



# Super-resolution data assimilation

Sébastien Barthélémy<sup>1</sup> · Julien Brajard<sup>2</sup> · Laurent Bertino<sup>2</sup> · François Counillon<sup>1,3</sup>

Received: 22 July 2021 / Accepted: 14 July 2022 / Published online: 11 August 2022  
© The Author(s) 2022

## Abstract

Increasing model resolution can improve the performance of a data assimilation system because it reduces model error, the system can more optimally use high-resolution observations, and with an ensemble data assimilation method the forecast error covariances are improved. However, increasing the resolution scales with a cubical increase of the computational costs. A method that can more effectively improve performance is introduced here. The novel approach called “Super-resolution data assimilation” (SRDA) is inspired from super-resolution image processing techniques and brought to the data assimilation context. Starting from a low-resolution forecast, a neural network (NN) emulates the fields to high-resolution, assimilates high-resolution observations, and scales it back up to the original resolution for running the next model step. The SRDA is tested with a quasi-geostrophic model in an idealized twin experiment for configurations where the model resolution is twice and four times lower than the reference solution from which pseudo-observations are extracted. The assimilation is performed with an Ensemble Kalman Filter. We show that SRDA outperforms both the low-resolution data assimilation approach and a version of SRDA with cubic spline interpolation instead of NN. The NN’s ability to anticipate the systematic differences between low- and high-resolution model dynamics explains the enhanced performance, in particular by correcting the difference of propagation speed of eddies. With a 25-member ensemble at low resolution, the SRDA computational overhead is 55% and the errors reduce by 40%, making the performance very close to that of the high-resolution system (52% of error reduction) that increases the cost by 800%. The reliability of the ensemble system is not degraded by SRDA.

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

## 1 Introduction

The quality of a forecast relates to the accuracy of the initial condition and to its dynamical consistency. Data assimilation (DA) methods estimate the initial condition based on observations, a dynamical model, and statistical information. The Ensemble Kalman Filter (EnKF, Evensen (2003)) is one of such method and proceeds in two steps: a Monte Carlo model integration and a linear analysis update based on the ensemble covariance.

When designing a data assimilation system, computational resources are limited and there is a trade-off between assigning computing resources to the model and increasing the cost of the data assimilation method, Lei and Whitaker (2017). Increasing the model resolution can (better) resolve small-scale processes that are parameterized otherwise, e.g., Gent et al. (1995), mitigate model error and better reproduce mechanisms of predictability, Hewitt et al. (2017); Langehaug et al. (2019). A side advantage of reducing model error is that it improves the performance of the data assimilation

---

Responsible Editor: Responsible Editor: Pierre F.J. Lermusiaux

✉ Sébastien Barthélémy  
sebastien.barthelemy@uib.no

Julien Brajard  
julien.brajard@nersc.no

Laurent Bertino  
laurent.bertino@nersc.no

François Counillon  
francois.counillon@nersc.no

<sup>1</sup> Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, Norway

<sup>2</sup> Nansen Environmental and Remote Sensing Center, Bergen, Norway

<sup>3</sup> Nansen Environmental and Remote Sensing Center and Bjerknes Centre for Climate Research, Bergen, Norway

for ensemble methods that rely heavily on the accuracy of the forecast error covariances, Counillon et al. (2021).

Another important added value of increasing resolution is that it makes a better use of high-resolution observations and removes the representativity errors, Janjić et al. (2018). Indeed, several observations (and even more in the upcoming future) are available at resolutions that are typically higher than that of basin-scale models. For example, several satellites (Sentinel-3, SUOMI, AVHRR, etc.) can provide SST observations well below 1-km resolution and the most recent altimeter L2 data (SARAL/Altika, Sentinel-3A, and CryoSat-2) have an along-track resolution below 7 km, Veron et al. (2015); Schneider et al. (2017). As a comparison, a few forced global data assimilation ocean models are run at a grid resolution of  $1/12^\circ$  (see Tonani et al. (2015); Lellouche et al. (2018); Metzger et al. (2014)), equivalent to 7.5 km at  $40^\circ\text{N}$  (with the exception of NRL at  $1/24^\circ$ , though not public yet). Ocean data assimilation is also used for climate predictions on subseasonal-to-seasonal and seasonal-to-decadal timescales. For seasonal-to-decadal predictions, Boer et al. (2016), most systems have an horizontal resolution of  $1^\circ$  in their ocean component, with few running at  $0.25^\circ$  (see e.g. table 1 in Smith et al. (2020)).

There are also pragmatic challenges with increasing resolution. First, increasing resolution of ocean models moves them into a gray zone of mixed parameterized/resolved, Hallberg (2013), that are difficult to handle. Second, the inverse cascade of kinetic energy spectrum lowers the predictability of small-scale processes, Sandery and Sakov (2017). Thus, increasing the resolution requires a higher resolution observation network, larger ensemble size, and higher frequency of assimilation to outperform lower resolution systems, Thoppil et al. (2021). Since increasing the resolution scales with a cubic computational cost, it prohibits the use of more advanced data assimilation methods that require several instances of the model, Tonani et al. (2015). The energy consumption of any operational ocean prediction (at for example submesoscale resolution) and climate simulations with models that can explicitly resolve some of the most important physical processes would be excessive. This study proposes alternate methods that can benefit from the high-resolution model and high-resolution observation at a reasonable cost.

It is not uncommon for operational or climate centers to run concurrently two or more consecutive versions of the

same system at different resolutions. The old version system (at lower resolution) being kept while a new one is being developed and its superiority thoroughly demonstrated. Some methods have tried to take advantage of the coexistence of a coarse version of the model that can afford an ensemble simulation and of a deterministic high-resolution. A study by Gao and Xue (2008) introduced a mixed-resolution data assimilation algorithm where the covariances are computed from the low-resolution dynamic ensemble, interpolated to the high-resolution grid to assimilate data within the high-resolution deterministic model. The success of the method was encouraging but only for small jumps of resolution (up to one-fourth). Rainwater and Hunt (2013) introduced a mixed-resolution scheme based on the local ensemble transform Kalman filter (LETKF, Hunt et al. (2007)), using two dynamic ensembles from high- and low-resolution versions of the Lorenz-96 model and combining linearly their covariance matrices to update both ensembles. The authors have shown that for a similar computational cost, the mixed-resolution scheme can achieve better results in terms of RMSE than the high-resolution standalone EnKF. Campin et al. (2011) coupled local fine-grained 2d models with a coarse-grained hydrostatic model. The resulting coupled model performs better than the coarse-grained hydrostatic model only, at a computational cost much lower than that of a full 3d non-hydrostatic simulation. Machine learning techniques provide a framework to explore the vast space of nonlinear combinations. Data assimilation and machine learning methods have recently been combined to benefit from both approaches: Brajard et al. (2021); Gupta and Lermusiaux (2021). Following that path, we will try to combine a technique from image processing with standard data assimilation technique.

Super-resolution schemes aim to increase the resolution of an image. These techniques have been applied lately using machine learning approaches in geosciences. For example, Rodrigues et al. (2018) used a convolutional neural network to provide high-resolution weather fields from low-resolution ones. The method was assessed over a region of South America and outperforms in terms of RMSE various linear combinations of models at different resolutions. Vandal et al. (2018) provided a generalized stacked super-resolution convolutional neural network framework for statistical downscaling of climate variables. It was tested with the super-resolution of precipitation fields over the contiguous United States and compared to other statistical downscaling methods. It was shown that this framework performed closely or better than the other selected methods, for example, in terms of daily predictability, extreme precipitation, and daily root mean square error.

In this work, we propose a data assimilation algorithm that integrates a low-resolution physical model to produce the ensemble forecast. Right before the assimilation step, the

**Table 1** Summary of different model configurations

Name	Grid point size	Time step	State size	Snapshot figure size
HR	1	1	$129 \times 129$	Fig. 1a
LR	4	2	$65 \times 65$	Fig. 1b
ULR	16	4	$33 \times 33$	Fig. 1c

ensemble is downscaled to the high-resolution space in order to perform the data assimilation step in that space and to take advantage from the high-resolution observations. The idealized experiment is set up unfavorably to the data assimilation system by using ten times stronger friction than the model that produced the observations. After the assimilation step, the ensemble is upscaled back by sub-sampling to the low-resolution space for the next model integration. A neural network and a cubic spline interpolation are used to map the low-resolution forecast to the high-resolution space. The neural network performs better than the cubic spline interpolation which demonstrates the efficiency of the approach. We emphasize that the technique is only enhancing the model output resolution (“soft” machine learning) that requires much less learning than learning a model from scratch (“hard” machine learning, Chantry et al. (2021)). If the mapping is accurate, our algorithm benefits from both a cheap model integration and a high-resolution analysis. By doing so, we aim to show that one can benefit from emulating the high-resolution version of a model from a lower resolution version so that one can improve the assimilation process. For doing so, the super-resolution data assimilation approach proposed in this study requires the availability of a free run (without data assimilation) of the high-resolution version of the model. Such long simulations are often produced to assess whether resolution is reducing model error and improving representation of observed modes of variability, Hewitt et al. (2017); Langehaug et al. (2019). As a comparison, the standard Decadal Climate Prediction Project, Boer et al. (2016), that aims to assess prediction on seasonal to decadal prediction requires about 8000 model years. Carrying such a long experiment with a high-resolution Earth System Model is out of reach, while several centers have simulations for several centuries with a high-resolution model at their disposal (e.g., HighResMIP (High Resolution Model Intercomparison Project), authname (year)).

The overview of this article is as follows: the Section 2 presents the assimilation scheme used in this study and its combination with super-resolution. Section 3 exposes the physical model used in this study, and the neural network and the training set used, as well as the setup of the data assimilation experiments. Section 4 displays the results while Section 5 discusses the results obtained. Section 6 provides the conclusions of this work together with some perspectives.

## 2 Methods

### 2.1 The deterministic ensemble Kalman filter

Let  $n$  be the model state dimension,  $\mathbf{E} \in \mathbb{R}^{n \times N}$  an ensemble of  $N$  model states  $(\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, \dots, \mathbf{X}^{(N)})$ ,  $\mathbf{x} \in \mathbb{R}^n$  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.

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

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

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 equations, the superscripts a and f stand respectively for the analyzed and forecast states of the mean and the anomalies.

The true state of the system is noted  $\mathbf{x}^t$ . Let  $p$  be the number of observations  $\mathbf{d}$ ; they are defined by Eq. 2:

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

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

In this study, the deterministic EnKF (DEnKF) introduced by Sakov and Oke (2008) is used. The DEnKF is a square-root (deterministic) formulation of the EnKF that solves the analysis without the need for perturbation of the observations. It inflates the errors by construction and is intended to perform well in operational applications where corrections are small, Sakov and Oke (2008). The system has been robustly tested in a vast range of systems, Sakov et al. (2012); Counillon et al. (2016); Bethke et al. (2021). The DEnKF uses the two steps of sequential data assimilation: a forecast step and an analysis step.

In the forecast step, each member  $i$  is integrated by the model from one assimilation cycle  $k - 1$  to the next,  $k$ :

$$\mathbf{X}_k^{f,i} = \mathcal{M}(\mathbf{X}_{k-1}^{a,i}), \quad i = 1, \dots, N, \tag{3}$$

where  $\mathcal{M}$  is an operator that stands for the model integration.

The analysis step of the DEnKF proceeds at assimilation cycle  $k$  in two stages, the update of the mean, Eq. 4a, and the update of the ensemble anomalies, Eq. 4b:

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

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

where :

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

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

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

The DEnKF method will be used in the following but will be referred to as EnKF since the algorithmic parts of SRDA are independent of the flavor of the EnKF analysis scheme.

## 2.2 Super-resolution data assimilation

The super-resolution method involves a model at two different resolutions—high (HR) and low (LR). In the following, the subscripts  $H$  and  $L$  will be used in the equations to denote if an object, matrix or vector, is in the HR or the LR space. The principle of super-resolution data assimilation is to perform the forecast step with the LR model in order to reduce its computational cost and to perform the analysis step in the HR space to benefit from HR observations.

Let  $\mathbf{E}_L$  be an ensemble of  $N$  LR model states  $(\mathbf{X}_L^{(1)}, \mathbf{X}_L^{(2)}, \dots, \mathbf{X}_L^{(N)})$ . Following Eq. 3, the forecast step from assimilation cycle  $k-1$  to assimilation  $k$  writes for each ensemble member:

$$\mathbf{X}_{L,k}^{f,i} = \mathcal{M}_L(\mathbf{X}_{L,k-1}^{a,i}), \quad i = 1, \dots, N, \quad (6)$$

where  $\mathcal{M}_L$  stands for the LR model integration.

At the end of the forecast step, every member  $\mathbf{X}_{L,k}^{f,i}$  is down-scaled from the LR to the HR grid to produce an emulated HR member  $\mathbf{X}_{H,k}^{f,i}$ :

$$\mathbf{X}_{H,k}^{f,i} = \mathcal{D}(\mathbf{X}_{L,k}^{f,i}), \quad i = 1, \dots, N, \quad (7)$$

where  $\mathcal{D}$  stands for a downscaling operator from the LR to the HR grid. In this study, two different downscaling operators are used: a cubic spline interpolation operator and a neural network (hereafter denoted NN).

The mean  $\mathbf{x}_{H,k}$  and the anomalies  $\mathbf{A}_{H,k}$  of the emulated HR ensemble at assimilation cycle  $k$  are updated based on Eqs. 4a and 4b:

$$\mathbf{x}_{H,k}^a = \mathbf{x}_{H,k}^f + \mathbf{K}_{H,k}(\mathbf{d}_{H,k} - \mathbf{H}_{H,k}\mathbf{x}_{H,k}^f), \quad (8a)$$

$$\mathbf{A}_{H,k}^a = \mathbf{A}_{H,k}^f - \frac{1}{2}\mathbf{K}_{H,k}\mathbf{H}_{H,k}\mathbf{A}_{H,k}^f, \quad (8b)$$

where  $\mathbf{K}_{H,k}$ ,  $\mathbf{H}_{H,k}$ , and  $\mathbf{d}_{H,k}$  are respectively the HR Kalman gain, the HR observation operator, and the HR observations at assimilation cycle  $k$ .

After the analysis step, every member of the HR ensemble is up-scaled from the HR grid back to the LR grid before the next forecast step:

$$\mathbf{X}_{L,k}^{a,i} = \mathcal{U}(\mathbf{X}_{H,k}^{a,i}), \quad i = 1, \dots, N, \quad (9)$$

where  $\mathcal{U}$  is the upscaling operator from the HR to the LR grid. In this study,  $\mathcal{U}$  is always a cubic spline interpolation operator.

In this study, we have used also an ultra-low-resolution version of the model, referred to as ULR (see Section 3.1), instead of the LR version. In this case, Eqs. 6, 9, and 7 still apply the same.

In the following, depending on the choice of the downscaling method, the super-resolution data assimilation scheme will be referred to as SRDA-NN or SRDA-cubic, or simply SRDA if there is no need to specify the downscaling method.

## 2.3 Rewriting the SRDA as a LR scheme

The goal of this section is to show that the SRDA, as presented in Section 2.2, can be rewritten as a LR scheme, in the sense that it provides a LR analyzed state and does not require to up-scale and down-scale the whole ensemble at each assimilation cycle.

It can be shown (see Appendix 1) that the SRDA can be formulated as follows:

$$\mathbf{x}_{L,k}^a = \tilde{\mathbf{x}}_{L,k}^f + \frac{\tilde{\mathbf{A}}_{L,k}^f \hat{\mathbf{A}}_{H,k}^f}{N-1} \left( \frac{\hat{\mathbf{A}}_{H,k}^f (\hat{\mathbf{A}}_{H,k}^f)^T}{N-1} + \mathbf{R} \right)^{-1} (\mathbf{d}_{H,k} - \tilde{\mathbf{x}}_{H,k}^f), \quad (10a)$$

$$\mathbf{A}_{L,k}^a = \tilde{\mathbf{A}}_{L,k}^f - \frac{1}{2} \frac{\tilde{\mathbf{A}}_{L,k}^f \hat{\mathbf{A}}_{H,k}^f}{N-1} \left( \frac{\hat{\mathbf{A}}_{H,k}^f (\hat{\mathbf{A}}_{H,k}^f)^T}{N-1} + \mathbf{R} \right)^{-1} \hat{\mathbf{A}}_{H,k}^f, \quad (10b)$$

where  $\tilde{\mathbf{x}}_{L,k}^f$  and  $\tilde{\mathbf{A}}_{L,k}^f$  are respectively the mean and the anomalies of a LR ensemble obtained after application of the operator  $\mathcal{Q}$  on the LR background ensemble  $\mathbf{E}_{L,k}^f$ . The operator  $\mathcal{Q}$  is defined (see Appendix 1) as the composition of a downscaling operator  $\mathcal{D}$ , from the LR to the HR grid, and the upscaling operator  $\mathcal{U}$ , from the HR to the LR grid. If the operator  $\mathcal{D}$  stands for the neural network, it corrects the LR model error and down-scales the fields to the HR grid, and the operator  $\mathcal{Q}$  (by composition of  $\mathcal{D}$  and  $\mathcal{U}$ ) stands for the correction of the LR model error. If the operator  $\mathcal{D}$  stands for the cubic spline interpolation, the operator  $\mathcal{Q}$  stands for the identity of the LR space. In the following, we will only consider the case of  $\mathcal{Q}$  as an operator that corrects the LR model error. We have:

$$\tilde{\mathbf{x}}_{L,k}^f = \overline{\mathcal{Q}(\mathbf{E}_{L,k}^f)}, \quad \tilde{\mathbf{A}}_{L,k}^f = \mathcal{Q}(\mathbf{E}_{L,k}^f) - \overline{\mathcal{Q}(\mathbf{E}_{L,k}^f)}. \quad (11)$$

Similarly,  $\tilde{\mathbf{x}}_{H,k}^f$  and  $\hat{\mathbf{A}}_{H,k}^f$  are respectively the mean and the anomalies of the LR ensemble mapped to the HR observations' space after application of a super-resolution

observation operator  $\mathcal{H}$  on the LR background ensemble  $\mathbf{E}_{L,k}^f$  (see Appendix 1). We have:

$$\hat{\mathbf{x}}_{H,k}^f = \mathcal{H}\left(\overline{\mathbf{E}_{L,k}^f}\right), \quad \hat{\mathbf{A}}_{H,k}^f = \mathcal{H}\left(\mathbf{E}_{L,k}^f\right) - \mathcal{H}\left(\overline{\mathbf{E}_{L,k}^f}\right). \quad (12)$$

Following Eqs. 10a–10b, the SRDA can be interpreted as the combination of the correction of the LR model error (with the operator  $\mathcal{Q}$ ) and the assimilation of HR observations with the super-resolution observation (operator  $\mathcal{H}$ ). As the operator  $\mathcal{Q}$  is the composition of the operator  $\mathcal{D}$  that corrects the LR model error and maps LR fields to the HR grid, and the operator  $\mathcal{U}$  that maps HR fields to the LR grid, it is equivalent to an operator that would only correct the LR model error. This is why in practice, the operator  $\mathcal{Q}$  could be defined and trained as an operator that only corrects the LR model error without the need to downscale and upscale the whole ensemble at each assimilation cycle. Thus, this formulation requires only the downscaling of the observed fields at the observation points. However, it still requires to train and use two different operators, the operator  $\mathcal{Q}$  that corrects the LR model error and the super-resolution operator  $\mathcal{H}$ . Additionally, the analysis is done in the LR space and not in the HR space. In the following, the HR formulation of the SRDA is used, except in the particular case of Fig. 13, Section 4, where we show the respective benefits of the model error correction and the super-resolution observation operator.

### 3 Models and data

#### 3.1 Quasi-geostrophic model

The version of the quasi-geostrophic (QG) model used in this study is the same as that of Sakov and Oke (2008). 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  $\psi$  is given by Eq. 13:

$$\partial_t q = -\beta\psi_x - \rho J(\psi, q) - \rho_b \zeta + \rho_h \Delta \zeta - \rho_{bh} \Delta^2 \zeta + 2\pi \sin(2\pi y), \quad (13)$$

where:

- $\zeta = \Delta\psi$  is the relative vorticity;
- $\beta$  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  $y$ );
- $\rho$  is a multiplying scalar;
- $\rho_b$  is the bottom friction;

- $\rho_h$  is the horizontal friction;
- $\rho_{bh}$  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.

We have followed the settings of Sakov and Oke (2008):  $\beta = 1$ ,  $\rho = 10^5$ ,  $\rho_b = \rho_h = 0$ ,  $\rho_{bh} = 2 \times 10^{-12}$  for the true run and  $\rho_{bh} = 2 \times 10^{-11}$  for the data assimilation experiments. In the following, we note  $\rho_{bht}$  and  $\rho_{bha}$  the bi-harmonic friction for the true run and for the data assimilation experiments respectively. The boundary condition is given by  $\psi = \Delta\psi = \Delta^2\psi = 0$ . An example of output of the QG model obtained with these settings is given in Fig. 1. For more details about the QG model, see Jelloul and Huck (2003); Sakov and Oke (2008). The QG model has been used in numerous data assimilation studies such as Sakov and Oke (2008); Dubinkina (2013); Attia and Sandu (2019); Gilbert et al. (2017); Counillon et al. (2009) among others. The QG model is also part of two data assimilation packages: EnKF-Matlab<sup>1</sup> and DAPPER<sup>2</sup> (Data Assimilation with Python: a Package for Experimental Research) and has been considered in the current configuration (observation network, assimilation frequency) as a good test case for realistic atmospheric or oceanic data assimilating system, Sakov and Oke (2008).

In this study, the QG model is used at three different resolutions that are referred to as high-resolution (HR), low-resolution (LR), and ultra-low-resolution (ULR). Since all models inevitably truncate small-scale variations, the true run is a free run of the HR version of the model with a bi-harmonic horizontal friction coefficient ( $\rho_{bht}$ ) ten times lower than that used for the ensemble ( $\rho_{bha}$ ), which makes the output of the ensemble much smoother than the free run. The higher friction also stabilizes the assimilation runs, consistently with Sakov and Oke (2008). The free run is used to generate the synthetic observations (see Section 3.4), and it is the same in every experiment. As the HR, the LR, and the ULR grids overlap, the true state can be upscaled by sub-sampling to the LR and the ULR grids to compute the scores on these grids which allows to compare the different approaches. The low- and ultra-low-resolutions are defined relatively to the high-resolution version of the model and have respectively twice as low and four times as low-resolution as the high-resolution. The three different resolutions are summarized in Table 1. In the table, the grid point size is expressed in terms of HR grid points, and the time step is expressed in terms of that of the HR model. The domain used in this study is  $129 \times 129$ . The computational cost of doubling the resolution from LR to HR results in an increase

<sup>1</sup> <https://enkf.nersc.no/Code/EnKF-Matlab/>

<sup>2</sup> <https://github.com/nansencenter/DAPPER>



**Fig. 1** Snapshot of the sea level elevation for **a** the HR model, **b** the LR model, and **c** the ULR model starting from the same initial condition and after 60,000 model time steps. The black points stand for the location of the observation points

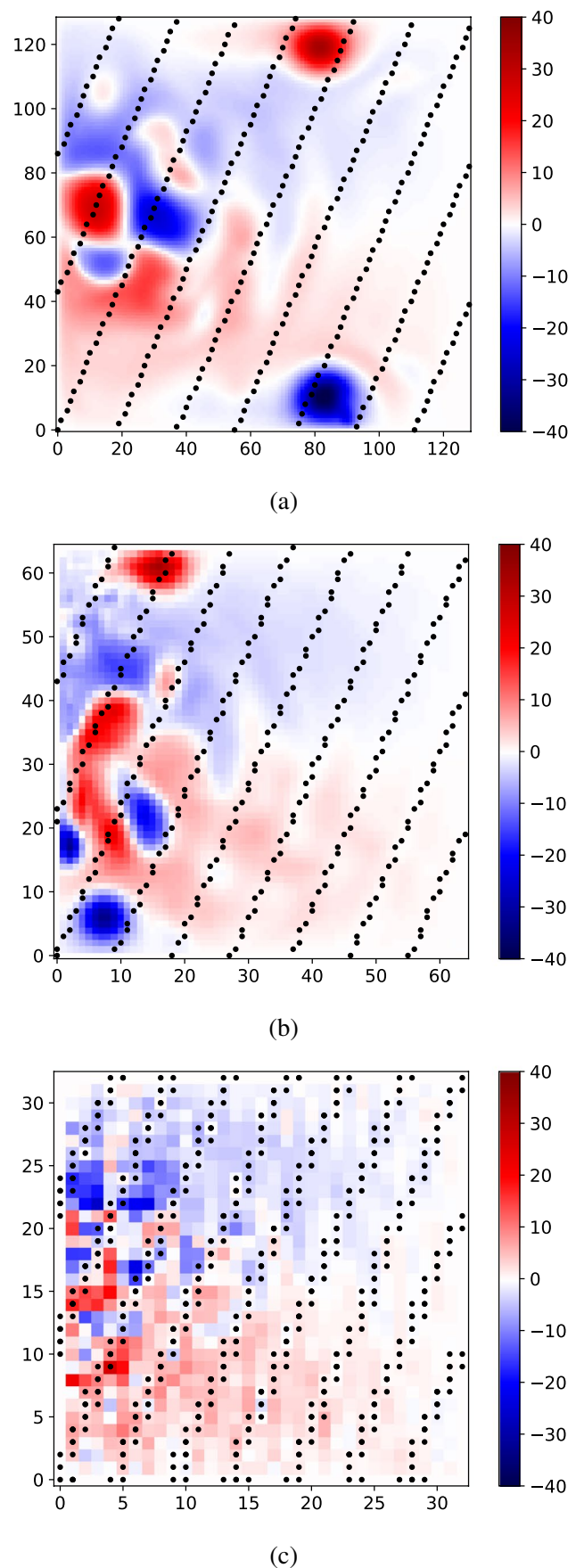
of the computational cost by a factor 8 as there are 4 times more points (in  $x$  and  $y$ ) and one needs to divide the time step by two in order to satisfy the Courant-Friedrichs-Lewy condition. In order to emphasize the differences between the 3 different resolutions of the model, we provide in Figs. 2 and 3, respectively, the plots of the mean kinetic energy (MKE) and of the eddy kinetic energy (EKE) for the three different resolutions with  $\rho_{bht} = 2 \times 10^{-11}$  (used for the data assimilation systems) but also for the high-resolution true run (with  $\rho_{bht} = 2 \times 10^{-12}$ ). In all three resolutions using  $\rho_{bht} = 2 \times 10^{-11}$ , both MKE and EKE are reduced near the northern and southern boundaries of the domain compared to the truth. The spatial pattern of MKE and EKE in the HR run (panel (b) in Figs. 2 and 3) resembles that of the truth, but is on the low side. As the resolution decreases, MKE and EKE are further reduced near the northern and southern boundaries (being absent in the ULR) but the eddy activity intensifies on the western boundary, most particularly for EKE that is getting larger than the truth (see panels c and d in Fig. 3). We can notice that some numerical instabilities start to emerge in the ultra-low-resolution.

### 3.2 Training dataset

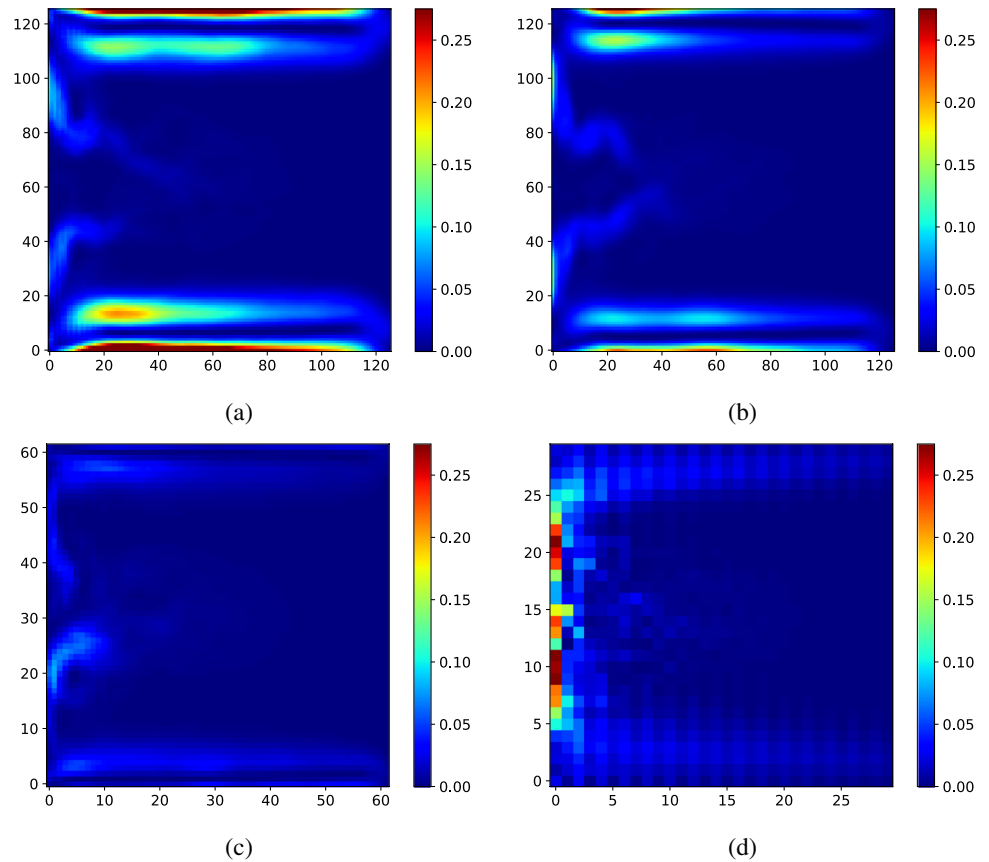
The neural network (NN) super-resolution operator  $\mathcal{D}_{HL}$  is trained using a dataset obtained from a 1,200,000 time step simulation of the HR model and with a bi-harmonic friction  $\rho_{bh} = 2.0 \times 10^{-11}$ . From this simulation,  $K$  ( $K = 10,001$  in our case) snapshots  $\mathbf{X}_{H,k}^f$  were regularly sampled every 120 time steps. At time  $t_{k-1}$ , the HR field is first scaled to low dimension using the operator  $\mathcal{U}$  as defined in Eq. 9, and then the LR model is integrated over  $\Delta_t$  ( $\Delta_t$  being the time between two assimilation cycles;  $\Delta_t = 12$  in our case), to produce a low-resolution field  $\mathbf{X}_{L,k}^f$  such as

$$\mathbf{X}_{L,k} = \mathcal{M}_L \circ \mathcal{U}(\mathbf{X}_{H,(k-1)}). \quad (14)$$

We have computed the time auto-correlation, and the correlation is reduced by 80% after 240 time steps. It implies that every second sample is assumed to be uncorrelated. The 10,000 couples  $(\mathbf{X}_{L,k}, \mathbf{X}_{H,k})$  can be used to calibrate a NN super-resolution operator. The first 8000 samples of the dataset constitute the training set, used to optimize the parameters of the neural network (see Section 3.3). To avoid correlation between the training set and the validation, 3 samples are ignored and the remaining 1997 samples are the validation set, used to evaluate the performances of the operator. The validation is thus considered independent from



**Fig. 2** Mean kinetic energy for **a** the truth, **b** the high-resolution, **c** the low-resolution, and **d** the ultra-low-resolution



the training set. We applied the same procedure to produce the dataset in the ULR case.

Note that, while this procedure aims at mimicking the low-resolution forecast members  $\mathbf{X}_{L,k}^f$  obtained in the SRDA approach, the statistical distribution of the forecast could be slightly different in the training phase as during the data assimilation. Indeed in the training/validation, the LR forecast is initialized with an upscaled HR simulation, whereas during the SRDA algorithm the forecast is initialized with an upscaled analysis. If the model contains bias, the analysis states can be statistically different from the states obtained by a HR model. If this is the case, an additional DA step could be performed to produce the training set, similarly to what is done in Brajard et al. (2021).

### 3.3 Super-resolution neural network

The neural network architecture is the enhanced deep super-resolution network (EDSR) adapted from Lim et al. (2017) in which super-resolution is applied to RGB photographic images. However, geophysical fields present differences with photographic images. First, the number of channels is not fixed to 3 (red, green, blue). In the case of the quasi-geostrophic model, there is only one channel corresponding to the sea surface elevation. Second, the pattern and texture of sea surface elevation are smoother and more homogeneous

than patterns in photographic images. For those reasons, the network chosen has been slightly simplified by specifying only one input channel and by reducing the number of degrees of freedoms (hereafter denoted *weights*) in comparison with Lim et al. (2017). A diagram of the network is presented in Fig. 4. We give hereafter a short description of all the blocks; a more complete description can be found in Lim et al. (2017).

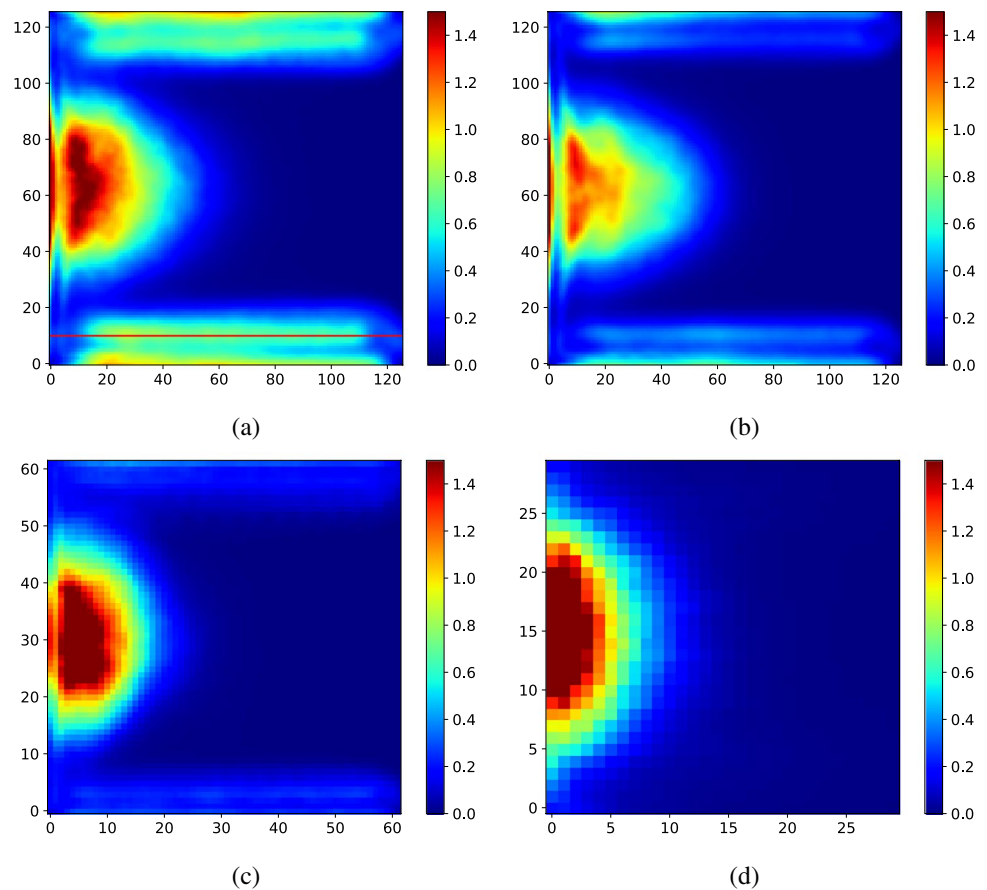
The **Scaling** block is a simple scaling function that scales the input feature images. In our case, it is a multiplicative factor of 0.04 so that the values are mainly between  $-1$  and  $1$ . There are no trainable weights in this block.

The **Conv** block is a convolutional block as introduced in LeCun et al. (1989). In our model, each convolutional block is composed of 16 filters of size  $3 \times 3$ . The weights of the filters are optimized during the training phase.

The **ReLU** block is a nonlinear function  $f$  defined by  $f(x) = \max(0, x)$  applied pointwise to each input feature. It enables the model to be non-linear.

The **ResBlock** is a non-linear convolutive block in which the input is added to the output of the block. In particular, it mitigates the gradient vanishing problem in deep learning architecture and it has been proved to be efficient in image processing problems, He et al. (2016). The weights to be optimized are the weights of the filter of the convolutional blocks.

**Fig. 3** Eddy kinetic energy for **a** the truth, **b** the high-resolution, **c** the low-resolution, and **d** the ultra-low-resolution



The **Shuffle** is where the scaling is realized in practice. The block takes as inputs  $F \times 2^2$  image features of size  $n_L \times n_L$  where  $n_L$  is the size of the low-resolution image ( $64 \times 64$  for LR and  $32 \times 32$  for ULR) and produce  $F$  image feature of size  $n_H \times n_H$  where  $n_H = 2 \times n_L$  by intertwines input features into spatial blocks. In other terms, this layer is a specific type of reshaping operator from a tensor of size  $(n_L, n_L, F \times 2^2)$  to a tensor of size  $(2n_L, 2n_L, F)$ . There are no trainable weights in this block. This procedure is detailed in Shi et al. (2016).

The total number of weights of the neural network is 22,273 for the model mapping LR to HR, and 23,361 for the model mapping ULR to HR.

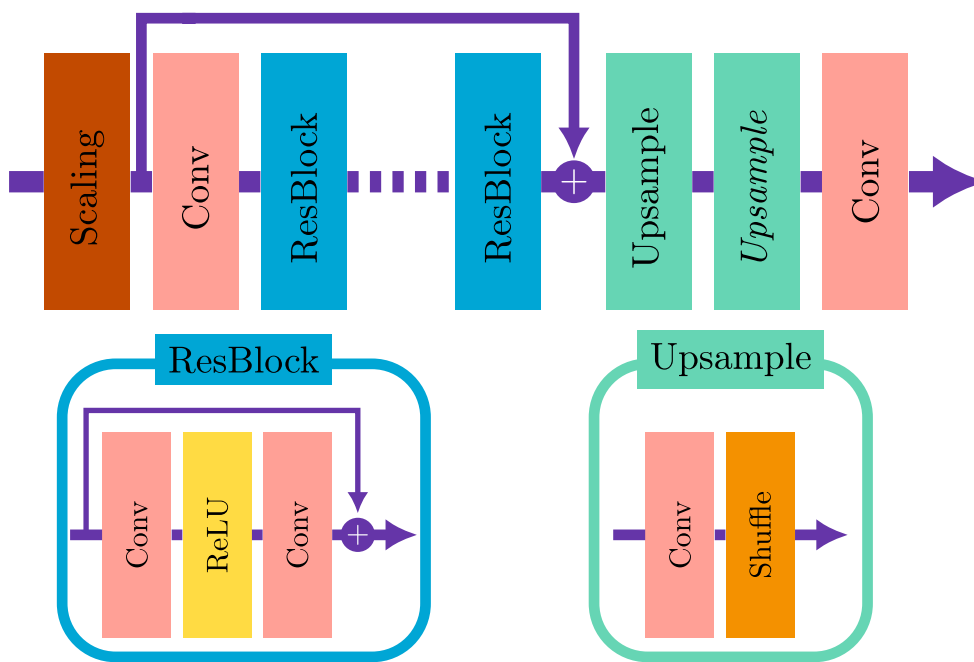
The training of the model consists of optimizing the weights of the convolutional blocks to minimize the mean absolute error between the output and the target contained in the training set. The optimizer chosen is Adam Kingma and Ba (2014), with a learning rate of  $10^{-4}$  and a batch size of 32. The training is stopped after 100 epochs (an epoch is when all the training samples have been presented to the model). After each epoch, the weights of the neural network have been modified several times. Figure 5 shows the evolution of the loss (mean absolute error) as a function of the epochs. It can be seen that the decreasing of the loss has been stabilized after 100 epochs for both models (ULR and LR) and

that errors on the validation and training sets are very close. Very little tuning of the training was performed because most of the settings were chosen from Lim et al. (2017); in particular, the validation set has not been used for an early stopping of the training. Fine-tuning would be possible and could help reduce the cost of the training and of the model computation, but it was found to be unnecessary here given the relative simplicity of the physical system considered.

### 3.4 Setup of the data assimilation experiments

We have used a twin experiment to assess the validity of the SRDA scheme. The experiment is carried over a time window of 6000 time steps with an analysis step every 12 time steps, which amounts to a total of 500 assimilation cycles. Twelve time steps correspond also to the frequency of the model outputs in the training of the neural network. The data assimilation experiments have been performed in a different setting than that of the true run. The experiment is not a perfect twin experiment as the bi-harmonic friction coefficient used for the true run,  $\rho_{bht} = 2.0 \times 10^{-12}$ , is different from that used to run the ensemble in the data assimilation experiments,  $\rho_{bha} = 2.0 \times 10^{-11}$  and it reflects on the dynamical behavior of the model (see Figs. 2 and 3). The initial ensemble at the





**Fig. 4** Diagram representing the NN model. The *Upsample* block is used only for upsample from ULR to HR, otherwise only one Upsample block is used

beginning of the simulation is independent from the initial state used in the true.

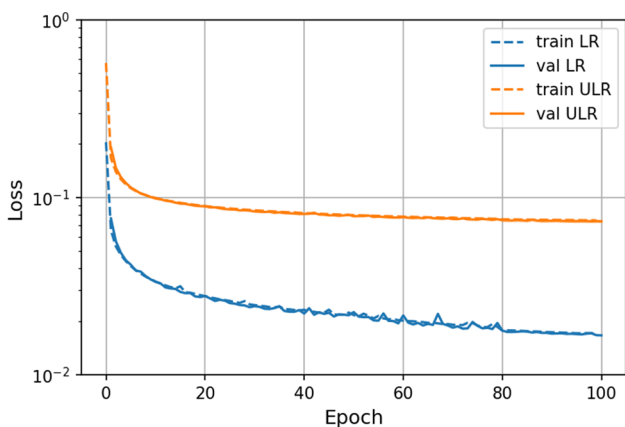
In order to evaluate the benefits from the SRDA scheme, we compare it to the standard EnKF at HR, LR, and ULR resolution (referred to hereafter as EnKF-HR, EnKF-LR, and EnKF-ULR respectively).

The synthetic observations are generated by adding a Gaussian noise with zero mean and standard deviation  $\sigma_{o,H} = 2$  to the true run. With a spacing between tracks of

400 km and an angle at the equator of  $66^\circ$ , the location of the observation points mimics altimeter satellite tracks like the reference satellites Topex-Poséidon and the Jason series. There are 300 observation points; thus, the number of observations is two orders of magnitude smaller than the size of the HR state vector. The LR (resp. ULR) observations are the same as those of the true run and their location is derived from that of the HR grid by shifting each observation point to the nearest LR (resp. ULR) grid point. In the case where two observations are shifted to the same grid point, the point with the highest initial ordinate is shifted to the point above (see Fig. 1b and c) in order to avoid two observations to overlap. This results in an increase of the observation error over the LR and the ULR grids, respectively:  $\sigma_{o,L} = 2.4$  and  $\sigma_{o,UL} = 3.7$ . The error observations are uncorrelated and the observation error matrix is diagonal,  $\mathbf{R} = \sigma_o^2 \mathbf{I}$  ( $\sigma_o$  being equal to  $\sigma_{o,H}$ ,  $\sigma_{o,L}$ , and  $\sigma_{o,UL}$  depending on the resolution of the model). For each observation point  $i$ , and each grid point  $j$ , the observation operator is defined as follows:  $\mathbf{H}_{ij} = 1$  if  $j$  corresponds to the observation point  $i$ , 0 otherwise.

In order to mitigate the impact of sampling errors from the finite size ensemble, a local analysis scheme using a Gaspari and Cohn function as a tapering function was used, Sakov and Bertino (2011). The optimal localization radius is empirically estimated for each scheme and for each ensemble size, see Section 4.1.

At each assimilation cycle  $k$ , the performance of the different assimilation schemes is estimated using the spatial



**Fig. 5** Convergence of the training of the NN models on log-scale for the loss. Solid line represents the loss of the validation set and the dashed line the loss of the training set. The blue line is the loss of the LR model and the orange line is the loss of the ULR model

root mean square error  $r_k$ , Eq. 15, the ensemble spread  $s_k$  Fortin et al. (2014), Eq. 16, and the Pearson correlation coefficient  $c_k$ , Eq. 17.

$$r_k = \sqrt{\frac{1}{n} \sum_{j=1}^n (\mathbf{x}_k^a(j) - \mathbf{x}_k^t(j))^2}, \tag{15}$$

$$s_k = \sqrt{\frac{1}{n} \sum_{j=1}^n \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_k^{a,i}(j) - \mathbf{x}_k^a(j))^2}, \tag{16}$$

$$c_k = \frac{\sum_{j=1}^n (\mathbf{x}_k^a(j) - \bar{\mathbf{x}}_k^a)(\mathbf{x}_k^t(j) - \bar{\mathbf{x}}_k^t)}{\sqrt{\sum_{j=1}^n (\mathbf{x}_k^a(j) - \bar{\mathbf{x}}_k^a)^2} \sqrt{\sum_{j=1}^n (\mathbf{x}_k^t(j) - \bar{\mathbf{x}}_k^t)^2}}, \tag{17}$$

where  $\bar{\mathbf{x}}_k^a = \frac{1}{n} \sum_{j=1}^n \mathbf{x}_k^a(j)$ ,  $\bar{\mathbf{x}}_k^t = \frac{1}{n} \sum_{j=1}^n \mathbf{x}_k^t(j)$ , and  $\mathbf{x}_k^a$  is the analyzed ensemble mean at assimilation cycle  $k$ .

The temporal means of these scores, computed over the assimilation cycles beyond cycle 10, are presented. Hence, the first 10 assimilation cycles correspond to the time needed for the system to converge to stable performance (assimilation spin up) and are not relevant.

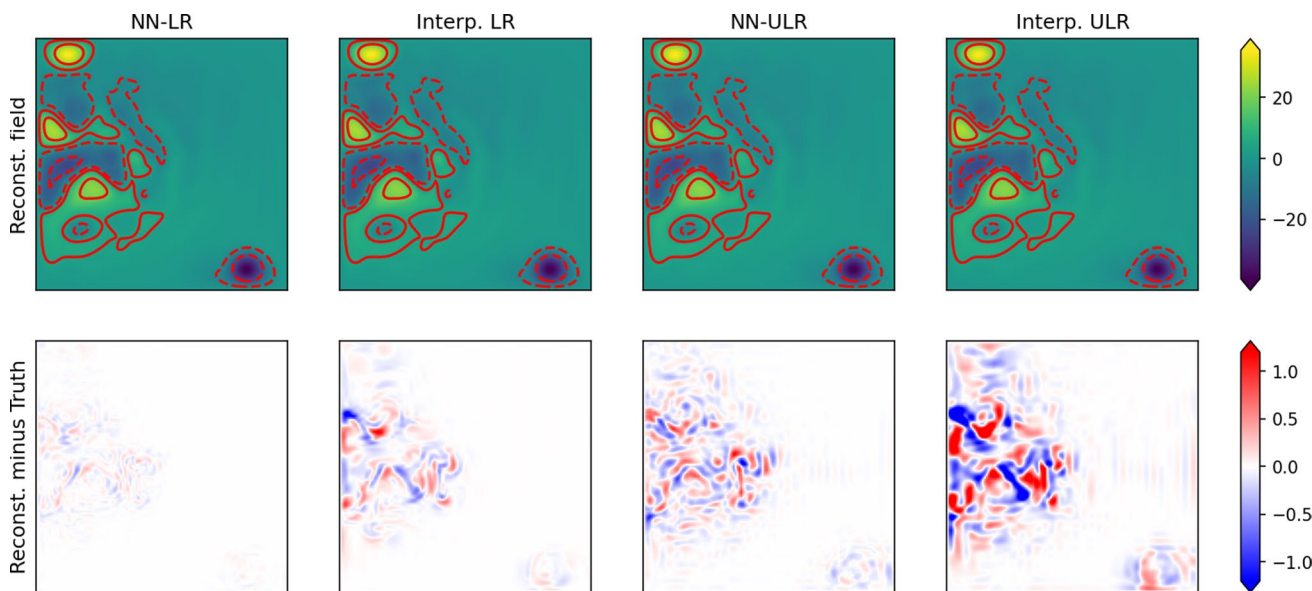
For the EnKF-HR, the SRDA-NN, and the SRDA-cubic schemes, the assimilation step is performed in the HR space. The scores  $r$ ,  $s$ , and  $c$  are thus computed in the HR space while for the EnKF-LR scheme (EnKF-ULR respectively), the scores are computed in the LR (ULR) space. As these scores are averaged over the size of the domain ( $r$  and  $s$ ) or do not depend on it ( $c$ ), it allows for a fair comparison of

the scores despite the difference in resolution of the different schemes.

## 4 Results

### 4.1 Super-resolution

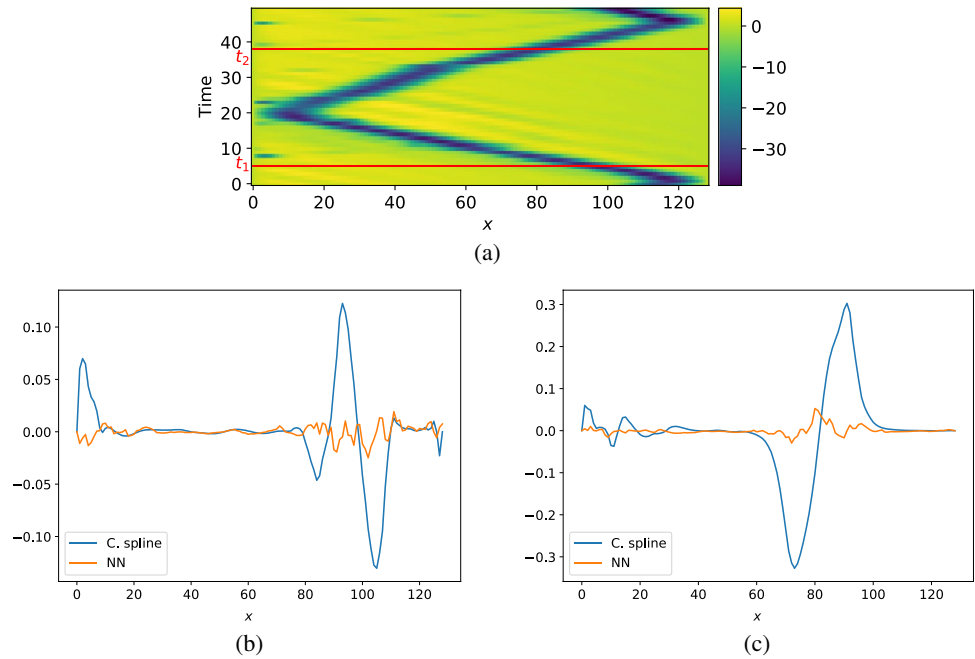
Examples of HR fields reconstructed from LR and ULR are shown in Fig. 6. The original high-resolution field has been arbitrarily chosen in the validation dataset. The reconstructions by a cubic-spline interpolation and the NN model are compared. Due to the spatial variability, the reconstructed fields are not easily distinguishable and appear very alike to the true HR field (represented in contour plot). Nevertheless, looking at the difference between the reconstructed field and the truth (second row in Fig. 6), we can see that NN reduces significantly the error of the interpolation. As expected, the error of the reconstruction is higher from ULR than from LR and the error, and it is more important in dynamical active regions. In Fig. 7a, we show the Hovmöller plot of the sea surface elevation along a fixed latitude ( $y = 10$ ). This section is a characteristic eddy pathway near the southern boundary (with EKE showing a maximum, see Fig. 3). We clearly see the signature of eddies travelling westward and eastward. With the cubic spline interpolation, there is a bimodal error pattern at the location of the eddy, which is characteristic of a displacement error—to the east of the true location (see Fig. 7b and c). It shows that NN can also mitigate some systematic biases (here eddy travelling speed) compared to cubic spline interpolation.



**Fig. 6** Snapshot (top row) of the HR-reconstructed sea level elevation from LR field and ULR field computed by the cubic spline interpolation and by the neural network. The contour (with values of  $-18$ ,  $-4$  in dotted lines and  $4$ ,  $18$  in solid lines) corresponds to the true state.

The bottom row represents the difference between the reconstructed field and the true reference field. The snapshot belongs to the validation set (not used for optimizing the NN)

**Fig. 7** **a** Hovmöller plot of sea level elevation along the southern boundary ( $y=10$ ; see red line in Fig. 3a). Error along that section at time **b**  $t_1$  and **c**  $t_2$

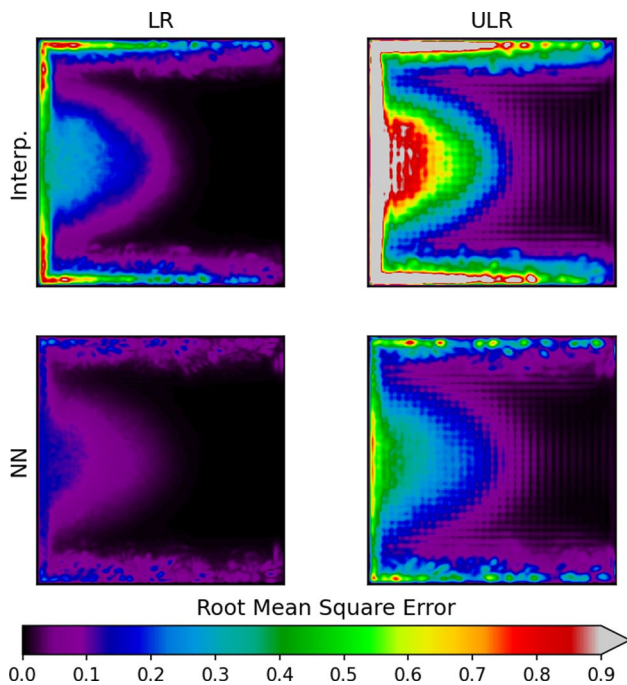


In Fig. 8 and in Table 2, the root-mean-squared error averaged over the whole validation period is presented. It confirms that there is a significant overall improvement in the reconstruction with NN. This result also confirms that the improvement is most noticeable in the high mesoscale activity region.

### 4.2 Data assimilation

The data assimilation experiments are performed with the different assimilation schemes (EnKF-HR, EnKF-LR, SRDA-NN, SRDA-cubic), and at different ensemble sizes. For each experiment, a sensitivity analysis was performed to identify the optimal inflation coefficient and localization radius (in terms of RMSE), which allows for a fair comparison of the different schemes. While the optimal inflation coefficient is roughly the same for all the schemes (not shown), the optimal localization radius of the EnKF-HR is larger than that of all the other schemes. In the following, the results exhibited are those obtained with the optimal parameters unless explicitly stated. This general sensitivity analysis was also performed for both the EnKF-LR and EnKF-ULR.

Figure 9 displays the mean RMSE of the different assimilation schemes for (a) the LR ensemble and (b) the ULR ensemble. For both ensembles, the EnKF-LR/ULR scheme (black bars) displays the worst results because of increased model error and sub-optimal use of observations. The SRDA-cubic scheme (blue bars) displays better results in terms of RMSE compared to those of the EnKF-LR/ULR with a reduction of the RMSE of approximately 18% for the LR ensemble and up to 34% for the ULR ensemble with



**Fig. 8** Root mean square error of the HR-reconstructed sea level elevation from the LR field (left column) and from the ULR field (right column) by the cubic spline interpolation (top row) and the neural network (bottom row). The error is computed on the validation set (not used for optimization)

**Table 2** Root-mean-squared error of the different super-resolution models

	LR	ULR
Interp.	$1.87 \times 10^{-1}$	$5.16 \times 10^{-1}$
NN	$5.51 \times 10^{-2}$	$1.96 \times 10^{-1}$

a 5-members ensemble. The only difference between the EnKF-LR/ULR and the SRDA-cubic occurs in the assimilation step (performed in the HR space for SRDA-cubic). It shows the benefits from assimilating the HR observations in the HR space. For example, with an ensemble size of 5 members, the EnKF-LR diverges while it converges with SRDA-cubic. The EnKF-ULR converges with 5 members but the error is very large. The SRDA-NN scheme (green bars) displays better results than the EnKF-LR/ULR and the SRDA-cubic. The neural network downscaling reduces the forecast error of the LR/ULR model and provides HR background states that are closer to the realistic HR fields than with the SRDA-cubic; see Section 4.1. This results in a better estimation of prior and as such of the analysis. The reduction of RMSE compared to the EnKF-LR/ULR is about 38–40% for the LR ensemble and ranges from 57 to 73% for the ULR ensemble. This is anticipated because the model error is larger in the ULR than in the LR case so that the NN can more efficiently improve the performance. In the particular case of the LR ensemble with an ensemble size of 5 members, the SRDA-NN, the mean RMSE is close to that of the EnKF-HR. Despite an important reduction of the LR/ULR model error, the emulated HR fields computed by the neural network are still filled with error, which results ultimately in worse results of the SRDA-NN compared to the EnKF-HR (red bars) but at a fraction of its computational cost; see Table 3. SRDA-NN shows an increased RMSE than EnKF-HR between 11 and 14% (depending on the ensemble size) for the LR ensemble, and between 28 and 49% for the ULR ensemble. For both the LR and the ULR ensembles, SRDA-NN converges for small ensembles,  $N = 5, 10$ , and increasing the ensemble size yields only slight improvements as for the EnKF-HR.

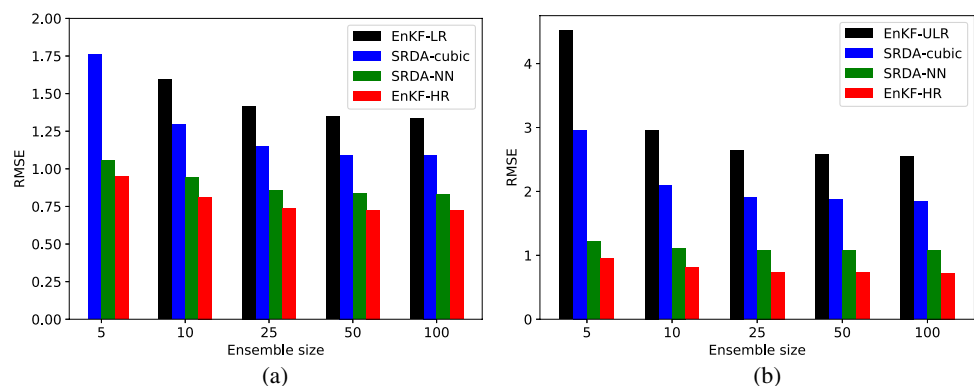
Figure 10 displays the time series of the RMSE for the different assimilation schemes with an ensemble size of  $N = 50$  for the LR ensemble (Fig. 10a) and for the ULR ensemble (Fig. 10b). At each assimilation cycle, the performance of the different schemes is well in line with the time average statistics (Fig. 9). Conclusion are very similar for other ensemble sizes (not shown). The EnKF-LR/ULR

performs worse than all the other schemes. The SRDA-cubic performs worse than the SRDA-NN while the EnKF-HR performs better than all the other schemes. The RMSE of the SRDA-NN scheme displays more stable performance than EnKF-LR/ULR and the SRDA-cubic schemes, in particular during challenging events (for example around assimilation cycles 150 or between assimilation cycles 300 and 350). This demonstrates the ability of the NN-scheme to better cope with challenging events. Hence, the reduction of RMSE of the SRDA-NN compared to the two previous schemes can reach up to 45% during challenging events (e.g., at assimilation cycle 300–350 in Fig. 10a), while the average improvement of the mean RMSE was 37% compared to the EnKF-LR and 24% compared to the SRDA-cubic with the LR ensemble.

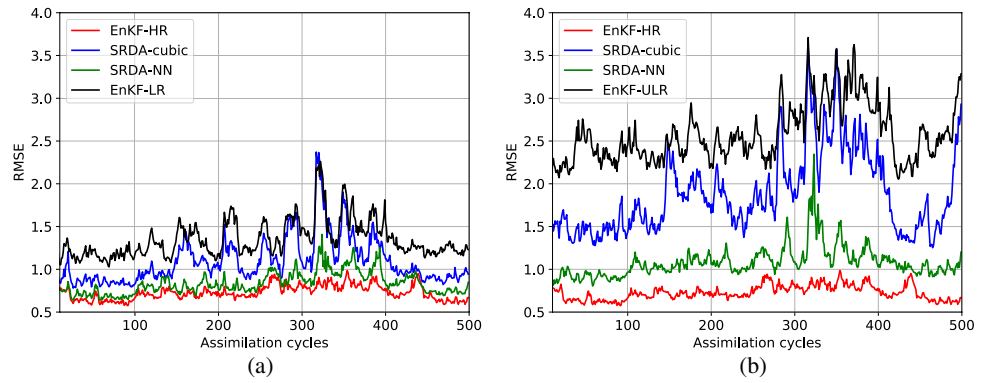
Figure 11 displays the time series of the Pearson correlation coefficient for ensemble size  $N = 50$ . The conclusions of the relative performances of the schemes remain the same as for the time series of the RMSE. The EnKF-LR/ULR and the EnKF-HR perform respectively the worst and the best while the SRDA-NN performs better than the SRDA-cubic. However, the relative improvement of the correlation is not the same as before, as the correlations of all the schemes are larger than 0.95 for the LR ensemble and 0.88 for the ULR ensemble, meaning that all the schemes perform well in terms of correlation. The temporal evolution of the correlation is as expected anti-correlated to that of RMSE. For example, between the assimilation cycles 300 and 350, there is a drop of the correlation for the EnKF-LR and the SRDA-cubic (concurrently to an increase in RMSE) schemes while it remains high for SRDA-NN (RMSE remains low).

Figure 12 displays the averaged RMSE versus the averaged ensemble spread for each ensemble size and each assimilation scheme, for the LR ensemble, Fig. 12a, and for the ULR ensemble, Fig. 12b. For each configuration (ensemble size, assimilation scheme), a point whose coordinates are the averaged RMSE and the averaged ensemble spread is plotted. In an ensemble prediction system, the ensemble spread should match the error of the ensemble mean, Fortin

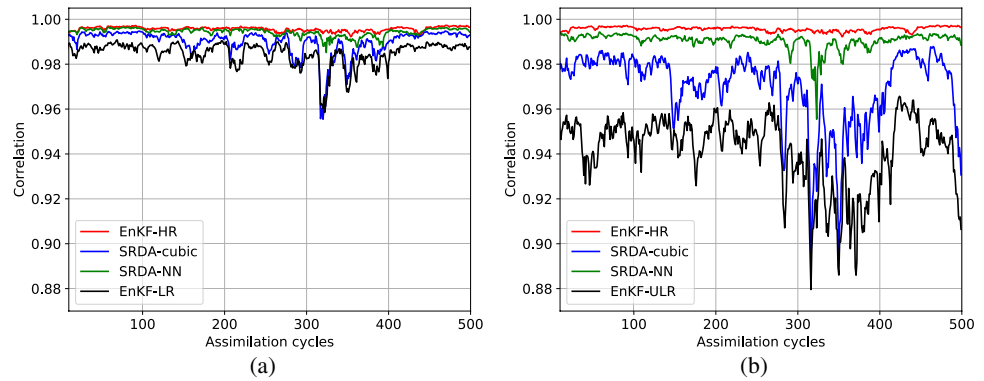
**Fig. 9** Mean RMSE for the different schemes at 5 different ensemble sizes (5, 10, 25, 50, 100) for the corresponding optimal inflation coefficient and localization radius. The EnKF-LR did not converge for ensemble size of 5 members



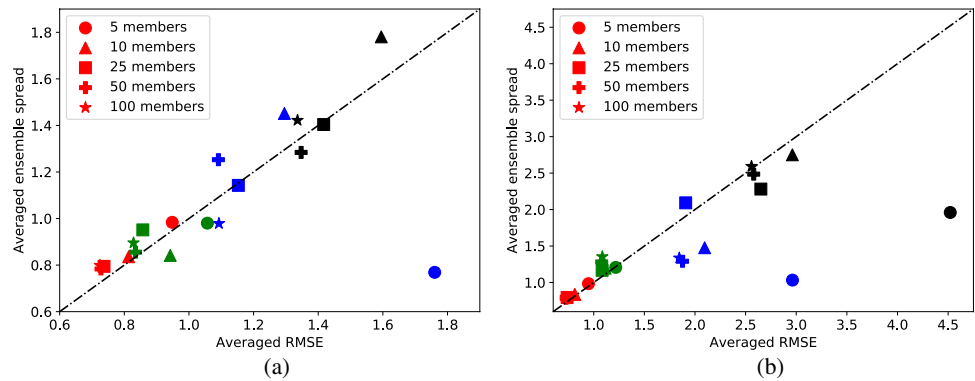
**Fig. 10** Time series of the RMSE for **a** the LR ensemble and **b** the ULR ensemble and ensemble size  $N = 50$



**Fig. 11** Time series of the Pearson correlation coefficient for **a** the LR ensemble and **b** the ULR ensemble and ensemble size  $N = 50$



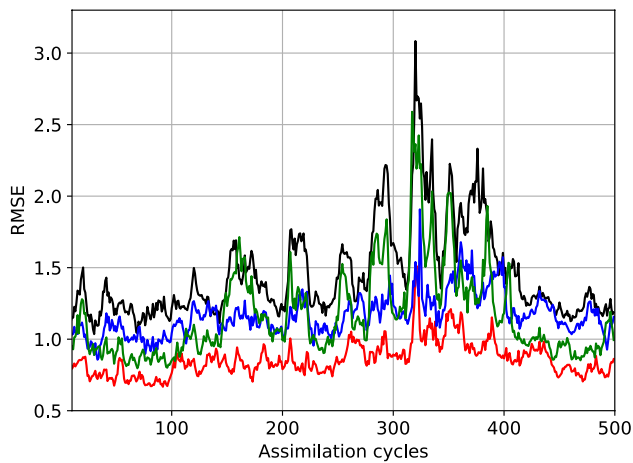
**Fig. 12** Averaged RMSE versus averaged ensemble spread for **a** the LR ensemble and **b** the ULR ensemble. The black symbols stand for the EnKF-LR, the blue ones for the SRDA-cubic, the green ones for the SRDA-NN, and the red ones for the EnKF-HR



et al. (2014); Rodwell et al. (2016), if a configuration displays a correct relation spread/error, the corresponding point should be close to the diagonal (black dashed line). For the optimal parameters, all the schemes display a good spread/error relation for the different ensemble sizes, except for the case of  $N = 5$  for the SRDA-cubic and the EnKF-LR/ULR where the spread of the ensemble is underestimated (in the sense that it is smaller than the RMSE). This is a common behavior for ensemble DA as spurious correlations are large with only five members, and it is hard to tune inflation to preserve reliability. On the contrary, the values of SRDA-NN are all on the diagonal, implying that it preserves well the reliability of the system for all ensemble sizes.

The LR version of the SRDA (Section 2.3), referred to hereafter in this paragraph as LR-SRDA, was also tested and compared to the initial formulation of the SRDA (Section 2.2). The SRDA and the LR-SRDA provided similar results (not shown). Rewriting the SRDA as a LR scheme allows estimating and comparing the respective benefits of the LR model error and the super-resolution observation operator. Figure 13 displays the time series of the RMSE with an ensemble size of 25 members, for the EnKF-LR (black line), the standard SRDA (red line), the LR-SRDA with only model error correction (blue line), and the LR-SRDA with only the super-resolution observation operator (green line). It shows that the benefits of SRDA come from





**Fig. 13** Time series of the RMSE for the EnKF-LR (black line), the SRDA (red line), Eqs. 10a–10b with only model correction (blue line) and with only the super-resolution observation operator (green line)

the combination of model error correction and the super-resolution observation operator. We note that the correction of the model error is particularly important for constraining the error during challenging events (for example between assimilation cycles 300 and 350) while the super-resolution observation operator achieves a continuous improvement over the standard EnKF but struggles during challenging events.

## 5 Discussion

In this section, we discuss the potential applicability of the method with realistic models and highlight some foreseen challenges.

The method aims at high-dimension systems as it provides accurate high-resolution analyses at a cost close to that of a low-resolution model. The costs of the algorithms are summarized in Table 3. All the experiments were done on an Intel Broadwell CPU with 64 GiB of memory except the training of the neural net that was computed on a NVIDIA V100 GPUs. The small overhead of SRDA-NN compared with the cost of SRDA-cubic is due to the NN step that is more costly than the cubic interpolation.

Note that the overall cost of SRDA is strongly reduced compared to that of the EnKF-HR because of the low cost of the LR ensemble. It should be noted that the overhead on the LR version of the system resides in the analysis step only. Thus, the additional cost depends on the respective share of model integration and the analysis step in the total cost. In the QG model, the cost of the model integration is about 80% of the total cost while for a realistic ocean

**Table 3** Cost of the different algorithm

Algorithm name	Number of model integrations	Wall clock time
EnKF-HR	I	27m20s
EnKF-LR	I/8	6m03s
SRDA-cubic (LR)	I/8	9m29s
SRDA-NN (LR)	I/8	9m32s
EnKF-ULR	I/64	2m26s
SRDA-cubic (ULR)	I/64	7m20s
SRDA-NN (ULR)	I/64	7m37s
HR simulation for training	1,200,000	2h49m12s
$\mathcal{D}_{HL}$ training (LR)	0	8min14s
$\mathcal{D}_{HL}$ training (ULR)	0	8min51s

prediction system, Sakov et al. (2012), and climate prediction system, Bethke et al. (2021), the analysis is usually much less than the total cost (2% in Sakov et al. (2012)). We thus expect that the additional cost of the SRDA would be lower for such applications. There are additional costs ahead of the assimilation, both to produce a high-resolution simulation and to train the neural network. The cost of training the neural network is generally smaller (it is only done once) and the cost of applying the NN is predictable, so it can be taken into account when designing the data assimilation system. The high-resolution run is most often already available from model performance testing and comparison experiments. The results presented here are obtained in the univariate case (the variable in the QG model defines the state unequivocally, Sakov and Oke (2008)). Using this method for more realistic setups would need a multivariate downscaling operator. In principle, this is straightforward either using the DA method, which handles well multivariate state vectors, or using an image super-resolution neural network designed for 3 output variables (R,G,B). In the latter case, the NN will need to satisfy the physical balances of multivariate models. For instance, even small differences of salinity or temperature can make the density profile unstable. To address this problem, there might be a need to add physical constraints in the neural network training (see, e.g., Beucler et al. (2021)). Note that physical constraints can also be imposed by an adequate choice of predictors specified as an input of the super-resolution operator. For example, it is expected that non-flat bathymetry can impact the downscaling procedure, so the high-resolution bathymetry should be specified as an input of the NN.

In the proposed configuration, the SRDA was applied to a high-resolution version of the same model, but we foresee that the technique could be also applied to a different model as made available from the destination Earth program,

Voosen (2020), or even directly to a high-resolution observational dataset if available.

Furthermore, we have yet only considered downscaling of the whole domain. However, the SRDA algorithm could be adapted to perform high-resolution analysis only in selected sub-regions of the domain where the error is expected to be high—e.g., in regions of high mesoscale activity or where the model is not eddy resolving, where high-resolution nested models are available and where high-resolution is available (e.g., HF radar). This potentially opens for a local version of SRDA where resolution is only locally increased. For such application, the emulator could use convolutional layers, Bocquet et al. (2019).

Finally, we have used the SRDA scheme with the EnKF data assimilation method but we foresee that it could be applied to several other sequential data assimilation methods such as 3DVAR, Tonani et al. (2015), or the EnOI, Oke et al. (2010); Evensen (2003). The EnKF has the advantage of providing flow-dependent covariance, but one can use static covariances to perform assimilation in the high-resolution space. With such methods, one does not need to perform an ensemble of low-resolution but a single instance would be sufficient. One could also combine the flow-dependent properties of the low-resolution ensemble with a high-resolution static covariance using hybrid covariance methods, Hamill and Snyder (2000); Counillon et al. (2009); Wang et al. (2007).

## 6 Conclusion

In this study, we presented a new data assimilation scheme derived from the Ensemble Kalman Filter that embeds a super-resolution neural network into a data assimilation system. This new scheme works as follows: the forecast step is performed at low resolution to limit the computational cost, but the analysis step is performed at high resolution to benefit optimally from the high-resolution observations. The scheme is called “super-resolution data assimilation” (SRDA) because the resolution of the background ensemble is enhanced to perform the assimilation step in the high-resolution space. Two different downscaling operators were studied and compared: a straightforward cubic spline interpolation operator and a more advanced convolutional neural network operator that is commonly used for enhancing photographic images. The method was tested in an idealized twin experiment with a factor of two and four in resolution.

The neural network was trained with a dataset of matching pairs between low- and high-resolution states and outperformed the cubic spline interpolation operator. In particular, the neural network was able to correct the position of the eddies of the reconstructed fields, reducing the impact of the

low-resolution model error. The downscaling performance was, as expected, better for a factor of two than a factor of four in resolution.

The downscaling performance proved important for the data assimilation performance. Both SRDA with cubic spline interpolation and with neural network perform better than the standard low-resolution EnKF. But the SRDA neural network, as it also reduces model error, performs better than its counterpart, and provides results close to those of the standard high-resolution EnKF at a much lower computational cost.

Nonetheless, this study was carried out in an idealized twin experiment framework, with a 2-dimensional monovariate model. The limitations of this method and its applicability to a more realistic model have been discussed and some tracks to overcome those limitations have been advanced.

## A: Rewriting the SRDA as a LR scheme

If the LR and the HR grids overlap, the upscaling operator  $\mathcal{U}$  is equivalent to a sub-sampling operator and is then linear. Applying  $\mathcal{U}$  to the system of Eqs. 8a–8b leads to:

$$\mathbf{x}_{L,k}^a = \mathcal{U}(\mathbf{x}_{H,k}^f) + \frac{\mathcal{U}(\mathbf{A}_{H,k}^f)(\mathbf{H}_{H,k}\mathbf{A}_{H,k}^f)^T}{N-1} \left( \frac{\mathbf{H}_{H,k}\mathbf{A}_{H,k}^f(\mathbf{H}_{H,k}\mathbf{A}_{H,k}^f)^T}{N-1} + \mathbf{R} \right)^{-1} (\mathbf{d}_{H,k} - \mathbf{H}_{H,k}\mathbf{x}_{H,k}^f), \tag{18a}$$

$$\mathbf{A}_{L,k}^a = \mathcal{U}(\mathbf{A}_{H,k}^f) - \frac{1}{2} \frac{\mathcal{U}(\mathbf{A}_{H,k}^f)(\mathbf{H}_{H,k}\mathbf{A}_{H,k}^f)^T}{N-1} \left( \frac{\mathbf{H}_{H,k}\mathbf{A}_{H,k}^f(\mathbf{H}_{H,k}\mathbf{A}_{H,k}^f)^T}{N-1} + \mathbf{R} \right)^{-1} \mathbf{H}_{H,k}\mathbf{A}_{H,k}^f, \tag{18b}$$

where  $\mathbf{x}_{L,k}^a = \mathcal{U}(\mathbf{x}_{H,k}^a)$  and  $\mathbf{A}_{L,k}^a = \mathcal{U}(\mathbf{A}_{H,k}^a)$ .

$\mathbf{x}_{H,k}^f$  and  $\mathbf{A}_{H,k}^f$  being respectively the mean and the anomalies of the downscaled background ensemble  $\mathbf{E}_{H,k}^f = \mathcal{D}(\mathbf{E}_{L,k}^f)$  and  $\mathcal{U}$  and  $\mathcal{H}$  being linear, we have:

$$\mathcal{U}(\mathbf{x}_{H,k}^f) = \overline{\mathcal{U} \circ \mathcal{D}(\mathbf{E}_{L,k}^f)}, \tag{19}$$

$$\mathcal{U}(\mathbf{A}_{H,k}^f) = \overline{\mathcal{U} \circ \mathcal{D}(\mathbf{E}_{L,k}^f)} - \overline{\mathcal{U} \circ \mathcal{D}(\mathbf{E}_{L,k}^f)}, \tag{20}$$

$$\mathbf{H}_{H,k}\mathbf{x}_{H,k}^f = \overline{\mathbf{H}_{H,k} \circ \mathcal{D}(\mathbf{E}_{L,k}^f)}, \tag{21}$$

$$\mathbf{H}_{H,k}\mathbf{A}_{H,k}^f = \overline{\mathbf{H}_{H,k} \circ \mathcal{D}(\mathbf{E}_{L,k}^f)} - \overline{\mathbf{H}_{H,k} \circ \mathcal{D}(\mathbf{E}_{L,k}^f)} \tag{22}$$

For the sake of simplicity, in the following, we define the operators  $\mathcal{Q}$  and  $\mathcal{H}$ :

$$\mathcal{Q} : \begin{cases} \mathbb{R}^{n_L \times N} & \rightarrow & \mathbb{R}^{n_L \times N} \\ \mathbf{E}_{L,k} & \mapsto & \tilde{\mathbf{E}}_{L,k} = \mathcal{U} \circ \mathcal{D}(\mathbf{E}_{L,k}) \end{cases} \quad (23a)$$

$$\mathcal{H} : \begin{cases} \mathbb{R}^{n_L \times N} & \rightarrow & \mathbb{R}^{p \times N} \\ \mathbf{E}_{L,k} & \mapsto & \hat{\mathbf{E}}_{H,k} = \mathbf{H}_{H,k} \circ \mathcal{D}(\mathbf{E}_{L,k}) \end{cases} \quad (23b)$$

$\mathcal{Q}$  is an operator that maps the LR background ensemble  $\mathbf{E}_{L,k}$  into the LR space and can be interpreted as an operator that corrects the LR model error; see Section 4.1. In Eq. 23a,  $\mathcal{Q}$  is defined as the composition of the operators  $\mathcal{U}$  and  $\mathcal{D}$ , but it could be defined as an operator that minimizes the mean absolute error between the output and the training set upscaled to the LR grid.  $\mathcal{U}(\mathbf{x}_{H,k}^f)$  and  $\mathcal{U}(\mathbf{A}_{H,k}^f)$  represent then the mean and the anomalies of the corrected ensemble  $\tilde{\mathbf{E}}_{L,k}$ . In the following, we note:

$$\mathcal{U}(\mathbf{x}_{H,k}^f) = \overline{\tilde{\mathbf{E}}_{L,k}} = \tilde{\mathbf{x}}_{L,k}^f, \quad (24)$$

$$\mathcal{U}(\mathbf{A}_{H,k}^f) = \tilde{\mathbf{E}}_{L,k} - \overline{\tilde{\mathbf{E}}_{L,k}} = \tilde{\mathbf{A}}_{L,k}^f. \quad (25)$$

$\mathcal{H}$  is an operator that maps the LR background ensemble  $\mathbf{E}_{L,k}$  into the HR observation space and can be interpreted as a super-resolution observation operator. In Eq. 23b,  $\mathcal{H}_{H,k}$  is defined using the operator  $\mathcal{D}$ , but it could also be defined by proceeding to the super-resolution of the LR ensemble  $\mathbf{E}_{L,k}$  at the observation points.  $\mathbf{H}_{H,k} \mathbf{x}_{H,k}^f$  and  $\mathbf{H}_{H,k} \mathbf{A}_{H,k}^f$  represent respectively the mean and the anomalies, at the HR observation points, of the downscaled LR ensemble  $\mathbf{E}_{L,k}$ . In the following, we note:

$$\mathbf{H}_{H,k} \mathbf{x}_{H,k}^f = \overline{\hat{\mathbf{E}}_{H,k}} = \hat{\mathbf{x}}_{H,k}^f, \quad (26)$$

$$\mathbf{H}_{H,k} \mathbf{A}_{H,k}^f = \hat{\mathbf{E}}_{H,k} - \overline{\hat{\mathbf{E}}_{H,k}} = \hat{\mathbf{A}}_{H,k}^f. \quad (27)$$

Replacing the terms  $\mathcal{U}(\mathbf{x}_{H,k}^f)$ ,  $\mathcal{U}(\mathbf{A}_{H,k}^f)$ ,  $\mathbf{H}_{H,k} \mathbf{x}_{H,k}^f$ , and  $\mathbf{H}_{H,k} \mathbf{A}_{H,k}^f$  in Eq. 18a–18b, respectively by  $\tilde{\mathbf{x}}_{L,k}^f$ ,  $\tilde{\mathbf{A}}_{L,k}^f$ ,  $\hat{\mathbf{x}}_{H,k}^f$ , and  $\hat{\mathbf{A}}_{H,k}^f$ , we get Eq. 10a–10b.

**Acknowledgements** The authors wish to thank Mao-Lin Shen (UiB) for the fruitful discussions.

**Funding** Open access funding provided by University of Bergen (incl Haukeland University Hospital) Julien Brajard and François Counillon are funded by the Norwegian Research Council (#270733 and #309562). Julien Brajard is also associate professor at Sorbonne Université. Sébastien Barthélémy received support from European Union's Horizon 2020 research and innovation program, Grant 727852, and the Trond Mohn Foundation, under project number BFS2018TMT01. This work has also received a grant for computer time from the Norwegian Program for supercomputing (NOTUR2, project number nn9039k) and a storage grant (NORSTORE, NS9039k).

**Data availability** The data that support the findings of this study are available from the corresponding author, upon reasonable request.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Attia A, Sandu A (2019) DATeS : a highly extensible data assimilation testing suite v1. 0. Geoscientific Model Development 12(2):629–649
- Bethke I, Wang Y, Counillon F, Keenlyside N, Kimmritz M, Fransner F, Samuelsen A, Langehaug H, Svendsen L, Chiu PG et al (2021) Norcpml and its contribution to cmip6 dcpp. Geoscientific Model Development 14(11):7073–7116
- Beucler T, Pritchard M, Rasp S, Ott J, Baldi P, Gentile P (2021) Enforcing analytic constraints in neural networks emulating physical systems. Physical Review Letters 126(9):098302
- Bocquet M, Brajard J, Carrassi A, Bertino L (2019) Data assimilation as a learning tool to infer ordinary differential equation representations of dynamical models. Nonlinear Processes in Geophysics 26(3):143–162
- Boer GJ, Smith DM, Cassou C, Doblas-reyes F, Danabasoglu G, Kirtman B, Kushnir Y, Kimoto M, Meehl GA, Msadek R, Mueller WA (2016) The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. Geoscientific Model Development 9:3751–3777. <https://doi.org/10.5194/gmd-9-3751-2016>
- Brajard J, Carrassi A, Bocquet M, Bertino L (2021) Combining data assimilation and machine learning to infer unresolved scale parametrization. Philosophical Transactions of the Royal Society A 379(2194):20200086
- Campin JM, Hill C, Jones H, Marshall J (2011) Super-parameterization in ocean modeling: Application to deep convection. Ocean Modelling 36(1–2):90–101
- Chantry M, Christensen H, Dueben P, Palmer T (2021) Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI. Philosophical Transactions of the Royal Society A 379(2194):20200083
- Counillon F, Sakov P, Bertino L (2009) Application of a hybrid EnKF-OI to ocean forecasting. Ocean Science 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, Toniazzo T, Koseki S, Demissie T, Bethke I, Wang Y (2021) Relating model bias and prediction skill in the equatorial atlantic. *Climate Dynamics* 56(7):2617–2630
- Dubinkina S (2013) Relevance of conservative numerical schemes for an Ensemble Kalman Filter. *Quarterly Journal of the Royal Meteorological Society* 10(2013):1–10
- 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, Ancitil F, Turcotte R (2014) Why Should Ensemble Spread Match the RMSE of the Ensemble Mean? *Journal of Hydrometeorology* 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 Review* 136(3):945–963. <https://doi.org/10.1175/2007MWR2120.1>
- Gent PR, Willebrand J, McDougall TJ, McWilliams JC (1995) Parameterizing Eddy-Induced Tracer Transports in Ocean Circulation Models. *Journal of physical oceanography* 25(4):463–474
- 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 and software* 32(2):237–249. <https://doi.org/10.1080/10556788.2016.1257616>
- Gupta A, Lermusiaux PF (2021) Neural closure models for dynamical systems. *Proceedings of the Royal Society A* 477(2252):20201004
- Haarsma RJ, Roberts MJ, Vidale PL, Senior CA, Bellucci A, Bao Q, Chang P, Corti S, Fučkar NS, Guemas V, et al. (2016) High resolution model intercomparison project (highresmip v1. 0) for cmip6. *Geoscientific Model Development* 9(11):4185–4208
- Hallberg R (2013) Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects. *Ocean Modelling* 72:92–103. <https://doi.org/10.1016/j.ocemod.2013.08.007>
- Hamill TM, Snyder C (2000) A Hybrid Ensemble Kalman Filter-3D Variational Analysis Scheme. *Monthly Weather Review* 128(8):2905–2919. [https://doi.org/10.1175/1520-0493\(2000\)128<3c2905:AHEKFFV>3e2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<3c2905:AHEKFFV>3e2.0.CO;2)
- He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
- Hewitt HT, Bell MJ, Chassignet EP, Czaja A, Ferreira D, Gri SM, Hyder P, Mcclean JL, New AL, Roberts MJ (2017) Will high-resolution global ocean models benefit coupled predictions on short-range to climate time scales? *Ocean Modelling* 120(July):120–136. <https://doi.org/10.1016/j.ocemod.2017.11.002>
- 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):0511236. <https://doi.org/10.1016/j.physd.2006.11.008>
- 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. *Quarterly Journal of the Royal Meteorological Society* 144(713):1257–1278. <https://doi.org/10.1002/qj.3130>
- Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*
- Langehaug HR, Sandø AB, Årthun M, Ilicak M (2019) Variability along the Atlantic water pathway in the forced Norwegian Earth System Model. *Climate Dynamics* 52(1):1211–1230. <https://doi.org/10.1007/s00382-018-4184-5>
- LeCun Y et al (1989) Generalization and network design strategies. *Connectionism in perspective* 19:143–155
- Lei L, Whitaker JS (2017) Journal of Advances in Modeling Earth Systems. *Journal of Advances in Modeling Earth Systems* 9:781–789
- Lellouche JM, Greiner E, Le Galloudec O, Garric G, Regnier C, Drevillon M, Benkiran M, Testut CE, Bourdalle-Badie R, Gasparin F et al (2018) Recent updates to the copernicus marine service global ocean monitoring and forecasting real-time 1/ 12° high-resolution system. *Ocean Science* 14(5):1093–1126
- 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*
- Jelloul MB, Huck T (2003) Basin-Mode Interactions and Selection by the Mean Flow in a Reduced-Gravity Quasigeostrophic Model. *Journal of physical oceanography* 33(11):2320–2332
- Metzger EJ, Smedstad OM, Thoppil PG, Hurlburt HE, Cummings JA, Wallcraft AJ, Zamudio L, Franklin DS, Posey PG, Phelps MW et al (2014) Us navy operational global ocean and arctic ice prediction systems. *Oceanography* 27(3):32–43
- Oke PR, Brassington GB, Griffin DA, Schiller A (2010) Ocean data assimilation: a case for ensemble optimal interpolation. *Australian Meteorological and Oceanographic Journal* 59(1SP):67–76
- Rainwater S, Hunt B (2013) Mixed-Resolution Ensemble Data Assimilation. *Monthly Weather Review* 141(9):3007–3021. <https://doi.org/10.1175/mwr-d-12-00234.1>
- Rodrigues ER, Oliveira I, Cunha RLF, Netto MAS (2018) DeepDownscale : a deep learning strategy for high-resolution weather forecast. *IEEE 14th International Conference on e-Science (e-Science)*. <https://doi.org/10.1109/eScience.2018.00130>
- Rodwell MJ, Lang STK, Ingleby NB, Bormann N, Holm E, Rabier F, Richardson DS, Yamaguchi M (2016) Reliability in ensemble data assimilation. *Quarterly Journal of the Royal Meteorological Society* 142(January):443–454. <https://doi.org/10.1002/qj.2663>
- 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, 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>
- Sakov P, Counillon F, Bertino L, Lisæter K, Oke P, Korabely A (2012) Topaz4: an ocean-sea ice data assimilation system for the north atlantic and arctic. *Ocean Science* 8(4):633–656
- Sandery PA, Sakov P (2017) Ocean forecasting of mesoscale features can deteriorate by increasing model resolution towards the sub-mesoscale. *Nature Communications* 8(1):1–8. <https://doi.org/10.1038/s41467-017-01595-0>
- Schneider R, Godiksen PN, Villadsen H, Madsen H, Bauer-Gottwein P (2017) Application of cryosat-2 altimetry data for river analysis and modelling. *Hydrology and Earth System Sciences* 21(2):751–764
- Shi W, Caballero J, Huszár F, Totz J, Aitken AP, Bishop R, Rueckert D, Wang Z (2016) Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 1874–1883
- Smith DM, Scaife AA, Eade R, Athanasiadis P, Bellucci A, Bethke I, Bilbao R, Borchert LF, Caron L, Counillon F, Danabasoglu G, Delworth T, Dunstone NJ, Flavoni S, Hermanson L, Keenlyside N, Kharin V, Kimoto M, Merryfield WJ, Mignot J, Mochizuki T, Modali K, Monerie P, Müller WA, Nicolì D, Ortega P, Pankatz K, Pohlmann H, Robson J, Ruggieri P, Swingedouw D, Wang Y, Wild S, Yeager S, Yang X, Zhang L (2020) North Atlantic climate far more predictable than models imply. *Nature* 583(7818):796–800. <https://doi.org/10.1038/s41586-020-2525-0>

- Thoppil PG, Frolov S, Rowley CD, Reynolds CA, Jacobs GA, Metzger EJ, Hogan PJ, Barton N, Wallcraft AJ, Smedstad OM, Shriver JF (2021) prediction horizon for ocean mesoscale variability. *Communications Earth & Environment* 2(1):1–9. <https://doi.org/10.1038/s43247-021-00151-5>
- Tonani M, Balmaseda M, Bertino L, Blockley E, Brassington G, Davidson F, Drillet Y, Hogan P, Kuragano T, Lee T, Mehra A, Paranathara F, Tanajura CAS, Wang H (2015) Status and future of global and regional ocean prediction systems. *J Oper Oceanogr* 8(sup2):s201–s220. <https://doi.org/10.1080/1755876X.2015.1049892>
- Vandal T, Kodra E, Ganguly S, Michaelis A, Nemani R, Ganguly AR (2018) DeepSD : Generating High Resolution Climate Change Projections through Single Image Super-Resolution. *International Joint Conferences on Artificial Intelligence Organization* pp 1663–1672
- Verron J, Sengenes P, Lambin J, Noubel J, Guillot A, Picot N, Coutin-faye S, Gairola RM, Murthy DVAR, Richman JG, Pascual A, Rémy F, Gupta PK, Verron J, Sengenes P, Lambin J, Noubel J, Guillot A, Picot N, Coutin-faye S, Sharma R, Gairola RM, Murthy DVAR, Richman JG, Griffin D, Pascual A, Rémy F, Verron J, Sengenes P, Noubel J, Steunou N, Guillot A, Picot N, Griffin D, Pascual A (2015) The SARAL / AltiKa Altimetry Satellite Mission The SARAL / AltiKa Altimetry Satellite Mission. *Marine Geodesy* 38(S1):2–21. <https://doi.org/10.1080/01490419.2014.1000471>
- Voosen P (2020) Europe builds 'digital twin' of earth to hone climate forecasts
- Wang X, Hamill TM, Whitaker JS, Bishop CH (2007) A comparison of hybrid ensemble transform kalman filter-optimum interpolation and ensemble square root filter analysis schemes. *Monthly weather review* 135(3):1055–1076