Super-resolution data assimilation

Sébastien Barthélémy[1,2], Julien Brajard[3,4], Laurent Bertino[2,3], François Counillon[1,2,3]

October 13, 2022

[1]University of Bergen,
[2]Bjerknes Centre for Climate Research,
[3]Nansen Environmental and Remote Sensing Centre,
[4]Sorbonne Université

Overview

1. Introduction
2. Super-resolution data assimilation
3. Hybrid covariance super-resolution data assimilation
4. Links between the geometrical features of P^{f} and super-resolution data assimilation
5. Conclusion and perspectives

Overview

1. Introduction
    - Objectives, motivation and method
    - Model used
    - Training and set-up of the neural network and downscaling performance
2. Super-resolution data assimilation
3. Hybrid covariance super-resolution data assimilation
4. Links between the geometrical features of $\mathbf{P}^{\mathrm{f}}$ and super-resolution data assimilation
5. Conclusion and perspectives

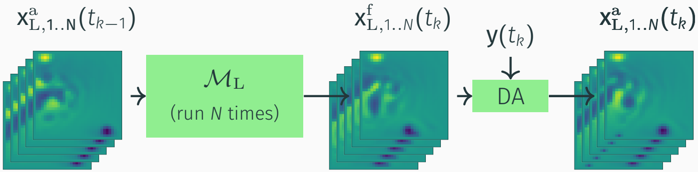
Objectives

The resolution of observations grows faster than model resolution.

1. Emulating a HR EnKF while running the forecast step with a LR model
2. Reduction of the computational cost of the EnKF
3. Taking advantage of HR observations with a LR model

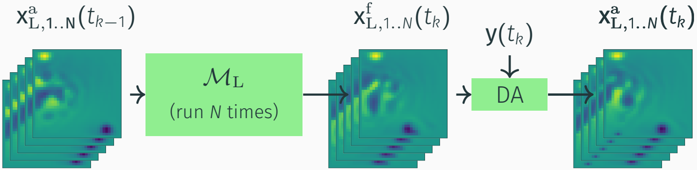
Motivation and method

EnKF - Low Resolution (EnKF-LR)



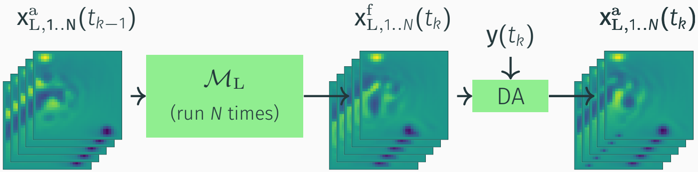
Motivation and method

EnKF - Low Resolution (EnKF-LR)

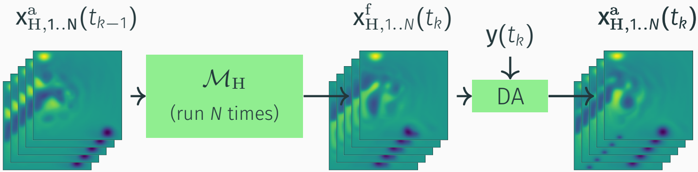


	EnKF-LR		
Observation error	High✔		
High-resolution processes	Poorly resolved✔		
Computational cost	Low✔		
Ensemble size	Big✔		

EnKF - Low Resolution (EnKF-LR)



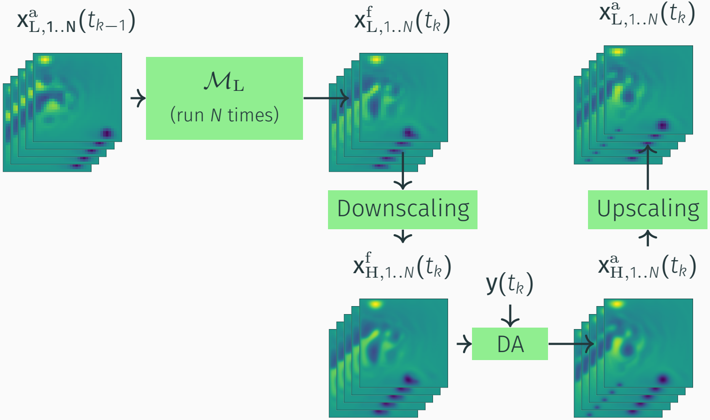
EnKF - High Resolution (EnKF-HR)



	EnKF-LR	EnKF-HR	
Observation error	High✘	Low✔	
High-resolution processes	Poorly resolved✘	Resolved✔	
Computational cost	Low✔	High, $\mathcal{O}(n^3)$✘	
Ensemble size	Big✔	Small✘	

Motivation and method

EnKF - Super-resolution data assimilation (SRDA)



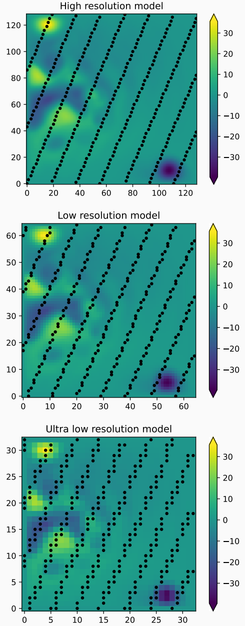
	EnKF-LR	EnKF-HR	SRDA
Observation error	High✘	Low✔	Low✔
High-resolution processes	Poorly resolved✘	Resolved✔	Emulated✔
Computational cost	Low✔	High, $\mathcal{O}(n^3)$✘	Low✔
Ensemble size	Big✔	Small✘	Big✔

Model used

- Model used: Quasi-geostrophic model [Sakov and Oke, 2008]

Configuration	State size	Cost
HR	129×129	C
LR	65×65	$C/8$
ULR	33×33	$C/64$

- Observations:
  - True value perturbed by a gaussian noise of standard deviation 2
  - Available every $\Delta t = 12$
  - Located along simulated satellite tracks (black dots on the figures)
  - Note the representativeness errors.



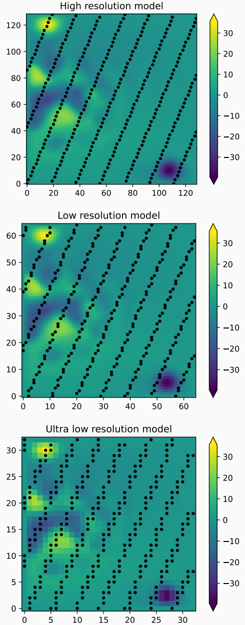
- Model used: Quasi-geostrophic model [Sakov and Oke, 2008]

Configuration	State size	Cost
HR	129×129	C
LR	65×65	$C/8$
ULR	33×33	$C/64$

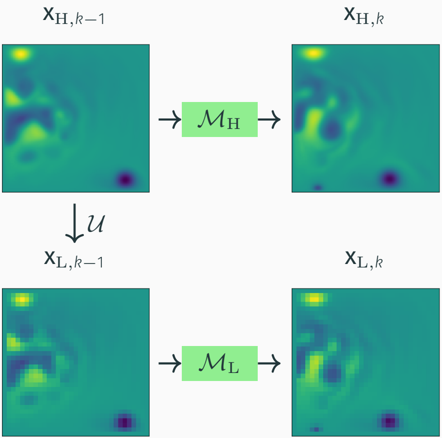
- Observations:
  - True value perturbed by a gaussian noise of standard deviation 2
  - Available every $\Delta t = 12$
  - Located along simulated satellite tracks (black dots on the figures)
  - Note the representativeness errors.

Downscaling operator?

- A simple cubic spline interpolation
- A neural network



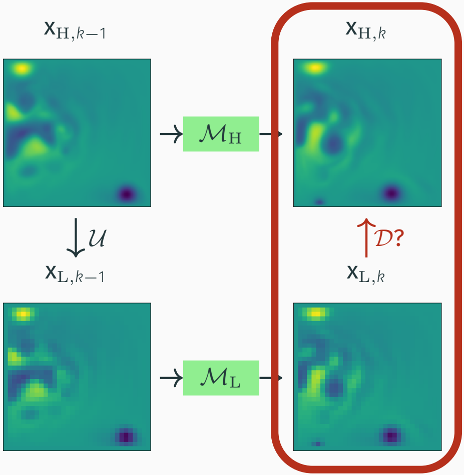
- Run one simulation of the HR model.
- Assemble matching pairs of (U)LR and HR states: $(\mathsf{x}_{\mathrm{L},k}, \mathsf{x}_{\mathrm{H},k})$



$\mathcal{U}$: Upscaling (subsampling operator)

Training set for the neural network

- Run one simulation of the HR model.
- Assemble matching pairs of (U)LR and HR states: $(\mathbf{x}_{\mathrm{L},k}, \mathbf{x}_{\mathrm{H},k})$

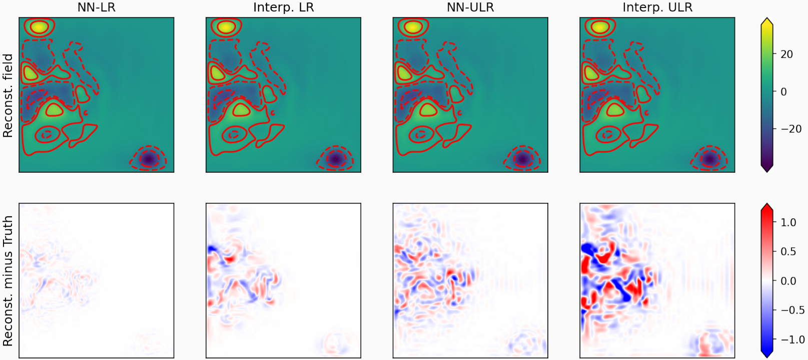


$\mathcal{U}$: Upscaling (subsampling operator)
$\mathcal{D}$: Downscaling (Neural network)

- Number of pairs: 10,000
- 8000 for training / 2000 for validation
- Architecture of the enhanced deep super-resolution network (EDSR) [Lim et al., 2017]
- Training: minimization of the mean absolute error

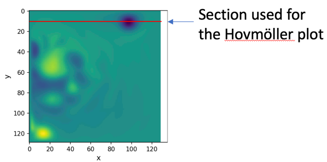
Downscaling performance

- Illustration with one typical sample

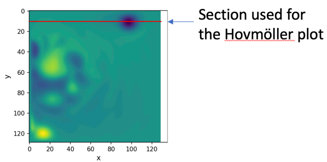


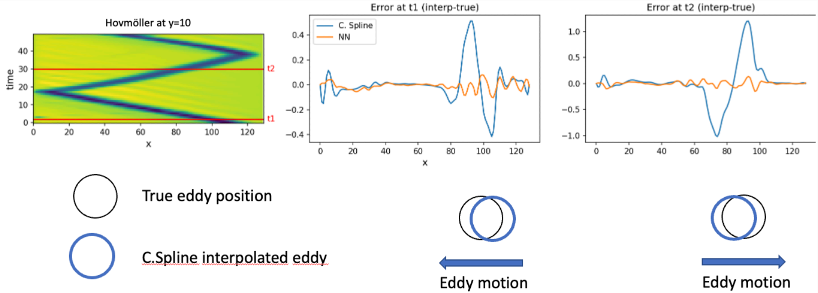
red lines: Contour of the true HR state

Model error correction



Model error correction





- Eddy propagation slower in the LR model
- The NN is smart enough to learn that

1. Introduction

2. Super-resolution data assimilation

Super-resolution data assimilation performance

Computing performance

Reformulation of the SRDA as a LR scheme

3. Hybrid covariance super-resolution data assimilation

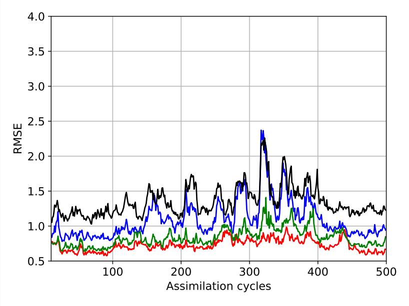
4. Links between the geometrical features of $\mathbf{P}^{\mathrm{f}}$ and super-resolution data assimilation

5. Conclusion and perspectives

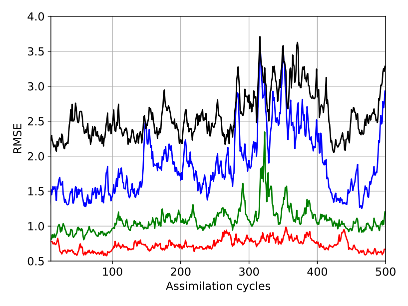
Super-resolution data assimilation performance

- Twin experiments with 500 assimilation cycles
- Sensitivity analysis to tune the optimal localisation and inflation
- Strong improvement irrespective of ensemble size
- Method able to predict uncertainties, same reliability as the EnKF

Performance with LR model



Performance with ULR model



EnKF-LR

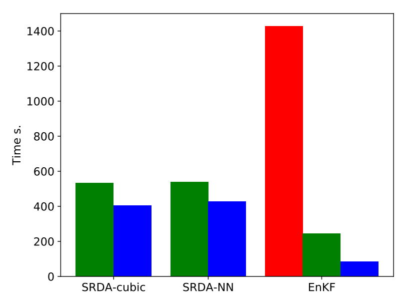
SRDA with cubic spline

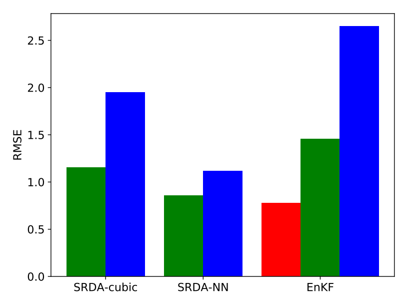
SRDA with NN

EnKF-HR

Computing performance - Total CPU time in Python

- With 25 members sequentially
- Same inflation and localization coefficients
- SRDA holds its promise: good value for the cost

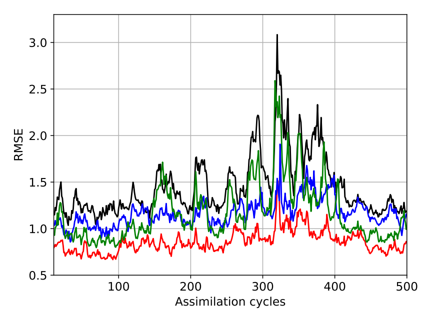




High resolution
Low resolution
Ultra low resolution

Reformulating the SRDA as a LR scheme

- We can reformulate the SRDA into LR EnKF equations so that we can separate the contributions from:
  1. the model error correction;
  2. the super-resolution observation operator (representativeness).



- Model error correction improves performance during challenging events
- Super-resolution obs. operator reduces error over the whole period

EnKF-LR
SRDA only with model error correction
Complete SRDA-NN
SRDA only with the super-res. observation operator

Overview

1. Introduction

2. Super-resolution data assimilation

3. Hybrid covariance super-resolution data assimilation

   Extension of SRDA to the hybrid covariance configuration

   Constant assimilation cost: fixed ensemble size

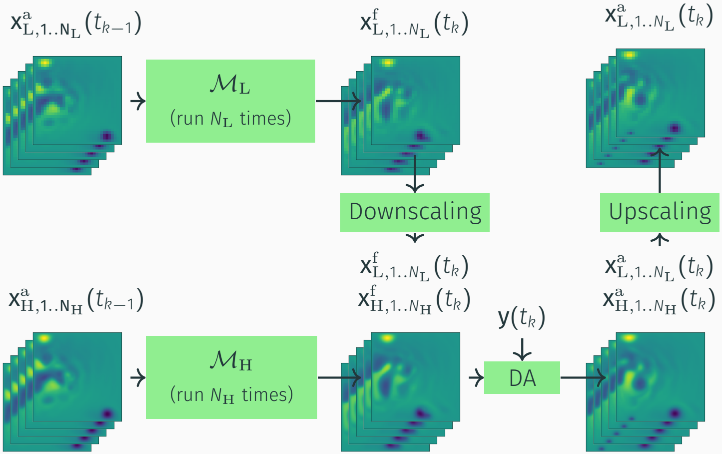
   Constant integration cost: trade off HR/LR ensembles

4. Links between the geometrical features of $\mathbf{P}^{\mathrm{f}}$ and super-resolution data assimilation

5. Conclusion and perspectives

Hybrid covariance SRDA

Hybrid covariance SRDA (Hybrid SRDA)



Hybrid covariance SRDA

The hybrid covariance matrix $\mathsf{P}^{\mathrm{f}}_{\mathrm{h}}$ is a linear combination of:

- $\mathsf{P}^{\mathrm{f}}_{\mathrm{HR}}$ computed from the HR ensemble;
- $\mathsf{P}^{\mathrm{f}}_{\mathrm{LR}}$ computed from the LR ensemble **downscaled** to the HR grid:

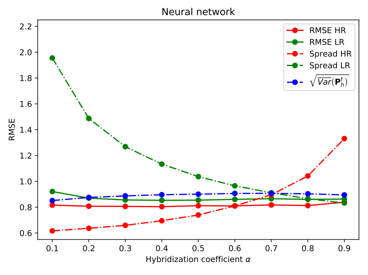
$$\mathsf{P}^{\mathrm{f}}_{\mathrm{h}} = (1-\alpha)\mathsf{P}^{\mathrm{f}}_{\mathrm{HR}} + \alpha\mathsf{P}^{\mathrm{f}}_{\mathrm{LR}}, \qquad 0 \leq \alpha \leq 1. \tag{1}$$

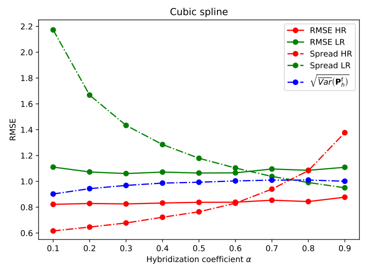
- $\alpha = 0$ full HR case $\rightarrow$ EnKF-HR
- $\alpha = 1$ full LR case $\rightarrow$ EnKF-LR
- downscaling method "cubic spline interpolation"

  $\Rightarrow$ [Rainwater and Hunt, 2013].
- Results computed over the HR ensemble unless otherwise stated.

DA experiments with fixed ensemble size: influence of α - LR case

- Twin experiments with 500 assimilation cycles;
- $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (5, 10)$ with localization $L = 5$ and inflation $\lambda = 1.05$
- $\alpha = 0, 0.1, 0.2, \ldots, 1$

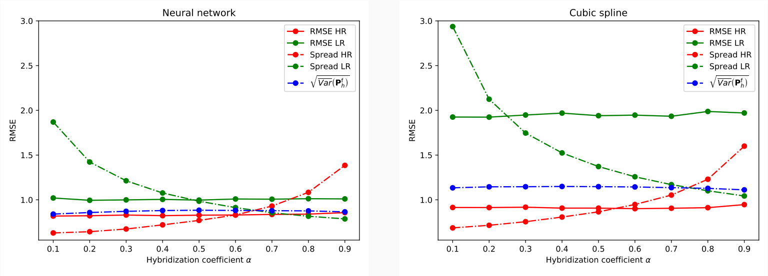




- Limited influence on the RMSE;
- Strong influence on the spread of the ensembles;
- No convergence for $\alpha = 0$, and $\alpha = 1$ (depends on L, λ and $(N_{\mathrm{H}}, N_{\mathrm{L}})$)

DA experiments with fixed ensemble size: influence of α - ULR case

- Twin experiments with 500 assimilation cycles;
- $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (5, 10)$ with localization $L = 5$ and inflation $\lambda = 1.05$
- $\alpha = 0, 0.1, 0.2, \ldots, 1$

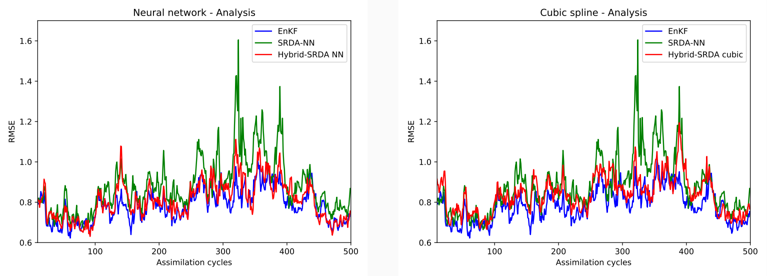


- Same conclusions as in the LR case.

Intercomparison of SRDA, hybrid-SRDA, and EnKF - LR case

For fixed parameters $L = 5$, $\lambda = 1.05$, $\alpha = 0.6$, $(N_\mathrm{H}, N_\mathrm{L}) = (5, 10)$, and $N = 15$:

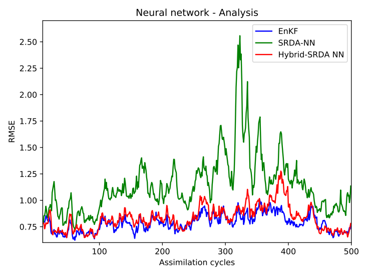
- same computational time of assimilation;
- the **hybrid-SRDA** outperforms the **SRDA-NN**;
- the **EnKF** outperforms the **hybrid-SRDA**.

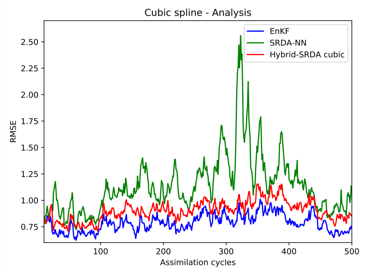


Intercomparison of SRDA, hybrid-SRDA, and EnKF - ULR case

For fixed parameters $L = 5$, $\lambda = 1.05$, $\alpha = 0.6$, $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (5, 10)$, and $N = 15$:

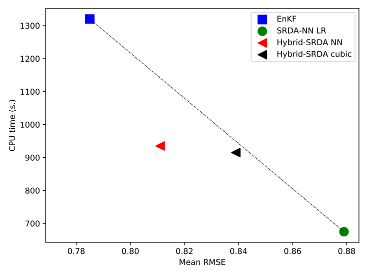
- same computational time of assimilation;
- the **hybrid-SRDA** outperforms the **SRDA-NN**;
- the **EnKF** outperforms the **hybrid-SRDA**.

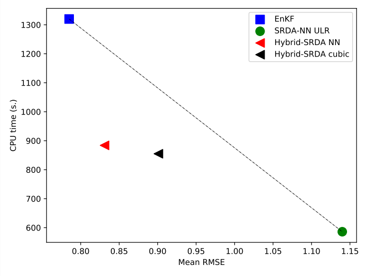




Computational effectiveness

- Same inflation and localization coefficients: $\lambda = 1.05$ and $L = 5$;
- $N = 15$ for the **EnKF** and the **SRDA**;
- $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (5, 10)$ for the **hybrid-SRDA**.





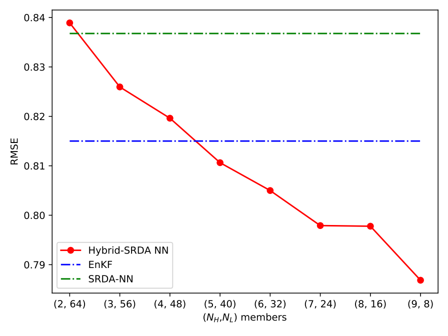
- The method is cost effective if it is under the black dashed line;
- The method improves the ratio ”Error increase/time reduction” if it is under the black dashed line compared to the SRDA-NN.

Trade off HR/LR ensemble sizes experiments

Trade off: 1 HR member $\approx$ 8 LR members (integration time)

Design of the experiments:

- Twin experiments with 500 assimilation cycles;
- **EnKF-HR** with $N_{\mathrm{H}} = 10$ members, **SRDA-NN** with $N_{\mathrm{L}} = 80$ members;
- **Hybrid-SRDA NN** with $(N_{\mathrm{H}}, N_{\mathrm{L}}) = (2, 64), (3, 56), \ldots, (9, 8)$;
- Optimal localization, inflation and hybridization coefficients.

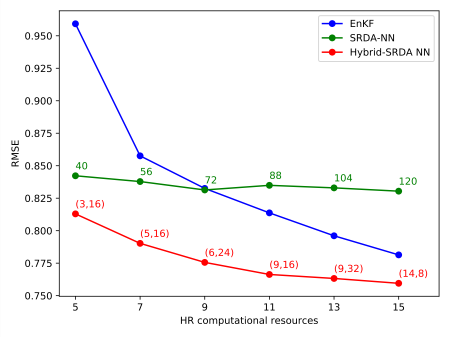


Trade off HR/LR ensemble sizes experiments

Trade off: 1 HR member $\approx$ 8 LR members (integration time)

Design of the experiments:

- Twin experiments with 500 assimilation cycles;
- Computational resources of running $\approx$ 5, 7, . . . , 15 HR members;
- Optimal localization, inflation and hybridization coefficients.



Overview

1. Introduction
2. Super-resolution data assimilation
3. Hybrid covariance super-resolution data assimilation
4. Links between the geometrical features of P^{f} and super-resolution data assimilation
5. Conclusion and perspectives

Characterization of the covariance functions

Characterization of the covariance functions with their **variance** and **correlation length scale**:

$$\mathsf{P}^{\mathrm{f}} = \Sigma \mathsf{C}^{\mathrm{f}} \Sigma^{T} \tag{2}$$

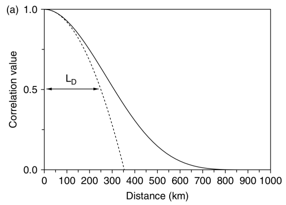
- C^{f} background error correlation matrix;
- Σ diagonal matrix of the square root of the variance of P^{f}.

Characterization of the covariance functions

Characterization of the covariance functions with their **variance** and **correlation length scale**:

$$\mathbf{P}^{\mathrm{f}} = \Sigma \mathbf{C}^{\mathrm{f}} \Sigma^{T} \quad (2)$$

- $\mathbf{C}^{\mathrm{f}}$ background error correlation matrix;
- Σ diagonal matrix of the square root of the variance of $\mathbf{P}^{\mathrm{f}}$.



Parabolic based approximation of the correlation length scale:

$$L_p = \frac{\delta x}{\sqrt{-2 \ln \left(\rho(\delta x) \right)}} \quad (3)$$

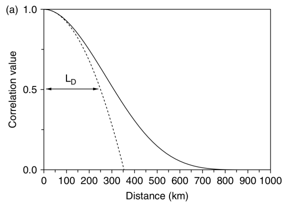
Source: [Pannekoucke et al., 2008]

Characterization of the covariance functions

Characterization of the covariance functions with their **variance** and **correlation length scale**:

$$\mathbf{P}^{\mathrm{f}} = \Sigma \mathbf{C}^{\mathrm{f}} \Sigma^{T} \quad (2)$$

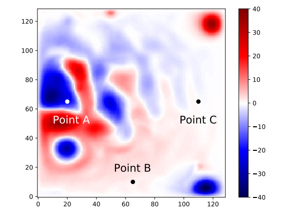
- $\mathbf{C}^{\mathrm{f}}$ background error correlation matrix;
- Σ diagonal matrix of the square root of the variance of $\mathbf{P}^{\mathrm{f}}$.



Parabolic based approximation of the correlation length scale:

$$L_p = \frac{\delta x}{\sqrt{-2\ln\left(\rho(\delta x)\right)}} \quad (3)$$

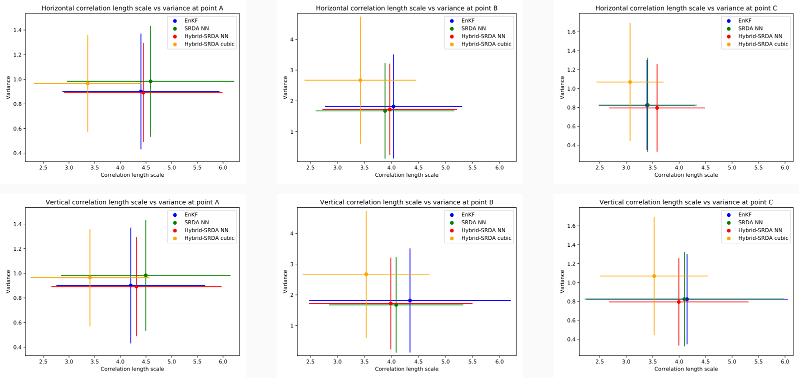
Source: [Pannekoucke et al., 2008]



- **Point A:** high temporal and spatial variability;
- **Point B:** passing of eddies;
- **Point C:** small temporal and spatial variability.

Characterization of the covariance functions

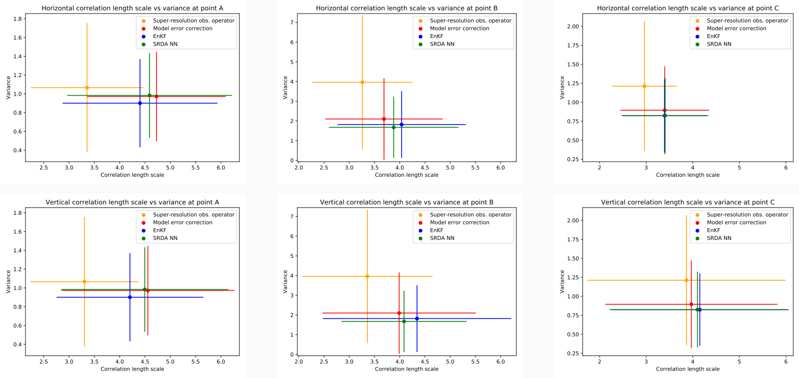
- ▶ Twin experiments with 500 assimilation cycles;
- ▶ Estimating the variance and the horizontal/vertical correlation length scales at points A, B, C;
- ▶ Results averaged over the whole period.



Point A Point B Point C

- ▶ The hybrid-SRDA cubic does not provide a correct modelling of $\mathbf{P}^f$;

- Twin experiments with 500 assimilation cycles;
- Respective influence of the model error correction and super-resolution observation operator on the variance and the correlation length scales.



Point A Point B Point C

Overview

1. Introduction

2. Super-resolution data assimilation

3. Hybrid covariance super-resolution data assimilation

4. Links between the geometrical features of P^{f} and super-resolution data assimilation

5. Conclusion and perspectives

Main results

- **SRDA** nearly as accurate as the **EnKF** with HR model but for a cost close to the **EnKF-LR**;
- The **hybrid-SRDA** outperforms the **SRDA** and performs almost as good as the **EnKF** but for a *reduced cost*;
- For limited integration resources the **hybrid SRDA** systemically outperforms the **EnKF**;
- The NN can correct systematic differences of eddy propagation caused by the low resolution;
- The NN allows for a better representation (on average) of the geometrical features of the covariance matrix;
- The results are stable in time (especially during challenging events);

Main results

- **SRDA** nearly as accurate as the **EnKF** with HR model but for a cost close to the **EnKF-LR**;
- The **hybrid-SRDA** outperforms the **SRDA** and performs almost as good as the **EnKF** but for a *reduced cost*;
- For limited integration resources the **hybrid SRDA** systemically outperforms the **EnKF**;
- The NN can correct systematic differences of eddy propagation caused by the low resolution;
- The NN allows for a better representation (on average) of the geometrical features of the covariance matrix;
- The results are stable in time (especially during challenging events);

Perspectives

- Application to a more realistic (multivariate) model $\rightarrow$ project Impetus;
- Application restricted to parts of the model domain,
- Use NN-downscaling for the initialization of HR forecasts.

SRDA (only!) paper available on *Ocean Dynamics*!

https://link.springer.com/article/10.1007/s10236-022-01523-x

Barthélémy, S., Brajard, J., Bertino, L., and Counillon, F. (2022).
Super-resolution data assimilation.
Ocean Dynamics, pages 1–18.

Lim, B., Son, S., Kim, H., Nah, S., and 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*.

Pannekoucke, O., Berre, L., and Desroziers, G. (2008).
Background error correlation length scale estimates and their sampling statistics.
Quarterly Journal of the Royal ..., 134(March):497–508.

Rainwater, S. and Hunt, B. (2013).
Mixed-Resolution Ensemble Data Assimilation.
Monthly Weather Review, 141(9):3007–3021.

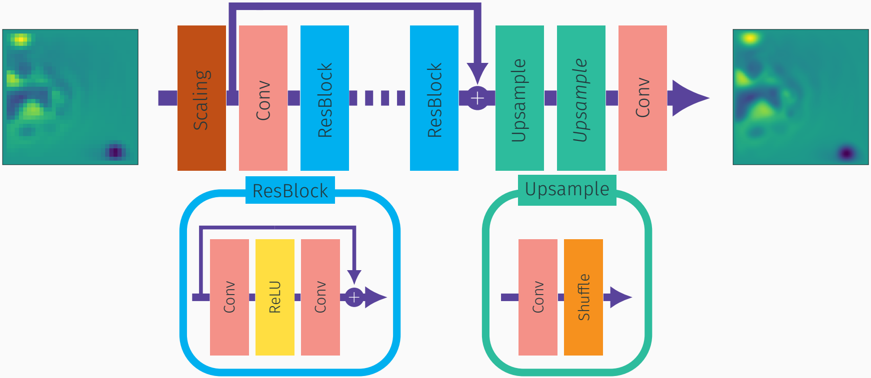
Sakov, P. and Oke, P. R. (2008).
A deterministic formulation of the ensemble Kalman filter: An alternative to ensemble square root filters.
Tellus, Series A: Dynamic Meteorology and Oceanography, 60 A(2):361–371.

sebastien.barthelemy@uib.no – julien.brajard@nersc.no

Acknowledgement:

NFR project SFE(#2700733)

Setup of the neural network



Architecture of the enhanced deep super-resolution network (EDSR) [Lim et al., 2017]

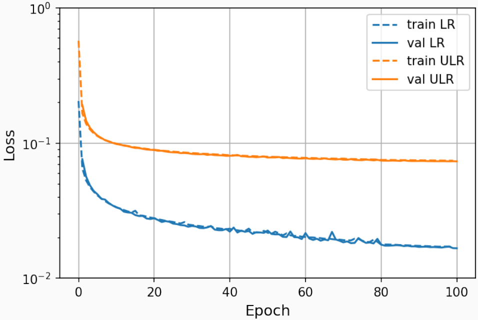
Training of the neural network

Minimize the mean absolute error (MAE):

$$L(\mathbf{w}) = \sum_{k=1}^{K} \sum_{i=1}^{S} \left| \mathcal{D}(\mathbf{x}_{\mathrm{L},k})_i - x_{\mathrm{H},k,i} \right| ,$$

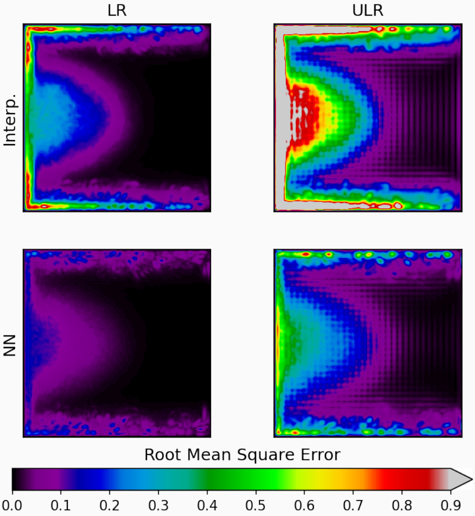
- i: the pixel index
- S: size of the state (129×129)
- K: size of the training set (K=8000)
- $\mathbf{w}$: weights of the neural network ($\sim 20,000$)

Training curve



Downscaling performance (2)

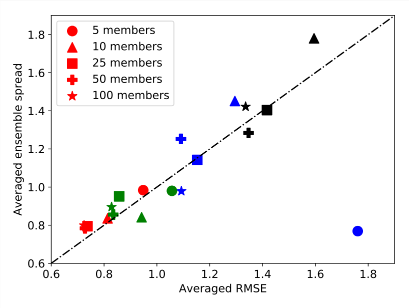
- Score on the validation dataset



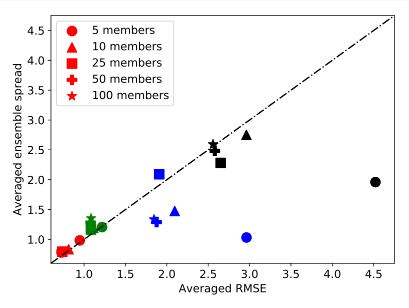
Super-resolution data assimilation performance

- Method able to preserve the reliability

Performance with LR model



Performance with ULR model



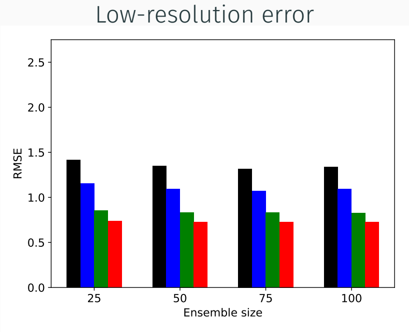
EnKF-LR

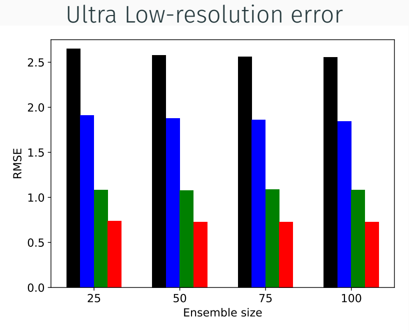
SRDA with cubic spline

SRDA with NN

EnKF-HR

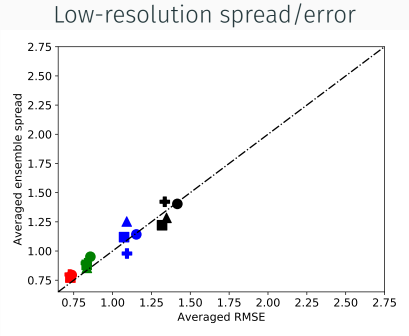
Super-resolution data assimilation performance

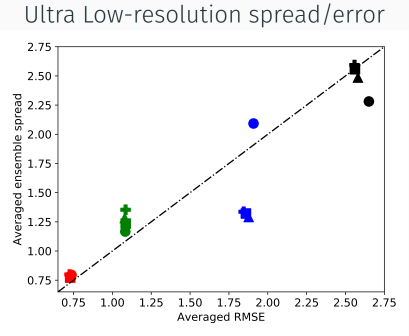




- DA in low-resolution
- SRDA with cubic spline interpolation
- SRDA with NN downscaling
- DA in high-resolution

Spread/error of the ensemble





DA in low-resolution

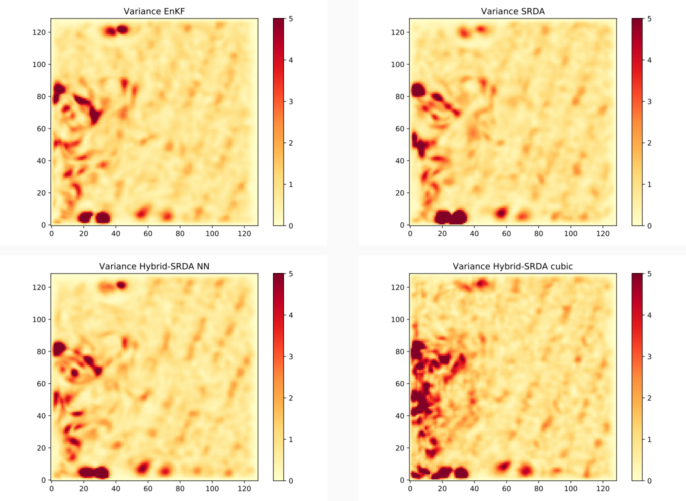
SRDA with cubic spline interpolation

SRDA with NN downscaling

DA in high-resolution

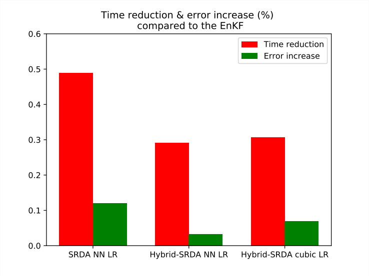
Spatial pattern of the variance

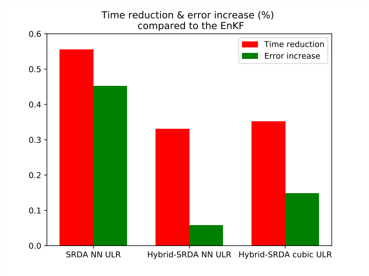
Pattern of the variance at assimilation cycles 321

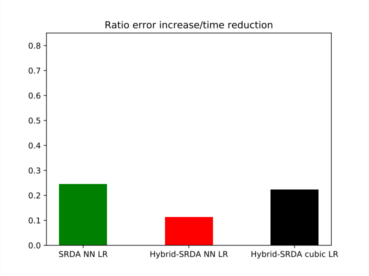


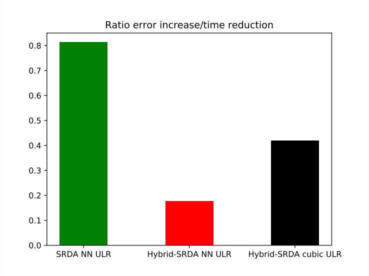
- **SRDA NN** and **hybrid-SRDA NN** have same pattern of variance as the **EnKF**.

Computational effectiveness









Hybrid covariance SRDA

HR ensemble update

$$\begin{cases} \mathsf{x}^{\mathrm{a}}_{\mathrm{H}} &= \mathsf{x}^{\mathrm{f}}_{\mathrm{H}} + \mathsf{K}_{\mathrm{h}} \left(\mathsf{d}_{\mathrm{H}} - \mathsf{H}_{\mathrm{H}} \mathsf{x}^{\mathrm{f}}_{\mathrm{H}}\right) \\ \mathsf{A}^{\mathrm{a}}_{\mathrm{H}} &= \mathsf{A}^{\mathrm{f}}_{\mathrm{H}} - \frac{1}{2} \mathsf{K}_{\mathrm{h}} \mathsf{H}_{\mathrm{H}} \mathsf{A}^{\mathrm{f}}_{\mathrm{H}} \end{cases}$$

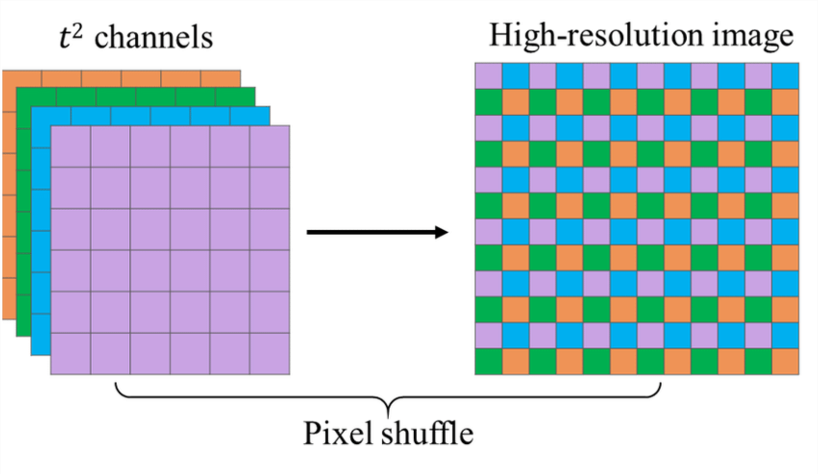
LR ensemble update (in the HR space)

$$\begin{cases} \mathsf{x}^{\mathrm{a}}_{\mathrm{L}\to\mathrm{H}} &= \mathsf{x}^{\mathrm{f}}_{\mathrm{L}\to\mathrm{H}} + \mathsf{K}_{\mathrm{h}} \left(\mathsf{d}_{\mathrm{H}} - \mathsf{H}_{\mathrm{H}} \mathsf{x}^{\mathrm{f}}_{\mathrm{L}\to\mathrm{H}}\right) \\ \mathsf{A}^{\mathrm{a}}_{\mathrm{L}\to\mathrm{H}} &= \mathsf{A}^{\mathrm{f}}_{\mathrm{L}\to\mathrm{H}} - \frac{1}{2} \mathsf{K}_{\mathrm{h}} \mathsf{H}_{\mathrm{H}} \mathsf{A}^{\mathrm{f}}_{\mathrm{L}\to\mathrm{H}} \end{cases}$$

where K_{h} is the hybrid Kalman gain:

$$\mathsf{K}_{\mathrm{h}} = \mathsf{P}^{\mathrm{f}}_{\mathrm{h}} \mathsf{H}^{T}_{\mathrm{H}} \left(\mathsf{H}_{\mathrm{H}} \mathsf{P}^{\mathrm{f}}_{\mathrm{h}} \mathsf{H}^{T}_{\mathrm{H}} + \mathsf{R}_{\mathrm{H}}\right)^{-1} \tag{4}$$

The shuffle operator



Qin, Mengjiao, et al. **"Remote Sensing Single-Image Resolution Improvement Using A Deep Gradient-Aware Network with Image-Specific Enhancement."** *Remote Sensing* 12.5 (2020): 758.