



HAL
open science

Sparse SPDE priors for uncertainty quantification of neural variational schemes

Maxime Beauchamp, Hugo Georgenthum, Ronan Fablet

► **To cite this version:**

Maxime Beauchamp, Hugo Georgenthum, Ronan Fablet. Sparse SPDE priors for uncertainty quantification of neural variational schemes. ISDA 2023: 23rd International Conference on Intelligent Systems Design and Applications, Oct 2023, Bologne (ITA), Italy. 2023. hal-04236738

HAL Id: hal-04236738

<https://imt.hal.science/hal-04236738>

Submitted on 11 Oct 2023

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.



Distributed under a Creative Commons Attribution 4.0 International License

Sparse SPDE priors for uncertainty quantification of neural variational schemes

Maxime Beauchamp¹, Hugo Georgenthum¹, Ronan Fablet¹,

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

maxime.beauchamp@imt-atlantique.fr



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) is the high-resolution 1/60° NATL60 configuration of the NEMO (Nucleus for European Modeling of the Ocean) model. A 10° × 10° GULFSTREAM region is used with downgraded resolution to 1/20°. Nadir pseudo-observations Y are sampled from GT based on real along-track satellites paths.

The 2021 ODC provides the dataset to test our model on real-world dataset. In the following experiments, we apply our ensemble-based 4DVarNet approach with 60 members on the 6 nadir constellation on January 2017.

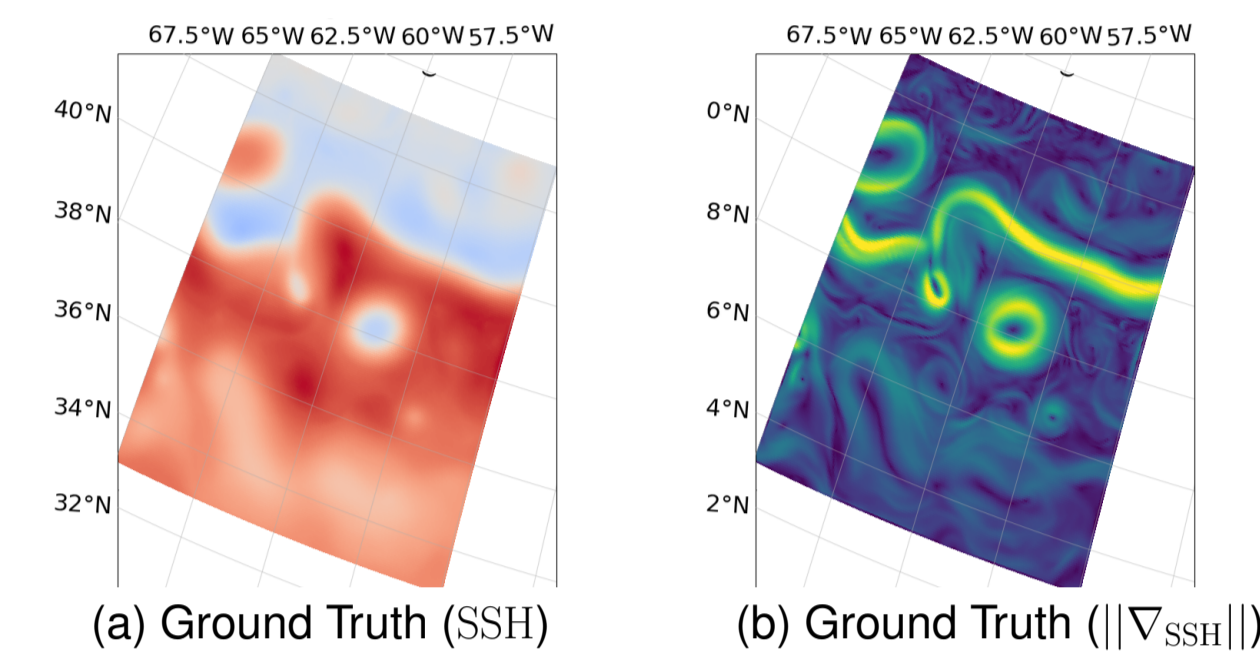


Figure 1: SSH (Ground Truth) and $\|\nabla_{SSH}\|$ on October 25, 2012

We investigate how the ensemble-based extension of 4DVarNet (En4DVarNet) helps to quantify the uncertainty of SSH spatio-temporal fields $\mathbf{x} = \{\mathbf{x}_{t-L}, \dots, \mathbf{x}_{t+L}\}$.

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 Φ which convolutional architecture. Then, we can exploit the automatic differentiation tools embedded in deep learning framework to consider the following iterative gradient-based solver Γ for the minimization of variational cost \mathcal{J}_Φ 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 α 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 Φ 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. τ_t acts as variance regularization over space and time and Θ 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}_Θ denotes the precision matrix of the prior state trajectory, and the joint learning of SPDE parametrization Θ with solver Γ 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.

Validation

Fig.5 displays the SSH spread from 2016-12-31 to 2017-01-25 (every 5 days). We can see how the observations help in decaying the reconstruction uncertainty. Then, when moving away from the observations at the center of the assimilation window, the uncertainty grows quickly, not only based on the distance from the nadir altimeters but also influenced by the SSH spatio-temporal dynamics.

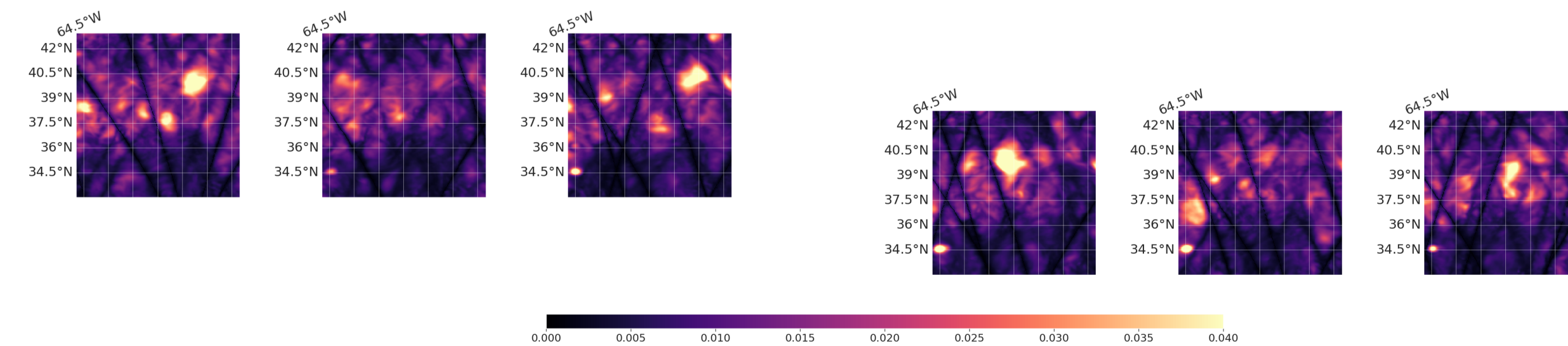


Figure 5: 4DVarNet SSH ensemble-based standard deviations from 2016-12-31 to 2017-01-25 (every 5 days)

Results

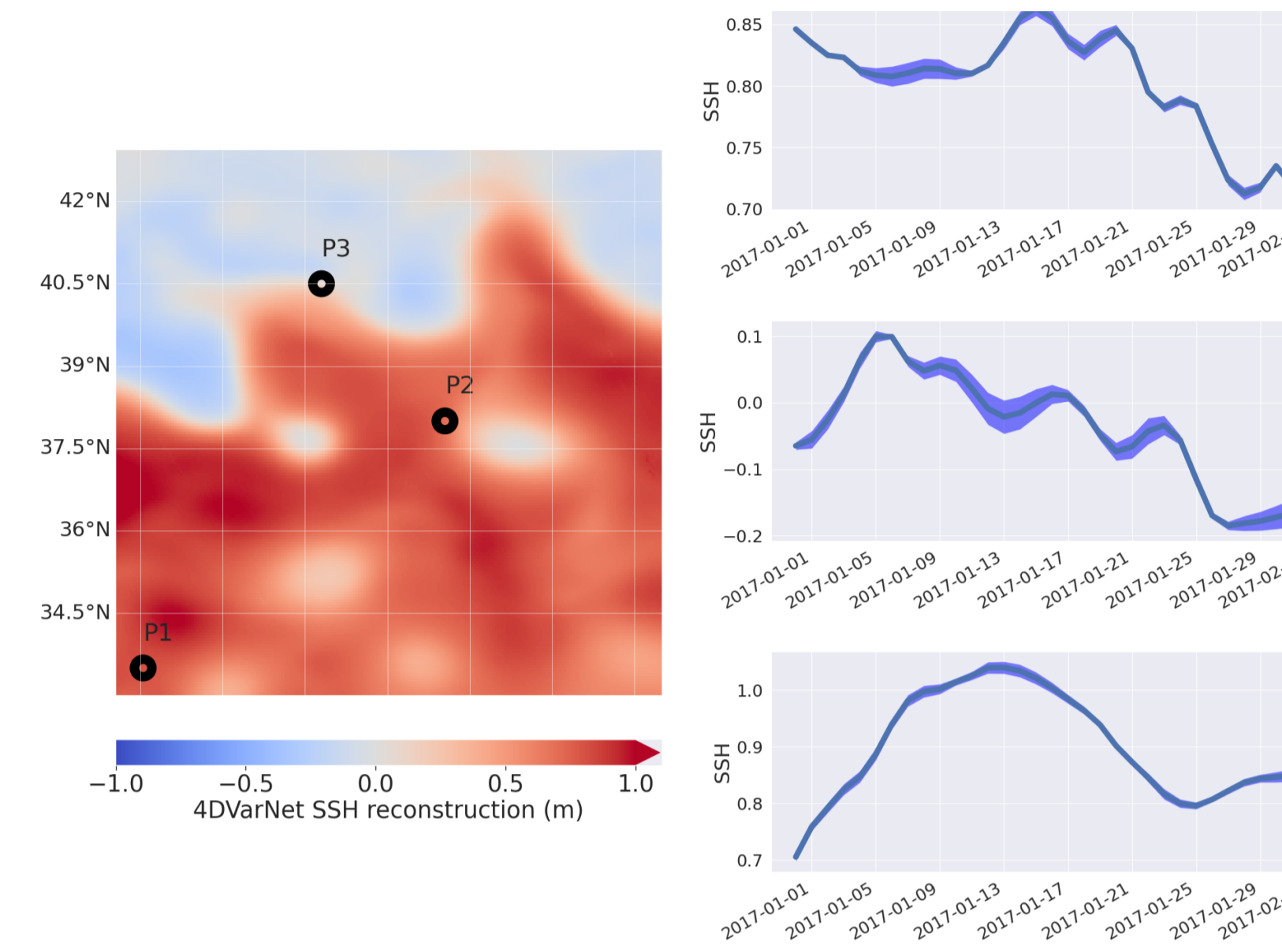


Figure 2: Example of ensemble-based 4DVarNet mean and spreads on the GulfStream domain (2017-01-10) based on the 6 nadirs aggregation dataset

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, 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 kriging.

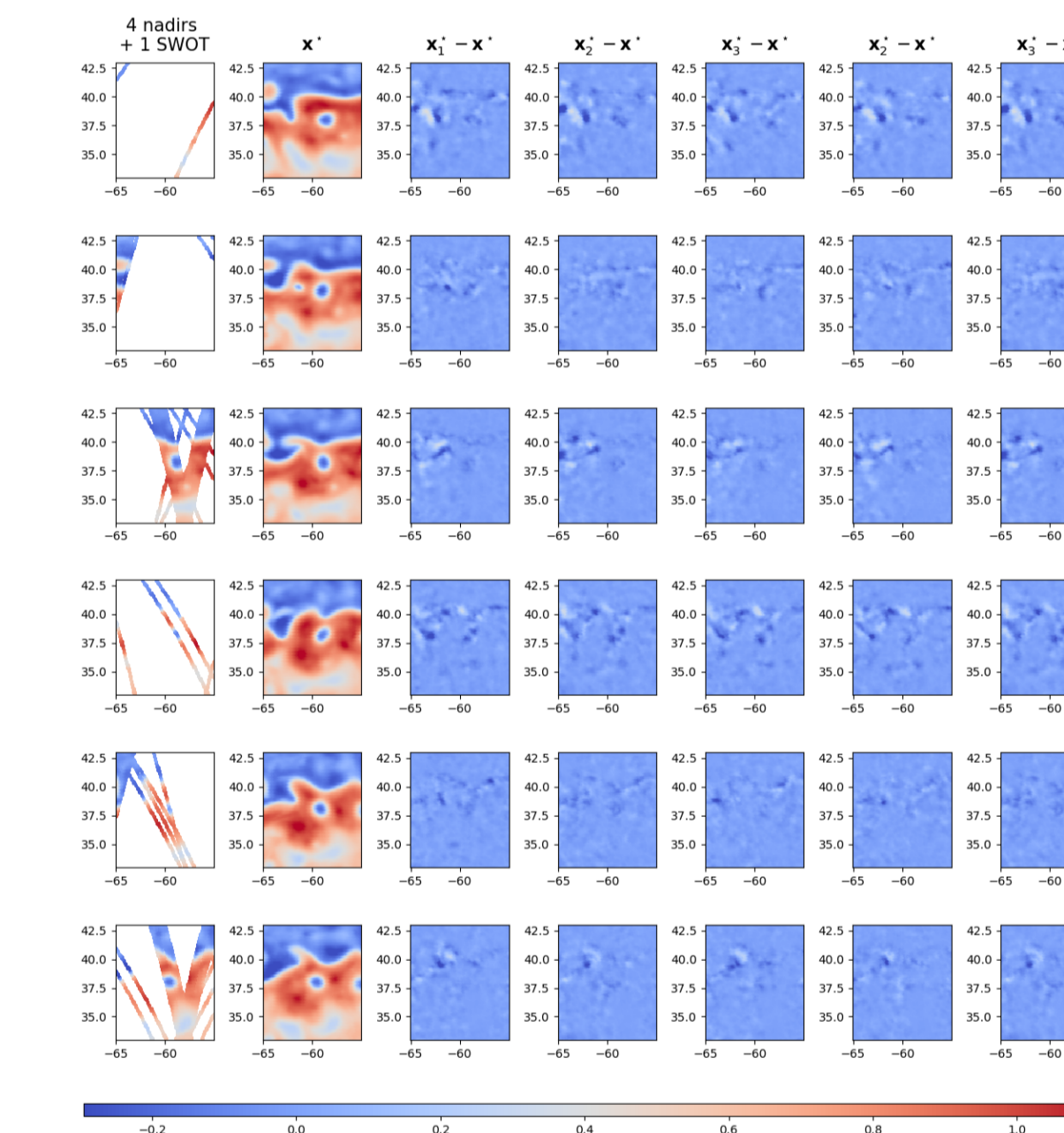


Figure 3: From left to right along the 2020 ODC 42 days test period and every 6 days: Pseudo-observations (nadir and SWOT), Neural variational scheme with SPDE prior (mean), and deviations to the mean for 5 conditional simulations

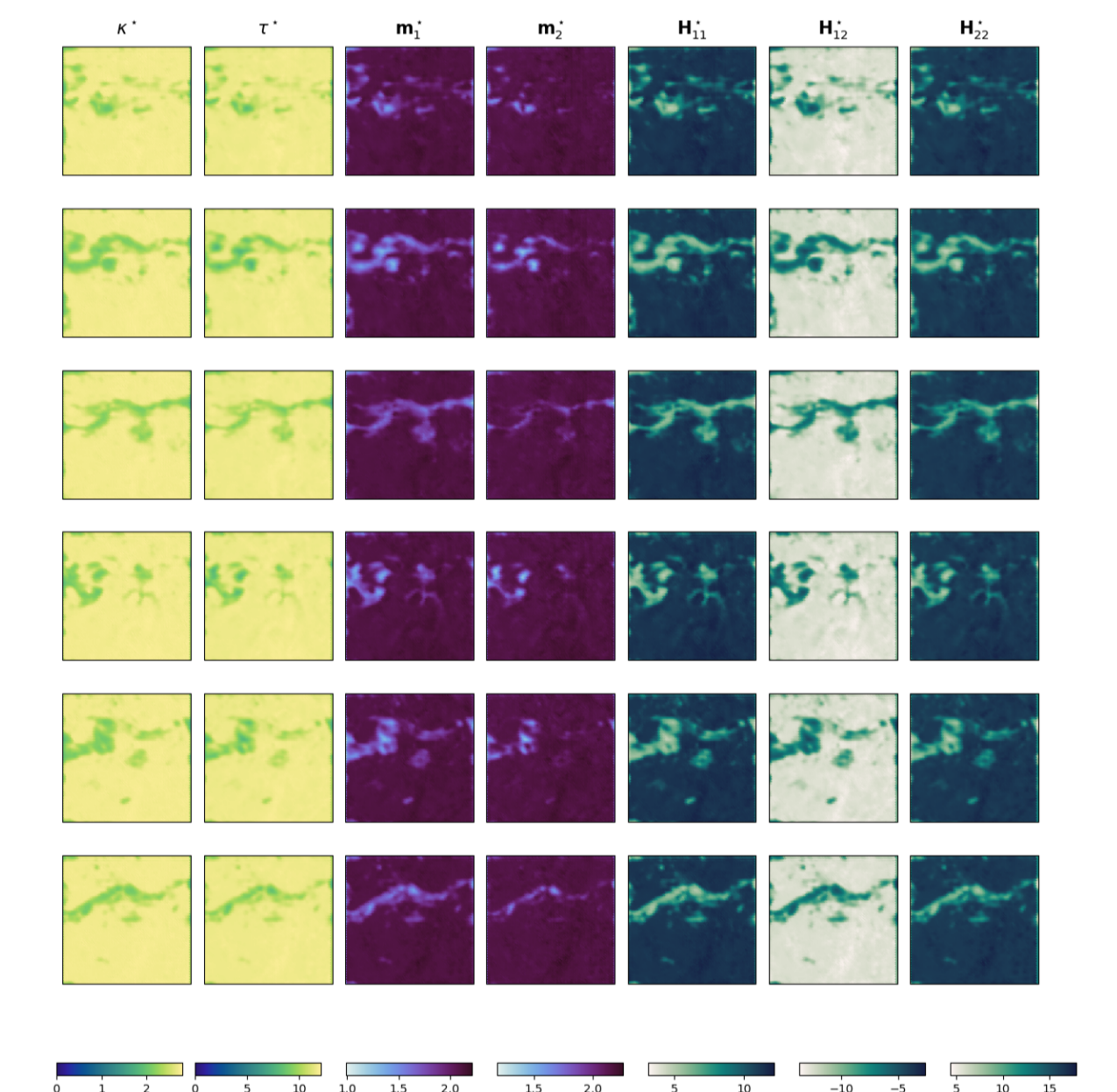


Figure 4: Parameter estimation of the SPDE prior in the neural scheme along the 2020 ODC 42 days test period and every 6 days. From left to right: κ , τ (variance regularization), m^1 , m^2 (advection fields), $H^{1,1}$, $H^{1,2}$ and $H^{2,2}$ (diffusion tensor).

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 978-3-662-05294-5.