Intercomparison of data-driven and learning-based interpolations of along-track Nadir and wide-swath Swot altimetry observations

Maxime Beauchamp, Ronan Fablet, Clément Ubelmann, Maxime Ballarotta, Bertrand Chapron

Article

Intercomparison of data-driven and learning-based interpolations of along-track Nadir and wide-swath Swot altimetry observations

Maxime Beauchamp [1], Ronan Fablet[1], Clément Ubelmann[2], Maxime Ballarotta[3] and Bertrand Chapron[4]

[1] IMT Atlantique Bretagne-Pays de la Loire, Brest, France; maxime.beauchamp@imt-atlantique.fr
[2] Ocean Next, Grenoble, France; clement.ubelmann@ocean-next.fr
[3] Collecte Localisation Satellites (CLS), Ramonville St-Agne, France ; mballarotta@groupcls.com
[4] IFREMER, Plouzané; bertrand.chapron@ifremer.fr
* Correspondence: maxime.beauchamp@imt-atlantique.fr
† This paper is an extended version of our paper published in Climate Informatics 2020.

Version September 28, 2020 submitted to Remote Sens.

Abstract: 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. In this paper, we investigate how they might help to solve for the oversmoothing of the state-of-the-art optimal interpolation (OI) techniques in the reconstruction of sea surface height (SSH) spatio-temporal fields. We focus on two small $10° \times 10°$ GULFSTREAM and $8° \times 10°$ OSMOSIS regions, part of the North-Atlantic basin: the GULFSTREAM area is mainly driven by energetic mesoscale dynamics while OSMOSIS is less energetic but with more noticeable small spatial patterns. Based on Observation System Simulation Experiments (OSSE), we will use the the NATL60 high resolution deterministic ocean simulation of the North Atlantic to generate two types of pseudo altimetric observational dataset: along-track nadir data for the current capabilities of the observation system and wide-swath SWOT data in the context of the upcoming SWOT mission. We briefly introduce the analog data assimilation (AnDA), an up-to-date version of the DINEOF algorithm, and a new neural networks-based end-to-end learning framework for the representation of spatio-temporal irregularly-sampled data. The main objective of this paper consists in providing a thorough intercomparison exercise with appropriate benchmarking metrics to assess if these approaches helps to improve the SSH altimetric interpolation problem and to identify which one performs best in this context. We demonstrate how the newly introduced NN method is a significant improvement with a plug-and-play implementation and its ability to catch up the small scales ranging up to 40km, inaccessible by the conventional methods so far. A clear gain is also demonstrated when assimilating jointly wide-swath SWOT and (agreggated) along-track nadir observations.

Keywords: Data-driven and learning-based approaches ; Interpolation ; Benchmarking ; Nadir & SWOT altimetric satellite data ; Sea surface height (SSH)

1. Introduction

Thanks to the ocean surface remote sensing data acquired by different altimetric missions (TOPEX/Poseidon, ERS-1, ERS-2, Geosat Follow-On, Jason-1, Envisat and OSTM/Jason-2), our understanding of the ocean circulation has been considerably improved over the last decades. But currently, the range of scales over 150km remains inaccessible to altimetric derived products because

of the limited number of altimetric missions and their spatio-temporal sampling [1]. In this context, a very active field of research now consists in taking advantage of the big amount of data and numerical simulations available to overcome these limits of conventional altimetric products, which motivate complementary developments combining high resolution remote sensing and numerical simulations. Over the last years, purely data-driven and artifical intelligence (AI)-based algorithms have just been proposed [2–6] to deal with problems directly related to data assimilation and operational oceanography. More specifically, promising preliminary results have been seen for the sea surface reconstruction and prediction from partial and noisy satellite observations.
In this paper, we propose an intercomparison exercise of several data-driven and learning-based approaches to help for the reconstruction of altimetric fields. As a baseline the DUACS operational processing tool based on well established optimal interpolation (OI) techniques will be considered [7]. In Section 2, we present the case study and its dataset, developed within the BOOST-SWOT project framework (https://meom-group.github.io/projects/boost-swot): the NATL60 high resolution deterministic ocean simulation of the North Atlantic [8] is used as reference to simulate Sea Surface Height (SSH) along-track observations collected by four nadir, which is typically representative of the current observational altimetric capabilities. As an additional feature for the upcoming 2021 SWOT mission, pseudo-SWOT wide-swath observations also following realistic orbits are generated based on the NATL60 simulation. In Section 3, we present the data-driven approaches used in the intercomparison: 1) AnDA, a purely data-driven data assimilation scheme combining a patch-based analog forecasting operator with Kalman-based ensemble data assimilation, 2) VE-DINEOF, an EOF-based iterative method to interpolate in space and time the missing data, and 3) learning-based innovative end-to-end learning techniques that aims to learn jointly the Neural Network (NN) representation of the dynamics coupled with a NN-based solver of the targeted minimization problem. In Section 4, we provide a detailed evaluation of the results obtained over two small regions, GULFSTREAM and OSMOSIS, part of the North-Atlantic basin and labeled with very different energetic dynamics. The GULFSTREAM area is mainly driven by mesoscale processes with large eddies and OSMOSIS is less energetic but the small spatial patterns are more noticeable making its reconstruction also challenging. Last, a discussion based on the evaluation is engaged to give synthetic key results and additional insights for future related works.

2. Case study and data

2.1. NATL60

The Nature Run (NR) used in this work corresponds to the NATL60 configuration [8] of the NEMO (Nucleus for European Modeling of the Ocean) model. It is one of the most advanced state-of-the-art basin-scale high-resolution (1/60°) simulation available today, whose surface field effective resolution is about 7km.
In this work, two specific $10^\circ \times 10^\circ$ GULFSTREAM and $8^\circ \times 10^\circ$ OSMOSIS domains are chosen (see Figure 1) to assess the performance of the data-driven interpolation methods. Over this regions, the Sea Surface Height (SSH), the resolution of the nature run is downgraded to 1/20°, which is enough to capture both the GULFSTREAM mesoscale dynamical regime and the OSMOSIS small scales, while avoiding unnecessary heavy computation time.
The NATL60 nature run will then be used as the reference Ground Truth (GT) in an observing system simulation experiments (OSSE). The pseudo-altimetric nadir and SWOT observational datasets will be generated by a realistic sub-sampling of satellite constellations.

2.2. Nadir

To provide the pseudo-nadir dataset, supposed to be representative of what is a current pre-SWOT observational altimetric dataset, the groundtracks of 4 altimetric missions (TOPEX/Poseidon, Geosat, Jason-1 and Envisat) picked up from the 2003 constellation, are used to interpolate the NATL60

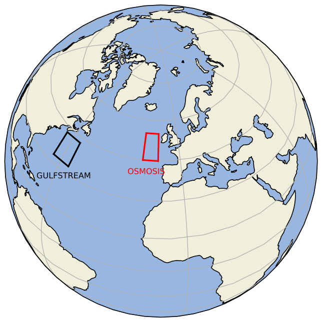
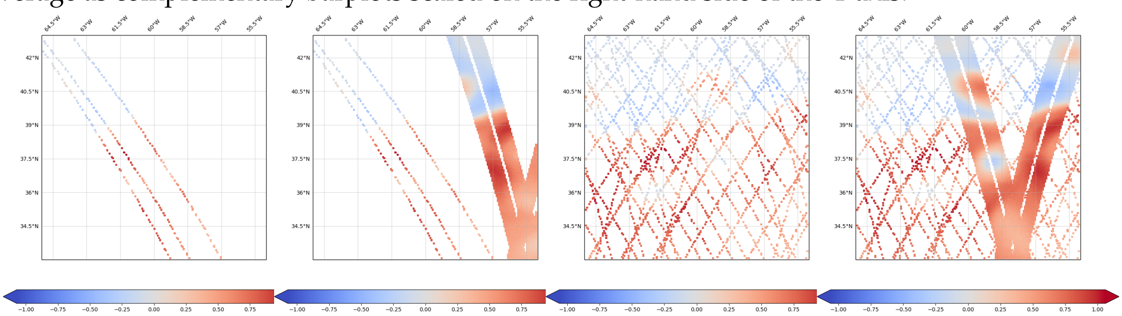


Figure 1. GULFSTREAM and OSMOSIS domain

simulation from October 1st, 2012 to September 29th, 2013, thus covering a whole year of data. A Gaussian white noise with variance $\sigma^2 = 4 - 9\text{cm}^2$ is then added to the interpolated NATL60 simulation by the SWOTsimulator tool to mimic a noise with a spectrum of error consistent with global estimates from the Jason-2 altimeter [9].
Because the space-time interpolations will focus on a daily-basis temporal resolution, we also build nadir pseudo-observations with an additional strategy by accumulating observations over a time window $t_k \pm d$ days centered at time t_k in order to increase the daily nadir spatial sampling. As in [5], we investigate the response of the different interpolation techniques when parameter d is either set to 0 or 5, see Figures 2a and 2c for the corresponding aggregation on August 4, 2013, and August 5, 2013.

2.3. SWOT

In the same line, SWOT-like pseudo observations are also produced by the swotsimulator tool [10] in its swath mode with an along-track and across-track 2km spatial resolution, the same theoretical resolution the upcoming SWOT mission derived products should be able to provide. The nadir mode of the generator also provide pseudo-nadir along track observations though they are not used here. The simulator also adds instrumental noise on the idealized pseudo-SWOT dataset [11,12]. This noise potentially exhibits strong space-time correlations. Thus, the pseudo-SWOT observations are first preprocessed [13] to filter out these correlated components and avoid major issues in the assimilation and/or learning process of the interpolation methods.
Let precise that over the low-latitude GULFSTREAM domain, the SWOT sampling is irregular leading to sequences of several days with only pseudo-nadir observations. This does not happen on the higher latitude OSMOSIS area where the SWOT temporal coverage is more regular. It can be seen along this paper on the time series evaluation figures embedding additional information about the daily spatial coverage as complementary barplots scaled on the right-hand side of the Y-axis.



(a) nadir (d=0) **(b)** nadir (d=0) + swot **(c)** nadir (d=5) **(d)** nadir (d=5) + swot

Figure 2. 0 and 5-days accumulated along-track nadir and wide-swath pseudo-observations on August 4, 2013 (a,b) and August 5, 2013 (c,d)

2.4. DUACS OI products

The DUACS system is an operational production of sea level products for the Marine (CMEMS) and Climate (C3S) services of the E.U. Copernicus program, on behalf of the CNES french space agency. It is mainly based on optimal interpolation techniques whose parameters are fully described in [7]. This methodology has been applied on the previously introduced pseudo along-track nadir and wide-swath SWOT data to generate regularly (0.25°x0.25°) daily gridded maps.

3. Methods

The data-driven methods we are investigating aims at solving smaller scales than operational OI products, more adapted to estimate large scale dynamics. Along this line, we are using in the following a multiscale decomposition:

$$\mathbf{x} = \overline{\mathbf{x}} + d\mathbf{x} + \epsilon \tag{1}$$

and all the interpolations methods used here will work on the anomaly field $d\mathbf{x}$, seen as the difference between the original field $\mathbf{x}$ and the large scales components provided by the OI. In the end, we hope the effective resolution estimated for the anomaly field $d\mathbf{x}$ to be better than the OI-based representation of the dynamics. In what follows, $\mathbf{y}(\Omega) = \{\mathbf{y}_k(\Omega_k)\}$ denotes the observational data corresponding to subdomain $\Omega = \{\Omega_k\} \subset \mathcal{D}$, $\overline{\Omega}$ denotes the gappy part of the SSH field and index k refers to time t_k.

3.1. AnDA

The Analog Data Assimilation (AnDA) is a purely data-driven data assimilation method introducing a statistical operator $\mathcal{A}$ as a substitute for the dynamical model $\mathcal{M}$, leading to the following state-space formulation :

$$\begin{cases} d\mathbf{x}_{k+1} & = \mathcal{A}_{k+1}(d\mathbf{x}_k) + \boldsymbol{\mu}_k \\ d\mathbf{y}_k & = \mathcal{H}_k(d\mathbf{x}_k) + \boldsymbol{\varepsilon}_k \end{cases} \tag{2}$$

The analog forecasting operator $\mathcal{A} : \mathbf{dx}^a_{k-1} \mapsto \mathbf{dx}^f_k$, where superscripts a and f respectively relies to analysis and forecast, is built from the K most similar states to $\mathbf{dx}^a_{k-1}$ in the available past state dynamics catalog, supposed to be large enough to describe the space-time evolution of the processes. More precisely, $\mathbf{dx}^f_k$ is sampled from the Gaussian prior $\mathbf{dx}^f_k|\mathbf{dx}^a_{k-1} \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, where the mean $\boldsymbol{\mu}_k$ and the covariance matrix $\boldsymbol{\Sigma}_k$ are estimated using the so-called locally linear model [2], i.e. a weighted linear regression between the K nearest analogs and their successors.
In the experiments, the diagonal of the observation error matrix $\mathbf{R}_k = Cov(\boldsymbol{\varepsilon}_k)$ is not assumed constant but its values increase according to a parametric function of the hourly time lag between the observation and the day to estimate:

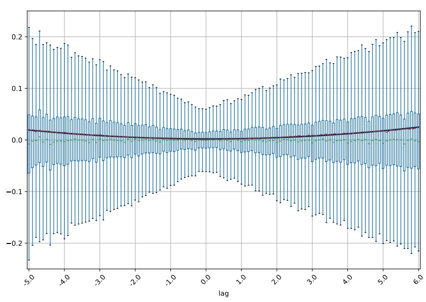


Figure 3. Variance of the observation error $\boldsymbol{\varepsilon}_k$ as a function of the hourly lag between the observation and the day to estimate

As in [5], a patch-based version of AnDA coupled with an EOF-based representation of the individual patches is used. The anomaly field $d\mathbf{x}$ is splitted into 169 vectorized patches $\mathbf{p}(\mathbf{s}, t)$ of sizes $1^\circ \times 1^\circ$, corresponding to 20 pixels $\times$ 20 pixels, with overlapping areas of 5 pixels. An EOF-based decomposition of each individual vectorized anomaly patches is then carried out to deal with the curse of dimensionality. Finally, the whole AnDA algorithm is performed at the patch-level, meaning that both the analog prediction and the assimilation are done onto the lower-dimensional space of their EOF-based representation. A final post-processing step (denoted as post-AnDA) is used to project the prediction onto the original space-time domain and average the overlapping patches to smooth out some blocky artefacts coming from the patch decomposition. On this last point, an improvement can be considered by using a convolutional neural network (CNN) to learn how to reconstruct the whole domain from the set of overlapping patches, as in [6].

3.2. VE-DINEOF

VE-DINEOF is a state-of-the-art interpolation approach [14] using an EOF-based iterative filling strategy. Typically the large-scale component provided by the OI is used (or 0 values if working on the anomaly) as a first guess to fill in the missing data over Ω. After each iteration and until convergence, the field is projected onto the N most significant EOF components of the lower dimensional space and new values for the missing data are used based on the updated reconstruction of the field. Finally, the VE-DINEOF algorithm is here proposed in its patch-based version, in the exact similar setting proposed for AnDA.

3.3. End-to-end NN-learning

An end-to-end learning representation has recently been introduced in [15] to deal with image sequences involving potentially large missing data rates. In this framework, an energy-based representation U_θ to minimize is introduced :

$$U_\psi(d\mathbf{x}) = \|d\mathbf{x} - \psi\left(d\mathbf{x}\right)\|^2 \tag{3}$$

where the operator $\psi = \psi_\theta$ denotes a NN-based representation of the underlying processes and $\|.\|^2_\Omega$ refers to the L2 norm evaluated on subdomain Ω. Within a Bayesian framework, the interpolator I_{U_ψ} of the irregular space-time dataset $\{d\mathbf{y}_k(\Omega_k)\}$, referred ad the hidden state in a classic data assimilation framework, can be obtained by solving the minimization statement:

$$\widehat{d\mathbf{x}_k} = I_{U_\psi}\left(d\mathbf{y}_k(\Omega_k)\right) = \arg\min_{d\mathbf{x}} U_\psi\left(d\mathbf{x}\right) \tag{4}$$

such that $I_{U_\psi}\left(d\mathbf{y}_k(\Omega_k)\right) = d\mathbf{x}_k$ if no observational error are considered.
Last, for a specific definition of interpolator I, the learning problem for optimizing parameters θ of the NN representation ψ can be stated as the minimization of the reconstruction error for the whole observed data time series:

$$\widehat{\theta} = \arg\min_\theta \sum_k \left\| d\mathbf{y}_k(\Omega_k) - I_{U_\psi}\left(d\mathbf{y}_k(\Omega_k)\right)\right\|^2_{\Omega_k} \tag{5}$$

3.3.1. Architecture

Typically, two NN-based energy parametrizations are considered:

1. First, a classic convolutional auto-encoders (ConvAE) representations $\psi(\cdot) = \phi_D(\phi_E(\cdot))$ where the encoding operator ϕ_E maps the anomaly state $d\mathbf{x}$ onto a lower-dimensional space and the decoder ϕ_D has to project this encoded representation in the original space. It involves the following encoder architecture: five consecutive blocks with a Conv2D layer, a ReLu layer and a 2x2 average pooling layer, the first one with 40 filters and the following four ones with two

times the number of filters of the previous Conv2D layer (i.e. 80, 160 and 320 filters), and a final linear convolutional layer with 20 filters. The output of the encoder is 5x5x40. The decoder involves a Conv2DTranspose layer with ReLu activation for an initial 20x20 upsampling stage a Conv2DTranspose layer with ReLu activation for an additional 2x2 upsampling stage, a Conv2D layer with 40 filters and a last Conv2D layer with 22 filters (the length of the image time series times the number of covariates - the OI here - used in the model). All Conv2D layers use 3x3 kernels. Overall, this model involves $\approx$ 600,000 parameters.

2. GE-NN: Second, NN-based Gibbs-Energy (GENN) representations where $d\mathbf{x}_s$, the anomaly observed at location $\mathbf{s} \in \mathcal{D}$, is supposed to be explained by the potential function $\psi(d\mathbf{x}_{\delta s})$ with δs a predefined neighbourhood of site $\mathbf{s}$, thus relating this representation to Markovian priors embedded in CNNs. A low energy-state $U_\psi(d\mathbf{x}) = \int_{\mathcal{D}} U_\psi(d\mathbf{x}_s) ds$ over the entire domain $\mathcal{D}$ ensures to provide a good state space reconstruction. Regarding the architecture involved, the following scheme is used: an initial 4x4 average pooling, a Conv2D layer with 40 filters, 11x11 kernel, ReLu activation and a zero-weight constraint on the center of the convolution window, a 1x1 Conv2D layer with 40 filters, a ResNet composed of an initial mapping to an initial 200x200x(5x40) space with a Conv2D+ReLu layer and a linear 1x1 Conv2D+ReLu layer with 40 filters. Last, a final 4x4 Conv2DTranspose layer with a linear activation for an upsampling to the input shape is considered. GE-NN involves 10 residual units for a total of $\approx$ 450,000 parameters.

We may point out that the considered GENN architecture is not applied to the initial 0.05° resolution but to downscaled grids by a factor of 4 through the introduced average pooling. First, this makes the comparison easier with the 0.25° DUACS OI resolution. Second, the application of GENNs to the finest resolution showed a lower performance, thus implying that considering a scale-selection problem when applying a given prior is mandatory. The upscaling involves the combination of a Conv2DTranspose layer with 11 filters, a Conv2D layer with a ReLu activation with 22 filters and a linear Conv2D layer with 11 filters.

3.3.2. Fixed-point solver

Based on this NN-parametrization of operator ψ and related energy/cost function U_ψ, an iterative fixed-point solver can be used to optimize parameters θ of the NN-model (ConvAE or GENN) ψ w.r.t cost U_ψ, see the corresponding sketch in Figure 4:

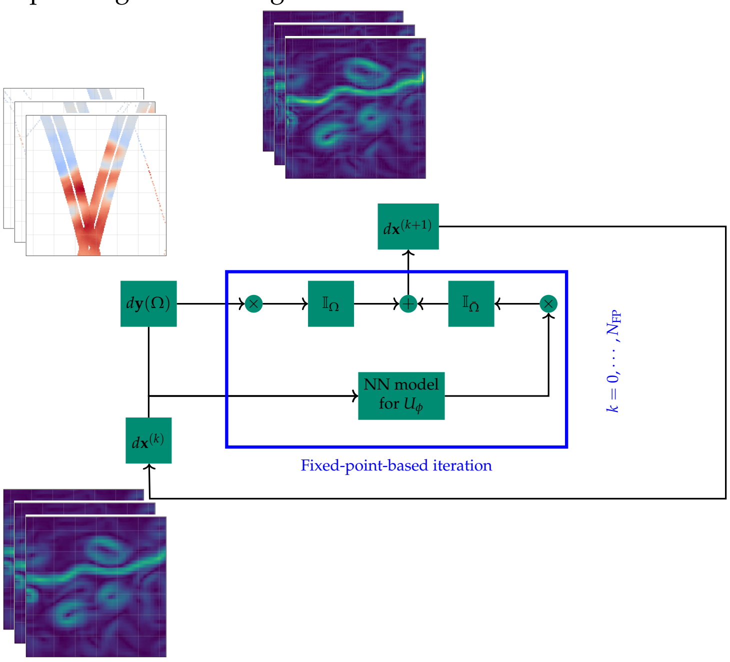


Figure 4. Sketch of the iterative fixed-point algorithm

The underlying idea is rather similar to the DINEOF approach, see Section 3.2, leading to the iterative update of the hidden state:

$$\begin{cases} \mathbf{x}^{(k+1)} & = \psi\left(\mathbf{x}^{(k)}\right) \\ \mathbf{x}^{(k+1)}\left(\Omega\right) & = \mathbf{y}\left(\Omega\right) \\ \mathbf{x}^{(k+1)}\left(\overline{\Omega}\right) & = \mathbf{x}^{(k+1)}\left(\overline{\Omega}\right) \end{cases}$$

It is parameter-free and easily implemented as a NN in a joint solution with the NN-parametrization of U_θ for the interpolation problem. The two NN-architectures are then referred as FP-ConvAE and FP-GENN. Let note that additional improvements are expected when using an iterative gradient-based formulation of the solver, where the gradient of U_ψ is replaced by a ConvNet or LSTM unit $G(\mathbf{x} - \psi(\mathbf{x}))$, thus enabling to solve jointly for the parametrization of ψ and G. Complementary results on SST datasets regarding this point can be found in [15]. Let precise that during the learning phase, anomaly image time series $d\mathbf{x}_{k \pm dT} = d\mathbf{x}_{k-dT:k+dT}$ are built with time window $dT = 5$, centered on time t_k, leading to image time series of length 11. Last, the above-mentioned works are generalized to establish a connection between 4DVAR variational data assimilation and joint learning of models and solvers in [16].

4. Evaluation

4.1. *Experimental/benchmarking setup*

A specific aspect of this work consists in the period of data available because the NATL60 native run is only one-year long which is relatively short in comparison with the training period typically used in the previous related work mentioned in Introduction. To get around this issue, we decide to build four 20-days long validation period homogeneously distributed along this one-year dataset (see the starting dates reported on Figures 5 and 6), supposed to be representative of the different seasonality effects that may be encountered during the year.

Regarding the metrics used in the intercomparison exercise, daily normalized RMSE (nRMSE) time series are first provided: they give a quick overview of the potential gain obtained with the data-driven interpolators. Additional correlation and variances scores are also computed, then all displayed together with the RMSE as Taylor diagrams. We also provide three other indicators, namely the global reconstruction score (R-score) for the known SSH field areas (Ω), the interpolation performance (I-score) for the missing data areas ($\overline{\Omega}$), and the reconstruction performance of the trained NN-based representation of the SSH dynamics for FP-ConvAE and FP-GENN when applied to gap-free SSH fields (AE-score). Last, signal-to-noise ratios are also computed in the spectral domain, in particular to assess up to which spatial scale the different interpolators are able to reproduce the ground truth. Table 1 provides all the formulas used to compute the above mentioned metrics used along Section 4.

	Name	Formula
Temporal domain	RMSE	$\text{RMSE}(t_k)=\sqrt{\frac{1}{\lvert\widetilde{\mathcal{D}}\rvert}\sum_{\widetilde{\mathcal{D}}}(\mathbf{x}_k-\hat{\mathbf{x}}_k)^2}$
	Error variance	$\sigma^2_{\mathbf{x}-\hat{\mathbf{x}}}(t_k)=\frac{1}{\lvert\widetilde{\mathcal{D}}\rvert}\sum_{\widetilde{\mathcal{D}}}\left[(\mathbf{x}_k-\hat{\mathbf{x}}_k)-\overline{(\mathbf{x}_k-\hat{\mathbf{x}}_k)}\right]^2$
	Correlation	$\text{COR}(t_k)=\frac{\text{Cov}(\mathbf{x}_k,\hat{\mathbf{x}}_k)}{\sigma(\mathbf{x}_k)\sigma(\hat{\mathbf{x}}_k)}$
	Reconstruction score	$\text{R-score} = 100\times\left(1-\frac{\sum_{\Omega}((\mathbf{x}-\overline{\mathbf{x}})-(\hat{\mathbf{x}}-\overline{\hat{\mathbf{x}}}))^2}{\sum_{\Omega}(\mathbf{x}-\overline{\mathbf{x}})^2}\right)$
	Interpolation score	$\text{I-score} = 100\times\left(1-\frac{\sum_{\overline{\Omega}}((\mathbf{x}-\overline{\mathbf{x}})-(\hat{\mathbf{x}}-\overline{\hat{\mathbf{x}}}))^2}{\sum_{\overline{\Omega}}(\mathbf{x}-\overline{\mathbf{x}})^2}\right)$
	Auto-encoder score	$\text{AE-score} = 100\times\left(1-\frac{\sum_{\widetilde{\mathcal{D}}}((\mathbf{x}-\overline{\mathbf{x}})-(\psi(\mathbf{x})-\overline{\psi(\mathbf{x})}))^2}{\sum_{\widetilde{\mathcal{D}}}(\mathbf{x}-\overline{\mathbf{x}})^2}\right)$
Spectral domain	RAPS	$\text{RAPS}(\lambda)=\text{DSP}(\hat{\mathbf{x}}_k)(\lambda)$
	Signal-to-Noise Ratio	$\text{SNR}(\lambda)=\frac{\text{DSP}(\mathbf{x}_k-\hat{\mathbf{x}}_k)(\lambda)}{\text{DSP}(\mathbf{x}_k)(\lambda)}$

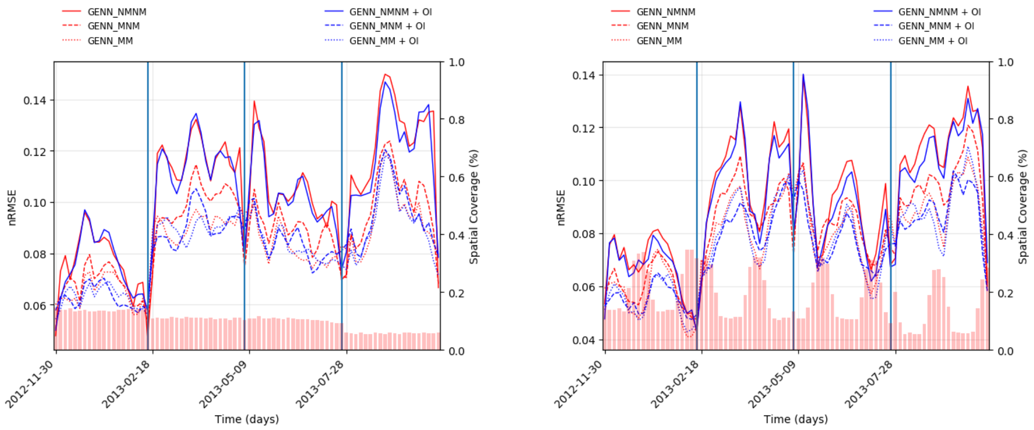
Table 1. Temporal and spectral statistics used to assess the performance of the interpolators in the Observation System Simulation Experiment

where $\widetilde{\mathcal{D}}$ denotes the gridded version of domain $\mathcal{D}$ and $\lvert\widetilde{\mathcal{D}}\rvert$ is then the number of grid nodes of $\widetilde{\mathcal{D}}$. DSP denotes the density power spectrum, as introduced by Welch [17].

4.2. GULFSTREAM

We first have to discuss the time window parameter d related to the aggregation of along-track data over a specific day t_k, see Section 2.2. A same value of this parameter may not be optimal for all the interpolators: AnDA exhibits a better performance when considering only along-track nadir data of the day ($d = 0$), thus contradicting the previous optimal results of $d = 5$ found by [5] over the Mediterranean sea, which may indicate AnDA responds differently to the along-track aggregation strategy depending on the energetic dynamical regime of the region. On the other hand, both FP-ConvAE and FP-GENN interpolators performs better (not shown here) by aggregating nadir data over a 5-day time window. As a consequence, the results presented in what follows will use value of $d = 0$ for AnDA and VE-DINEOF and $d = 5$ for FP-ConvAE and FP-GENN.
Next, to evaluate the behaviour of the different interpolators on both along-track nadir samplings and their fusion with wide-swath SWOT datasets and make the comparison possible, we have to preliminary define if the NN-based interpolator will be used under a supervised or unsupervised learning strategy. Figure 5 depicts how the FP-GENN interpolator performs using nadir data (a) or their joint use with SWOT (b), according to the input and target data used for the training. Six possible configurations have been tested. Two supervised versions using the gap-free NATL60 simulations as target, and either the pseudo-observations or the gap-free maps as input, respectively denoted as FP-GENN-MNM and FP-GENN-NMNM. A fully unsupervised FP-GENN-MM version is also used in which both input and target are only made of the pseudo-observations. These three configurations are also tested when adding the DUACS OI product as a covariate for input data, because we think that this may give a prior information about how the anomaly field $d\mathbf{x}$ is distributed. Within this part-GULFSTREAM domain, we clearly see the best performance is obtained by the unsupervised

configuration of FP-GENN: it is a keypoint result because the learning network abilities seems to be better when it is fully data-driven, meaning that it benefits from its knowledge of the spatio-temporal location and occurence of the data, which is a fairly new avenue for data assimilation related problems. The use of the OI as a covariate improve the FP-GENN behaviour but not systematically.

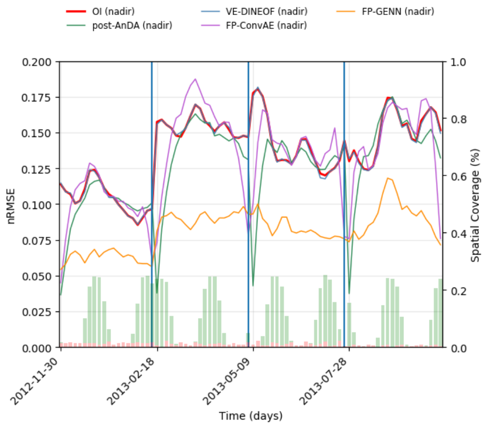


(a) nadir **(b)** nadir+swot

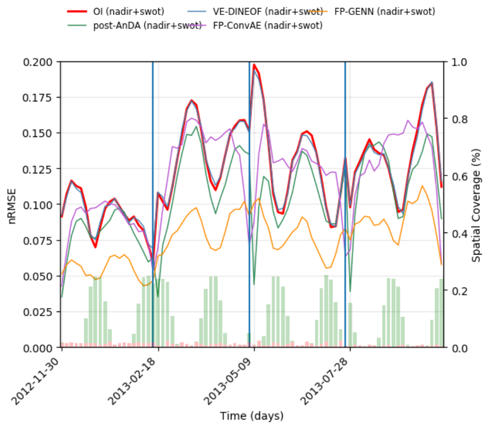
Figure 5. Daily spatial nRMSE computed on the 80-days non-continuous validation period for the six supervised/unsupervised FP-GENN configurations. The spatial coverage of 0-days accumulated along-track nadir (a) expanded with wide-swath SWOT data (b) is provided by the red-colored barplot

Intriguingly, if the joint use of nadir and SWOT data generally improves the results, using only nadir in the unsupervised FP-GENN may yield to a better reconstruction the days where no SWOT data is available. We hope that a longer training period could help the network to learn from the masking periodicity of 2D wide-swath data. Based on these first results, the FP-GENN interpolator is used in its unsupervised configuration with OI used as a covariate. Because FP-ConvAE generally shows lower performance, probably because auto-encoders may not be relevant for the reconstruction of fine-scale processes, it will be used in the following in its mid-supervised configuration (FP-ConvAE-MNM) as a low-rated NN-scheme among the NN-based interpolators.

Figure 6 presents the daily nRMSE of the different interpolators: it can be seen how FP-GENN significantly outperforms the conventional OI-based interpolator, but also the other data-driven algorithms used in the experiment. In addition, the FP-GENN mapping error seems to be more stable along time than the OI, meaning that in case of a missing altimetric mission, the error would also remain more stable. AnDA still remains quite efficient at the very beginning of the four 20-days validation period, which is probably related to a strong persistence of the mesoscale dynamics of the SSH over the region. In other words, the one-year catalog (minus the 80 validation days) obviously enable to build a good analog forecasting operator when knowing the short-term dynamics, but its accuracy quickly decays afterwards, which may not be fair for AnDA that probably requires longer simulations-based catalog in this low-latitude GULFSTREAM region with large Rossby radius of deformation.



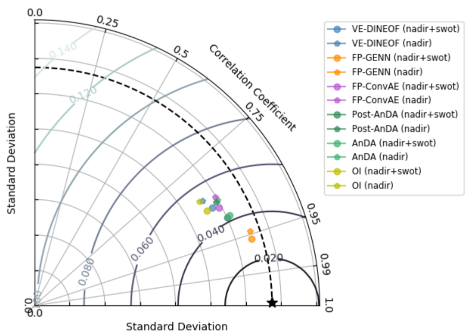
(a) nadir



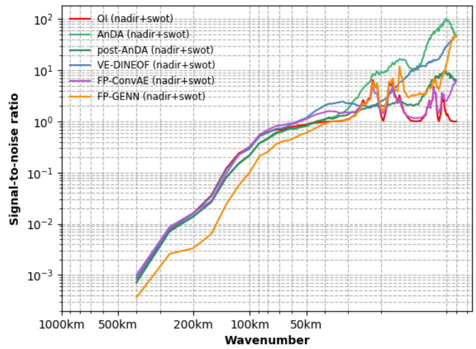
(b) nadir+swot

Figure 6. Daily spatial nRMSE computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN. The spatial coverage of 0-days accumulated along-track nadir and wide-swath SWOT data are respectively provided by the red and green-colored barplots

The Taylor diagram in Figure 7a, here calculated over the 80 validation days and focusing only on small-scale structures by applying a high-pass filter that spectrally separates the horizontal scales ranging in the order of 150km, also confirms our first findings.



(a) Taylor diagram



(b) Signal-to-noise ratio

Figure 7. Taylor diagram and Signal-to-noise ratio computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN computed for both nadir use only and joint assimilation/learning with wide-swath SWOT data

In Table 2, R/I/AE-scores are applied to both SSH (after application of a retrieving high-pass filter to keep only the small scales information) and its gradient (module). Regarding the R-scores, AnDA and VE-DINEOF are often the best way to keep track of the known areas, which is not surprising since these two methods makes an explicit use of the observational altimetric data into their mapping process. When looking at the I-scores, where no data is available, FP-GENN now clearly stands out from the other interpolators, which motivate its future use for irregularly-sampled data with large missing data rates. In addition, because its reconstruction scores remain overall satisfactory, in particular when considering the joint learning on nadir and SWOT data, these results are supplementary arguments on account of this markovian-related NN-based formulation.

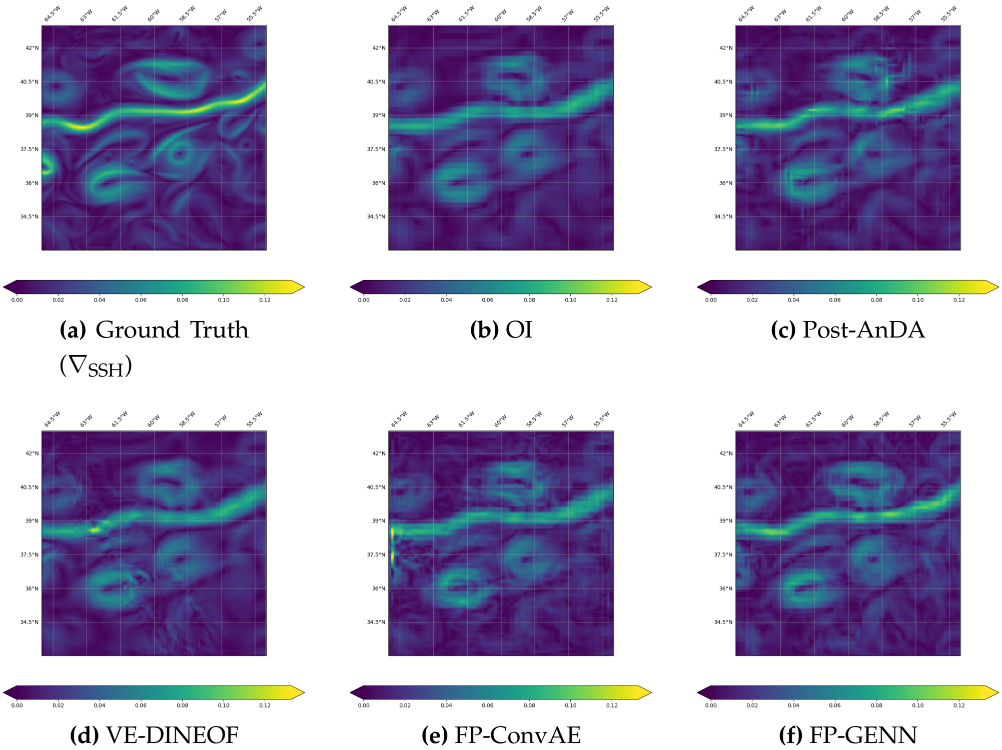
	Model type	R-score	I-score	AE-score
nadir	OI	87.32	72.17	–
	AnDA	94.85	77.91	–
	VE-DINEOF	**96.11**	72.72	–
	FP-ConvAE	87.82	76.32	82.85
	FP-GENN	91.78	**84.56**	93.15
nadir + SWOT	OI	93.25	74.25	–
	AnDA	96.05	83.55	–
	VE-DINEOF	**97.13**	75.28	–
	FP-ConvAE	80.63	77.51	83.26
	FP-GENN	96.49	**90.13**	95.58

	Model type	R-score	I-score	AE-score
nadir	∇_{OI}	78.03	75.97	–
	∇_{AnDA}	**85.56**	79.14	–
	$\nabla_{\mathrm{VE-DINEOF}}$	82.69	75.61	–
	$\nabla_{\mathrm{FP-ConvAE}}$	77.80	76.81	75.89
	$\nabla_{\mathrm{FP-GENN}}$	81.05	**80.56**	84.24
nadir + SWOT	∇_{OI}	73.83	75.78	–
	∇_{AnDA}	**89.89**	82.88	–
	$\nabla_{\mathrm{VE-DINEOF}}$	88.19	76.69	–
	$\nabla_{\mathrm{FP-ConvAE}}$	76.20	76.49	75.84
	$\nabla_{\mathrm{FP-GENN}}$	86.96	**85.33**	88.23

Table 2. SSH and SSH gradient field R/I/AE-scores computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for both nadir use only and joint assimilation/learning with wide-swath SWOT data

Last, when computing the radially averaged power spectra as a spatial domain averaged over the 80-days validation period and the associated signal-to-noise ratio for joint use of along-track nadir with SWOT data (Figure 7b), we observe that AnDA and FP-GENN lead to a better constraint of the SSH spectrum compared to the actual OI capabilities. In particular, FP-GENN produces a spectrum closer to the ground truth real spectrum, by catching up the submesoscale range up to 60km (when picking up signal-to-noise ratio equals to 0.5) when considering a joint learning from along-track nadir and additional wide-swath SWOT data. Let note on Figure 7b the importance of the patch-based AnDA post-processing on its performance which clearly appears on the spectra: its overestimation by the blocky patch-based AnDA rough outputs is partly mitigated thanks to the smoothing produced by averaging the patches overlapping areas. This result may certainly be further improved, for instance by training a CNN rather than using a simple average-based smoothing.

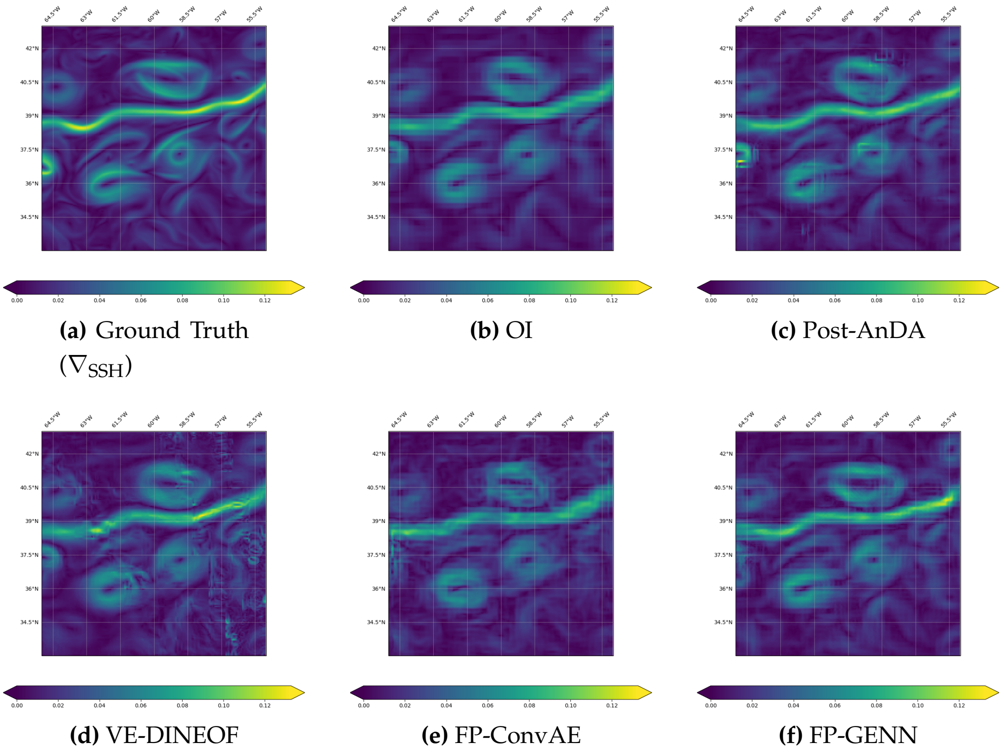
To further enhance the vizualisation of the improvements brought by the different interpolators, Figures 8 and Figure 9 depict the velocity ground truth as well as its global reconstruction based on OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN with both single along-track nadir data and joint use with wide-swath pseudo-observations on August 4, 2013. In Appendix A, complementary figures are provided for the SSH on the same day. To support what has already been said through the performance analysis previously discussed, FP-GENN using 5-days accumulated nadir observations appears closer to the groud truth SSH field than the reconstruction obtained with FP-ConvAE using a similar solver but a simple auto-encoder representation of the dynamics. The latter clearly oversmoothes the true field and also exhibits some unnecessary artefacts on the SSH gradient thus explaining the noisy-related small scale energies on the spectra. The same artefacts appears on the VE-DINEOF mapping which exhibits discontinuities between the known wide-swath-informed areas and the filled missing data. Last, AnDA also behaves well, especially because the wide-swath SWOT data coverage on this specific day is important, getting its performance closer to FP-GENN than the day without the 2D-SWOT information.



(a) Ground Truth (∇_{SSH}) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

Figure 8. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only



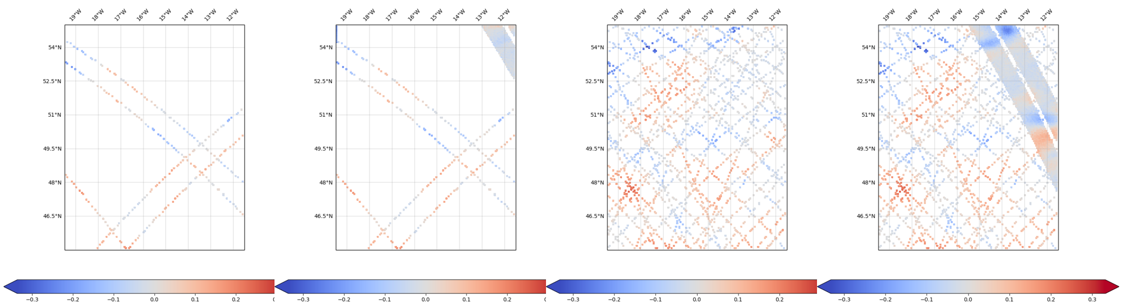
(a) Ground Truth (∇_{SSH}) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

Figure 9. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

4.3. OSMOSIS

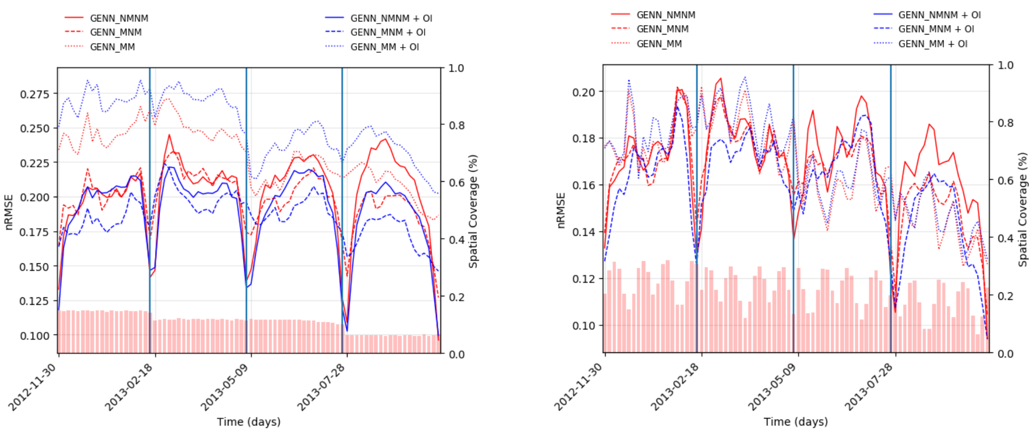
As already been done for the GULFSTREAM domain, we investigate how the different interpolation techniques behaves when varying nadir aggregation parameter d Figures 10a and 10c for the corresponding aggregation on August 4, 2013, and August 5, 2013.



(a) nadir (d=0) **(b)** nadir (d=0) + swot **(c)** nadir (d=5) **(d)** nadir (d=5) + swot

Figure 10. 0 and 5-days accumulated along-track nadir and wide-swath pseudo-observations on August 4, 2013 (a,b) and August 5, 2013 (c,d)

The daily nRMSE as a function of the along-track nadir time window parameter d (not shown here) leads to the same GULFSTREAM-related optimal values, namely ANDA behaves best when considering only the data restained to the targetted day t_k and both FP-ConvAE and FP-GENN performs better with $d = 5$.
Regarding the GENN configuration, the fully unsupervised FP-GENN-MM + OI configuration, the one using only the observations as both target and input (with OI as additional covariate) does not seem to perform well on the OSMOSIS domain, while it was the best option in the GULFSTREAM region.



(a) nadir **(b)** nadir+swot

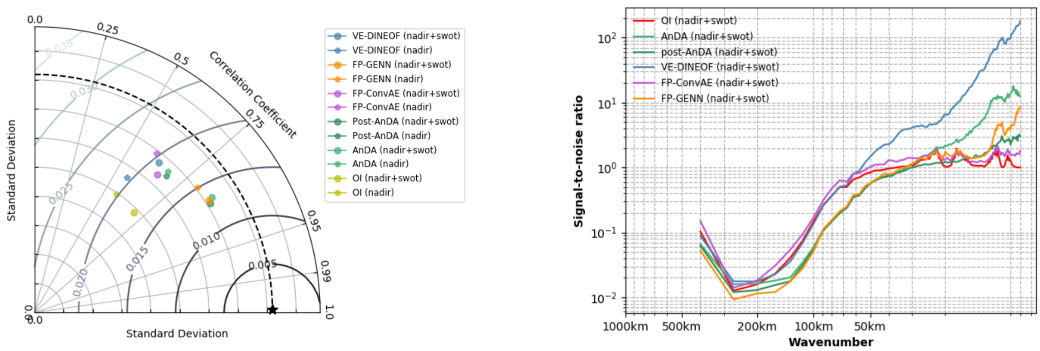
Figure 11. Daily spatial nRMSE computed on the 80-days non-continuous validation period for the six supervised/unsupervised FP-GENN configurations. The spatial coverage of 0-days accumulated along-track nadir (a) expanded with wide-swath SWOT data (b) is provided by the red-colored barplot

It is especially noticeable on the 20-days long time series, see Figure 11. However, this result should be qualified because when replicating the same preliminary work to find the best FP-GENN configuration but with no observation errors, see Figure A12 in Appendix B, the unsupervised configuration is again the best solution. Thus, on this less energetic OSMOSIS domain, but with more discernable fine scales, the observational errors seems to have much more consequences than when considering a domain mainly driven by mesoscale energies. In this Section, we then selected the supervised configuration FP-GENN-MNM + OI, in which the gap-free ground truth is used as target in the learning process, which does not prevent its use for future operational context, since the GENN inputs are still made of purely observational data: along this line, this type of configuration is here similar to the AnDA setup that needs both observation data and gap-free data to be operated.

On Figure 12 of the daily nRMSE obtained with our set of data-driven interpolators along the validation period, it can be seen that using AnDA with along-track nadir data and wide-swath SWOT observations gets the best scores, which is confirmed on the Taylor diagram (Figure 13a) and also with R/I/AE-scores in Table 3. Still, FP-GENN performs in a very similar way and and the single use of nadir data is largely favorable to FP-GENN-MNM + OI.

(a) nadir **(b)** nadir+swot

Figure 12. Daily spatial nRMSE computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN. The spatial coverage of 0-days accumulated along-track nadir and wide-swath SWOT data are respectively provided by the red and green-colored barplots



(a) Taylor diagram **(b)** Signal-to-noise ratio

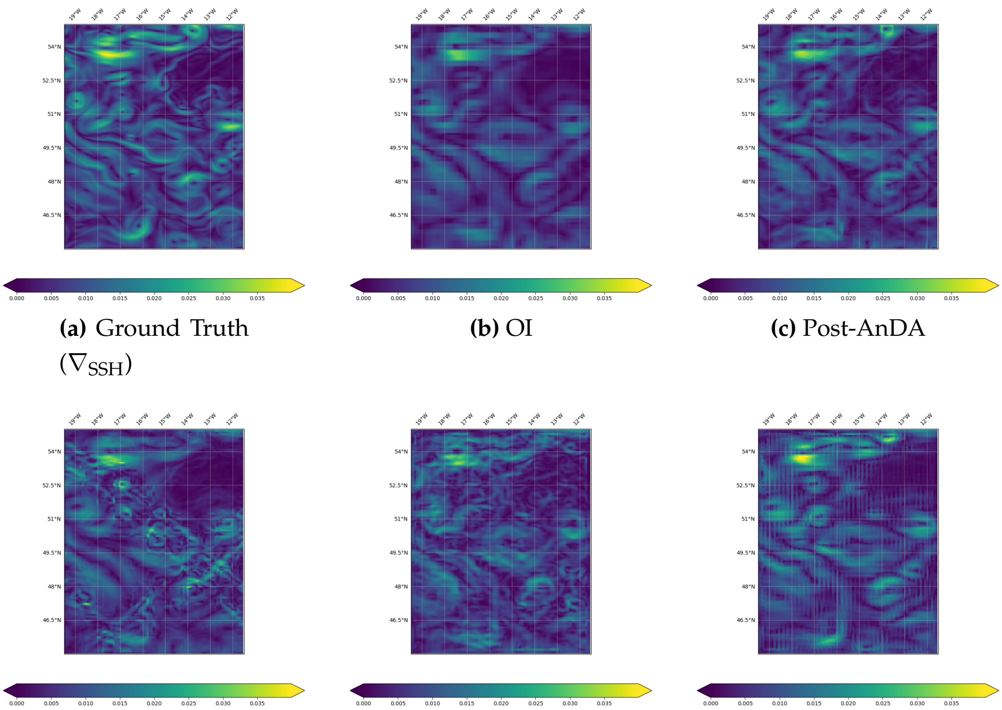
Figure 13. Taylor diagram and Signal-to-noise ratio computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN computed for both nadir use only and joint assimilation/learning with wide-swath SWOT data

On the spectral analysis in Figure 13b, the signal-to-noise ratio of FP-GENN and AnDA indicates a capability to retrieve spatial scales up to 50-60km, while the OI clearly only catches again the spatial scales over 100km. Again, let remain that when no observational errors are introduced, see Figure A14b in Appendix B, the fully unsupervised configuration of FP-GENN still behaves better. The single use of along-track nadir data clearly downgrades the performance of interpolations even if the gain remains significant for FP-GENN.

	Model type	R-score	I-score	AE-score
nadir	OI	42.05	32.11	_
	AnDA	58.85	47.02	_
	VE-DINEOF	26.29	30.61	_
	FP-ConvAE	37.20	31.67	47.77
	FP-GENN	**67.94**	**62.52**	80.40
nadir + SWOT	OI	54.21	47.75	_
	AnDA	**81.15**	**70.91**	_
	VE-DINEOF	69.08	32.98	_
	FP-ConvAE	45.15	42.70	47.93
	FP-GENN	77.16	69.56	83.08

	Model type	R-score	I-score	AE-score
nadir	∇_{OI}	48.83	47.57	_
	∇_{AnDA}	**58.78**	**55.17**	_
	$\nabla_{\mathrm{VE-DINEOF}}$	33.11	35.28	_
	$\nabla_{\mathrm{FP-ConvAE}}$	32.15	35.87	41.24
	$\nabla_{\mathrm{FP-GENN}}$	50.53	52.12	60.41
nadir + SWOT	∇_{OI}	36.83	47.30	_
	∇_{AnDA}	**72.35**	**67.59**	_
	$\nabla_{\mathrm{VE-DINEOF}}$	22.08	24.90	_
	$\nabla_{\mathrm{FP-ConvAE}}$	38.22	43.13	42.03
	$\nabla_{\mathrm{FP-GENN}}$	56.29	59.21	67.69

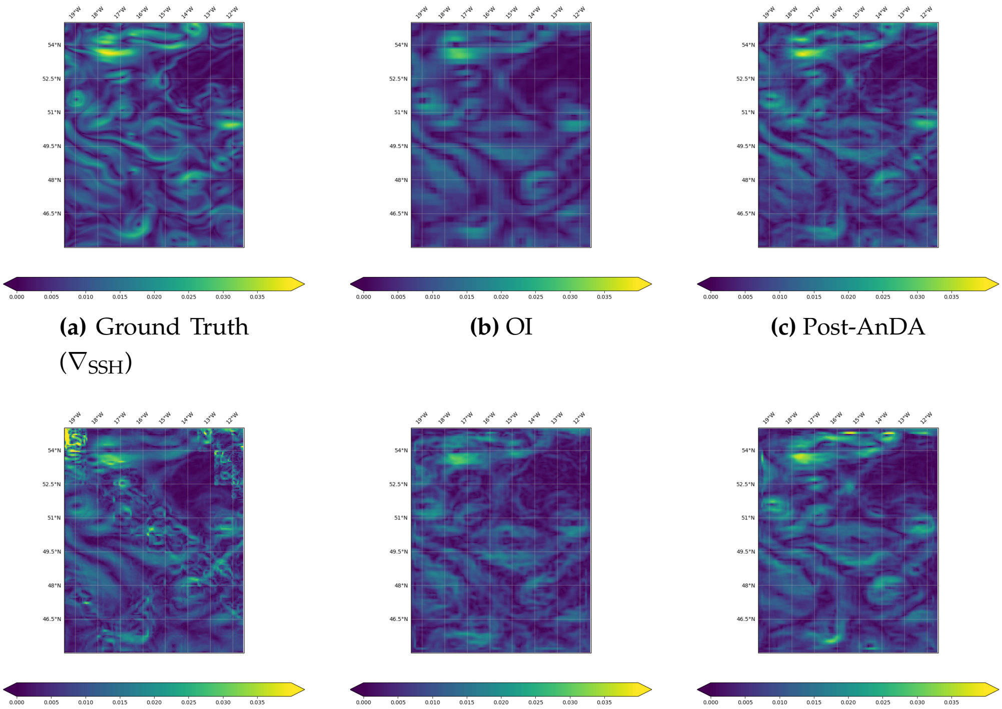
Table 3. SSH and SSH gradient field R/I/AE-scores computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for both nadir use only and joint assimilation/learning with wide-swath SWOT data



(a) Ground Truth (∇_{SSH}) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

Figure 14. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only



(a) Ground Truth (∇_{SSH}) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

Figure 15. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

5. Discussion

In this study focusing on how data-driven and learning-based algorithms may help to improve the reconstruction performance of altimetric fields generally given by a state-of-the-art optimal interpolation (OI) baseline, here provided by the DUACS processing chain, we used two small areas with different energetic dynamics: the $10° \times 10°$ GULFSTREAM domain mainly driven by mesoscale processes and the $8° \times 10°$ OSMOSIS domain, less energetic but labelled with more small scale structures. Based on the NATL60 numerical simulations [8], some experiments were designed in which pseudo observational along-track nadir and wide-swath SWOT realistic datasets are generated. Because the DUACS OI [7] of these pseudo-observations is used as the reference, all the investigated methods are applied in a multi-scale decomposition framework where the anomaly $d\mathbf{x}$ is seen as the difference between the original field $\mathbf{x}$ and the large-scale component $\overline{\mathbf{x}}$ provided by the OI.

Knowing the underlying reality, it was possible to precisely assess the reconstruction abilities of both AnDA and DINEOF data-driven methodologies, already consolidated with numerous experiences and methodological developments reported in the literature [2,5,6,14]. As a new competitive learning-based approach, we proposed to apply specifically interpolation-designed neural networks involving a joint interpolation and representation learning for irregularly-sampled satellite-derived geophysical fields [15]. As a short synthesis of these evaluations reported in Sections 4.2 and 4.3, some key points can be retrieved:

- A significant gain from data-driven methods compared to the OI-based DUACS baseline: up to 40% relative gain on the SSH daily root mean squared error, in particular on the GULFSTREAM domain where the small scale spatial patterns structures are less noticeable compared to OSMOSIS ;
- A better reconstruction performance of the learning-based GENN introducing a GMRF representation closely related to Gibbs energy concepts compared to AnDA and DINEOF ;

- A significant contribution from the 2D spatial information provided by the additional SWOT sampling to improve the reconstruction of altimetric fields with a relative gain up to 30% on the SSH daily mean squared error, when comparing to the single use of along-track nadir 1D information. Within this combined use of the two datasets, the spectral analysis indicates the new capability to reconstruct spatial scales up to 50-60km which is an important improvement compared to the scales that OI is handling by now; on the other hand, the temporal sampling being less important than nadir tracks, in particular on the GULFSTREAM domain where periods of several days without any SWOT information appears, the reconstruction on these specific periods is sometimes better when learning only with along-track nadir as inputs: we believe that a longer training period (not available here) should improve the behaviour of the NN on this specific issue;
- The possibility of neural network methods to learn from the single observations, without requiring any numerical simulations, which is of particular interest on low latitude areas where the Rossby radius of deformation is large, thus requesting an important catalog to efficently retrieve the SSH dynamics over the year.

As it stands, the results obtained are very encouraging: FP-GENN is a "plug-and-play" algorithm whose conceptual use easily enables its implementation on new datasets. Many perspectives have to be considered in the short and medium terms.
The configuration of FP-GENN used here aims at minimizing the difference between the true anomaly state of the system $d\mathbf{x}$ and its representation $\psi(d\mathbf{x})$ through energy form $||d\mathbf{x} - \psi(d\mathbf{x})||^2$. Alternate energy forms have to be investigated, considering extremes or more generally the whole pdf. In addition, the fixed-point solver used in the joint interpolation approach with GENN never goes too far from the observations, even though they are noisy, which can be an issue in the case of a strong noise including spatial and/or temporal correlations, which was already seen when using SWOT data without any preprocessing (not shown here).
From a methodological point of view, the next developments are expected in the coming related works to increase the gain already observed with FP-GENN:

- use a joint learning of the dynamical representation ψ and the solver Γ minimizing its reconstruction error. A significant gain on the reconstruction performance is expected according to preliminary results obtained with toy models [16];
- a stochastic extension of GENN for including in the NN-based framework an estimation of the uncertainties, thus enabling this new reconstruction method to fully compete with the other interpolators in a "data assimilation" context, with a possible link whith Gaussian Processes and the related Stochastic PDE formalism [18,19].

Besides methodological aspects, new applications are also promising. If we focused here on small North-Atlantic subdomains, the transfer of the NN-based interpolators to an operational process chain will be to reproduce a similar work on the whole basin where the computational constraints in this learning-based setting with large number of parameters is still a challenge. Using a Deep Learning multi-GPU framework and build a pre-operational demonstrator should be of great interestest in the community, as are other SWOT use cases, e.g. using a pre-learning on SWOT data to produce a new interpolation of historical along-track nadir datasets, or taking advantage of the SWOT fast-sampling phase data as inputs for learning prior to its use with SWOT upcoming "operational" data. Last, because the 2D information brought by SWOT showed a significant gain in the reconstruction, a natural extension of this work would be to consider pseudo-observations SKIM datasets [20], whose swath width is more than twice larger (110km vs 270km), and also to propose multivariate analyses including complementary datasets (SST/SSS), already existing in other data-driven schemes like AnDA with an easy extension as additional channels in a neural networks framework.

Supplementary Materials: The code is available on https://github.com/CIA-Oceanix/DINAE_keras with additional informations provided in the ReadMe file to describe the architecture of the code and how to use it

Author Contributions: “Conceptualization, Ronan Fablet and Maxime Beauchamp; methodology, Ronan Fablet and Maxime Beauchamp; software, Ronan Fablet and Maxime Beauchamp; validation, Maxime Beauchamp;

investigation, Maxime Beauchamp, Ronan Fablet, Clément Ubelmann, Maxime Ballarotta and Bertrand Chapron ; writing–original draft preparation, Maxime Beauchamp;

Funding: Funding for the authors was provided by the National Centre for Space Studies (CNES), the French government space agency

References

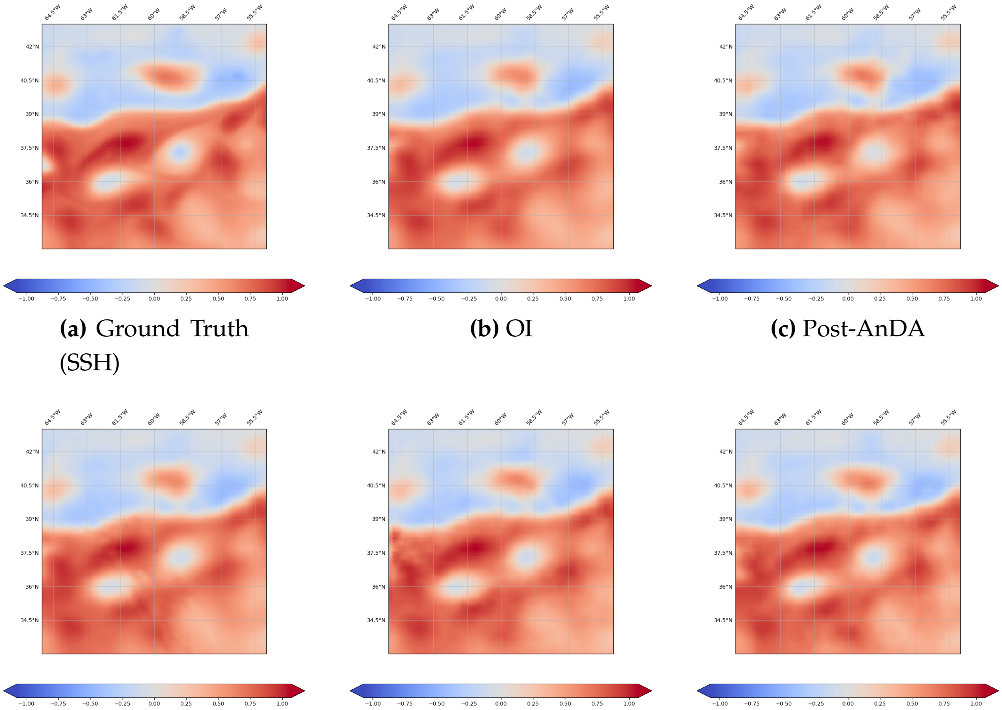
1. Ballarotta, M.; Ubelmann, C.; Pujol, M.I.; Taburet, G.; Fournier, F.; Legeais, J.F.; Faugère, Y.; Delepoulle, A.; Chelton, D.; Dibarboure, G.; Picot, N. On the resolutions of ocean altimetry maps. *Ocean Science* **2019**, *15*, 1091–1109. doi:10.5194/os-15-1091-2019.
2. Lguensat, R.; Tandeo, P.; Aillot, P.; Fablet, R. The Analog Data Assimilation. *Monthly Weather Review* **2017**.
3. Lguensat, R.; Huynh Viet, P.; Sun, M.; Chen, G.; Fenglin, T.; Chapron, B.; Fablet, R. Data-driven Interpolation of Sea Level Anomalies using Analog Data Assimilation **2017**.
4. Fablet, R.; Viet, P.H.; Lguensat, R. Data-Driven Models for the Spatio-Temporal Interpolation of Satellite-Derived SST Fields. *IEEE Trans. on Computational Imaging* **2017**, *3*, 647–657. doi:10.1109/TCI.2017.2749184.
5. Lopez-Radcenco, M.; Pascual, A.; Gomez-Navarro, L.; Aissa-El-Bey, A.; Chapron, B.; Fablet, R. Analog Data Assimilation of Along-Track Nadir and Wide-Swath SWOT Altimetry Observations in the Western Mediterranean Sea. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2019**, pp. 1–11. doi:10.1109/JSTARS.2019.2903941.
6. Ouala, S.; Fablet, R.; Herzet, C.; Chapron, B.; Pascual, A.; Collard, F.; Gaultier, L. Neural Network Based Kalman Filters for the Spatio-Temporal Interpolation of Satellite-Derived Sea Surface Temperature. *Remote Sensing* **2018**, *10*, 1864. doi:10.3390/rs10121864.
7. Taburet, G.; Sanchez-Roman, A.; Ballarotta, M.; Pujol, M.I.; Legeais, J.F.; Fournier, F.; Faugere, Y.; Dibarboure, G. DUACS DT2018: 25 years of reprocessed sea level altimetry products. *15*, 1207–1224. Publisher: Copernicus GmbH, doi:https://doi.org/10.5194/os-15-1207-2019.
8. Molines, J.M. meom-configurations/NATL60-CJM165: NATL60 code used for CJM165 experiment, 2018. doi:10.5281/zenodo.1210116.
9. Dufau, C.; Orsztynowicz, M.; Dibarboure, G.; Morrow, R.; Le Traon, P.Y. Mesoscale resolution capability of altimetry: Present and future. *Journal of Geophysical Research: Oceans* **2016**, *121*, 4910–4927, [https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015JC010904]. doi:10.1002/2015JC010904.
10. Gaultier, L.; Ubelmann, C.; Fu, L.L. The Challenge of Using Future SWOT Data for Oceanic Field Reconstruction. *Journal of Atmospheric and Oceanic Technology* **2015**, *33*, 119–126. doi:10.1175/JTECH-D-15-0160.1.
11. Esteban-Fernandez, D. SWOT project mission performance and error budget document. Technical report, JPL, NASA, 2014.
12. Gaultier, L.; Ubelmann, C. SWOT Simulator Documentation. Technical report, JPL, NASA, 2010.
13. Metref, S.; Cosme, E.; Le Guillou, F.; Le Sommer, J.; Brankart, J.M.; Verron, J. Wide-Swath Altimetric Satellite Data Assimilation With Correlated-Error Reduction. *Front. Mar. Sci.* **2020**, *6:822*. doi:doi: 10.3389/fmars.2019.00822.
14. Ping, B.; Su, F.; Meng, Y. An Improved DINEOF Algorithm for Filling Missing Values in Spatio-Temporal Sea Surface Temperature Data. *PLOS ONE* **2016**, *11*, e0155928. doi:10.1371/journal.pone.0155928.
15. Fablet, R.; Drumetz, L.; Rousseau, F.; Beauchamp, M. Joint interpolation and representation learning for irregularly-sampled satellite-derived geophysical fields.
16. Fablet, R.; Drumetz, L.; Rousseau, F. Joint learning of variational representations and solvers for inverse problems with partially-observed data, 2020, [arXiv:cs.LG/2006.03653].
17. Welch, P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics* **1967**, *15*, 70–73. doi:10.1109/TAU.1967.1161901.
18. Lindgren, F.; Rue, H.; Lindström, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation

approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **2011**, *73*, 423–498, [https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2011.00777.x]. doi:10.1111/j.1467-9868.2011.00777.x.

19. Sidén, P.; Lindsten, F. Deep Gaussian Markov random fields, 2020, [arXiv:stat.ML/2002.07467].
20. Ardhuin, F.; Brandt, P.; Gaultier, L.; Donlon, C.; Battaglia, A.; Boy, F.; Casal, T.; Chapron, B.; Collard, F.; Cravatte, S.; Delouis, J.M.; De Witte, E.; Dibarboure, G.; Engen, G.; Johnsen, H.; Lique, C.; Lopez-Dekker, P.; Maes, C.; Martin, A.; Marié, L.; Menemenlis, D.; Nouguier, F.; Peureux, C.; Rampal, P.; Ressler, G.; Rio, M.H.; Rommen, B.; Shutler, J.D.; Suess, M.; Tsamados, M.; Ubelmann, C.; Van Sebille, E.; Van Den Oever, M.; Stammer, D. SKIM, a Candidate Satellite Mission Exploring Global Ocean Currents and Waves, 2019. doi:https://doi.org/10.3389/fmars.2019.00209.

Appendix A. Complementary figures for SSH interpolations

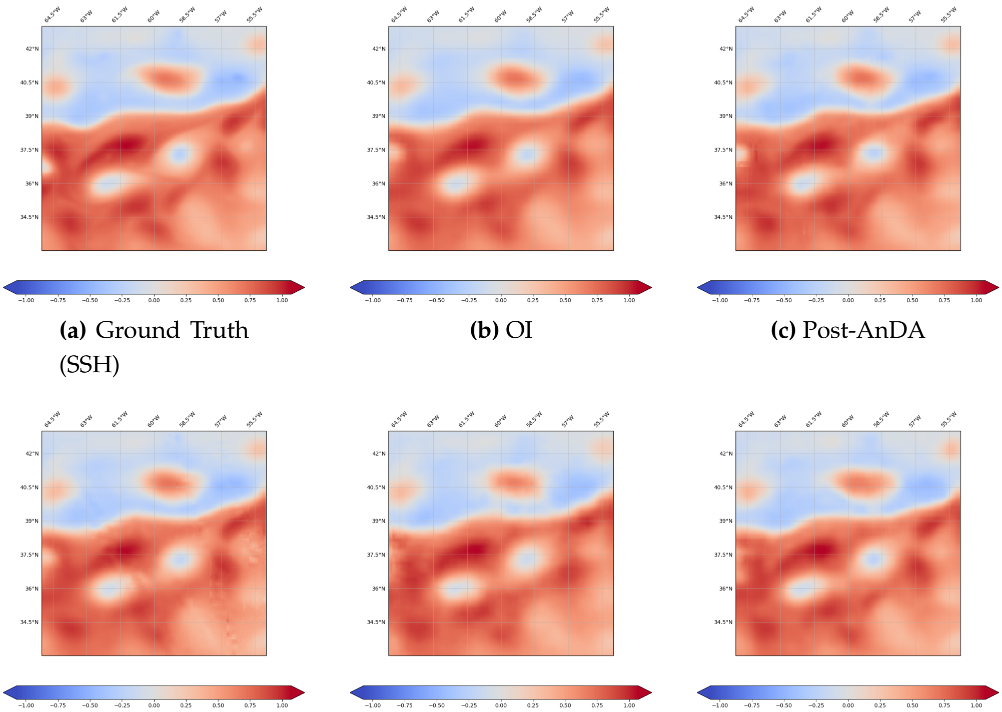
Appendix A.1. GULFSTREAM



(a) Ground Truth (SSH) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

Figure A1. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

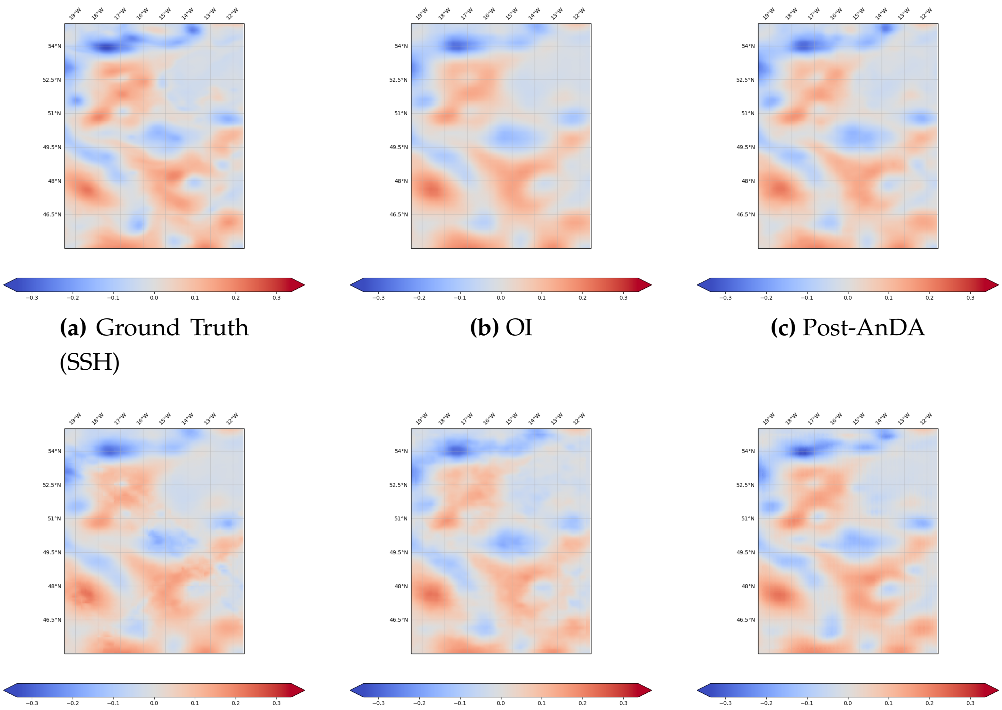


(a) Ground Truth (SSH) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

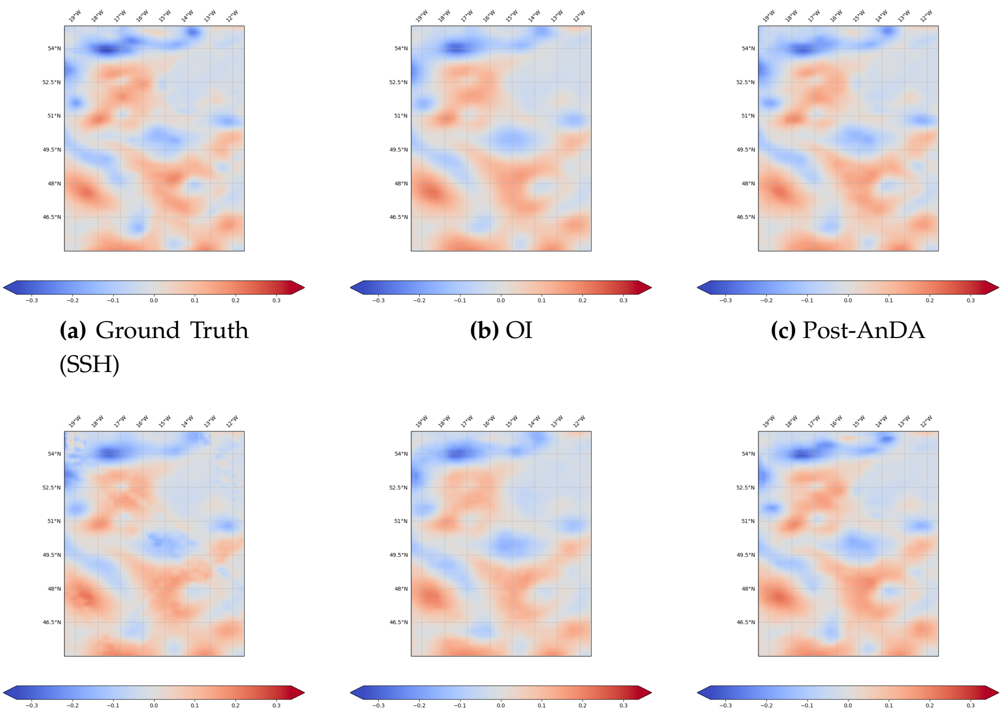
Figure A2. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

463 *Appendix A.2. OSMOSIS*



(a) Ground Truth (SSH) **(b)** OI **(c)** Post-AnDA
(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

Figure A3. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

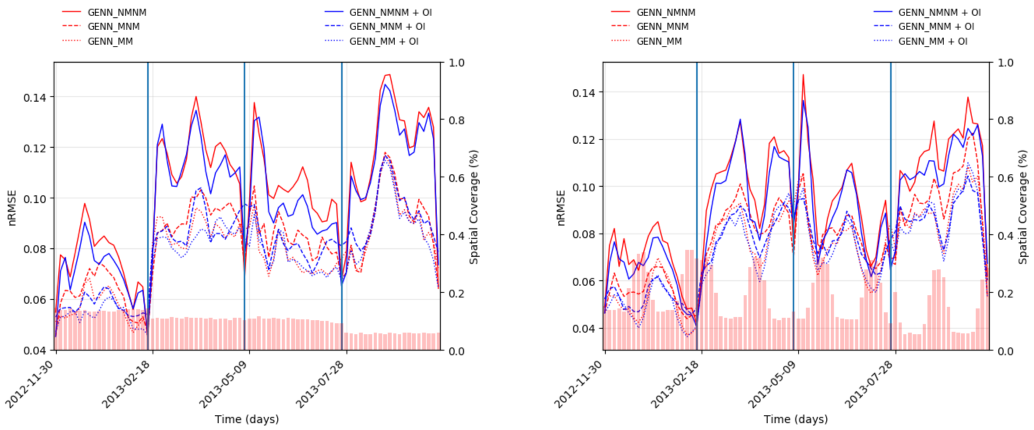


(a) Ground Truth (SSH) **(b)** OI **(c)** Post-AnDA
(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

Figure A4. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

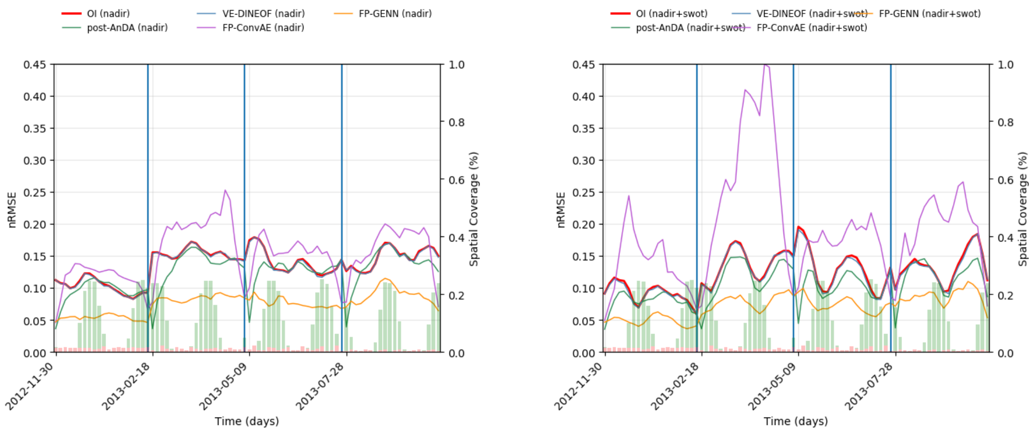
464 # Appendix B. OSSE without observation errors

465 ## *Appendix B.1. GULFSTREAM*



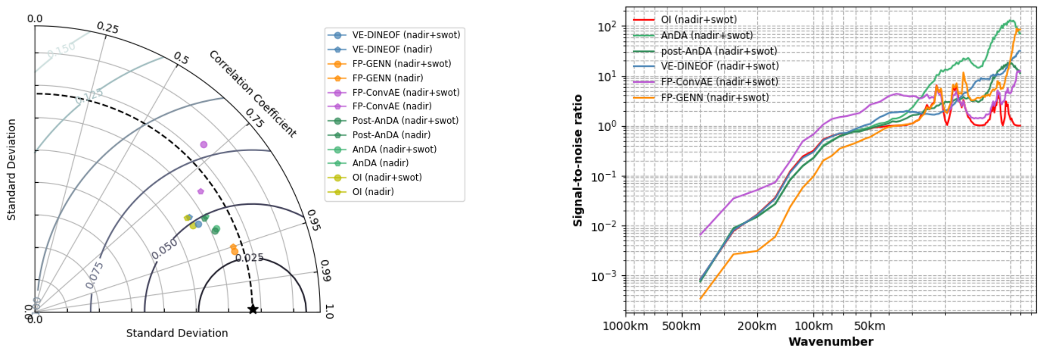
(a) nadir **(b)** nadir+swot

Figure A5. Daily spatial nRMSE computed on the 80-days non-continuous validation period for the six supervised/unsupervised FP-GENN configurations. The spatial coverage of 0-days accumulated along-track nadir (a) expanded with wide-swath SWOT data (b) is provided by the red-colored barplot



(a) nadir **(b)** nadir+swot

Figure A6. Daily spatial nRMSE computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN. The spatial coverage of 0-days accumulated along-track nadir and wide-swath SWOT data are respectively provided by the red and green-colored barplots



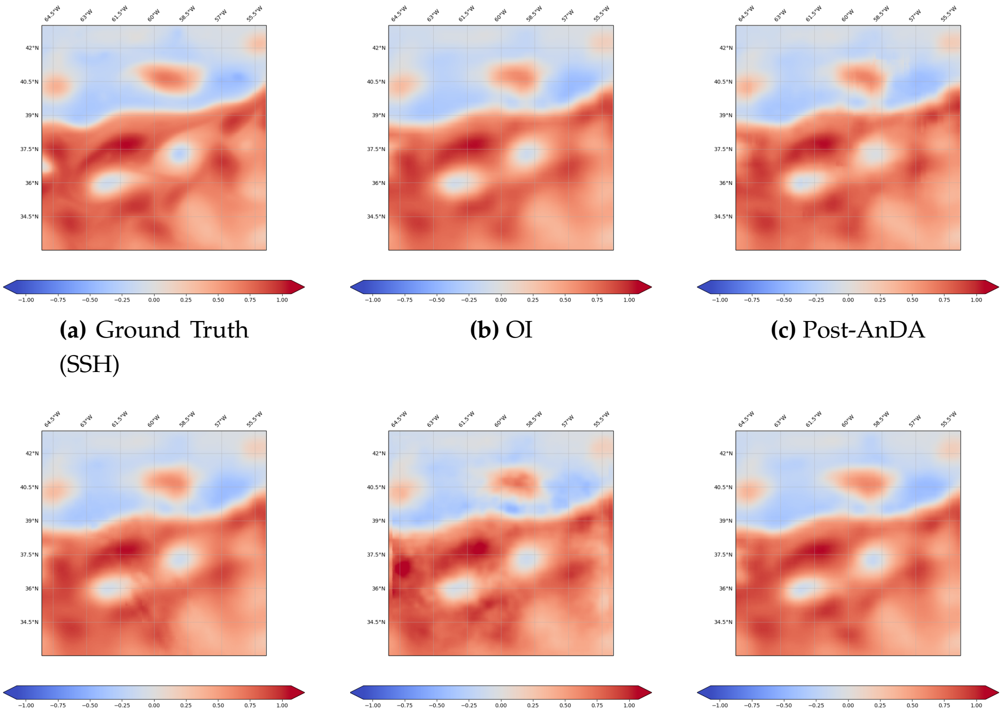
(a) Taylor diagram **(b)** Signal-to-noise ratio

Figure A7. Taylor diagram and Signal-to-noise ratio computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN computed for both nadir use only and joint assimilation/learning with wide-swath SWOT data

	Model type	R-score	I-score	AE-score
nadir	OI	86.53	72.25	_
	AnDA	90.56	76.81	_
	VE-DINEOF	91.33	72.58	_
	FP-ConvAE	69.46	63.82	79.86
	FP-GENN	**95.15**	**91.28**	96.32
nadir + SWOT	OI	91.76	75.30	_
	AnDA	91.72	82.43	_
	VE-DINEOF	92.47	76.00	_
	FP-ConvAE	42.78	34.96	79.93
	FP-GENN	**97.31**	**91.45**	96.87

	Model type	R-score	I-score	AE-score
nadir	∇_{OI}	76.14	72.41	_
	∇_{AnDA}	81.81	76.15	_
	$\nabla_{VE-DINEOF}$	80.09	72.07	_
	$\nabla_{FP-ConvAE}$	58.30	59.79	70.14
	$\nabla_{FP-GENN}$	**84.75**	**84.63**	88.05
nadir + SWOT	∇_{OI}	71.41	72.31	_
	∇_{AnDA}	85.85	79.80	_
	$\nabla_{VE-DINEOF}$	84.73	73.36	_
	$\nabla_{FP-ConvAE}$	31.78	36.48	69.72
	$\nabla_{FP-GENN}$	**87.75**	**85.35**	89.50

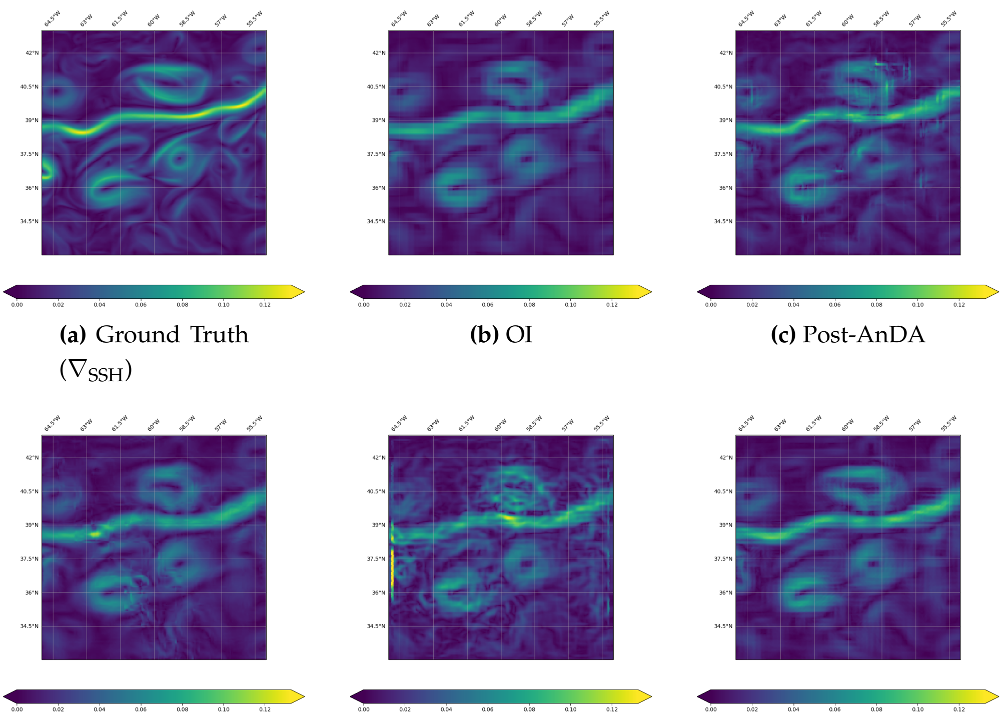
Table A1. SSH and SSH gradient field R/I/AE-scores computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for both nadir use only and joint assimilation/learning with wide-swath SWOT data



(a) Ground Truth (SSH) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

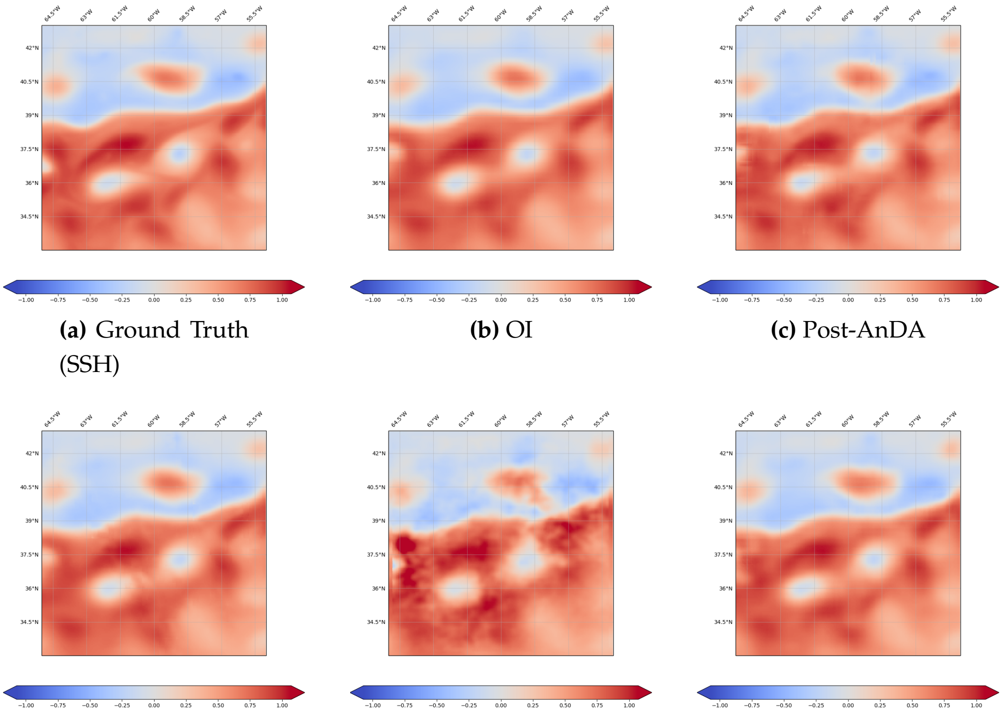
Figure A8. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only



(a) Ground Truth (∇_{SSH}) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

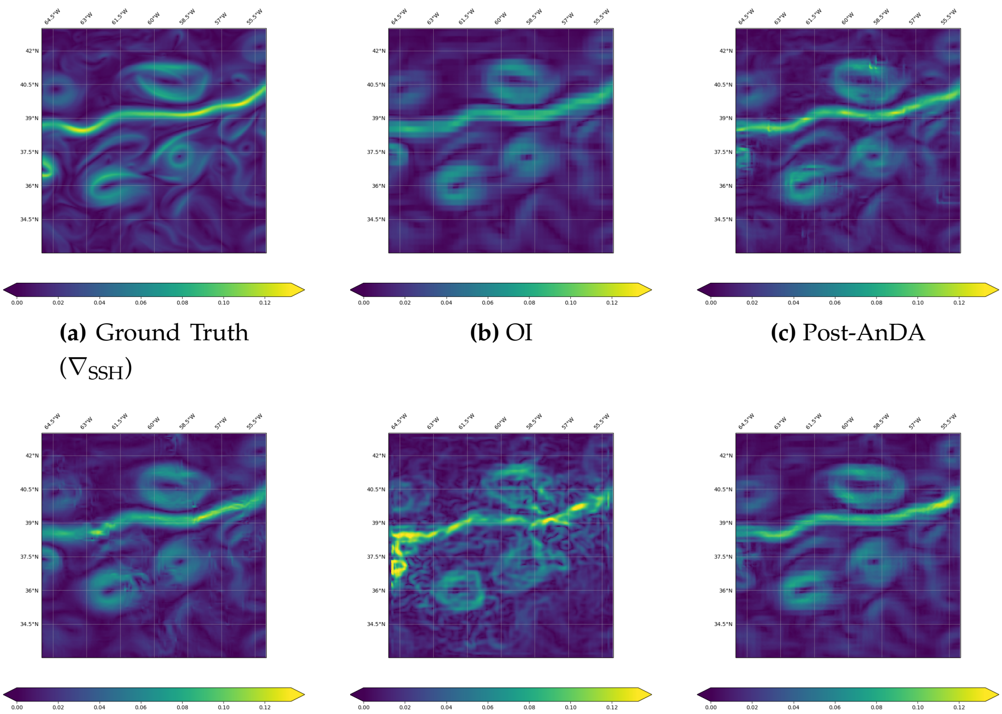
Figure A9. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only



(a) Ground Truth (SSH) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

Figure A10. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

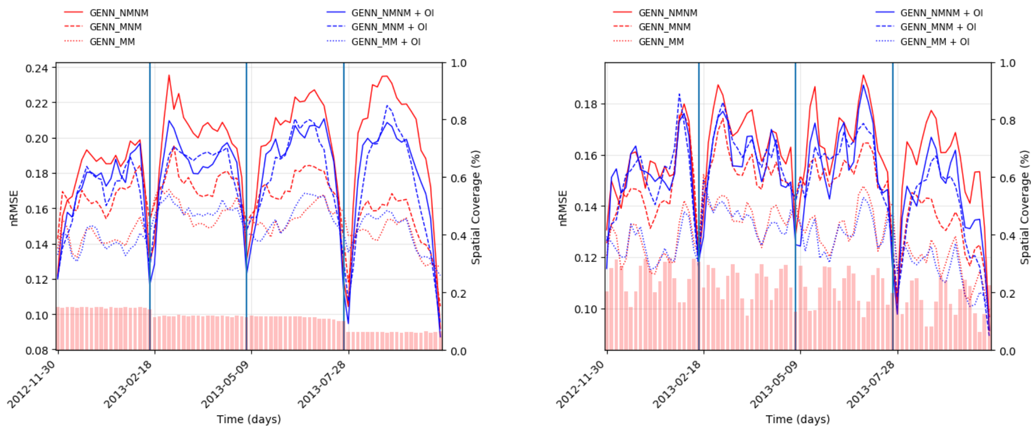


(a) Ground Truth (∇_{SSH}) **(b)** OI **(c)** Post-AnDA

(d) VE-DINEOF **(e)** FP-ConvAE **(f)** FP-GENN

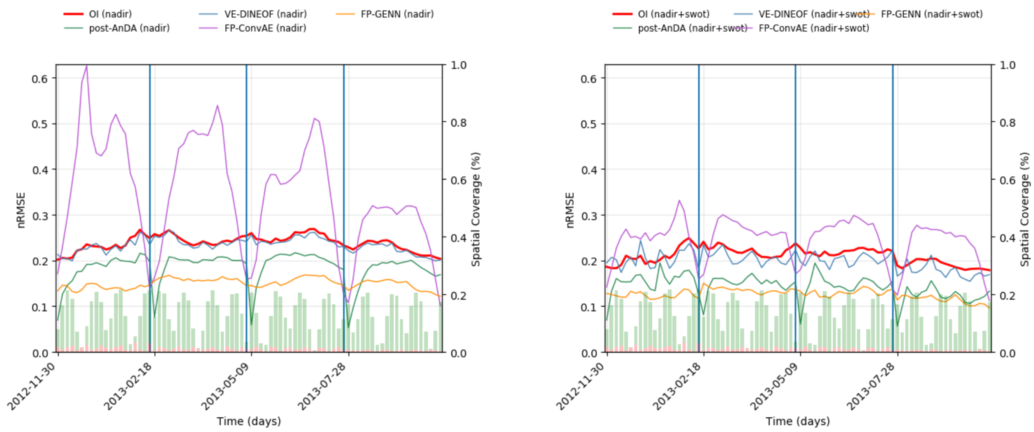
Figure A11. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

466 *Appendix B.2. OSMOSIS*



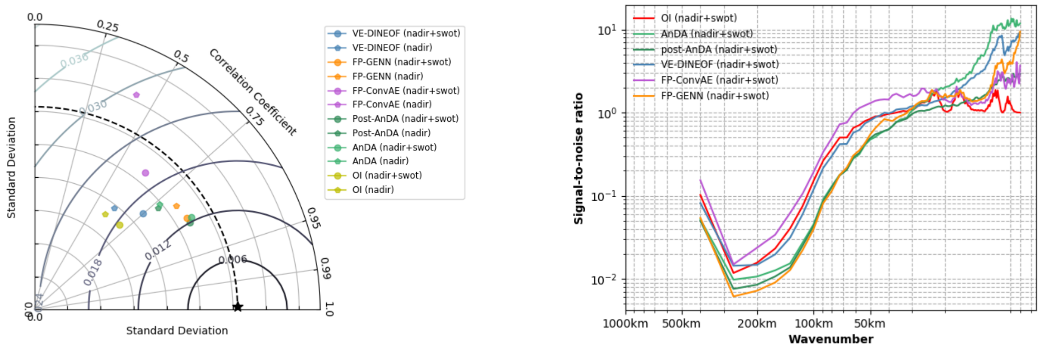
(a) nadir **(b)** nadir+swot

Figure A12. Daily spatial nRMSE computed on the 80-days non-continuous validation period for the six supervised/unsupervised FP-GENN configurations. The spatial coverage of 0-days accumulated along-track nadir (a) expanded with wide-swath SWOT data (b) is provided by the red-colored barplot



(a) nadir **(b)** nadir+swot

Figure A13. Daily spatial nRMSE computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN. The spatial coverage of 0-days accumulated along-track nadir and wide-swath SWOT data are respectively provided by the red and green-colored barplots



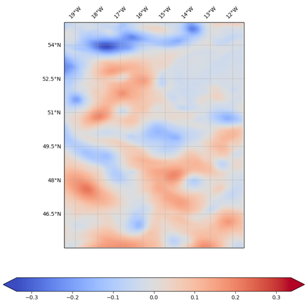
(a) Taylor diagram **(b)** Signal-to-noise ratio

Figure A14. Taylor diagram and Signal-to-noise ratio computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN computed for both nadir use only and joint assimilation/learning with wide-swath SWOT data

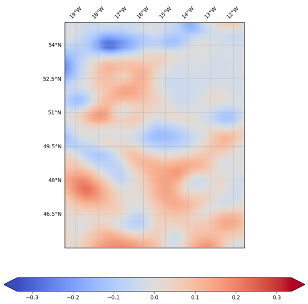
	Model type	R-score	I-score	AE-score
nadir	OI	44.63	34.93	_
	AnDA	76.60	59.42	_
	VE-DINEOF	77.17	37.66	_
	FP-ConvAE	28.39	17.00	42.94
	FP-GENN	**84.35**	**76.17**	86.30
nadir + SWOT	OI	54.31	47.87	_
	AnDA	83.07	74.95	_
	VE-DINEOF	83.47	51.50	_
	FP-ConvAE	36.80	33.37	47.56
	FP-GENN	**90.67**	**81.35**	88.04

	Model type	R-score	I-score	AE-score
nadir	∇_{OI}	49.53	48.20	_
	∇_{AnDA}	**64.56**	59.88	_
	$\nabla_{\mathrm{VE-DINEOF}}$	58.71	45.61	_
	$\nabla_{\mathrm{FP-ConvAE}}$	22.47	19.12	36.66
	$\nabla_{\mathrm{FP-GENN}}$	62.47	**61.64**	64.88
nadir + SWOT	∇_{OI}	37.55	47.93	_
	∇_{AnDA}	75.13	**70.22**	_
	$\nabla_{\mathrm{VE-DINEOF}}$	**79.31**	49.32	_
	$\nabla_{\mathrm{FP-ConvAE}}$	30.85	35.06	39.06
	$\nabla_{\mathrm{FP-GENN}}$	67.99	67.47	69.21

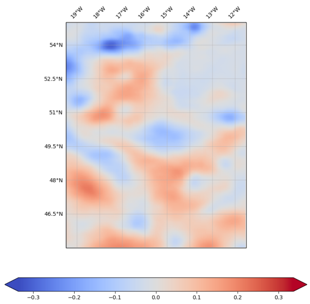
Table A2. SSH and SSH gradient field R/I/AE-scores computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for both nadir use only and joint assimilation/learning with wide-swath SWOT data



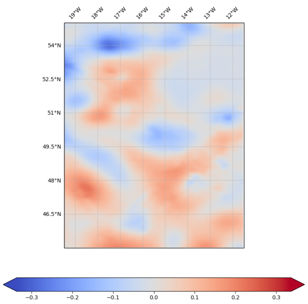
(a) Ground Truth (SSH)



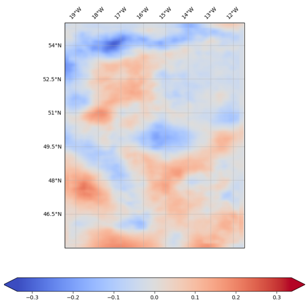
(b) OI



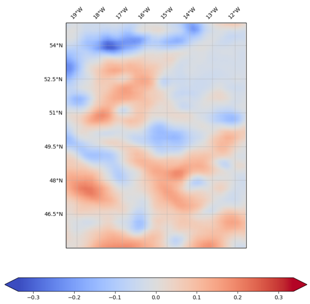
(c) Post-AnDA



(d) VE-DINEOF

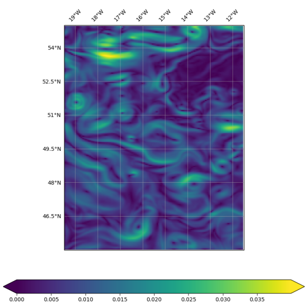


(e) FP-ConvAE

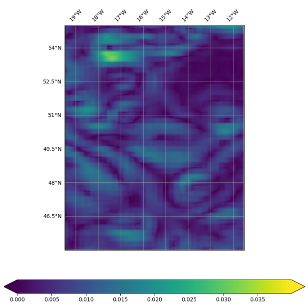


(f) FP-GENN

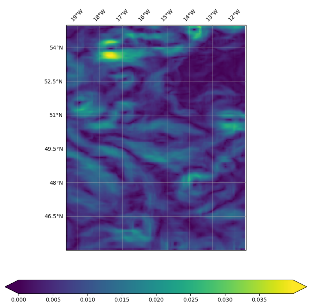
Figure A15. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only



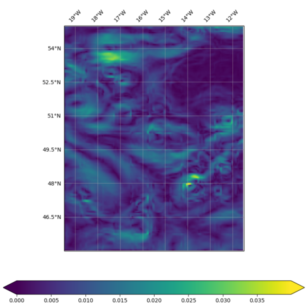
(a) Ground Truth (∇_{SSH})



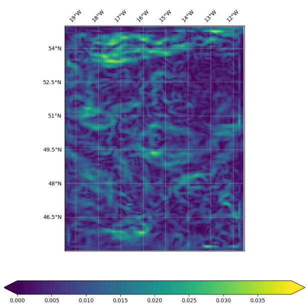
(b) OI



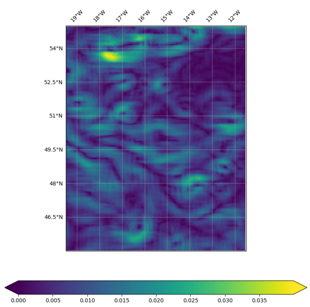
(c) Post-AnDA



(d) VE-DINEOF

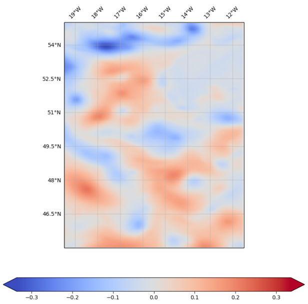


(e) FP-ConvAE

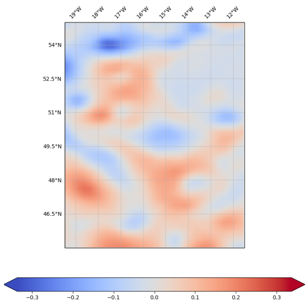


(f) FP-GENN

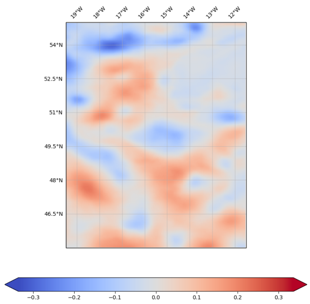
Figure A16. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only



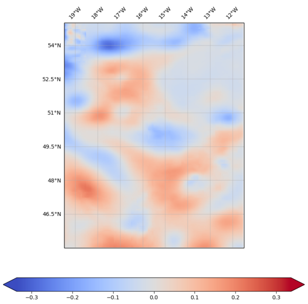
(a) Ground Truth (SSH)



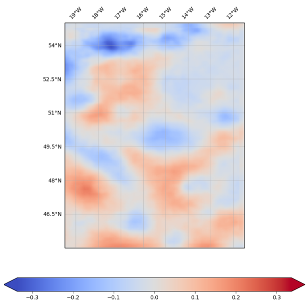
(b) OI



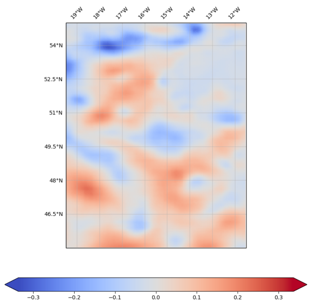
(c) Post-AnDA



(d) VE-DINEOF

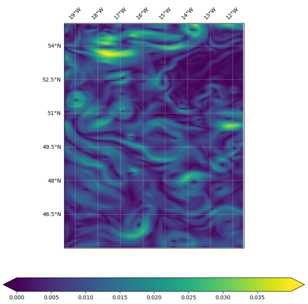


(e) FP-ConvAE

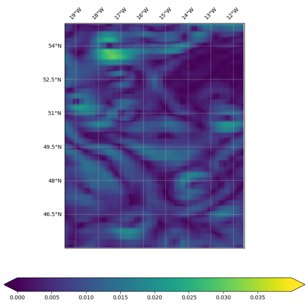


(f) FP-GENN

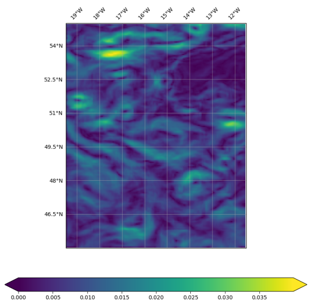
Figure A17. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data



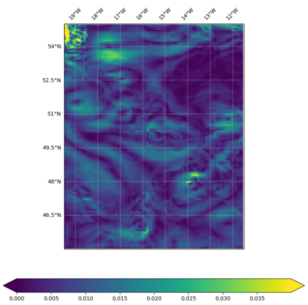
(a) Ground Truth (∇_{SSH})



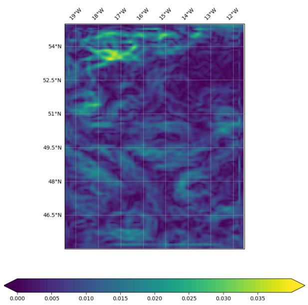
(b) OI



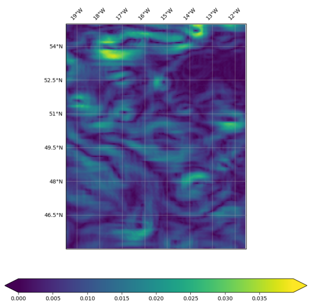
(c) Post-AnDA



(d) VE-DINEOF



(e) FP-ConvAE



(f) FP-GENN

Figure A18. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data