



**HAL**  
open science

# Neural SPDE solver for uncertainty quantification in high-dimensional ocean dynamics

Maxime Beauchamp, Pierre Tandeo, Nicolas Desassis, Ronan Fablet

► **To cite this version:**

Maxime Beauchamp, Pierre Tandeo, Nicolas Desassis, Ronan Fablet. Neural SPDE solver for uncertainty quantification in high-dimensional ocean dynamics. 55th International Liege Colloquium on Ocean Dynamics, May 2024, Liège (B), Belgium. hal-04675414

**HAL Id: hal-04675414**

**<https://imt.hal.science/hal-04675414>**

Submitted on 22 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Neural SPDE solver for uncertainty quantification in high-dimensional ocean dynamics

Maxime Beauchamp<sup>1</sup>, Pierre Tandeo<sup>1</sup>, Nicolas Desassis<sup>2</sup> and Ronan Fablet<sup>1</sup>

1. IMT Atlantique Bretagne-Pays de la Loire, Brest, France, 2. Mines ParisTech

maxime.beauchamp@imt-atlantique.fr



## I. Context

Over the last years, a very active field of research aims at exploring new data-driven and learning-based methodologies to propose computationally efficient strategies able to benefit from the large amount of observational remote sensing and numerical simulations for the reconstruction, interpolation and prediction of high-resolution derived products of geophysical fields. We use the BOOST-SWOT ocean data challenges (ODC) as playground datasets.

The 2020 ODC provides the dataset to train our models: the Ground Truth (GT)  $\mathbf{x}$  is the high-resolution  $1/60^\circ$  NATL60 configuration of the NEMO (Nucleus for European Modeling of the Ocean) model. A  $10^\circ \times 10^\circ$  GULFSTREAM region is used with downgraded resolution to  $1/20^\circ$ . Along-track nadir + wide-swath SWOT pseudo-observations  $\mathbf{y}$  are sampled from GT based on real along-track satellites paths. Last, a baseline is provided by the operational DUACS Optimal Interpolation (OI) product.

We investigate how the ensemble-based extension of 4DVarNet, denoted here as **4DVarNet-Gen** helps to quantify the uncertainty of Sea Level spatio-temporal fields  $\mathbf{x} = \{\mathbf{x}_{t-L}, \dots, \mathbf{x}_{t+L}\}$ .

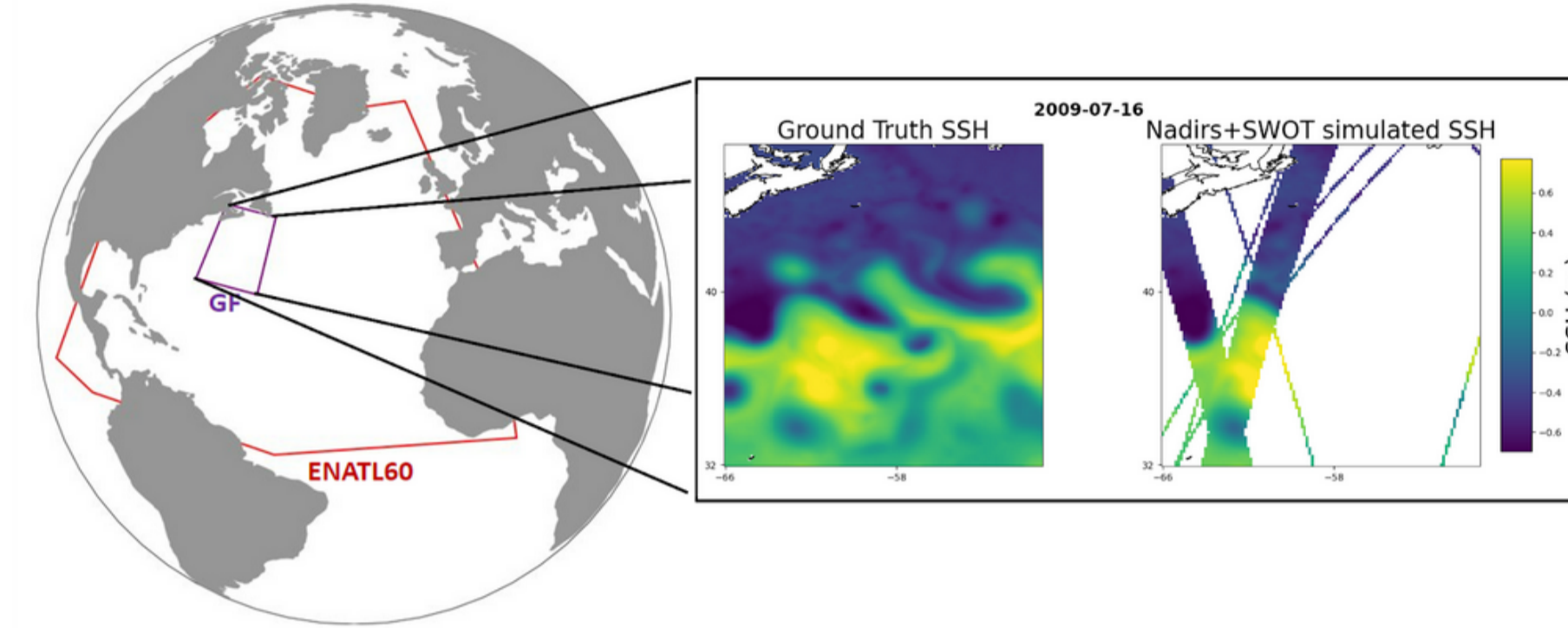


Figure 1: Ground Truth (SSH) and pseudo-observations

## II. Neural variational scheme with SPDE priors

### Basic concepts

From a methodological point of view, 4DVarNet [Fablet et al., 2021] derives an end-to-end neural architecture from an underlying variational data assimilation formulation:

$$\mathcal{J}_\Phi(\mathbf{x}, \mathbf{y}, \Omega) = \lambda_1 \|\mathbf{y} - \mathcal{H}(\mathbf{x})\|_\Omega^2 + \lambda_2 \|\mathbf{x} - \Phi(\mathbf{x})\|^2, \quad (1)$$

where  $\lambda_{1,2}$  are predefined or tunable scalar weights. In the regularization term, we substitute to the traditional dynamical prior  $\mathcal{M}$  a neural operator  $\Phi$  which convolutional architecture. Then, we can exploit the automatic differentiation tools embedded in deep learning framework to consider the following iterative gradient-based solver  $\Gamma$  for the minimization of variational cost  $\mathcal{J}_\Phi$  w.r.t. state  $\mathbf{x}$ :

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - \mathcal{K} \left[ \alpha \cdot \nabla_{\mathbf{x}} \mathcal{J}_\Phi(\mathbf{x}^{(i)}, \mathbf{y}, \Omega), h(i), c(i) \right] \quad (2)$$

where  $\mathcal{K}$  is a convolutional LSTM model and  $\alpha$  a normalization scalar. This iterative rule based on a trainable LSTM operator is similar to that classically used in meta-learning schemes [Andrychowicz et al., 2016]. Overall, a 4DVarNet scheme defines a neural architecture with a joint learning of operators  $\{\Phi, \Gamma\}$ , stated as the minimization of a reconstruction cost  $\mathcal{L}$ , which typically stands for the MSE (mean squared error) w.r.t. the Ground Truth.

### Stochastic extension

Prior operator  $\Phi$  is not easily interpretable, simply acting as a projection of state  $\mathbf{x}$  to help in the gradient-based minimization process. Here, we bring both interpretability and stochasticity in the neural scheme by considering as prior surrogate model a linear SPDE:

$$\mathbf{x}_{t+1} = \mathbf{M}_{t+1} \mathbf{x}_t + \mathbf{z}_{t+1}, \quad \mathbf{M}_{t+1} = [\mathbf{I} + \mathbf{L}_t(\Theta_t)]^{-1}, \quad \mathbf{z}_{t+1} \sim \mathcal{N}(\mathbf{0}, \tau_{t+1}^2 \mathbf{I}) \quad (3)$$

Operator  $\mathbf{L}(\Theta)$  states as the finite difference discretization scheme of a fractional advection-diffusion operator  $\{\kappa^2(\mathbf{s}, t) - \nabla \cdot \mathbf{m}(\mathbf{s}, t) - \nabla \cdot \mathbf{H}(\mathbf{s}, t) \nabla\}^{\alpha/2}$  with  $\mathbf{m}$  and  $\mathbf{H}$  resp. the advection vector and diffusion tensor.  $\tau_t$  acts as variance regularization over space and time and  $\Theta$  encompasses the SPDE parameters. Back to the OI variational problem, the prior regularization cost simply becomes  $\mathbf{x}^T \mathbf{Q}_\Theta \mathbf{x}$ , where  $\mathbf{Q}_\Theta$  denotes the precision matrix of the prior state trajectory, and the joint learning of SPDE parametrization  $\Theta$  with solver  $\Gamma$  for the reconstruction  $\mathbf{x}^*$  states as the minimization of:

$$\arg \min_{\Theta, \Gamma} \mathcal{L}(\mathbf{x}, \Theta^*, \mathbf{x}^*) \text{ s.t. } \mathbf{x}^* = \Psi_{\Theta, \Gamma}(\mathbf{x}^{(0)}, \mathbf{y}, \Omega) \quad (4)$$

with  $\mathcal{L}(\mathbf{x}, \Theta^*, \mathbf{x}^*) = \lambda_1 \mathcal{L}_1(\mathbf{x}, \mathbf{x}^*) + \lambda_2 \mathcal{L}_2(\mathbf{x}, \Theta^*)$ .  $\mathcal{L}_1(\mathbf{x}, \mathbf{x}^*) = \|\mathbf{x} - \mathbf{x}^*\|^2$  is the reconstruction cost, i.e. the MSE w.r.t Ground Truth and  $\mathcal{L}_2(\Theta^*|\mathbf{x})$  is the negative log-likelihood  $-\mathcal{L}(\Theta^*|\mathbf{x})$  with:

$$\mathcal{L}(\Theta^*|\mathbf{x}) = -|\mathbf{Q}(\Theta^*)| + \mathbf{x}^T \mathbf{Q}(\Theta^*) \mathbf{x} \quad (5)$$

used as the prior regularization cost.

## III. SPDE estimation and Generative modeling

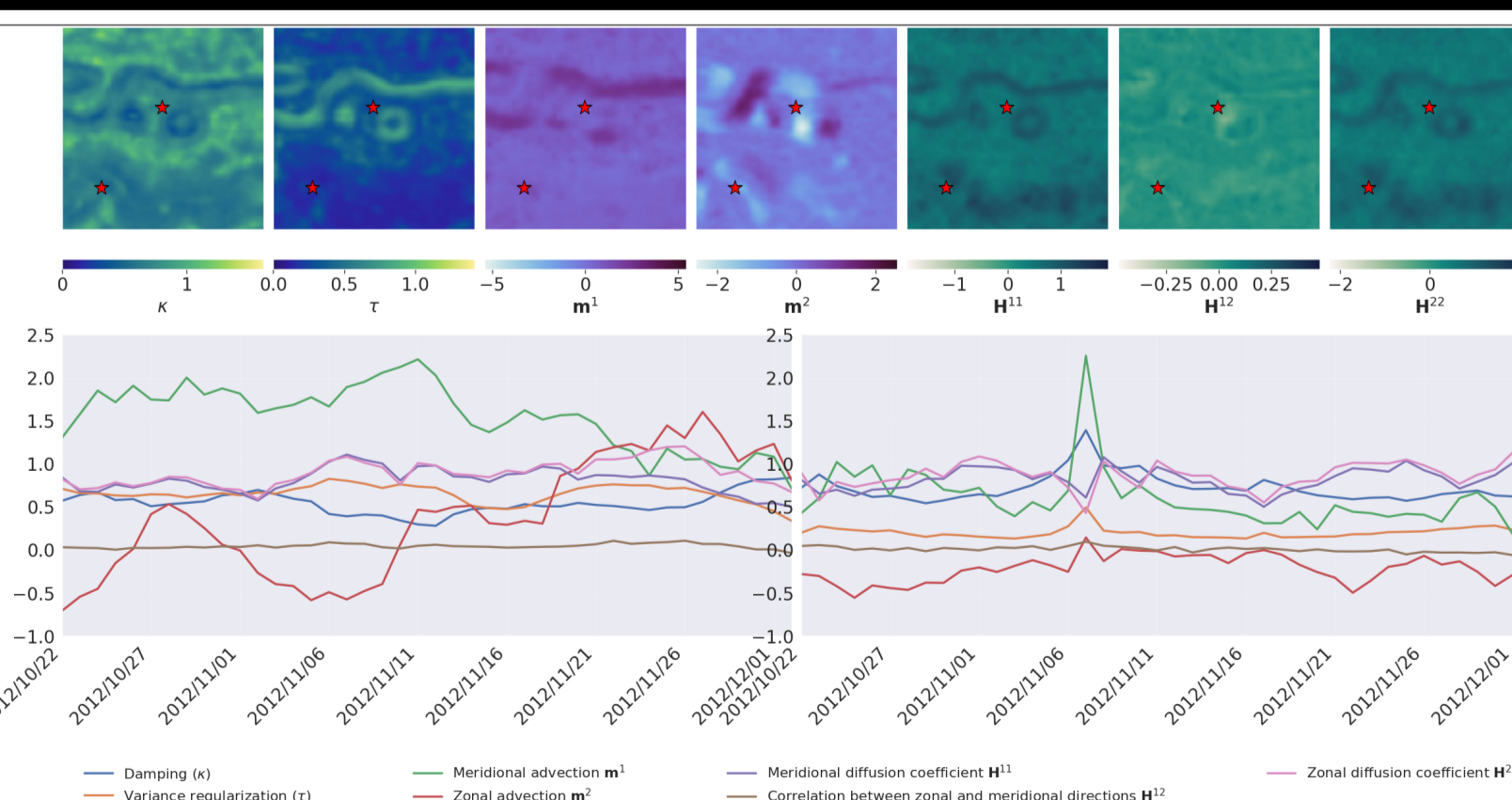


Figure 3: SPDE parameter estimations: [top] first day of the test period; [bottom] time series along the GF meandrum (left) and in the less energetic left-lower part of the domain (right).

On the top panel of Figure 3, the maps of all the parameters are shown on the first day of the test period, while on the bottom panel, the time series of these parameters along the test period on the two locations identified by the red stars are displayed. The first one (left) is located along the Gulf Stream meander and the second one (right) is in the less energetic left-lower part of the domain. It clearly indicates how the parameters are distributed according to the main meandrum and smaller eddies. Interestingly, we can see from the time series how the 7 parameters are not perfectly correlated along the test period, which indicates capabilities of the model to adapt each parameter independently given the observations available. Such a methodological development can be used not only for the initial mapping problem targeted, but also for near-real-time (NRT) and short-term forecast problems (see Figure 4), where we can see how the surrogate SPDE model can be used as a generative model for prior sampling in the DA scheme.

## IV. Results

Table 1: Performance metrics (RMSE scores mean and standard deviations, minimal spatial and temporal scales resolved) of DUACS OI and 4DVarNet-Gen for mapping, NRT, and forecast (lead times 2 and 4) configurations.

	name	mean(RMSE)	std(RMSE)	spatial scale (°)	temporal scale (days)
	OI (1 swot + 4 nadirs)	0.92	0.02	1.37	11.36
	4DVarNet-Gen (mapping)	0.94	0.01	1.15	6.98
	4DVarNet-Gen (NRT)	0.91	0.02	1.18	9.87
	4DVarNet-Gen (Forecast, leadtime=2)	0.87	0.03	1.6	12.63
	4DVarNet-Gen (Forecast, leadtime=4)	0.78	0.04	2.32	28.84

### Conditional simulations

Such SPDE formulation enables to draw samples  $\mathbf{x}^{s,i}$  in the prior distribution  $p(\mathbf{x})$ . To make these simulations compliant with the observations, we draw from traditional geostatistics and kriging-based conditioning [Wackernagel, 2003] and replace the prior model and kriging solver, also known as Optimal Interpolation (OI) or BLUE (Best Linear Unbiased Estimator) in the DA community, see e.g. [Asch et al., 2016], by our neural scheme. A conditional simulation  $\mathbf{x}^{*,i}$  writes:

$$\mathbf{x}^{*,i} = \mathbf{x}^* + \{\mathbf{x}^{s,i} - \mathbf{x}^{*,s,i}\} \quad (6)$$

It was shown that 4DVarNet estimation [Beauchamp et al., 2022; Fablet et al., 2021] enables to improve the OI, see also Table 1, i.e. it is still unbiased but with lower MSE w.r.t the ground truth, i.e. lower error variance. Running an ensemble of  $N$  simulations conditioned by 4DVarNet will provide an approximation of the probability distribution function  $p_{\mathbf{x}|\mathbf{y}}$  with both improvements on the two first moments  $\mathbf{x}^*$  and  $\mathbf{P}^*$  compared to simulations conditioned by traditional Gaussian solvers.

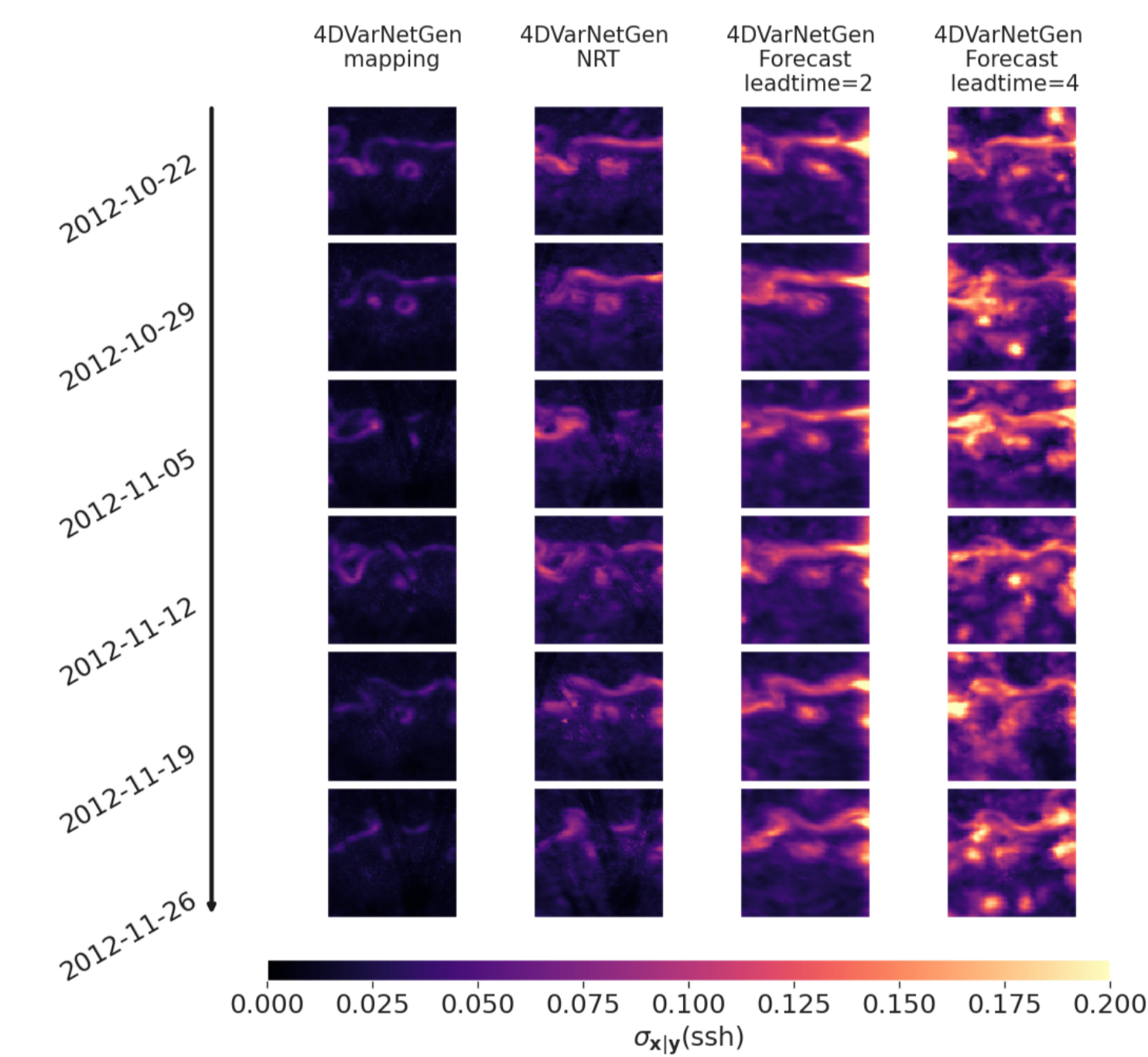


Figure 2: Standard deviations of 200 4DVarNet-Gen conditional simulations for mapping, NRT, and forecast (lead times 2 and 4) configurations along the test period.

## References

- Marcin Andrychowicz, Misha Denil, Sergio Gomez, Matthew W Hoffman, David Pfau, Tom Schaul, Brendan Shillingford, and Nando De Freitas. Learning to learn by gradient descent by gradient descent. In *Advances in neural information processing systems*, pages 3981–3989, 2016.
- M. Asch, M. Bocquet, and M. Nodet. *Data Assimilation. Fundamentals of Algorithms*. Society for Industrial and Applied Mathematics, December 2016. ISBN 978-1-61197-453-9. doi: 10.1137/1.9781611974546. URL <https://doi.org/10.1137/1.9781611974546>.
- M. Beauchamp, Q. Febvre, H. Georgenthum, and R. Fablet. 4dvarnet-ssh: end-to-end learning of variational interpolation schemes for nadir and wide-swath satellite altimetry. *Geoscientific Model Development Discussions*, 2022:1–37, 2022. doi: 10.5194/gmd-2022-241. URL <https://gmd.copernicus.org/preprints/gmd-2022-241/>.
- R. Fablet, M. Beauchamp, L. Drumetz, and F. Rousseau. Joint interpolation and representation learning for irregularly sampled satellite-derived geophysical fields. *Frontiers in Applied Mathematics and Statistics*, 7:25, 2021. ISSN 2297-4687. doi: 10.3389/fams.2021.655224. URL <https://www.frontiersin.org/article/10.3389/fams.2021.655224>.
- H. Wackernagel. *Multivariate Geostatistics. An Introduction with Applications*. Springer-Verlag Berlin Heidelberg, New York, 2003. doi: 10.1007/978-3-662-05294-5. ISBN 978-3-540-44142-7.

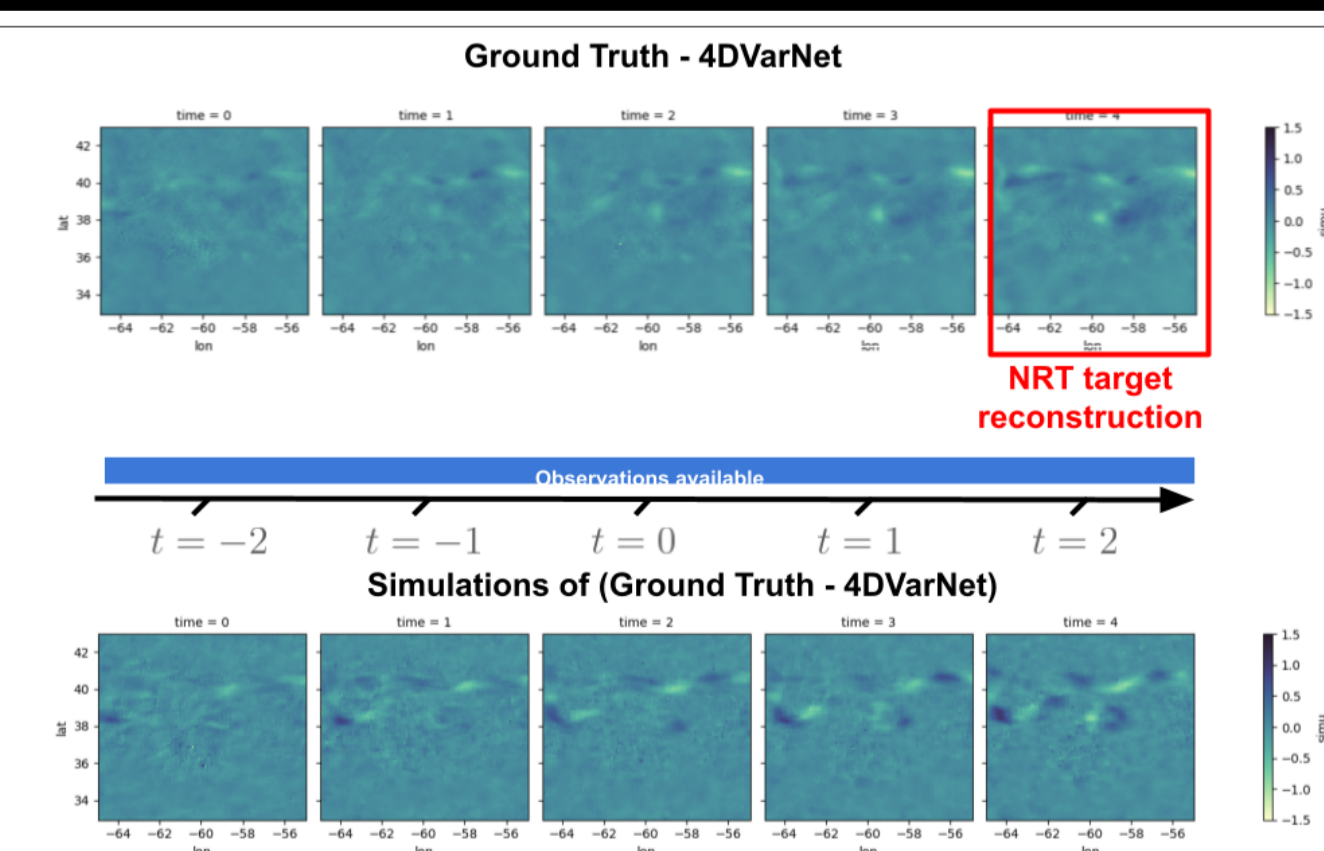


Figure 4: An example of the GT anomaly (GT-4DVarNet) for near-real-time (NRT) application along the data assimilation window and one simulation after training of the SPDE parameters