Ocean Dynamics
https://doi.org/10.1007/s10236-024-01643-6

Hybrid covariance super-resolution data assimilation

Sébastien Barthélémy[1,2] · François Counillon[1,2,3] · Julien Brajard[2,3] · Laurent Bertino[3]

Received: 21 May 2024 / Accepted: 2 October 2024

Abstract
The super-resolution data assimilation (SRDA) enhances a low-resolution (LR) model with a Neural Network (NN) that has learned the differences between high and low-resolution models offline and performs data assimilation in high-resolution (HR). The method enhances the accuracy of the EnKF-LR system for a minor computational overhead. However, performance quickly saturates when the ensemble size is increased due to the error introduced by the NN. We therefore combine the SRDA with the mixed-resolution data assimilation method (MRDA) into a method called "Hybrid covariance super-resolution data assimilation" (Hybrid SRDA). The forecast step runs an ensemble at two resolutions (high and low). The assimilation is done in the HR space by performing super-resolution on the LR members with the NN. The assimilation uses the hybrid covariance that combines the emulated and dynamical HR members. The scheme is extensively tested with a quasi-geostrophic model in twin experiments, with the LR grid being twice coarser than the HR. The Hybrid SRDA outperforms the SRDA, the MRDA, and the EnKF-HR at a given computational cost. The benefit is the largest compared to the EnKF-HR for small ensembles. However, even with larger computational resources, using a mix of high and low-resolution members is worth it. Besides, the Hybrid SRDA, the EnKF-HR, and the SRDA, unlike the MRDA, prevent the smoothing of dynamical structures of the background error covariance matrix. The Hybrid SRDA method is also attractive because it is customizable to available resources.

Keywords Super-resolution · Neural network · Ensemble data assimilation · Hybrid covariance · Quasi-geostrophic model

1 Introduction

Data assimilation (DA) methods estimate the initial condition of a forecast based on observations, a dynamical model, and statistical information on the error terms. They have played a vital role in the evolution of numerical weather prediction (Bauer et al. 2015) and are applied in many geophysical applications (Carrassi et al. 2018). The Ensemble Kalman Filter (EnKF) (Evensen 2003; Carrassi et al. 2018) has emerged in the past decades as a widespread data assimilation method because of its ease of implementation. The EnKF is inspired by Monte Carlo Markov Chain (MCMC) methods and proceeds into two steps: the forecast step, where an ensemble of realizations of the model is integrated forward in time, and the analysis step, where the ensemble is updated by linear regression, making use of the ensemble covariances to exploit the observations available. On the one hand, the computational costs of integration of the EnKF being proportional to the size of the ensemble can be deterring, especially in the case of geophysical models. However, on the other hand, the ensemble provides a time-evolving background error covariance matrix. The background error covariance matrix describes the spatial covariances between different variables and allows for the multivariate correction of the non-observed variables. However, in geophysical applications, the ensemble's size is necessarily limited, resulting in sampling errors, low rank of

Responsible Editor: Y. Miyazawa

✉ Sébastien Barthélémy
sebastien.barthelemy@uib.no

François Counillon
francois.counillon@nersc.no

Julien Brajard
julien.brajard@nersc.no

Laurent Bertino
laurent.bertino@nersc.no

[1] Universitetet i Bergen, Bergen, Norway

[2] Bjerknes Center for Climate Research, Bergen, Norway

[3] Nansen Environmental and Remote Sensing Center, Bergen, Norway

Published online: 14 October 2024

the background error covariance matrix, and ultimately deteriorated the system's performance. Several developments circumvent these issues, such as inflation (Anderson 2007), localization (Hamill et al. 2001) and hybrid covariance methods (Hamill and Snyder 2000).

While inflation and localization are routinely used, there is still room for developing and improving the original formulation of the EnKF. One research axis has been focusing on the benefits of using models at different resolutions. For example, Gao and Xue (2008) integrated and updated in low-resolution (LR) an ensemble and used its background error covariance matrix to update a single high-resolution (HR) state. The method, tested with different LR grids, provided results close to those of a single-resolution EnKF. The method of Gao and Xue (2008) was generalized by Rainwater and Hunt (2013) to the case of mixed-resolution data assimilation (referred to hereafter as MRDA). Two ensembles at two different resolutions are integrated, and their respective covariance matrices are linearly combined to update both ensembles in the HR space. The method was applied on different flavors of the Lorenz-96 toy model, Lorenz (1996), and performed much better than a full low-resolution EnKF, and compared with a complete HR EnKF that has more than twice the computational cost. Combining different resolutions has also percolated to hybrid variational-ensemble methods; see, for example, Kleist and Ide (2015); Buehner et al. (2010).

A different way to exploit LR models, named "Super-resolution data assimilation" (SRDA), was introduced in Barthélémy et al. (2022). If we consider an EnKF in LR, the computational cost is small, allowing for a large ensemble size. However, the error to the HR asymptotic covariance matrix can remain potentially high because of the bias introduced by the LR. On the other hand, an EnKF in HR does not allow for a large ensemble because of the computational cost of the model. However, it provides, for the same ensemble size, a better estimation of the background error covariance matrix compared to an EnKF in LR but at the expense of an increase of the computational cost. The SRDA was designed to take advantage of the best of both the high and low resolution. It was developed on top of an EnKF and consisted of emulating an HR EnKF by performing the ensemble integration in the LR similarly to Gao and Xue (2008), but downscaling to the HR grid with a neural network (NN) trained with offline free simulations. The assimilation is performed in the HR space to benefit from HR observations, and after assimilation, the HR fields are upscaled back to the LR grid. The authors showed that the SRDA performed closely to the EnKF, but for a fraction of the cost. It reduced the errors when the LR EnKF performed poorly and preserved the reliability of the ensemble. However, we will show here that for extensive computational resources, the downscaling errors of the SRDA remain and that the usual HR EnKF outperforms the SRDA. Besides, while the SRDA method was explicitly designed in an ensemble framework, it could also be beneficial in a variational one (Yasuda and Onishi 2023).

Nevertheless, some questions still need to be answered: is there some information in the HR model that cannot be emulated? Reintroducing at least one HR member in the ensemble would be a practical way to enrich the ensemble covariance. However, is this new approach cost-effective? In this work, we build on the work of Barthélémy et al. (2022) and Rainwater and Hunt (2013) by hybridizing the SRDA with an HR ensemble. The forecast step is accomplished in both the HR and LR spaces. As in the SRDA, the LR members are downscaled to the HR space with a NN before the assimilation step. The assimilation for both ensembles is performed in the HR space using the hybrid covariance from the HR and the downscaled LR members. After assimilation, the LR ensemble is upscaled back to the LR space. This flavor of the SRDA is called "Hybrid covariance super-resolution data assimilation" (Hybrid SRDA).

We extensively validate the approach with the quasi-geostrophic model run at two resolutions and compare the performance of the new scheme with the standard EnKF, the MRDA, and the SRDA. We will show that the Hybrid SRDA outperforms all other schemes at equivalent computational resources and better preserves the system's dynamical properties than the MRDA. It is also easily tunable and customizable to available resources, making it applicable to operational systems.

The overview of the paper is as follows: in Section 2, we present the ensemble data assimilation method used, the MRDA and the Hybrid SRDA schemes, the formulation of the Hybrid SRDA as a low-resolution scheme, and the super-resolution NN used in this work. In Sections 3, 4, and 5, we present the quasi-geostrophic model used in this study, the setting-up of the experiments, and the validation metrics. Section 6 shows the results of the Hybrid SRDA, and how they compare with the EnKF, the MRDA, and the SRDA. In the last part, Section 7, we give some conclusions and perspectives to this work.

2 Methods

We introduce the different methods that will be tested in the study. We start with a square root filter variant of the Ensemble Kalman Filter and continue with a suite of methods derived to improve its computational efficiency.

2.1 The deterministic ensemble Kalman filter

Let us consider a model whose dimension is n, and let $\mathbf{E} \in \mathbb{R}^{n \times N}$ an ensemble of N model realisations

$\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(N)}\right)$. We note $\mathbf{x} \in \mathbb{R}^n$ the ensemble mean and $\mathbf{A} \in \mathbb{R}^{n \times N}$ the ensemble anomalies; $\mathbf{x}$ and $\mathbf{A}$ are given by expressions (1a) and (1b) respectively:

$$\mathbf{x} = \frac{1}{N}\mathbf{E1}, \tag{1a}$$

$$\mathbf{A} = \mathbf{E}\left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right), \tag{1b}$$

where $\mathbf{I} \in \mathbb{R}^{N \times N}$ is the identity matrix and $\mathbf{1} \in \mathbb{R}^N$ is a vector with all elements equal to 1. In the following, $(.)^{\mathrm{f}}$ and $(.)^{\mathrm{a}}$ stand for the forecast and the analyzed states, respectively, and $(.)^T$ is the matrix transpose.

We note the true state $\mathbf{x}^{\mathrm{t}}$. In this study, we assimilate synthetic observations noted $\mathbf{d}$ that are defined by adding a Gaussian noise $\boldsymbol{\varepsilon}$ to the true state, see Eq. 2:

$$\mathbf{d} = \mathbf{H}\mathbf{x}^{\mathrm{t}} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(0, \mathbf{R}), \tag{2}$$

where $\mathbf{H} \in \mathbb{R}^{p \times n}$ is the observation operator, p is the number of observations, $\mathbf{R} \in \mathbb{R}^{p \times p}$ is the observation error covariance matrix.

In this study, we used the deterministic EnKF, DEnKF, introduced by Sakov and Oke (2008). The DEnKF is a square-root deterministic formulation of the EnKF Evensen (2003), which, unlike the EnKF, solves the analysis without the need for perturbation of the observations. It was shown that the two schemes behave similarly in the framework tested here (Sakov and Oke 2008), but the DEnKF performs slightly better than the standard EnKF. Furthermore, it inflates the errors by construction, which performs well in operational applications where corrections are minor (Sakov and Oke 2008). The system has been robustly tested in many systems (Sakov et al. 2012; Counillon et al. 2016; Bethke et al. 2021). In this work, we will refer to the DEnKF as EnKF for simplicity.

The DEnKF decomposes into two steps: the forecast step and the analysis step. In the forecast step, each member i of the ensemble is integrated forward in time by the model from the previous assimilation cycle to the current one:

$$\mathbf{x}^{\mathrm{f},i} = \mathcal{M}\left(\mathbf{x}^{\mathrm{a},i}\right), \quad i = 1, \ldots, N, \tag{3}$$

where $\mathcal{M}$ stands for the integration operator of the dynamical model.

The analysis step updates separately the ensemble mean, Eq. 4a, and the ensemble anomalies, Eq. 4b:

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{f}} + \mathbf{K}\left(\mathbf{d} - \mathbf{H}\mathbf{x}^{\mathrm{f}}\right), \tag{4a}$$

$$\mathbf{A}^{\mathrm{a}} = \mathbf{A}^{\mathrm{f}} - \frac{1}{2}\mathbf{K}\mathbf{H}\mathbf{A}^{\mathrm{f}}, \tag{4b}$$

where:

$$\mathbf{K} = \mathbf{P}^{\mathrm{f}}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^{\mathrm{f}}\mathbf{H}^T + \mathbf{R}\right)^{-1}, \tag{5a}$$

$$\mathbf{P}^{\mathrm{f}} = \frac{\mathbf{A}^{\mathrm{f}}\left(\mathbf{A}^{\mathrm{f}}\right)^T}{N-1}, \tag{5b}$$

are respectively the Kalman gain matrix and the background error covariance matrix estimated from the ensemble anomalies.

2.2 Super-resolution data assimilation

"Super-resolution data assimilation" (SRDA), introduced by Barthélémy et al. (2022), aims at taking advantage of the LR version of a dynamical model and HR observations to emulate an HR EnKF at a reduced computational cost.

We consider an ensemble $\mathbf{E}_{\mathrm{L}}$ of N_{L} LR model realisations $\left(\mathbf{x}_{\mathrm{L}}^{(1)}, \mathbf{x}_{\mathrm{L}}^{(2)}, \ldots, \mathbf{x}_{\mathrm{L}}^{(N_{\mathrm{L}})}\right)$. In the SRDA, the forecast step is performed in the LR space:

$$\mathbf{x}_{\mathrm{L}}^{\mathrm{f},i} = \mathcal{M}_{\mathrm{L}}\left(\mathbf{x}_{\mathrm{L}}^{\mathrm{a},i}\right), \quad i = 1, \ldots, N_{\mathrm{L}}, \tag{6}$$

where $\mathcal{M}_{\mathrm{L}}$ stands for the LR dynamical model operator. After the forecast step and before the data assimilation step, the LR members are downscaled to the HR space:

$$\mathbf{x}_{\mathrm{H}}^{\mathrm{f},i} = \mathcal{D}\left(\mathbf{x}_{\mathrm{L}}^{\mathrm{f},i}\right), \quad i = 1, \ldots, N_{\mathrm{L}}, \tag{7}$$

where $\mathcal{D}$ is a non-linear downscaling operator from the LR space to the HR space, which is trained with a neural network (NN) to learn the mismatch between the LR model and the HR model (Section 2.5). The training data set is made of pairs of HR and LR states that are constructed by repeating several times the following procedure: the HR model state is spatially interpolated to the LR grid, and the state of the LR and HR models are compared after a model integration equivalent to the assimilation window. It was shown that the operator reduces interpolation error and mitigates errors of the dynamical model (Barthélémy et al. 2022).

The mean and the anomalies of the ensemble are updated in the HR space following (4a-4b):

$$\mathbf{x}_{\mathrm{H}}^{\mathrm{a}} = \mathbf{x}_{\mathrm{H}}^{\mathrm{f}} + \mathbf{K}\left(\mathbf{d} - \mathbf{H}\mathbf{x}_{\mathrm{H}}^{\mathrm{f}}\right), \tag{8a}$$

$$\mathbf{A}_{\mathrm{H}}^{\mathrm{a}} = \mathbf{A}_{\mathrm{H}}^{\mathrm{f}} - \frac{1}{2}\mathbf{K}\mathbf{H}\mathbf{A}_{\mathrm{H}}^{\mathrm{f}}, \tag{8b}$$

where $\mathbf{K}$, $\mathbf{H}$, and $\mathbf{d}$ stand respectively for the HR Kalman gain, observation operator, and observations at the current assimilation cycle.

After the assimilation step, the members are upscaled back to the LR space for the next forecast:

$$\mathbf{x}_{\mathrm{L}}^{\mathrm{a,i}} = \mathbf{U}\,\mathbf{x}_{\mathrm{H}}^{\mathrm{a,i}}, \qquad i = 1, \ldots, N, \tag{9}$$

where $\mathbf{U}$ is a linear upscaling operator.

2.3 Mixed-resolution ensemble data assimilation

The "mixed-resolution ensemble data assimilation" (hereafter MRDA) was introduced by Rainwater and Hunt (2013) and builds the background error covariance matrix from the linear combination of the covariance matrix of ensembles at two different resolutions. In Rainwater and Hunt (2013), the mixed-resolution scheme was implemented with the LETKF, Hunt et al. (2007), which is another square root filter version of the EnKF.

In the following, we will use the subscripts:

- H $\leftarrow$ L to denote an object that has been downscaled from the LR to the HR space;
- L $\leftarrow$ H to denote an object that has been upscaled spatially from the HR to the LR space.

Note that in the MRDA and unlike in the SRDA, the downscaling operator is a simple spatial interpolation, i.e. linear or cubic spline.

Let us consider now two ensembles, $\mathbf{E}_{\mathrm{L}}$ and $\mathbf{E}_{\mathrm{H}}$, where N_{L} and N_{H} are the ensemble size of the LR and the HR respectively, $\mathbf{E}_{\mathrm{L}} = \left(\mathbf{x}_{\mathrm{L}}^{(1)}, \mathbf{x}_{\mathrm{L}}^{(2)}, \ldots, \mathbf{x}_{\mathrm{L}}^{(N_{\mathrm{L}})}\right)$ and $\mathbf{E}_{\mathrm{H}} = \left(\mathbf{x}_{\mathrm{H}}^{(1)}, \mathbf{x}_{\mathrm{H}}^{(2)}, \ldots, \mathbf{x}_{\mathrm{H}}^{(N_{\mathrm{H}})}\right)$. The forecast ensembles are integrated from the analysis ensemble at the previous assimilation cycle:

$$\mathbf{x}_{\mathrm{L}}^{\mathrm{f},i} = \mathcal{M}_{\mathrm{L}}\left(\mathbf{x}_{\mathrm{L}}^{\mathrm{a},i}\right), \quad i = 1, \ldots, N_{\mathrm{L}}, \tag{10a}$$

$$\mathbf{x}_{\mathrm{H}}^{\mathrm{f},i} = \mathcal{M}_{\mathrm{H}}\left(\mathbf{x}_{\mathrm{H}}^{\mathrm{a},i}\right), \quad i = 1, \ldots, N_{\mathrm{H}}, \tag{10b}$$

where $\mathcal{M}_{\mathrm{L}}$ and $\mathcal{M}_{\mathrm{H}}$ stand for the LR and HR dynamical models integration respectively.

For the analysis step, every LR member $\mathbf{x}_{\mathrm{L}}^{\mathrm{f},i}$ is downscaled from the LR to the HR grid, Eq. 11. We note $\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f},i}$ the resulting HR fields, and $\mathbf{E}_{\mathrm{H}\leftarrow\mathrm{L}}$ the associated HR ensemble.

$$\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f},i} = \mathbf{D}\,\mathbf{x}_{\mathrm{L}}^{\mathrm{f},i}, \quad i = 1, \ldots, N_{\mathrm{L}}. \tag{11}$$

The mean and the anomalies of the HR ensemble are noted respectively $\mathbf{x}_{\mathrm{H}}$ and $\mathbf{A}_{\mathrm{H}}$, while the mean and the anomalies of the LR ensemble downscaled to the HR space are noted $\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}$ and $\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}$ respectively. $\mathbf{x}_{\mathrm{H}}$, $\mathbf{A}_{\mathrm{H}}$, $\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}$, and $\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}$ are updated based on Eqs. 12a, 12b, 12c, and 12d:

$$\mathbf{x}_{\mathrm{H}}^{\mathrm{a}} = \mathbf{x}_{\mathrm{H}}^{\mathrm{f}} + \mathbf{K}^{\mathrm{h}}\left(\mathbf{d} - \mathbf{H}\mathbf{x}_{\mathrm{H}}^{\mathrm{f}}\right), \tag{12a}$$

$$\mathbf{A}_{\mathrm{H}}^{\mathrm{a}} = \mathbf{A}_{\mathrm{H}}^{\mathrm{f}} - \frac{1}{2}\mathbf{K}^{\mathrm{h}}\mathbf{H}\mathbf{A}_{\mathrm{H}}^{\mathrm{f}}, \tag{12b}$$

$$\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{a}} = \mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} + \mathbf{K}^{\mathrm{h}}\left(\mathbf{d} - \mathbf{H}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}\right), \tag{12c}$$

$$\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{a}} = \mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} - \frac{1}{2}\mathbf{K}^{\mathrm{h}}\mathbf{H}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}, \tag{12d}$$

where $\mathbf{H}$ and $\mathbf{d}$ are respectively the HR observation operator and the HR observations at the current assimilation cycle. In the following, $(.)^{\mathrm{h}}$ stands for a hybrid object. The hybrid Kalman gain $\mathbf{K}^{\mathrm{h}}$ is computed as follows:

$$\mathbf{K}^{\mathrm{h}} = \mathbf{P}^{\mathrm{h}}\mathbf{H}^{T}\left(\mathbf{H}\mathbf{P}^{\mathrm{h}}\mathbf{H}^{T} + \mathbf{R}\right)^{-1}, \tag{13}$$

where $\mathbf{R}$ is the HR observations error covariance matrix and $\mathbf{P}^{\mathrm{h}}$ is the hybrid background error covariance matrix:

$$\mathbf{P}^{\mathrm{h}} = (1-\alpha)\mathbf{P}_{\mathrm{H}}^{\mathrm{f}} + \alpha\mathbf{P}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}, \qquad 0 \leq \alpha \leq 1, \tag{14a}$$

$$\mathbf{P}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \frac{\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}\left(\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}\right)^{T}}{N_{\mathrm{L}} - 1}. \tag{14b}$$

The scalar α is called the hybridization coefficient.

In practice, $\mathbf{P}^{\mathrm{h}}$ can be expressed in terms of the hybrid ensemble anomalies $\mathbf{A}^{\mathrm{h}}$, Counillon et al. (2009); Rainwater and Hunt (2013), as the dimensions of the covariance matrices are often too large for it to be explicitly computed:

$$\mathbf{P}^{\mathrm{h}} = \frac{\mathbf{A}^{\mathrm{h}}\left(\mathbf{A}^{\mathrm{h}}\right)^{T}}{N_{\mathrm{H}} + N_{\mathrm{L}} - 1}, \tag{15}$$

with:

$$\mathbf{A}^{\mathrm{h}} = \sqrt{N_{\mathrm{H}} + N_{\mathrm{L}} - 1}\left[\sqrt{\frac{1-\alpha}{N_{\mathrm{H}}-1}}\mathbf{A}^{\mathrm{f}}_{\mathrm{H}}, \sqrt{\frac{\alpha}{N_{\mathrm{L}}-1}}\mathbf{A}^{\mathrm{f}}_{\mathrm{H}\leftarrow\mathrm{L}}\right], \tag{16}$$

where the notation [., .] stands for the concatenation of matrices with the same number of lines.

After the analysis step, the ensemble $\mathbf{E}^{\mathrm{a}}_{\mathrm{H}\leftarrow\mathrm{L}}$, is upscaled back linearly to the LR grid with the operator $\mathbf{U}$ before the next forecast step:

$$\mathbf{x}^{\mathrm{a},i}_{\mathrm{L}} = \mathbf{U}\,\mathbf{x}^{\mathrm{a},i}_{\mathrm{H}\leftarrow\mathrm{L}}, \quad i = 1, \ldots, N_{\mathrm{L}}. \tag{17}$$

2.4 Hybrid covariance super-resolution data assimilation

The "Hybrid covariance super-resolution data assimilation" scheme, hereafter referred to as "Hybrid SRDA", combines the SRDA scheme and the MRDA. In the perspective of the MRDA, the downscaling operator $\mathbf{D}$, Eq. 11, is replaced by the operator used in the SRDA $\mathcal{D}$ – i.e., a NN that learns the difference due to increased resolution during a model integration from one assimilation cycle to another. A schematic representation of the MRDA and the Hybrid SRDA is provided in Fig. 1.

It was demonstrated in Barthélémy et al. (2022) that the analysis of the SRDA can be performed in the LR space with identical performance. Similarly, the Hybrid SRDA can also be reformulated as a LR scheme. We refer to this formulation as "LR Hybrid SRDA" in the following. This reformulation can be advantageous when the training of the NN requires a lot of heavy HR model outputs. The LR Hybrid SRDA requires HR model outputs only at the observation location (on top of the LR samples) and not all HR model fields. Furthermore, it allows the decomposition of the super-resolution operator as a term that corrects for resolution model error and a spatial observation operator term, which can be helpful in understanding the relative benefits of the method (Section 6.2). A schematic representation of the LR Hybrid SRDA is given in Fig. 2. Full detail on the reformulation of the Hybrid SRDA as a LR scheme is provided in Appendix A.

2.5 Super-resolution neural network

The NN super-resolution operator $\mathcal{D}$ is identical to the one used in Barthélémy et al. (2022) with the same parameters since no additional training was conducted. A summary of the main properties of the NN operator is provided below, and the reader is referred to Barthélémy et al. (2022) for further details.

We use the enhanced deep super-resolution network (EDSR) model, adapted from Lim et al. (2017). The model consists of four residual convolutional blocks, one upsampling block for the LR model and a final convolutional layer.

The NN was trained using a dataset where the initial condition of the high-resolution (HR) model is upscaled to the LR space, the LR model integrated for 12-time steps (the length of the assimilation cycle), and the resulting state compared to that obtained with the HR model. The dataset consists of 10,001 snapshots of the HR fields, with the first 8,000 used as the training set, and the last 1,997 as the validation set.

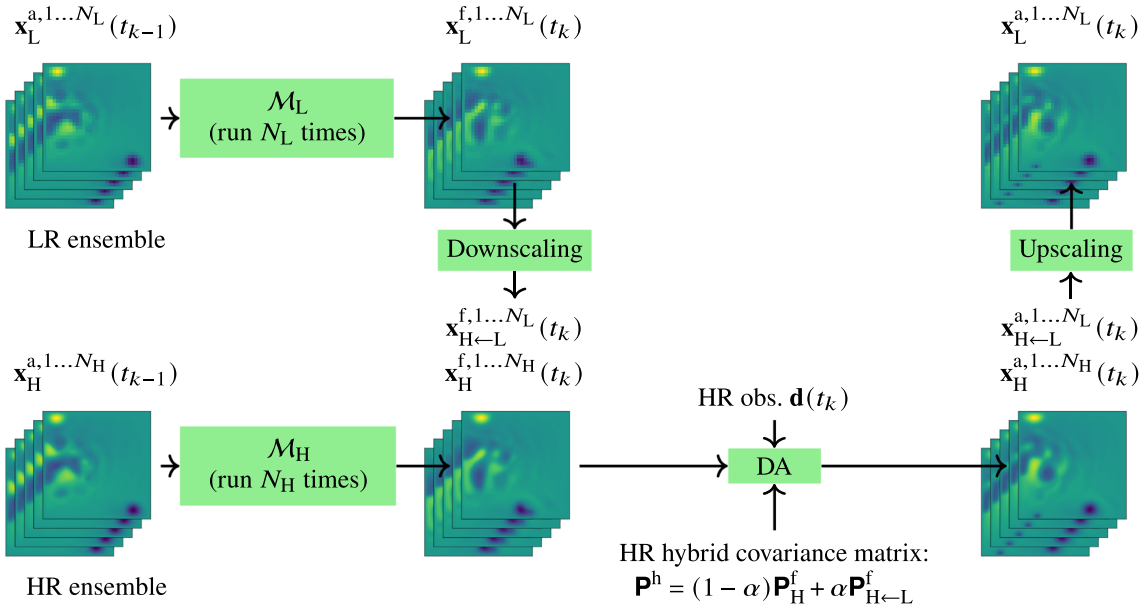


Fig. 1 Schematic representation of the MRDA and the Hybrid SRDA

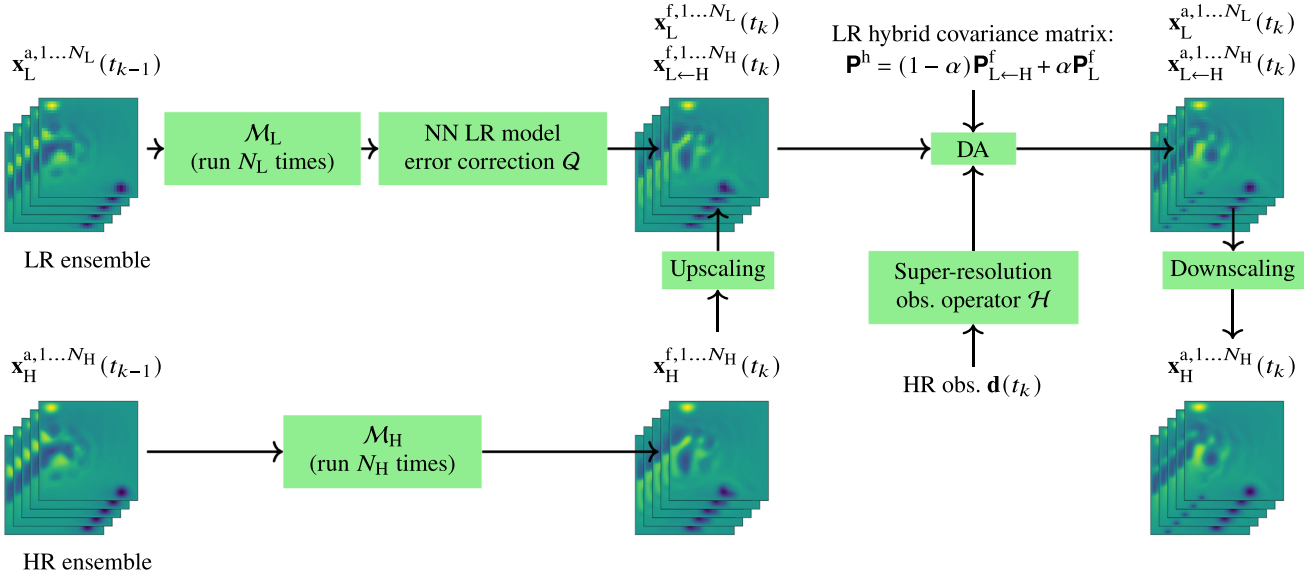


Fig. 2 Schematic representation of the LR version of the Hybrid SRDA

Three samples were discarded between the training and validation sections to ensure independence between the training and validation sets.

3 Models and data

3.1 Quasi-geostrophic model

The quasi-geostrophic (QG) model is a good test case for atmospheric and oceanic data assimilating systems, and we use the specific implementation (model parameter, observation network, assimilation frequency) that has been widely used for data assimilation schemes intercomparison (Sakov and Oke 2008; Dubinkina 2013; Attia and Sandu 2019; Gilbert et al. 2017; Counillon et al. 2009). This test case was also used in Barthélémy et al. (2022) to demonstrate the validity of the SRDA approach. The QG model is included in two data assimilation packages: DAPPER[1] (Data Assimilation with Python: a Package for Experimental Research), and NEDAS[2] (NERSC Ensemble Data Assimilation System).

It is a 1.5-layer reduced-gravity QG model on a square domain with a double-gyre wind forcing and bi-harmonic friction. The time evolution of the sea surface elevation ψ is given by Eq. 18:

$$\partial_t q = -\beta\psi_x - \rho J(\psi, q) - \rho_b\zeta + \rho_h\Delta\zeta - \rho_{bh}\Delta^2\zeta + 2\pi\sin(2\pi y), \quad (18)$$

where:

- $\zeta = \Delta\psi$ is the relative vorticity;
- β is the gradient of the Coriolis parameter with respect to the latitude;
- $q = \zeta - F\psi$ is the potential vorticity;
- F is the Froude number;
- $J(\psi, q) = \psi_y q_x - \psi_x q_y$ (where the subscripts x and y stand for the derivative with respect to the directions x and y);
- ρ is a multiplying scalar;
- ρ_b is the bottom friction;
- ρ_h is the horizontal friction;
- ρ_{bh} is the bi-harmonic horizontal friction;
- $2\pi\sin(2\pi y)$ represents a wind stress forcing term, with y being a location on a line of longitude.

with $\beta = 1$, $\rho = 10^5$, $\rho_b = \rho_h = 0$, and $\rho_{bh} = 2 \times 10^{-12}$ for the true run. For the data assimilation experiments $\rho_{bh} = 2 \times 10^{-11}$ is used. In the following, we refer to the bi-harmonic friction of the truth as ρ_{bht} and that of the data assimilation experiment as ρ_{bha}). The bi-harmonic friction of the assimilation run is higher, which makes the model used in assimilation non-identical to the truth. Hence, we are in the case where there is an error in the dynamical model used for the data assimilation which is more realistic and more challenging. The boundary condition is given by $\psi = \Delta\psi = \Delta^2\psi = 0$. An example of an output of the QG model obtained with these settings is given in Fig. 3. For more details about the QG model, see Jelloul and Huck (2003); Sakov and Oke (2008).

The QG model is used at two different resolutions that are referred to as high-resolution (HR, 129×129 grid cells),

[1] https://github.com/nansencenter/DAPPER

[2] https://github.com/nansencenter/NEDAS/

and low-resolution (LR, 64×64 grid cells), see Table 1. The differences between mean kinetic energy and eddy kinetic energy (deteriorated in the LR) are presented in Barthélémy et al. (2022) – Figs. 2 and 3. The computational cost of doubling the resolution from LR to HR results in an increase of the computational cost by a factor 8 as there are four times more points (in x and y), and one needs to divide the time step by two to satisfy the Courant-Friedrichs-Lewy condition.

4 Practical implementation of the data assimilation experiments

The EnKF, the MRDA, the SRDA, and the Hybrid SRDA are compared using twin experiments over 6000-time steps and with an analysis performed every 12-time steps (i.e. 500 assimilation cycles). At each assimilation cycle, we assimilate 300 synthetic observations generated by adding a white Gaussian noise with standard deviation $\sigma_{\mathrm{o,H}} = 2$ to the true run. Note that the true run is independent of the one used to train the super-resolution operator. The relative location of the observations mimics altimeter satellite tracks such as Topex-Poseidon and Jason, Le Traon and Ogor (1998); Chambers et al. (2003), with a spacing between tracks of 400 km between the along-track and an angle at the equator of approximately 66°. The observation operator is defined as the nearest neighbor (i.e., one for the nearest grid cell and 0 otherwise). The synthetic observations are constructed in the HR grid and, as such, are ideally collocated with the HR model. However, the LR grid introduces a shift, which results in an increased observation error term, (representativeness error term, Janjić et al. (2018)), which is estimated empirically at $\sigma_{\mathrm{o,L}} = 2.4$. The observations errors are by construction uncorrelated ($\mathbf{R}$ is diagonal). The initial ensemble at the start of the simulation is entirely independent from the truth. As the HR grid is a refinement of the LR grid, the two grids are overlapping, and the upscaling operator $\mathbf{U}$ in Eq. 17 is a sub-sampling of the HR grid.

We used localization – a local analysis framework with Gaspari and Cohn tapering, Gaspari and Cohn (1999) – and multiplicative inflation Li et al. (2009) to counteract sampling errors. The localization radius, inflation, and the hybridization coefficient α for the hybrid covariance schemes are tuned empirically to optimal performance for each ensemble size, i.e. minimizing the error while matching the spread.

5 Validation metrics

The performance of the different assimilation schemes is estimated at each assimilation cycle from the analyzed state, with the spatial root mean square error e, Eq. 19.

$$e = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathbf{x}^{\mathrm{a}}(j) - \mathbf{x}^{\mathrm{t}}(j)\right)^2}, \tag{19}$$

The spatial root mean square ensemble spread s, Eq. 20, was also considered to assess the reliability of the system (Fortin et al. 2014).

$$s = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\frac{1}{N-1}\sum_{i=1}^{N}\left(\mathbf{x}^{\mathrm{a},i}(j) - \mathbf{x}^{\mathrm{a}}(j)\right)^2}, \tag{20}$$

where $\mathbf{x}^{\mathrm{a}}$ and $\mathbf{x}^{\mathrm{t}}$ are respectively the mean analyzed state and the true state at the current assimilation cycle. In a perfectly reliable ensemble data assimilation system, $e = s$ Anderson and Anderson (1999). The general conclusions of the paper remain the same if we compute these scores on the prior state rather than on the analyzed state (not shown).

We have also considered the "spread reduction factor" (SRF) that allows to evaluate the relative efficiency of different assimilation schemes with different ensemble sizes. The SRF is defined as follows:

$$\mathrm{SRF} = \sqrt{\frac{\mathbf{H}\mathbf{P}^{\mathrm{f}}_{\mathrm{H}}\mathbf{H}^{T}}{\mathbf{H}\mathbf{P}^{\mathrm{a}}_{\mathrm{H}}\mathbf{H}^{T}}} - 1. \tag{21}$$

The SRF represents the reduction of the spread in the observation space and can be interpreted as a measure of the magnitude of the data assimilation correction, see Sakov and Bertino (2011). The SRF is between 0 (*i.e.* the data assimilation step has no impact on the ensemble) and ∞ (*i.e.* all the analyzed members are the same). As the EnKF performs a linear analysis update, an ideal ensemble data assimilation system would achieve optimal performance while performing the smallest updates possible.

The temporal average of these scores is presented beyond 10 assimilation cycles – the time needed for the system to

Table 1 Summary of different model configurations

Name	Grid point size	Time step	State size	Snapshot figure
HR	1	1	129×129	Figure 3-(a)
LR	4	2	65×65	Figure 3-(b)

Grid point size and time step are expressed in terms of that of the HR model

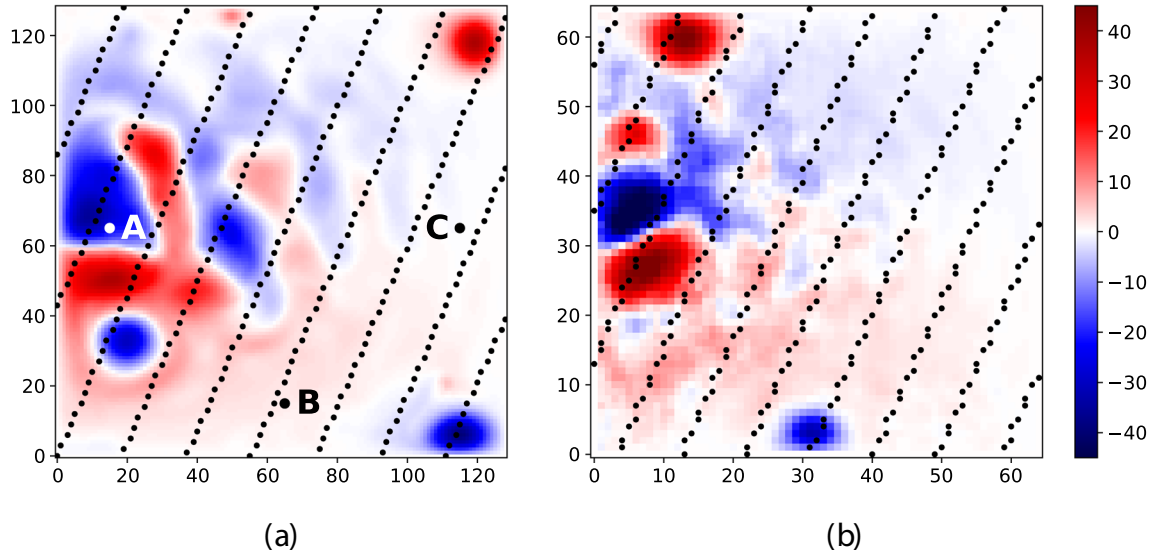


Fig. 3 Snapshot of the sea level for (a) the HR model (b) the LR model. The black points denote the location of the observations

converge to stable performance – to discard the assimilation spin-up phase.

For the EnKF-HR, Hybrid SRDA, MRDA, and SRDA, the assimilation step is performed in the HR space, and the scores e and s are computed in the HR space. In the case of the MRDA and the Hybrid SRDA, we follow Rainwater and Hunt (2013) and compute e, s, and the SRF over the HR ensemble. For example, if we use $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (5, 10)$ members, e, s, and the SRF are computed over the $N_{\mathrm{H}} = 5$ HR members. We tested computing these scores over the combination of the HR and the LR ensembles, but the LR members penalize the scores, especially for the MRDA (not shown).

6 Results

6.1 Methods intercomparison

We aim to compare first the performance of the different methods at equivalent computational costs. For large operational systems, the computational time of integrating the ensemble is often much larger than that of the assimilation step.

We start by assuming that the assimilation cost is negligible (and thus the additional cost of Hybrid SRDA and MRDA over EnKF) and compare the performance of the methods at an equivalent computational cost of model integration, which depends on the cost of running the ensemble of HR and LR members (having a ratio of 8 LR members to 1 HR member). We also discard the cost of training the NN operator (Section 2.5), which has to be done only once before the experiments. For example, we compare the methods for a computational cost equivalent to the EnKF-HR with 9 members, which is equivalent to using 72 LR members, and the MRDA and the Hybrid SRDA for a combination of $(N_{\mathrm{H}}, N_{\mathrm{L}})$ equivalent to 9 HR members, for example $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (3, 48)$. We show in Fig. 4-(a) the optimal RMSE for the different schemes for computational resources equivalent to integrating 5, 7, 9, …, 15 HR members. It is striking that the RMSE of the SRDA is stable or only marginally reducing with increasing computational resources and is poorer than the EnKF-HR beyond the cost of 9 HR members. On the contrary, the performance of the other scheme improves with increasing computational resources, and we anticipate they saturate around a computational cost equivalent to 25 HR members. For an equivalent computational cost of 5 HR members, SRDA already uses 40 LR members, and the sampling error is already low. It implies that the SRDA scheme is only beneficial for "low computational" resources and that its performance saturates quickly. The MRDA performs better than the standard EnKF, in agreement with Rainwater and Hunt (2013). The Hybrid SRDA is best for all computational resources and better than the MRDA, especially for minimal computational resources (roughly 14% of reduction of the RMSE). Another way to look at the plot is to check for which computational resources the system can achieve a given RMSE. For example, the Hybrid SRDA reaches an RMSE of 0.8 at half the computational resources required with the EnKF (6 vs 13). Except in the case of the EnKF for HR computational resources equal to 5, all schemes display small SRF between 0.17 and 0.23, Fig. 4-(b). This corresponds to relatively small corrections. This shows the superiority of the Hybrid SRDA as it achieves the best performance in terms of RMSE while, on

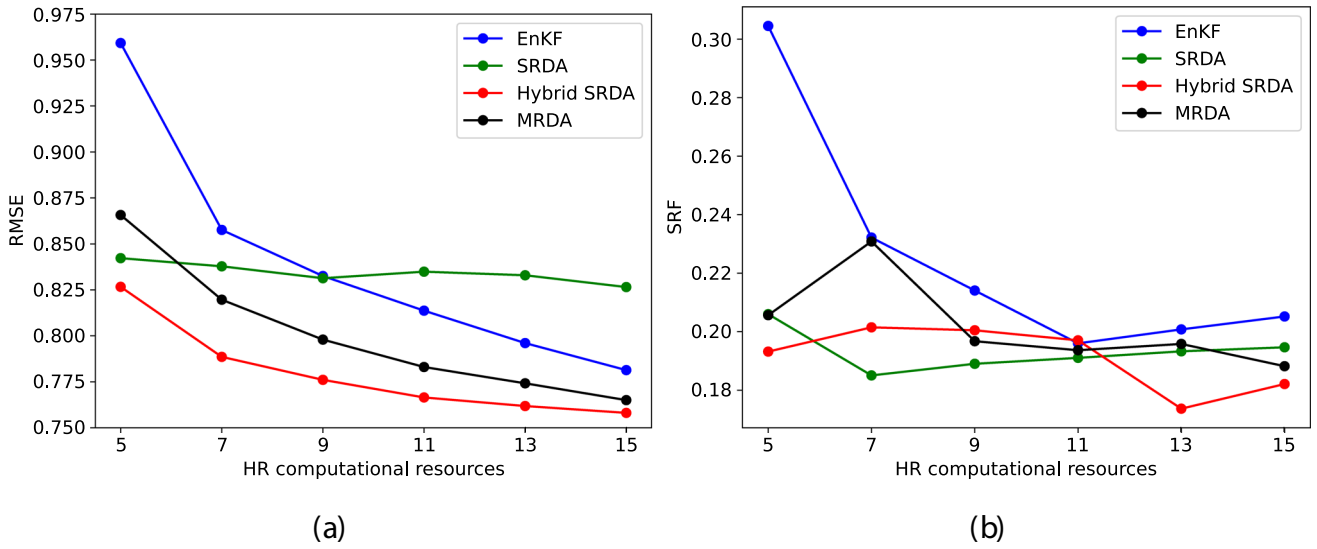


Fig. 4 (a) Optimal averaged RMSE and (b) associated averaged SRF for the different schemes as a function of the computational resources expressed in HR model integrations

average, the magnitude of the corrections is comparable to the other schemes. Besides, having a small SRF limits the risk of shocks due to the assimilation correction.

While the above shows performance where the integration of the forecast ensemble dominates the computational cost, we are investigating here the computational time reduction allowed by the SRDA, the Hybrid SRDA, and the MRDA. We compare, for a given ensemble size, the computational cost and the skill of an ensemble composed of 15 HR members (EnKF-HR), 15 emulated LR members (SRDA), and members of mixed resolution with $(N_H, N_L) = (5, 10)$ members (MRDA and Hybrid SRDA). Figure 5-(a) shows, for each of the schemes, the decomposition of the total CPU time according to the phase of the algorithm: integration, downscaling, assimilation, and upscaling. It shows that, for that ensemble size, the Hybrid SRDA and the MRDA reduce the computational time by one-third, and the SRDA by half, compared to the EnKF-HR. Additionally, Fig. 5-(a) shows that the computational time of the downscaling step with the NN remains limited and accounts for 7.9% of the total computational time for the Hybrid SRDA and 16.7% in the case of the SRDA. However, the relative computational cost of the

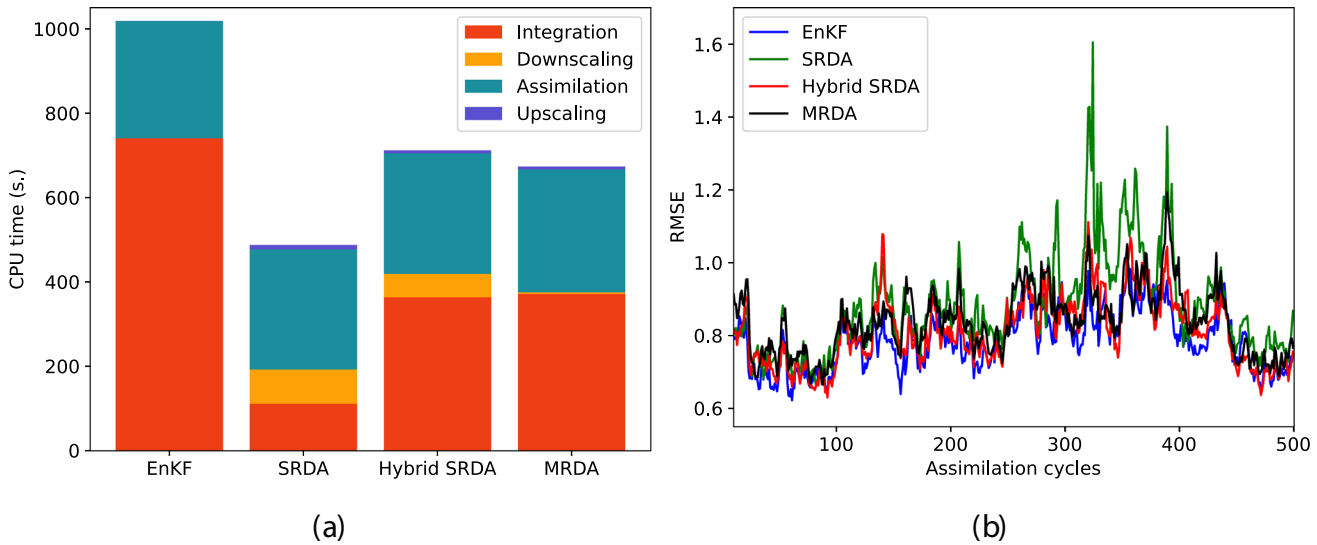


Fig. 5 (a) Bar diagram of the CPU time (s.) for the different schemes as a function of the steps of the algorithm: integration, downscaling, assimilation, and upscaling. (b) Time series of the RMSE for the EnKF with $N_H = 15$, the SRDA with $N_L = 15$, the Hybrid SRDA and the MRDA with $(N_H, N_L) = (5, 10)$

Table 2 Average SRF for the EnKF, the SRDA, the Hybrid SRDA, and the MRDA

	EnKF	SRDA	Hybrid SRDA	MRDA
SRF	0.228	0.227	0.228	0.246

NN downscaling step would become comparatively much smaller with realistic oceanic/atmospheric models because the cost of integration of these models would drastically increase. Figure 5-(b) shows the time series of the RMSE for the different schemes. The Hybrid SRDA displays minor differences with the EnKF, with an increase of the RMSE of approximately 3.1%, but with a reduction of the total computational cost of roughly one-third due to the trade-off between HR and LR members. The MRDA also shows performance very close to the EnKF, albeit poorer than the Hybrid SRDA. The SRDA with a 15-members ensemble displays a more significant error despite the NN emulator. The largest RMSE differences occur during challenging events, between assimilation cycles 300 and 400, where the model error is more significant. The SRDA results in a deterioration of the mean RMSE by approximately 12% compared to the EnKF-HR. In that configuration, all the schemes, except the MRDA, display a SRF close to 0.227, see Table 2. The MRDA displays a SRF of 0.246, which is approximately 7.9% larger than the other schemes. This shows that the MRDA needs to extract more information from the observations to perform as well as the Hybrid SRDA.

Forecast experiments were conducted for each scheme with the optimal ensemble size of Fig. 4-(a) for 7 HR computational resources. For each scheme and at each assimilation cycle, a 120 HR times steps forecast was performed (equivalent to 10 assimilation cycles). Figure 6-(a) shows the RMSE (as a function of the forecast lead time) averaged over the last 490 forecasts. It should be highlighted that while the RMSE is computed from 7 HR members in the EnKF we only use the HR members for the other scheme; i.e., 4 members in the Hybrid SRDA and 5 in the MRDA. This highlights that the benefit of the Hybrid SRDA and MRDA compared to the EnKF found at analysis time is also sustained at a longer lead time. Figure 6-(b) shows the averaged forecasted spread, we can notice that the spread in all schemes underestimates the amplitude of the RMSE as a consequence of model and sampling error, Carrassi et al. (2018). In an ensemble data assimilation system, this is counteracted with inflation, which is not used when running the forecast. In the particular case of the SRDA, the forecast is performed in the LR space, which explains why we observe such a drastic increase in RMSE at a longer lead time.

To further investigate the behavior of the different schemes, we analyze the spatial covariance structure from the forecast ensemble. For that, we compute the correlation length scale L_p and the variance of the covariance functions of the different schemes at three characteristic points (Fig. 3-(a)):

- **Point A:** in the left part of the domain, where spatial and temporal variability is high;
- **Point B:** in the lower part of the domain with the passing eddies;
- **Point C:** in the right part of the domain with low variability.

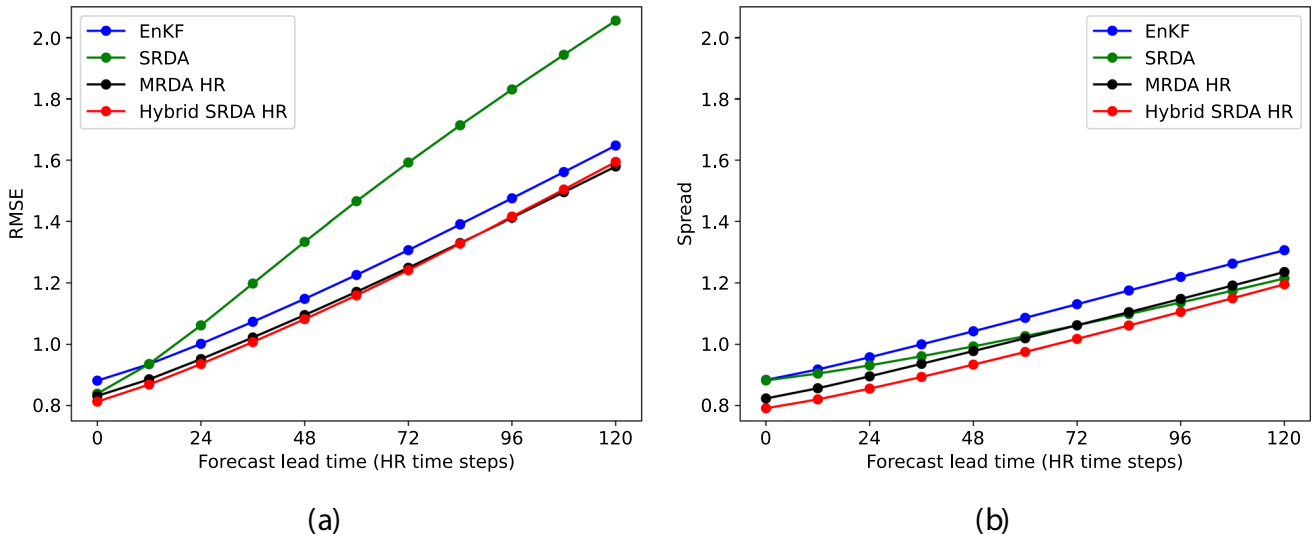


Fig. 6 (a) RMSE and (b) spread for the EnKF, the SRDA, and the HR ensemble of the MRDA and the Hybrid SRDA

The correlation length scale L_p is derived from the correlation matrix $\mathbf{C}^{\mathrm{f}}$, which is calculated from the covariance matrix $\mathbf{P}^{\mathrm{f}}$, as follows:

$$\mathbf{C}^{\mathrm{f}} = \mathbf{\Sigma}^{-1}\mathbf{P}^{\mathrm{f}}\mathbf{\Sigma}^{-1}, \tag{22}$$

where $\mathbf{\Sigma}$ is a diagonal matrix with the standard-deviation of $\mathbf{P}^{\mathrm{f}}$ on the diagonal. The correlation length scale L_p can be computed from the correlation matrix $\mathbf{C}^{\mathrm{f}}$ using the Gaussian-based formula in Pannekoucke et al. (2007):

$$L_p = \frac{\delta x}{\sqrt{-2\ln\left(\rho\left(\delta x\right)\right)}}, \tag{23}$$

where ρ is the correlation function at the point that we consider (i.e. ρ is the column of the correlation matrix $\mathbf{C}^{\mathrm{f}}$ corresponding to that point) and δx is the considered "separation distance"; in this study, δx corresponds to the HR grid resolution.

The correlation length scales are averaged along the horizontal and vertical axis to account for possible anisotropy in the covariance functions. The variance and the correlation length scales are computed in the HR space with 15 members for all methods – again, Hybrid SRDA and MRDA use $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (5, 10)$ distribution. Correlation length scales and variance in the EnKF-HR are our targets. The results for the horizontal length scales are displayed in Fig. 7.

At each point the correlation length scale in the MRDA is lower than that of the EnKF, and the variance is overestimated at point B. The variance at this point is driven by eddies passing along the domain's southern border. The propagation speed of the eddies, with the LR members, is slower than in the HR and in the true run Barthélémy et al. (2022), and their spatial dispersion increases the variance. The introduction of noise with the integration of some members in LR results in a reduction of the spatial correlation, and therefore, in a reduction of the correlation length scales.

The mean values of the variance and the correlation length scales in the SRDA and the Hybrid SRDA are much closer to that of the EnKF. It was shown in Barthélémy et al. (2022) that the NN is able to correct, to some extent, the LR model error and, in particular, to correct the propagation speed of the eddies. This results in a better estimation of the background error covariance matrix, and therefore, we show that the NN operator is also able to correct the correlation length scales and variance of the covariance functions.

6.2 Properties of the hybrid SRDA

In Section 2.4 and Appendix A, we show that it is possible to write down the Hybrid SRDA directly in the LR space (LR Hybrid SRDA). Instead of having one NN operator that interpolates the LR members to HR, this formulation uses 2 NN operators: one, $\mathcal{Q}$, that corrects the model error in the LR space, and another one, $\mathcal{H}$, that stands for a non-linear super-resolution observation operator.

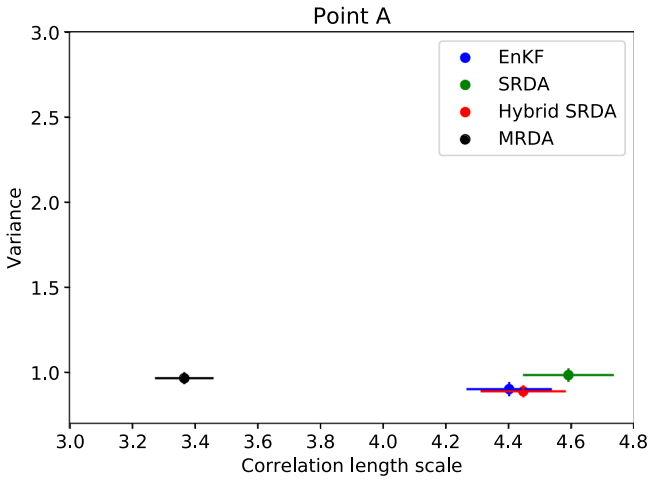
If we compare the two versions of the Hybrid SRDA (in HR and LR) at one assimilation cycle, their analysis error in the LR space is nearly identical (up to the machine error of the order of magnitude of 10^{-16}), but there are differences in the HR space analysis, of the order of 10^{-2}. These are the consequences of inaccuracy in the NN downscaling step from the LR to HR space after assimilation. These discrepancies impact only marginally the performance of the LR Hybrid SRDA, which shows overall somewhat comparable mean RMSE with the Hybrid SRDA see Table 3. In the following, they will be considered equivalent.

An advantage of the LR form of the Hybrid SRDA is that it allows the disentangling of the model error correction term and the observation operator term. Mathematically speaking, we can replace the operator $\mathcal{Q}$ in the equations (Appendix A) by identity, in which case only the super-resolution observation operator gets effective (we note Hybrid SRDA$_{\mathcal{H}}$), or conversely only applying the model error correction term and using linear spatial interpolation (noted Hybrid SRDA$_{\mathcal{Q}}$) for the observation operator. The standard Hybrid SRDA and the Hybrid SRDA$_{\mathcal{Q}}$ perform nearly equivalently, while Hybrid SRDA$_{\mathcal{H}}$ displays a larger RMSE in line with the MRDA, which implies that the reduction of error between Hybrid SRDA and MRDA is primarily due to the correction of the LR dynamical model, see Table 3.

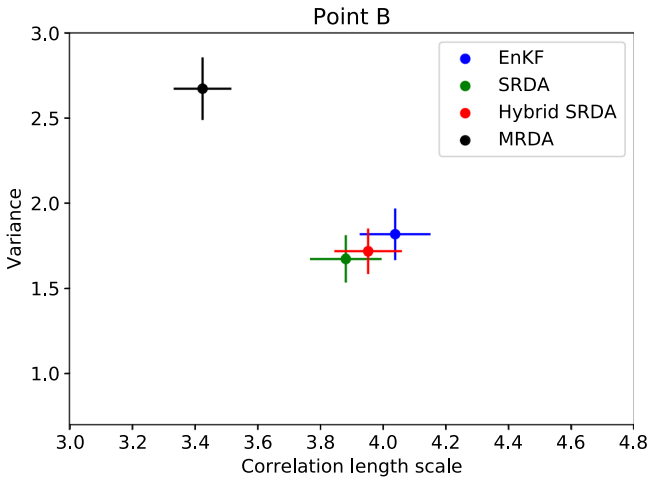
6.3 Tuning and robustness of the method

The Hybrid SRDA method requires tuning two additional parameters compared to the standard EnKF: the ratio of N_{H} and N_{L} and the hybrid coefficient α. Tuning data assimilation parameters can become a bottleneck with expensive operational models. We aim here to assess the sensitivity of the method to the choice of the method and try to guide on how to estimate the optimal value.

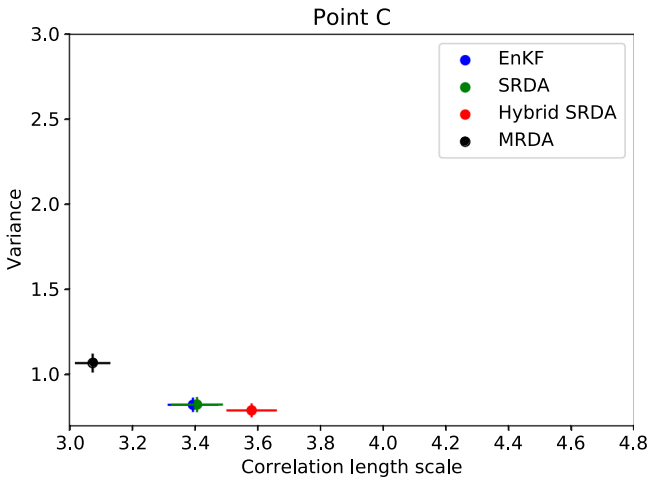
As shown in Fig. 8 for computational resources equivalent to 9 HR members, the Hybrid SRDA shows minor sensitivity to the distribution of $(N_{\mathrm{H}}, N_{\mathrm{L}})$ and less than the MRDA. The method shows very stable performance if the ratio $N_{\mathrm{H}}/\left(N_{\mathrm{H}} + N_{\mathrm{L}}\right)$ is selected between 10 and 45% (Fig. 9). This is the direct consequence of the reduction of the LR model error with the NN, see Barthélémy et al. (2022), that results in a reduction of the sampling error and limits the sensitivity of the scheme to the ratio $N_{\mathrm{H}}/\left(N_{\mathrm{H}} + N_{\mathrm{L}}\right)$. In such a configuration, the Hybrid SRDA consistently outperforms the MRDA with optimal settings. For low computational resources, a ratio of about 10% is best to have sufficient LR members to reduce sampling error. For more considerable computational resources, a higher ratio of [30;45] % is optimal. As such, we can conclude that the method is not sensitive



(a)



(b)



(c)

Fig. 7 Diagram showing the correlation length scale versus variance. The points represent the averaged values of the correlation length scales and the variance, and the bars depict the confidence interval at 95%

Table 3 RMSE of the Hybrid SRDA, the different flavors of the LR Hybrid SRDA, and the MRDA

	Hybrid SRDA	LR Hybrid SRDA	LR Hybrid SRDA$_{\mathcal{Q}}$	LR Hybrid SRDA$_{\mathcal{H}}$	MRDA
e	0.788	0.799	0.802	0.839	0.832

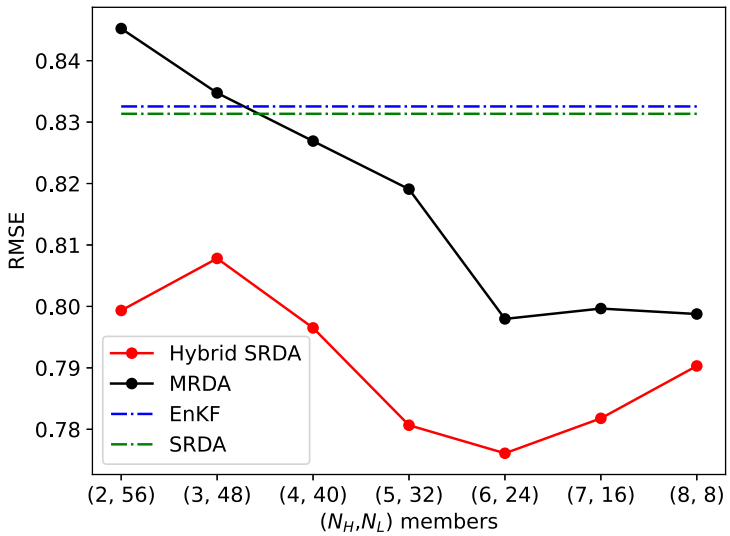


Fig. 8 Time averaged RMSE of the Hybrid SRDA and MRDA with various combinations of HR and LR members, all equivalent to 9 HR members. The EnKF-HR and the SRDA at equivalent computational resources are added for reference

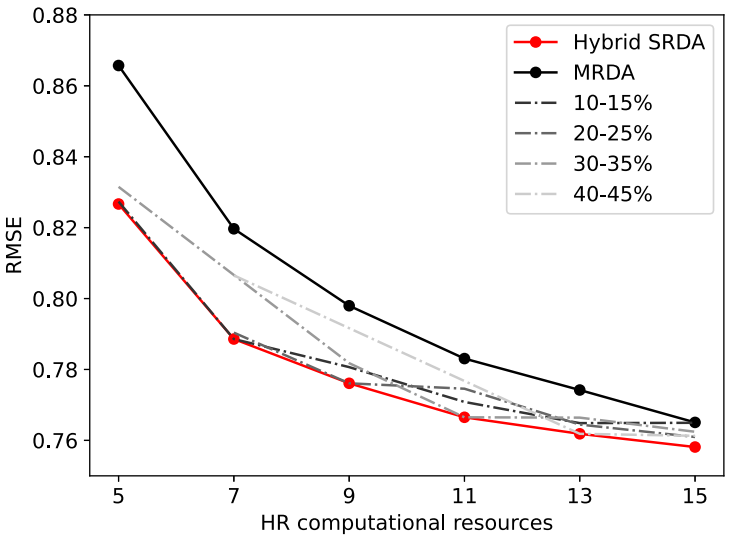


Fig. 9 Mean RMSE for the Hybrid SRDA for a different ratio $N_{\mathrm{H}}/\left(N_{\mathrm{H}}+N_{\mathrm{L}}\right)$. The optimal Hybrid SRDA and optimal MRDA are added for reference

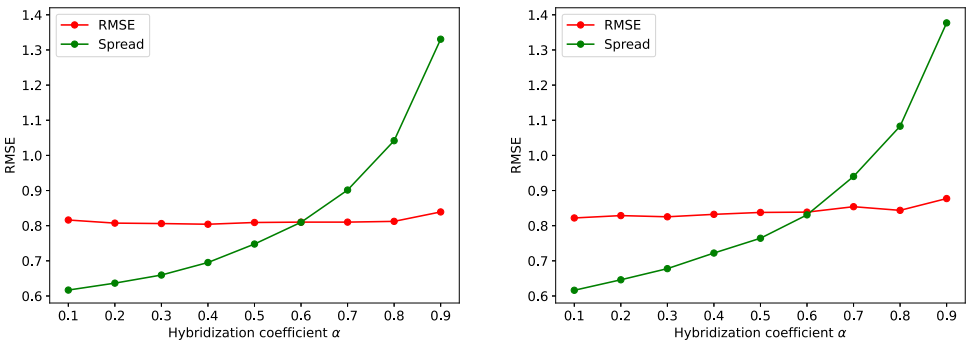


Fig. 10 RMSE and spread computed for the HR ensemble with different values of the hybridization coefficient α, for (a) the Hybrid SRDA and (b) the MRDA. In both cases $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (5, 10)$

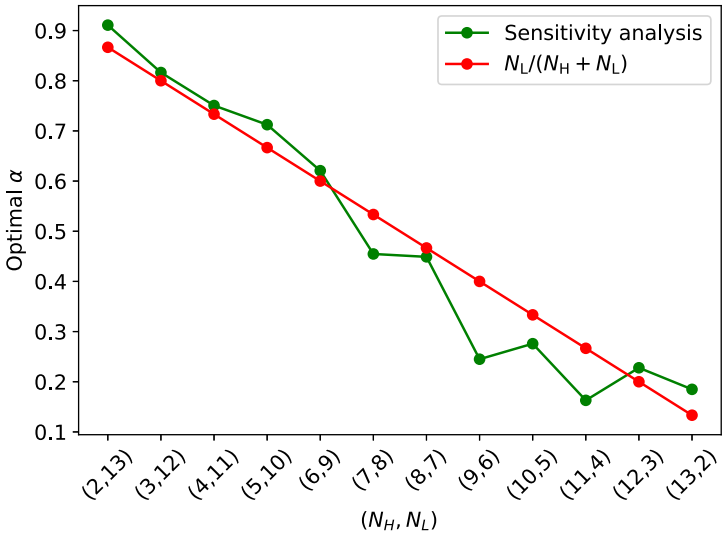


Fig. 11 Optimal hybridization coefficient from a sensitivity analysis (green line) and the ratio $N_L/(N_H + N_L)$ (red line)

to fine-tuning the ratio of (N_H, N_L) and that picking a value within the interval [10; 45] % provides nearly optimal results.

The hybridization coefficient α is another critical parameter of the Hybrid SRDA method and weighs the relative importance of the background error covariance from the HR and the LR ensembles (14a). However, Fig. 10 shows for computational resources equivalent to 7 HR members that the hybridization coefficients have limited influence over the RMSE for values between 0.1 and 0.9 for both the HR and the LR ensembles. Outside of the interval [0.1; 0.9], the RMSE increases until the scheme eventually diverges (not shown), similarly to Fig. 2 in Rainwater and Hunt (2013). Still, we can notice that the hybridization coefficient has more influence on the ensemble spread than the RMSE. In this study, the optimal value of the hybridization was set to ensure the DA system's reliability – i.e. that the spread matches the error of the ensemble mean Anderson and Anderson (1999); Fortin et al. (2014). For example, in the case of Fig. 10, the optimal value of α is 0.6 for both the Hybrid SRDA and the MRDA. However, another approach for estimating α would be to compute it as the ratio $N_L/(N_H + N_L)$ – this assumes that HR and LR emulated members have the same quality and that sampling error is the dominant source of error. Figure 11 displays the values of α computed with that formula, which matches the empirically tuned optimal values of α. It shows that the formula is a good approximation of the optimal hybridization coefficient α. We, therefore, again can conclude that the Hybrid SRDA is robust to the setting of the α values and that values that optimize the reliability of the system can be estimated based on the ratio of (N_H, N_L). Adaptive time-varying formulations for estimating α (El Gharamti 2021; Ménétrier and Auligné 2015; Barthélémy et al. 2024) have been proposed and were shown to improve the performance of the hybrid covariance DA methods. The use of adaptive parameters has not been explored here.

7 Conclusion

In this work, we have extended the concept of super-resolution data assimilation (SRDA, introduced by Barthélémy et al. 2022) building on the mixed-resolution ensemble data assimilation (MRDA) (Rainwater and Hunt 2013). We have shown that including a few HR simulations in the super-resolution ensemble mix is cost-effective. After the assimilation step, the LR emulated members are scaled back to the LR grid for the next integration cycle. While in Rainwater and Hunt (2013), the downscaling step is performed with a cubic spline interpolation operator, the Hybrid SRDA uses a neural network that learns the mismatch between the LR and HR models.

We have compared the DA methods' performance at an equivalent cost of model integration (i.e., the most common configuration in operational data assimilation systems). While the SRDA works well for low computational resources, its performance quickly saturates with increasing computational cost due to inaccuracy in the NN operator and rapidly performs poorer than the EnKF. The Hybrid SRDA, on the contrary, shows a pronounced improvement as computational resources increase and outperforms all other schemes (EnKF, MRDA, and SRDA). It shows that while the NN operator substantially improves the LR members, it is still preferable to use even a fraction of the available resources to HR members when formulating the covariance and performing the forecast. We demonstrate that the benefit of the Hybrid SRDA over the MRDA is mainly related to the correction of the low-resolution dynamical model error. We also show that the Hybrid SRDA performs best at an equivalent analysis computational cost. In particular, it does not reduce the correlation length scale of the covariance as in the MRDA. The Hybrid SRDA requires tuning two new parameters compared to the standalone EnKF – the ratio of the number of HR

members versus LR members and the hybridization coefficient –, which can be cumbersome to estimate with expensive systems. However, we show that the method is robust (with little sensitivity to the parameters) and that the optimal value of these parameters is easily predictable a priori.

As for the SRDA, the method's applicability to a realistic model remains to be demonstrated. One question is how to balance the computing load on parallel architecture since the runs of models at different resolutions are not equally fast. We see that as a manageable challenge as the model's code parallelization is typically scaling well with increasing computational resources, so it is possible to allocate more CPU in the HR model so that it clocks at the same speed as the LR model. For instance, we use this approach for the supermodel Counillon et al. (2023), where we run in parallel different ESMs with very different speeds. Another question concerns multivariate systems: the method was tested with an univariate QG model in this study. Will the NN operator work just as well for a high-dimensional multivariate system with irregular coastlines? Should we develop separate NNs for the emulation of each field (temperature, salinity...) and/or vertical layer, or should we only develop one NN that emulates all the variables together? Will the NN preserve the covariances between the different variables? The success of NNs emulating some of the most advanced forecasting models in NWP suggests that this is doable, Lam et al. (2023). We are therefore confident that the Hybrid SRDA has the potential to become applicable and competitive in future operational forecasting systems, and we are currently testing the approach in the field of climate predictions with the Norwegian Climate Prediction Model Bethke et al. (2021), and sea ice forecasting Xie et al. (2017).

Appendices

A Rewriting the Hybrid SRDA as a LR scheme

A.1 Statement

We assume that there exists two operators $\mathcal{Q}$ and $\mathcal{H}$ such that:

- $\mathcal{Q} : \mathbb{R}^{n_L \times N} \rightarrow \mathbb{R}^{n_L \times N}$ corrects the model error of the LR ensemble $\mathbf{E}_L^f$;
- $\mathcal{H} : \mathbb{R}^{n_L \times N} \rightarrow \mathbb{R}^{p \times N}$ is a super-resolution observation operator that maps the LR ensemble to the location of the HR observations.

The LR formulation of the Hybrid SRDA can be written as follows:

$$\mathbf{x}^a_{L\leftarrow H} = \mathbf{x}^f_{L\leftarrow H} + \mathbf{K}^h \left(\mathbf{d} - \mathbf{H}\mathbf{x}^f_H\right), \tag{24a}$$

$$\mathbf{A}^a_{L\leftarrow H} = \mathbf{A}^f_{L\leftarrow H} - \frac{1}{2}\mathbf{K}^h\mathbf{H}\mathbf{A}^f_H, \tag{24b}$$

$$\mathbf{x}^a_L = \tilde{\mathbf{x}}^f_L + \mathbf{K}^h \left(\mathbf{d} - \hat{\mathbf{x}}^f_{H\leftarrow L}\right), \tag{24c}$$

$$\mathbf{A}^a_L = \tilde{\mathbf{A}}^f_L - \frac{1}{2}\mathbf{K}^h\hat{\mathbf{A}}^f_{H\leftarrow L}, \tag{24d}$$

where $\mathbf{K}^h$ is the Kalman gain computed from the LR hybrid anomalies $\tilde{\mathbf{A}}^h$ and $\hat{\mathbf{A}}^h$:

$$\mathbf{K}^h = \frac{\tilde{\mathbf{A}}^h\left(\hat{\mathbf{A}}^h\right)^T}{N_H + N_L - 1}\left(\frac{\hat{\mathbf{A}}^h\left(\hat{\mathbf{A}}^h\right)^T}{N_H + N_L - 1} + \mathbf{R}\right)^{-1}, \tag{25}$$

and where the hybrid anomalies are defined as:

$$\tilde{\mathbf{A}}^h = \sqrt{N_H + N_L - 1}\left[\sqrt{\frac{1-\alpha}{N_H - 1}}\mathbf{A}^f_{L\leftarrow H}, \sqrt{\frac{\alpha}{N_L - 1}}\tilde{\mathbf{A}}^f_L\right], \tag{26a}$$

$$\hat{\mathbf{A}}^h = \sqrt{N_H + N_L - 1}\left[\sqrt{\frac{1-\alpha}{N_H - 1}}\mathbf{H}\mathbf{A}^f_H, \sqrt{\frac{\alpha}{N_L - 1}}\hat{\mathbf{A}}^f_{H\leftarrow L}\right]. \tag{26b}$$

Additionally, $\mathbf{x}_{L\leftarrow H}$ and $\mathbf{A}_{L\leftarrow H}$ stand for the mean and the anomalies, respectively, of the HR ensemble upscaled to the LR space. $\hat{\mathbf{x}}^f_L$, $\hat{\mathbf{A}}^f_L$, $\tilde{\mathbf{x}}^f_{H\leftarrow L}$, and $\tilde{\mathbf{A}}^f_{H\leftarrow L}$ stand for the mean and the anomalies of the ensembles obtained by application of the operators $\mathcal{Q}$ and $\mathcal{H}$ to the LR ensemble $\mathbf{E}^f_L$:

- $\tilde{\mathbf{x}}^f_L = \overline{\mathcal{Q}\left(\mathbf{E}^f_L\right)}$, $\quad \tilde{\mathbf{A}}^f_L = \mathcal{Q}\left(\mathbf{E}^f_L\right) - \overline{\mathcal{Q}\left(\mathbf{E}^f_L\right)}$;
- $\hat{\mathbf{x}}^f_{H\leftarrow L} = \overline{\mathcal{H}\left(\mathbf{E}^f_L\right)}$, $\quad \hat{\mathbf{A}}^f_{H\leftarrow L} = \mathcal{H}\left(\mathbf{E}^f_L\right) - \overline{\mathcal{H}\left(\mathbf{E}^f_L\right)}$.

A.2 Demonstration

If the upscaling operator $\mathbf{U}$ is linear or a cubic spline interpolation, applying $\mathbf{U}$ to the system of Eqs. 12a-12b-12c-12d leads to the following set of equations:

$$\mathbf{x}^a_{L\leftarrow H} = \mathbf{x}^f_{L\leftarrow H} + \frac{\mathbf{U}\mathbf{A}^h\left(\mathbf{H}\mathbf{A}^h\right)^T}{N_H + N_L - 1}\left(\frac{\mathbf{H}\mathbf{A}^h\left(\mathbf{H}\mathbf{A}^h\right)^T}{N_H + N_L - 1} + \mathbf{R}\right)^{-1}\left(\mathbf{d} - \mathbf{H}\mathbf{x}^f_H\right), \tag{27a}$$

$$\mathbf{A}^a_{L\leftarrow H} = \mathbf{A}^f_{L\leftarrow H} - \frac{1}{2}\frac{\mathbf{U}\mathbf{A}^h\left(\mathbf{H}\mathbf{A}^h\right)^T}{N_H + N_L - 1}\left(\frac{\mathbf{H}\mathbf{A}^h\left(\mathbf{H}\mathbf{A}^h\right)^T}{N_H + N_L - 1} + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{A}^f_H, \tag{27b}$$

$$\mathbf{x}_{\mathrm{L}}^{\mathrm{a}} = \mathbf{U}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} + \frac{\mathbf{U}\mathbf{A}^{\mathrm{h}}\left(\mathbf{H}\mathbf{A}^{\mathrm{h}}\right)^{T}}{N_{\mathrm{H}}+N_{\mathrm{L}}-1}\left(\frac{\mathbf{H}\mathbf{A}^{\mathrm{h}}\left(\mathbf{H}\mathbf{A}^{\mathrm{h}}\right)^{T}}{N_{\mathrm{H}}+N_{\mathrm{L}}-1}+\mathbf{R}\right)^{-1}\left(\mathbf{d}-\mathbf{H}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}\right), \tag{27c}$$

$$\mathbf{A}_{\mathrm{L}}^{\mathrm{a}} = \mathbf{U}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} - \frac{1}{2}\frac{\mathbf{U}\mathbf{A}^{\mathrm{h}}\left(\mathbf{H}\mathbf{A}^{\mathrm{h}}\right)^{T}}{N_{\mathrm{H}}+N_{\mathrm{L}}-1}\left(\frac{\mathbf{H}\mathbf{A}^{\mathrm{h}}\left(\mathbf{H}\mathbf{A}^{\mathrm{h}}\right)^{T}}{N_{\mathrm{H}}+N_{\mathrm{L}}-1}+\mathbf{R}\right)^{-1}\mathbf{H}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}, \tag{27d}$$

where $\mathbf{x}_{\mathrm{L}\leftarrow\mathrm{H}}^{\mathrm{a}} = \mathbf{U}\mathbf{x}_{\mathrm{H}}^{\mathrm{a}}$, $\mathbf{A}_{\mathrm{L}\leftarrow\mathrm{H}}^{\mathrm{a}} = \mathbf{U}\mathbf{A}_{\mathrm{H}}^{\mathrm{a}}$, $\mathbf{x}_{\mathrm{L}\leftarrow\mathrm{H}}^{\mathrm{f}} = \mathbf{U}\mathbf{x}_{\mathrm{H}}^{\mathrm{f}}$, $\mathbf{A}_{\mathrm{L}\leftarrow\mathrm{H}}^{\mathrm{f}} = \mathbf{U}\mathbf{A}_{\mathrm{H}}^{\mathrm{f}}$, $\mathbf{x}_{\mathrm{L}}^{\mathrm{a}} = \mathbf{U}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{a}}$, $\mathbf{A}_{\mathrm{L}}^{\mathrm{a}} = \mathbf{U}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{a}}$, and $\mathbf{A}^{\mathrm{h}}$ stands for the hybrid anomalies as defined in Eq. 16.

$\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$ and $\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$ are respectively the mean and the anomalies of the LR ensemble $\mathbf{E}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$ downscaled to the HR grid, we have: $\mathbf{E}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \mathcal{D}\left(\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}\right)$. By linearity of the operators $\mathbf{U}$ and $\mathbf{H}$, we have:

$$\mathbf{U}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \overline{\mathbf{U}\mathcal{D}\left(\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}\right)}, \tag{28}$$

$$\mathbf{U}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \mathbf{U}\mathcal{D}\left(\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}\right) - \overline{\mathbf{U}\mathcal{D}\left(\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}\right)}, \tag{29}$$

$$\mathbf{H}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \overline{\mathbf{H}\mathcal{D}\left(\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}\right)}, \tag{30}$$

$$\mathbf{H}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \mathbf{H}\mathcal{D}\left(\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}\right) - \overline{\mathbf{H}\mathcal{D}\left(\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}\right)}. \tag{31}$$

Following Barthélémy et al. (2022), we define the operators $\mathcal{Q}$ and $\mathcal{H}$:

$$\mathcal{Q}: \begin{cases} \mathbb{R}^{n_{\mathrm{L}}\times N} \rightarrow & \mathbb{R}^{n_{\mathrm{L}}\times N} \\ \mathbf{E}_{\mathrm{L}} \mapsto & \tilde{\mathbf{E}}_{\mathrm{L}} = \mathbf{U}\mathcal{D}\left(\mathbf{E}_{\mathrm{L}}\right) \end{cases} \tag{32a}$$

$$\mathcal{H}: \begin{cases} \mathbb{R}^{n_{\mathrm{L}}\times N} \rightarrow & \mathbb{R}^{p\times N} \\ \mathbf{E}_{\mathrm{L}} \mapsto & \hat{\mathbf{E}}_{\mathrm{H}\leftarrow\mathrm{L}} = \mathbf{H}\mathcal{D}\left(\mathbf{E}_{\mathrm{L}}\right) \end{cases} \tag{32b}$$

$\mathcal{Q}$ is an operator that maps the LR background ensemble $\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}$ to the LR space and can be interpreted as an operator that corrects the LR model error, as a composition of the operators $\mathcal{D}$ and $\mathbf{U}$, plus some interpolation error. Note that if the LR and the HR grids are overlapping then the operator $\mathbf{U}$ acts as a sub-sampling operator and then the interpolation error is null. In Eq. 32a, $\mathcal{Q}$ is defined as the composition of the operators $\mathcal{D}$ and $\mathbf{U}$, but it could be defined as an operator that minimizes the mean absolute error between the output of the LR version of the model and the training set upscaled to the LR grid. $\mathbf{U}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$ and $\mathbf{U}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$ represent the mean and the anomalies of the corrected ensemble $\tilde{\mathbf{E}}_{\mathrm{L}}^{\mathrm{f}}$, we note:

$$\mathbf{U}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \overline{\tilde{\mathbf{E}}_{\mathrm{L}}^{\mathrm{f}}} = \tilde{\mathbf{x}}_{\mathrm{L}}^{\mathrm{f}}, \tag{33}$$

$$\mathbf{U}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \tilde{\mathbf{E}}_{\mathrm{L}}^{\mathrm{f}} - \overline{\tilde{\mathbf{E}}_{\mathrm{L}}^{\mathrm{f}}} = \tilde{\mathbf{A}}_{\mathrm{L}}^{\mathrm{f}}. \tag{34}$$

$\mathcal{H}$ is an operator that maps the LR background ensemble $\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}$ to the HR observation space and can be interpreted as a super-resolution observation operator. In Eq. 32b, $\mathcal{H}$ is defined using the operator $\mathcal{D}$, but it could also be defined by proceeding to the super-resolution of the LR ensemble $\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}$ at the observation points. $\mathbf{H}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$ and $\mathbf{H}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$ represent respectively the mean and the anomalies, at the HR observation points, of the downscaled LR ensemble $\mathbf{E}_{\mathrm{L}}^{\mathrm{f}}$. In the following, we note:

$$\mathbf{H}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \overline{\hat{\mathbf{E}}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}} = \hat{\mathbf{x}}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}, \tag{35}$$

$$\mathbf{H}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} = \hat{\mathbf{E}}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}} - \overline{\hat{\mathbf{E}}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}} = \hat{\mathbf{A}}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}. \tag{36}$$

On the other hand, $\mathbf{U}$ and $\mathbf{H}$ act by linearity on the hybrid anomalies as follows:

$$\mathbf{U}\mathbf{A}^{\mathrm{h}} = \sqrt{N_{\mathrm{H}}+N_{\mathrm{L}}-1}\left[\sqrt{\frac{1-\alpha}{N_{\mathrm{H}}-1}}\mathbf{U}\mathbf{A}_{\mathrm{H}}^{\mathrm{f}}, \sqrt{\frac{\alpha}{N_{\mathrm{L}}-1}}\mathbf{U}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}\right], \tag{37}$$

$$= \sqrt{N_{\mathrm{H}}+N_{\mathrm{L}}-1}\left[\sqrt{\frac{1-\alpha}{N_{\mathrm{H}}-1}}\mathbf{A}_{\mathrm{L}\leftarrow\mathrm{H}}^{\mathrm{f}}, \sqrt{\frac{\alpha}{N_{\mathrm{L}}-1}}\tilde{\mathbf{A}}_{\mathrm{L}}^{\mathrm{f}}\right], \tag{38}$$

$$\mathbf{H}\mathbf{A}^{\mathrm{h}} = \sqrt{N_{\mathrm{H}}+N_{\mathrm{L}}-1}\left[\sqrt{\frac{1-\alpha}{N_{\mathrm{H}}-1}}\mathbf{H}\mathbf{A}_{\mathrm{H}}^{\mathrm{f}}, \sqrt{\frac{\alpha}{N_{\mathrm{L}}-1}}\mathbf{H}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}\right], \tag{39}$$

$$= \sqrt{N_{\mathrm{H}}+N_{\mathrm{L}}-1}\left[\sqrt{\frac{1-\alpha}{N_{\mathrm{H}}-1}}\mathbf{H}\mathbf{A}_{\mathrm{H}}^{\mathrm{f}}, \sqrt{\frac{\alpha}{N_{\mathrm{L}}-1}}\hat{\mathbf{A}}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}\right]. \tag{40}$$

For the sake of simplicity, we note:

$$\mathbf{U}\mathbf{A}^{\mathrm{h}} = \tilde{\mathbf{A}}^{\mathrm{h}} \tag{41}$$

$$\mathbf{H}\mathbf{A}^{\mathrm{h}} = \hat{\mathbf{A}}^{\mathrm{h}} \tag{42}$$

Replacing the terms $\mathbf{U}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$, $\mathbf{U}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$, $\mathbf{H}\mathbf{x}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$, $\mathbf{H}\mathbf{A}_{\mathrm{H}\leftarrow\mathrm{L}}^{\mathrm{f}}$, $\mathbf{U}\mathbf{A}^{\mathrm{h}}$, and $\mathbf{H}\mathbf{A}^{\mathrm{h}}$ in Eqs. 27a-27b-27c-27d, we get Eqs. 24a-24b-24c-24d.

Acknowledgements The authors want to express their sincere gratitude to the reviewers for their work, which helped improve the quality of the manuscript. Julien Brajard is also associate professor at Sorbonnes Université.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by S. B., F. C. and J. B.. The first draft of the manuscript was written by S. B. and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by University of Bergen (incl Haukeland University Hospital). This study was supported by Horizon Europe (#101081555, #101081273, and #101093293), the Research Council of Norway (Grant No. 301396) the Trond Mohn Foundation under project number: BFS2018TMT01. This work has also received a grant for computer time from the Norwegian Program for super-

computing (NOTUR2, project number NN9039K) and a storage Grant (NORSTORE, NS9039K).

Data Availibility The simulation results that support the findings of this study are available on Zenodo from: https://doi.org/10.5281/zenodo.11235146

Declarations

Competing Interests The authors declare no competing interests

Ethical Approval Not applicable.

Graphic program The figures were made using Matplotlib version 3.5.2 [Software], Hunter (2007), available from: https://doi.org/10.5281/zenodo.6982547

References

Anderson JL (2007) An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus, Series A: Dynamic Meteorology and Oceanography 59:210–224. https://doi.org/10.1111/j.1600-0870.2006.00216.x

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

Attia A, Sandu A (2019) DATeS : a highly extensible data assimilation testing suite v1 . 0. Geoscientific Model Develop 12(2):629–649

Barthélémy S, Brajard J, Bertino L, Counillon F (2022) Super-resolution data assimilation. Ocean Dynamics 72(8):661–678

Barthélémy S, Counillon F, Wang Y (2024) Adaptive covariance hybridization for the assimilation of sst observations within a coupled earth system reanalysis. J Adv Model Earth Syst 16(6):e2023MS003888

Bauer Hs, Schwitalla T, Wulfmeyer V, Ehret U, Neuper M, Caumont O (2015) Quantitative precipitation estimation based on high-resolution numerical weather prediction and data assimilation with WRF - a performance test. Tellus, Series A: dynamic Meteorology and Oceanography 0870(67.1):25047, https://doi.org/10.3402/tellusa.v67.25047

Bethke I, Wang Y, Counillon F, Keenlyside N, Kimmritz M, Fransner F, Samuelsen A, Langehaug H, Svendsen L, Chiu PG et al (2021) NorCPM1 and its contribution to CMIP6 DCPP. Geoscientific Model Development 14(11):7073–7116

Buehner M, Houtekamer P, Charette C, Mitchell HL, He B (2010) Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: description and single-observation experiments. Monthly Weather Rev 138(5):1550–1566

Carrassi A, Bocquet M, Bertino L, Evensen G (2018) Data assimilation in the geosciences: an overview of methods, issues, and perspectives. Wiley Interdisciplinary Reviews: climate change 9(5):e535

Chambers D, Ries J, Urban T (2003) Calibration and verification of Jason-1 using global along-track residuals with TOPEX special issue: jason-1 calibration/validation. Marine Geodesy 26(3–4):305–317

Counillon F, Sakov P, Bertino L (2009) Application of a hybrid EnKF-OI to ocean forecasting. Ocean Sci 5(4):389–401. https://doi.org/10.5194/os-5-389-2009

Counillon F, Keenlyside N, Bethke I, Wang Y, Billeau S, Shen ML, Bentsen M (2016) Flow-dependent assimilation of sea surface temperature in isopycnal coordinates with the Norwegian Climate Prediction Model. Tellus, Series A: dynamic meteorology and oceanography 68(1). https://doi.org/10.3402/tellusa.v68.32437

Counillon F, Keenlyside N, Wang S, Devilliers M, Gupta A, Koseki S, Shen ML (2023) Framework for an ocean-connected supermodel of the earth system. J Adv Model Earth Syst 15(3):e2022MS003310

Dubinkina S (2013) Relevance of conservative numerical schemes for an Ensemble Kalman Filter. Q J Royal Meteorological Soc 10(2013):1–10

El Gharamti M (2021) Hybrid Ensemble-Variational Filter: a spatially and temporally varying adaptive algorithm to estimate relative weighting. Monthly Weather Rev 149(1):65–76

Evensen G (2003) The ensemble kalman filter: Theoretical formulation and practical implementation. Ocean Dynamics 53(4):343–367. https://doi.org/10.1007/s10236-003-0036-9

Fortin V, Abaza M, Anctil F, Turcotte R (2014) Why should ensemble spread match the RMSE of the ensemble mean? J Hydrometeorol 15(2010):1708–1714. https://doi.org/10.1175/JHM-D-14-0008.1

Gao J, Xue M (2008) An efficient dual-resolution approach for ensemble data assimilation and tests with simulated doppler radar data. Monthly Weather Rev 136(3):945–963. https://doi.org/10.1175/2007MWR2120.1, http://journals.ametsoc.org/doi/abs/10.1175/2007MWR2120.1

Gaspari G, Cohn SE (1999) Construction of correlation functions in two and three dimensions. Tech Rep February 1996

Gilbert RC, Trafalis TB, Richman MB, Leslie LM, Trafalis TB, Richman MB, Leslie LMA (2017) A data-driven kernel method assimilation technique for geophysical modelling. Optimization Methods Softw 32(2):237–249. https://doi.org/10.1080/10556788.2016.1257616

Hamill TM, Snyder C (2000) A hybrid ensemble kalman filter-3D variational analysis scheme. Monthly Weather Rev 128(8):2905–2919. https://doi.org/10.1175/1520-0493(2000)128<2905:AHEKFV>2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282000%29128%3C2905%3AAHEKFV%3E2.0.CO%3B2

Hamill TM, Whitaker JS, Snyder C (2001) Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. Monthly Weather Rev 129:2776–2790. https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2

Hunt BR, Kostelich EJ, Szunyogh I (2007) Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. Physica D: Nonlinear Phenomena 230:112–126. https://doi.org/10.1016/j.physd.2006.11.008 0511236

Hunter JD (2007) Matplotlib: a 2d graphics environment. Comput Sci Eng 9(3):90–95. https://doi.org/10.1109/MCSE.2007.55

Janjić T, Bormann N, Bocquet M, Carton JA, Cohn SE, Dance SL, Losa SN, Nichols NK, Potthast R, Waller JA, Weston P (2018) On the representation error in data assimilation. Q J Royal Meteorological Soc 144(713):1257–1278. https://doi.org/10.1002/qj.3130

Jelloul MB, Huck T (2003) Basin-Mode Interactions and Selection by the Mean Flow in a Reduced-Gravity Quasigeostrophic Model. J Physical Oceanography 33(11):2320–2332

Kleist DT, Ide K (2015) An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. Monthly Weather Rev 143(2):452–470

Lam R, Sanchez-Gonzalez A, Willson M, Wirnsberger P, Fortunato M, Alet F, Ravuri S, Ewalds T, Eaton-Rosen Z, Hu W et al (2023) Learning skillful medium-range global weather forecasting. Science 382(6677):1416–1421

Le Traon PY, Ogor F (1998) ERS-1/2 orbit improvement using TOPEX/POSEIDON: the 2 cm challenge. J Geophysical Res: oceans 103(C4):8045–8057

Li H, Kalnay E, Miyoshi T, Danforth CM (2009) Accounting for model errors in ensemble data assimilation. Monthly Weather Rev 137(10):3407–3419. https://doi.org/10.1175/2009mwr2766.1

Lim B, Son S, Kim H, Nah S, Mu Lee K (2017) Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops

Lorenz EN (1996) Predictability: a problem partly solved. In: Proc. Seminar on predictability, vol 1

Ménétrier B, Auligné T (2015) Optimized localization and hybridization to filter ensemble-based covariances. Monthly Weather Rev 143(10):3931–3947. https://doi.org/10.1175/mwr-d-15-0057.1

Pannekoucke O, Berre L, Desroziers G (2007) Filtering properties of wavelets for local background-error correlations. Quarterly journal of the royal meteorological society: a journal of the atmospheric sciences, applied meteorology and physical oceanography 133(623):363–379

Rainwater S, Hunt B (2013) Mixed-resolution ensemble data assimilation. Monthly Weather Rev 141(9):3007–3021. https://doi.org/10.1175/mwr-d-12-00234.1

Sakov P, Bertino L (2011) Relation between two common localisation methods for the EnKF. Computational Geosciences 15(2):225–237. https://doi.org/10.1007/s10596-010-9202-6

Sakov P, Counillon F, Bertino L, Lisæter K, Oke P, Korablev A (2012) TOPAZ4: an ocean-sea ice data assimilation system for the North Atlantic and Arctic. Ocean Science 8(4):633–656

Sakov P, Oke PR (2008) A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters. Tellus, Series A: dynamic meteorology and oceanography 60 A(2):361–371, https://doi.org/10.1111/j.1600-0870.2007.00299.x

Xie J, Bertino L, Counillon F, Lisæter KA, Sakov P (2017) Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013. Ocean Sci 13(1):123–144

Yasuda Y, Onishi R (2023) Spatio-temporal super-resolution data assimilation (srda) utilizing deep neural networks with domain generalization. J Adv Model Earth Syst 15(11):e2023MS003658