evolves.

Applying vertical localisation directly to the ensemble-derived error modes via the Schur product results in alterations in the vertical gradient of the $\tilde{\rho}'$ field, depending on the kurtosis of the correlation curve applied. We can consider the following scenario: Assuming that the ensemble forecasts are hydrostatically balanced on the large scales, one could expect that assimilating a single $\tilde{\rho}'$ observation without vertical localisation would result in hydrostatically balanced $\tilde{\rho}'$ and b' increments. However, with vertical localisation, the sharpness of the correlation curve superimposes on the $\tilde{\rho}'$ fields in the ensemble-derived error modes and results in increments that decrease more rapidly with distance (sharper gradient) from the point of observation. Thus, a larger b' increment is required in order to maintain hydrostatic balance, but the actual b' increments are also reduced by the Schur product. In this scenario, the resulting b' increments would be sub-hydrostatic.

During the spin-up configuration testing with vertical localisation applied, it was noted that the root-mean-square value of the w field was gradually increasing throughout the earlier stages of the spin-up process. Since there exists a restoring A^2w source term in the prognostic b' equation, the root-mean-square value of the w field does not increase indefinitely because of corresponding induced changes in the b' field.

Author contributions. JL, JA and RB designed the experiments. JL developed the model code and conducted the simulations and analysis. JL prepared the manuscript which was vetted by RB and JA.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. To be completed.

References

Amezcua, J., Ide, K., Bishop, C. H., and Kalnay, E.: Ensemble clustering in deterministic ensemble Kalman filters, Tellus A: Dynamic Meteorology and Oceanography, 64, 18 039, 2012.

Asch, M., Bocquet, M., and Nodet, M.: Data Assimilation: Methods, Algorithms, and Applications, Fundamentals of Algorithms, SIAM, Society for Industrial and Applied Mathematics, https://books.google.co.uk/books?id=A3Q6vgAACAAJ, 2016.

Balci, N., Mazzucato, A. L., Restrepo, J. M., and Sell, G. R.: Ensemble dynamics and bred vectors, Monthly weather review, 140, 2308–2334, 2012.

Bannister, R.: A review of operational methods of variational and ensemble-variational data assimilation, Quarterly Journal of the Royal Meteorological Society, 143, 607–633, 2017.

Bannister, R. N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 134, 1951–1970, 2008a.

Bannister, R. N.: A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 134, 1971–1996, 2008b.

Bannister, R. N.: 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, Geoscientific Model Development, 13, 3789–3816, 2020.

Bannister, R. N.: Balance conditions in variational data assimilation for a high-resolution forecast model, Quarterly Journal of the Royal Meteorological Society, 2021.

Bédard, J., Caron, J.-F., Buehner, M., Baek, S.-J., and Fillion, L.: Hybrid Background Error Covariances for a Limited-Area Deterministic Weather Prediction System, Weather and Forecasting, 35, 1051–1066, 2020.

Berre, L., Ştefaănescu, S. E., and Pereira, M. B.: The representation of the analysis effect in three error simulation techniques, Tellus A: Dynamic Meteorology and Oceanography, 58, 196–209, 2006.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, Monthly weather review, 129, 420–436, 2001.

Buizza, R., Tribbia, J., Molteni, F., and Palmer, T.: Computation of optimal unstable structures for a numerical weather prediction model, Tellus A, 45, 388–407, 1993.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99, 10 143–10 162, 1994.

Evensen, G.: Data Assimilation: The Ensemble Kalman Filter, Springer Berlin Heidelberg, https://books.google.co.uk/books?id=VJ2oOecHhOYC, 2006.

Fortin, V., Abaza, M., Anctil, F., and Turcotte, R.: Why should ensemble spread match the RMSE of the ensemble mean?, Journal of Hydrometeorology, 15, 1708–1713, 2014.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological Society, 125, 723–757, 1999.

Goodliff, M., Amezcua, J., and Van Leeuwen, P. J.: Comparing hybrid data assimilation methods on the Lorenz 1963 model with increasing non-linearity, Tellus A: Dynamic Meteorology and Oceanography, 67, 26 928, 2015.

Hamill, T. M. and Snyder, C.: A hybrid ensemble Kalman filter–3D variational analysis scheme, Monthly Weather Review, 128, 2905–2919, 2000.

Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, Monthly Weather Review, 129, 2776–2790, 2001.

Heng, B. P., Tubbs, R., Huang, X.-Y., Macpherson, B., Barker, D. M., Boyd, D. F., Kelly, G., North, R., Stewart, L., Webster, S., et al.: SINGV-DA: A data assimilation system for convective-scale numerical weather prediction over Singapore, Quarterly Journal of the Royal Meteorological Society, 146, 1923–1938, 2020.

Houtekamer, P. and Derome, J.: Methods for ensemble prediction, Monthly Weather Review, 123, 2181–2196, 1995.

Houtekamer, P. L. and Mitchell, H. L.: A sequential ensemble Kalman filter for atmospheric data assimilation, Monthly Weather Review, 129, 123–137, 2001.

Ingleby, N. B.: The statistical structure of forecast errors and its representation in the Met. Office global 3-D variational data assimilation scheme, Quarterly Journal of the Royal Meteorological Society, 127, 209–231, 2001.

Kalnay, E.: Atmospheric modeling, data assimilation and predictability, Cambridge university press, 2003.

Leutbecher, M.: Diagnosis of ensemble forecasting systems, in: Seminar on Diagnosis of Forecasting and Data Assimilation Systems, pp. 235–266, 2009.

Lorenc, A. C.: The potential of the ensemble Kalman filter for NWP — A comparison with 4D-Var, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 129, 3183–3203, 2003.

Magnusson, L., Nycander, J., and Källén, E.: Flow-dependent versus flow-independent initial perturbations for ensemble prediction, Tellus A: Dynamic Meteorology and Oceanography, 61, 194–209, 2009.

Montmerle, T., Michel, Y., Arbogast, E., Ménétrier, B., and Brousseau, P.: A 3D ensemble variational data assimilation scheme for the limited-area AROME model: Formulation and preliminary results, Quarterly Journal of the Royal Meteorological Society, 144, 2196–2215, 2018.

Parrish, D. F. and Derber, J. C.: The National Meteorological Center's spectral statistical-interpolation analysis system, Monthly Weather Review, 120, 1747–1763, 1992.

Penny, S. G.: The hybrid local ensemble transform Kalman filter, Monthly Weather Review, 142, 2139–2149, 2014.

Petrie, R. E., Bannister, R. N., and Cullen, M. J. P.: The ABC model: a non-hydrostatic toy model for use in convective-scale data assimilation investigations, Geoscientific Model Development, 10, 4419–4441, 2017.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NMC: The generation of perturbations, Bulletin of the american meteorological society, 74, 2317–2330, 1993.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NCEP and the breeding method, Monthly Weather Review, 125, 3297–3319, 1997.

Wang, X., Snyder, C., and Hamill, T. M.: On the theoretical equivalence of differently proposed ensemble–3DVAR hybrid analysis schemes, Monthly Weather Review, 135, 222–227, 2007.

Wang, X., Barker, D. M., Snyder, C., and Hamill, T. M.: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment, Monthly Weather Review, 136, 5116–5131, 2008a.

Wang, X., Barker, D. M., Snyder, C., and Hamill, T. M.: A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part II: Real observation experiments, Monthly Weather Review, 136, 5132–5147, 2008b.

Whitaker, J. S. and Loughe, A. F.: The relationship between ensemble spread and ensemble mean skill, Monthly weather review, 126, 3292–3302, 1998.

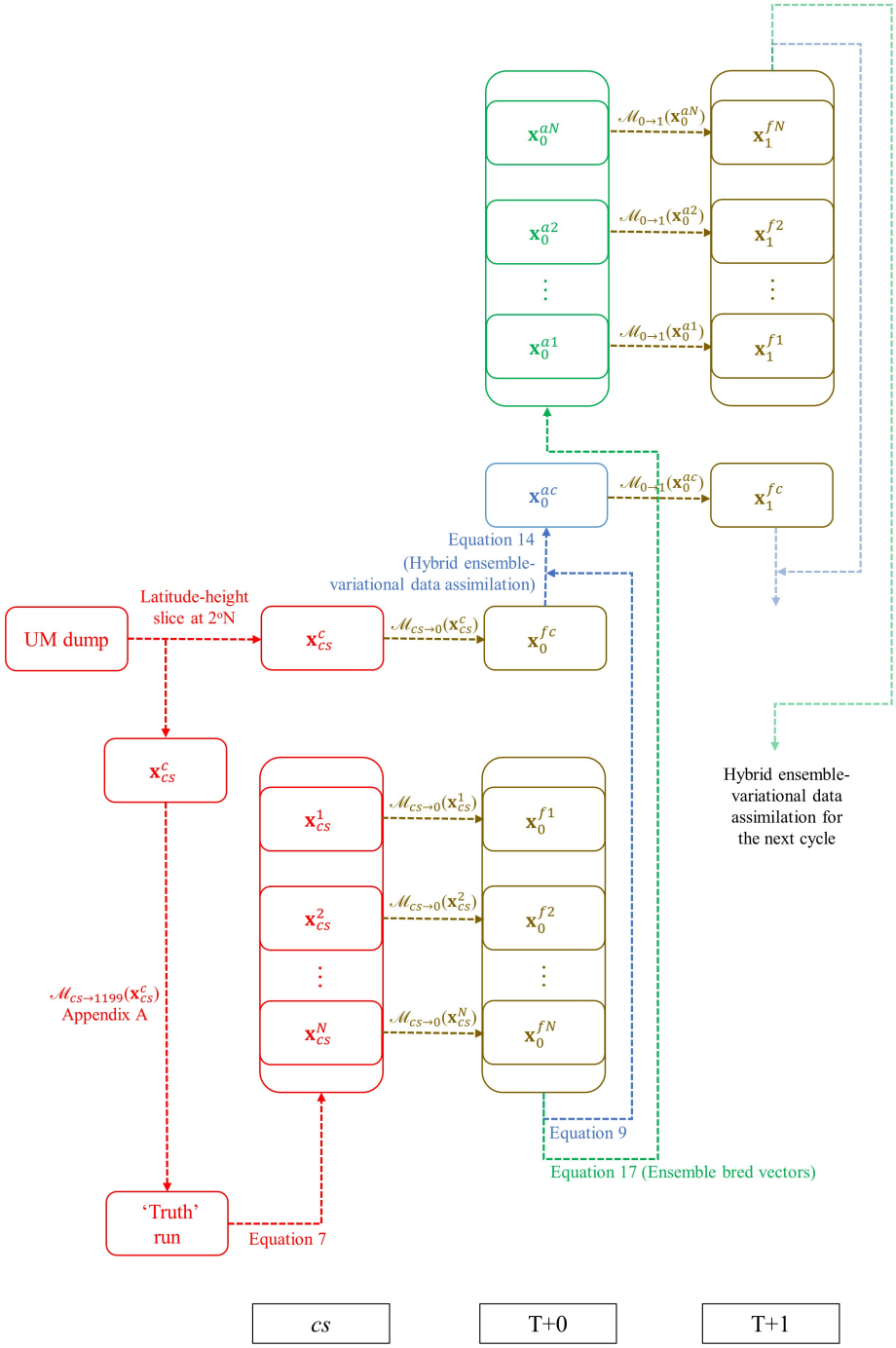

Figure 1. Schematic diagram of the ensemble and deterministic workflow for the hybrid-EnVar scheme in the ABC-DA system, illustrated for an hourly-cycling setup over the first cycle from a cold start. The subscripts refer to the validity time; cs refers to cold start. The superscripts fk and fc refer to the kth member of the forecast ensemble and the control forecast respectively, ak and ac refer to the kth member of the analysis ensemble and the hybrid control analysis respectively.

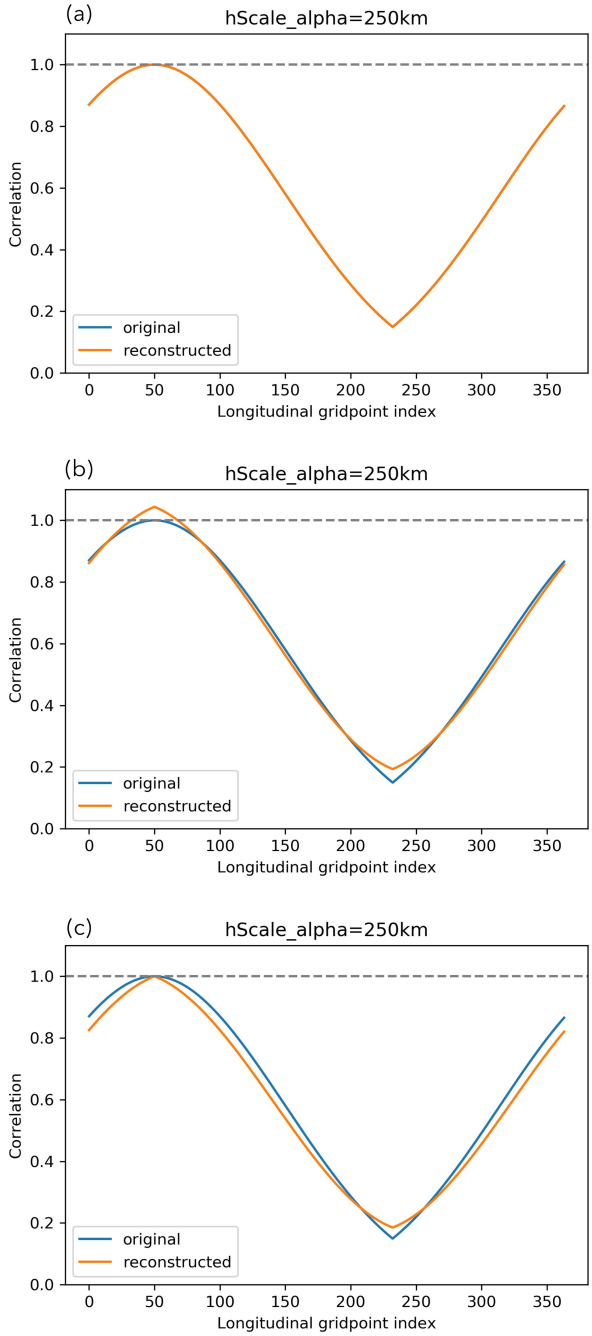

Figure 2. Correlation functions ($h^{\alpha} = 250$ km) with respect to longitudinal gridpoint 50, for an ABC-DA system with 364 longitudinal gridpoints and 1.5 km horizontal grid. The implied correlation functions (orange) are reconstructed from (a) all eigenvectors and eigenvalues of the eigen-decomposition of $\mathbf{L}_{horiz}$, (b) only eigenvectors with non-negative eigenvalues (c) only eigenvectors with non-negative eigenvalues that are scaled to restore initial total variance, and compared with the original Gaspari-Cohn function (blue).

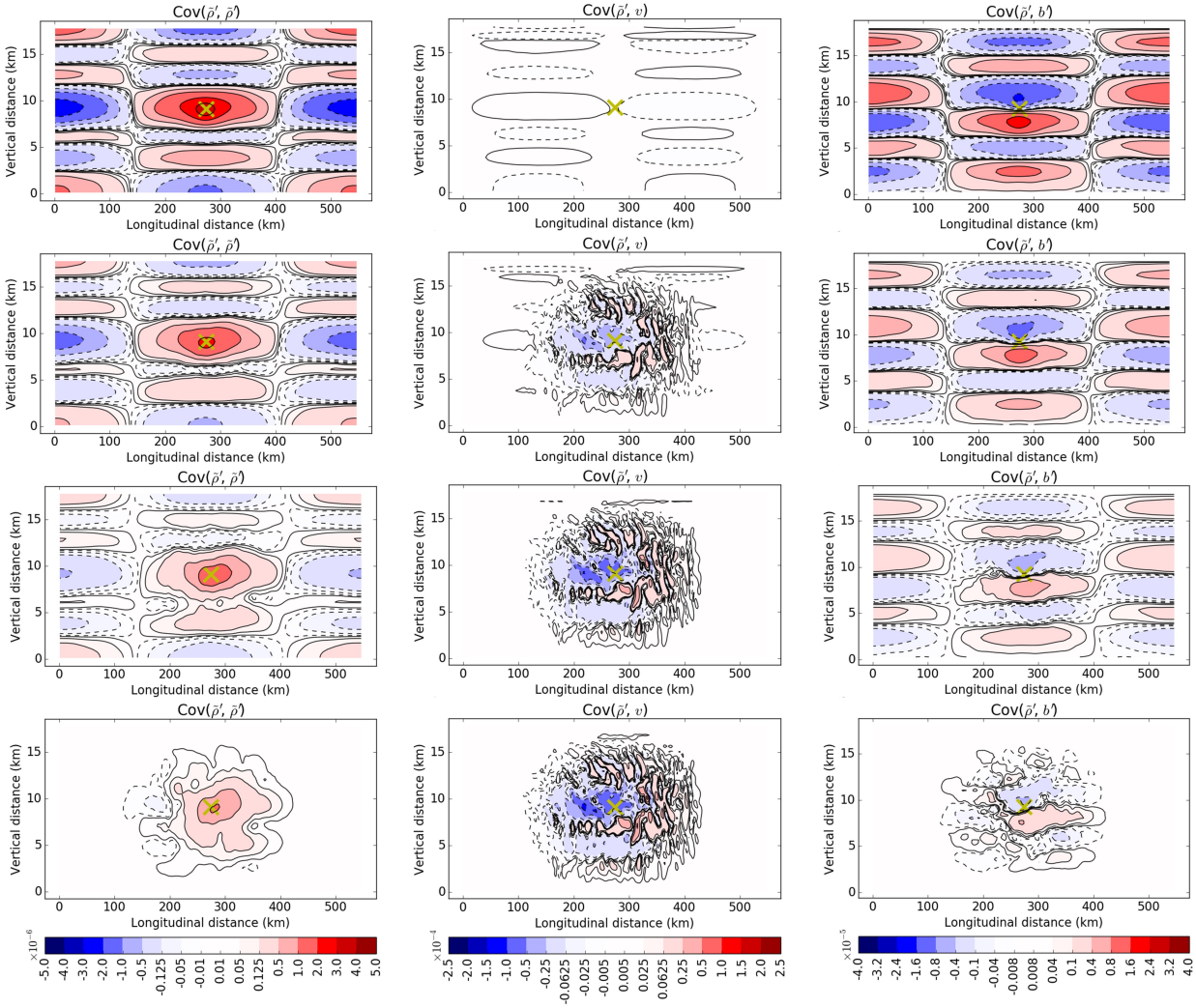

Figure 3. Implied background error covariances of $\tilde{\rho}'$ (leftmost column; Cov($\tilde{\rho}'$, $\tilde{\rho}'$)), v (middle column; Cov($\tilde{\rho}'$, v)) and b' (rightmost column; Cov($\tilde{\rho}'$, b')) with respect to a $\tilde{\rho}'$ point (yellow cross) near the centre of the domain for the first cycle after cold start. The rows represent configurations (a), (b), (c), and (d) respectively (see the list near the start of Section 4). Negative values have contours that are dashed.

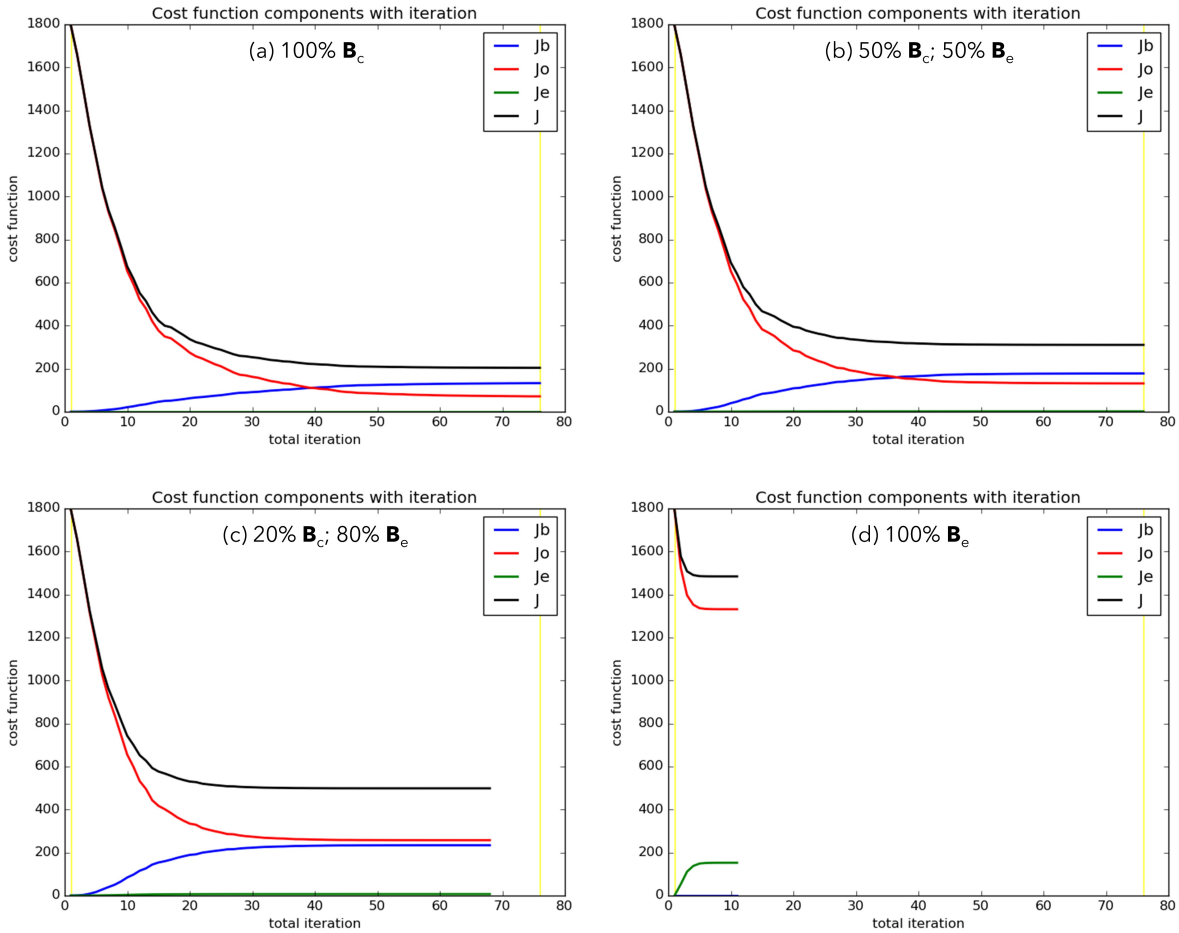

Figure 4. Total penalty (black) from the climatological background (blue), ensemble background (green) and observation (red) penalty contributions over the 75 inner loops for the first cycle of the EBV experiments, labelled (a) to (d) accordingly. Early termination of inner loops occurs when convergence criteria is satisfied, in (c) and (d). At convergence, ensemble penalty (green) in (b) and (c) is around 1.5 and 7 respectively.

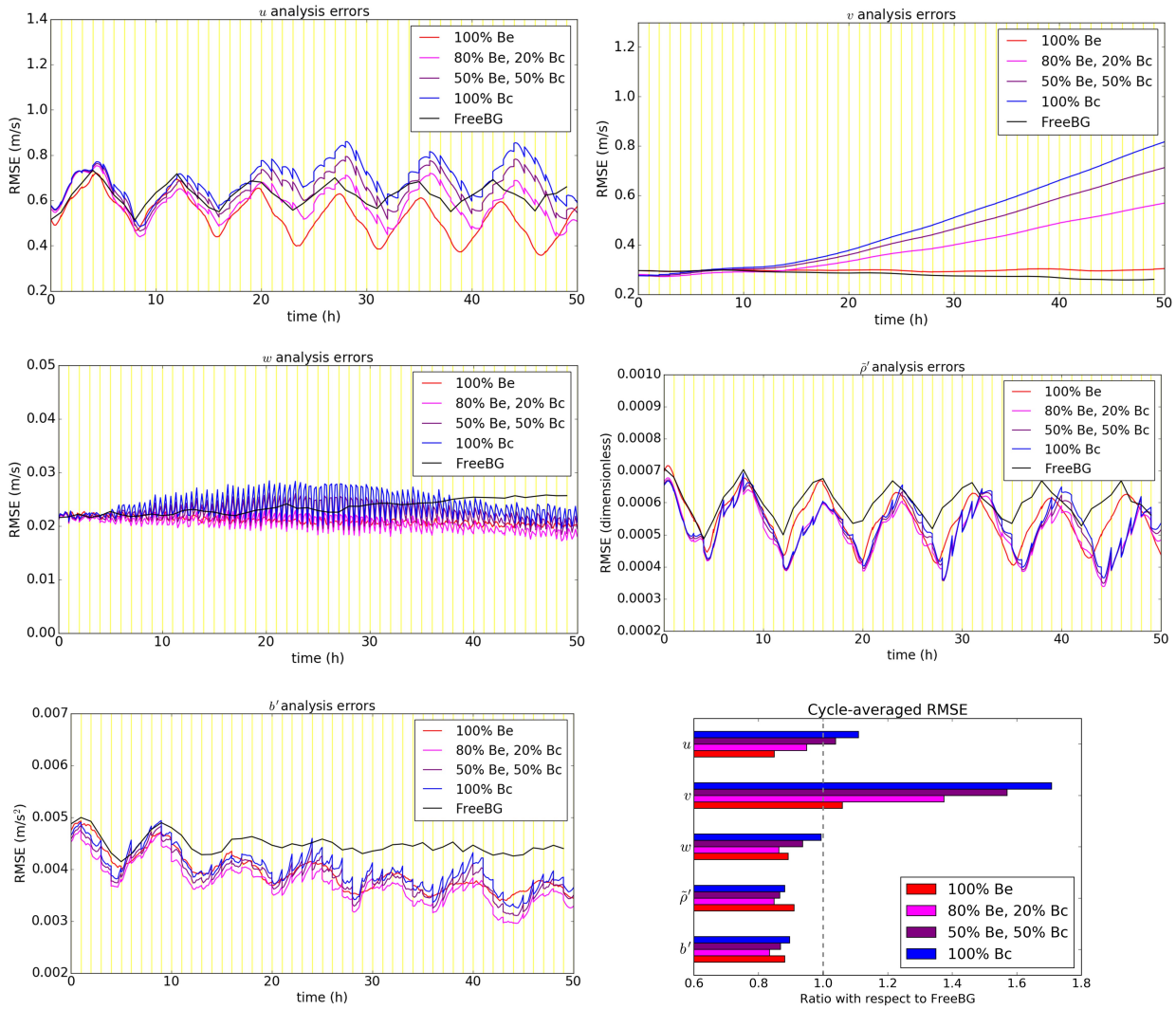

Figure 5. All panels except bottom right: time series of root-mean-square analysis errors for the EBV experiments (100% $\mathbf{B}_c$, configuration (a); 50% $\mathbf{B}_e$, 50% $\mathbf{B}_c$, (b); 80% $\mathbf{B}_e$, 20% $\mathbf{B}_c$, (c); 100% $\mathbf{B}_e$, (d)) and the free background run (FreeBG). The vertical yellow lines are the analysis times. Analysis errors are defined with respect to the "truth" run, computed every 10 minutes within the respective assimilation windows for EBV experiments and every hour for FreeBG. Bottom right: the ratio of the cycle-averaged RMSE of the EBV experiments with respect to FreeBG for the five ABC model variables.

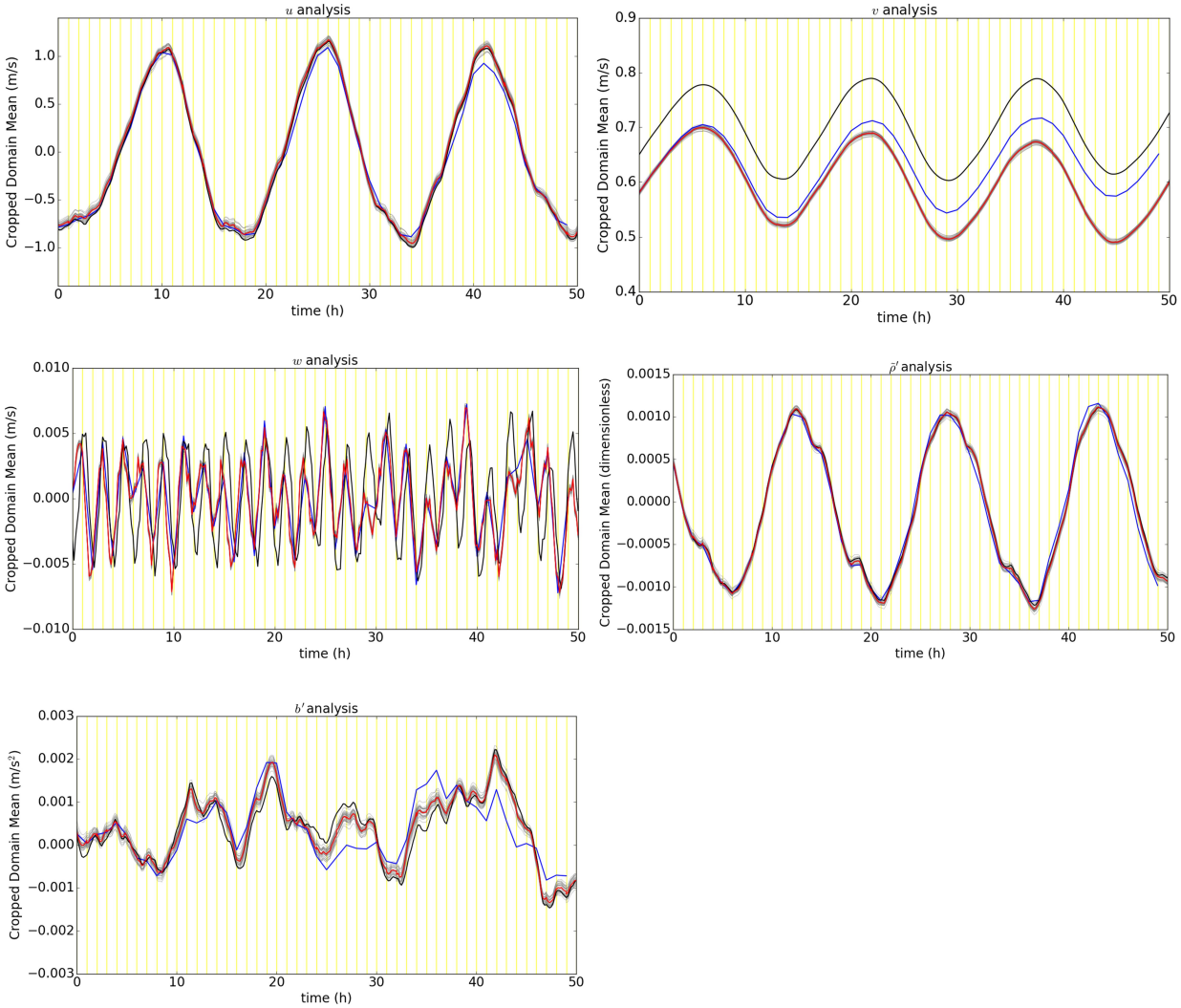

Figure 6. EBV(d) (100% $\mathbf{B}_e$) ensemble trajectories derived from gridpoint-averaged analysis fields and their forecasts over a subset of the full domain (a box located at the centre of the domain, model levels 25 to 35, longitudinal gridpoints 127 to 237). The corresponding ensemble mean (red), free background (blue) and "truth" (black) trajectories for the same subset domain are plotted alongside the individual ensemble member (grey) trajectories. Values for the free background are indicated every hour, and every 10 minutes for the other trajectories.

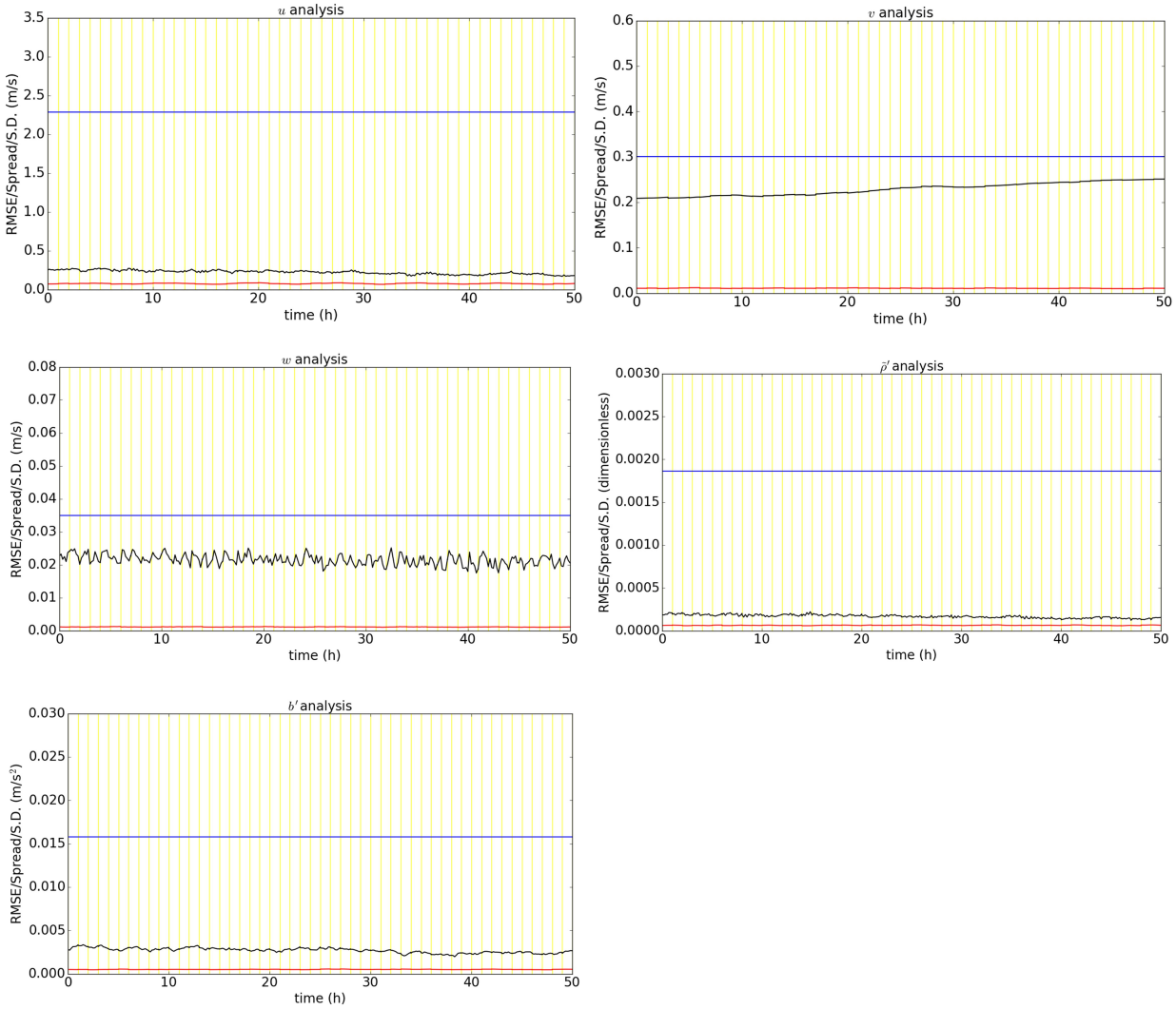

Figure 7. Time series of root-mean-square analysis errors (RMSE; black) and ensemble spread (Spread; red) for the EBV(d) (100% $\mathbf{B}_e$) ensemble, computed over a subset of the domain (a box located at the centre of the domain, model levels 25 to 35, longitudinal gridpoints 127 to 237). The implied (time-stationary) background error standard deviation at model level 30 is also included (S.D.; blue).

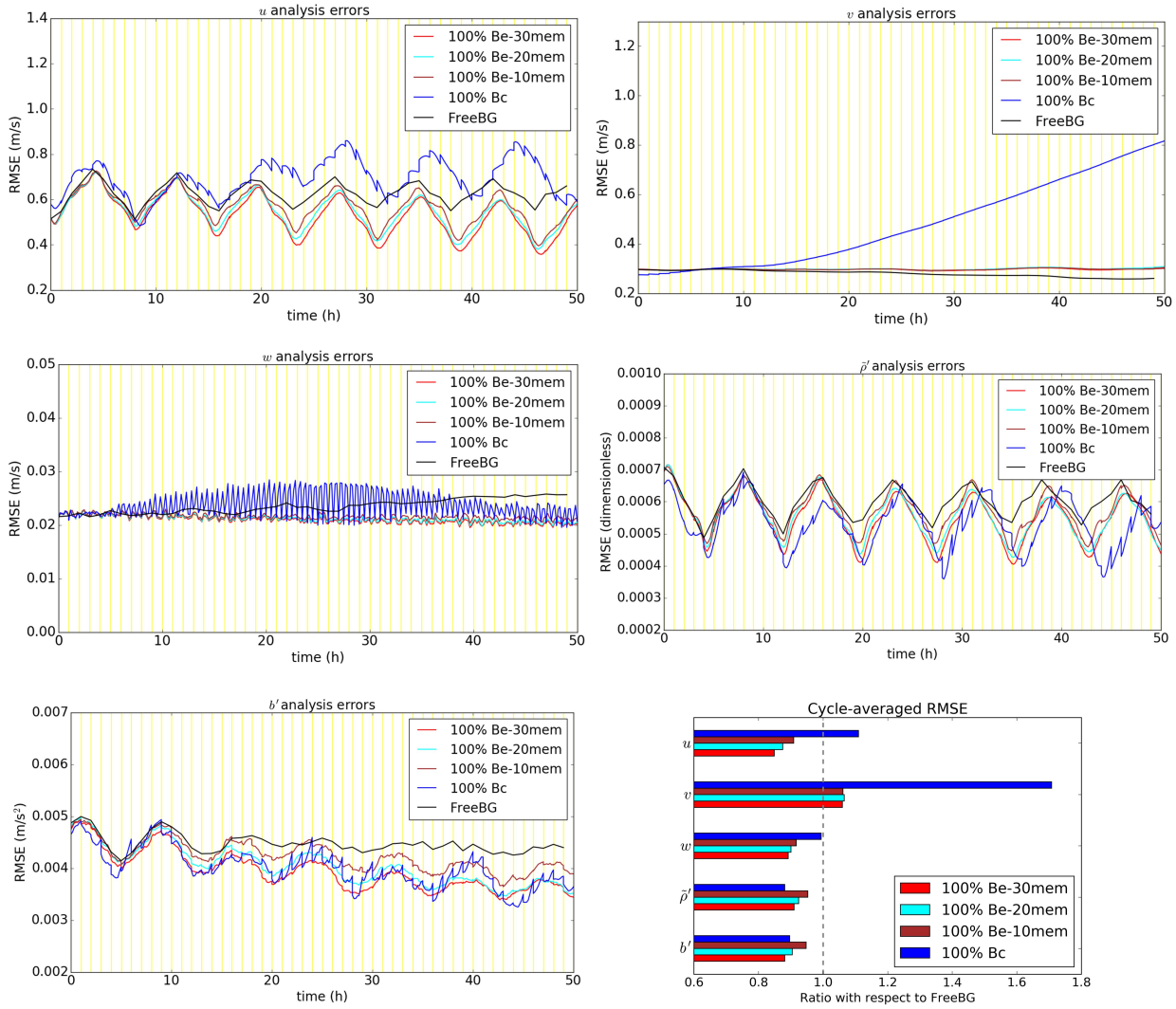

Figure 8. As in Fig. 5, but for EBV(d), EBV(d20) and EBV(d10) experiments (100% $\mathbf{B}_e$ with 30, 20, and 10 ensemble members respectively).