



HAL
open science

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

► To cite this version:

Maxime Beauchamp, Ronan Fablet, Clément Ubelmann, Maxime Ballarotta, Bertrand Chapron. Intercomparison of data-driven and learning-based interpolations of along-track Nadir and wide-swath Swot altimetry observations. In press. hal-02931892v1

HAL Id: hal-02931892

<https://imt-atlantique.hal.science/hal-02931892v1>

Preprint submitted on 7 Sep 2020 (v1), last revised 29 Sep 2020 (v2)

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.

Article

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

Maxime Beauchamp¹ , Ronan Fablet¹, Clément Ubelmann², Maxime Ballarotta³ and Bertrand Chapron⁴

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 7, 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 $10^\circ \times 10^\circ$ GULFSTREAM and $8^\circ \times 10^\circ$ 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 (aggregated) 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. 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 artificial intelligence (AI)-based algorithms have just been proposed [1–5] 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. 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 [6] 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 [6] 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 ($1/60^\circ$) 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 $1/20^\circ$, 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

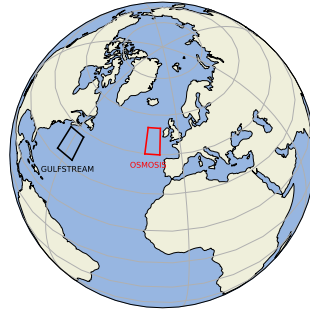


Figure 1. GULFSTREAM and OSMOSIS domain

77 simulation from October 1st, 2012 to September 29th, 2013, thus covering a whole year of data. A
 78 Gaussian white noise with variance $\sigma^2 = 30\text{cm}$ is then added to the interpolated NATL60 simulation
 79 on the nadir coordinates to simulate the instrumental acquisition noise [7].

80 Because the space-time interpolations will focus on a daily-basis temporal resolution, we also build
 81 nadir pseudo-observations with an additional strategy by accumulating observations over a time
 82 window $t_k \pm d$ days centered at time t_k in order to increase the daily nadir spatial sampling. As in [4],
 83 we investigate the response of the different interpolation techniques when parameter d is either set to
 84 0 or 5, see Figures 2a and 2c for the corresponding aggregation on August 4, 2013, and August 5, 2013.

85 2.3. SWOT

86 In the same line, SWOT-like pseudo observations are also produced by the swotsimulator tool [8]
 87 in its swath mode with an along-track and across-track 2km spatial resolution, the same theoretical
 88 resolution the upcoming SWOT mission derived products should be able to provide. The nadir mode
 89 of the generator also provide pseudo-nadir along track observations though they are not used here.
 90 The simulator also adds instrumental noise on the idealized pseudo-SWOT dataset [9,10]. This noise
 91 potentially exhibits strong space-time correlations. Thus, the pseudo-SWOT observations are first
 92 preprocessed [11] to filter out these correlated components and avoid major issues in the assimilation
 93 and/or learning process of the interpolation methods.

94 Let precise that over the low-latitude GULFSTREAM domain, the SWOT sampling is irregular leading
 95 to sequences of several days with only pseudo-nadir observations. This does not happen on the higher
 96 latitude OSMOSIS area where the SWOT temporal coverage is more regular. It can be seen along this
 97 paper on the time series evaluation figures embedding additional information about the daily spatial
 98 coverage as complementary barplots scaled on the right-hand side of the Y-axis.

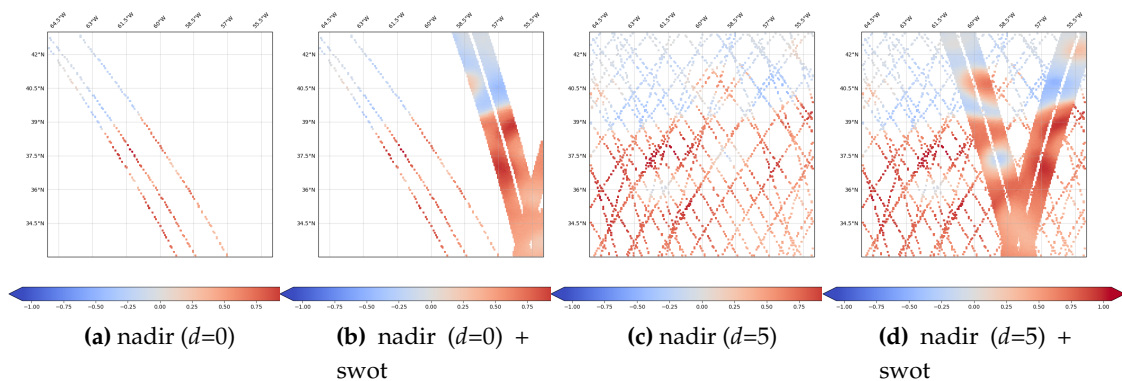


Figure 2. 0 and 5-days accumulated along-track nadir and wide-swath pseudo-observations on August 4, 2013 (a,b) and August 5, 2013 (c,d)

99 2.4. DUACS OI products

100 The DUACS system is an operational production of sea level products for the Marine (CMEMS)
 101 and Climate (C3S) services of the E.U. Copernicus program, on behalf of the CNES french space
 102 agency. It is mainly based on optimal interpolation techniques whose parameters are fully described
 103 in [12]. This methodology has been applied on the previously introduced pseudo along-track nadir
 104 and wide-swath SWOT data to generate regularly ($0.25^\circ \times 0.25^\circ$) daily gridded maps.

105 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:

$$\mathbf{x} = \bar{\mathbf{x}} + d\mathbf{x} + \epsilon \quad (1)$$

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

111 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 :

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

112 The analog forecasting operator $\mathcal{A} : \mathbf{dx}_{k-1}^a \mapsto \mathbf{dx}_k^f$, where superscripts a and f respectively relies
 113 to analysis and forecast, is built from the K most similar states to \mathbf{dx}_{k-1}^a in the available past state
 114 dynamics catalog, supposed to be large enough to describe the space-time evolution of the processes.
 115 More precisely, \mathbf{dx}_k^f is sampled from the Gaussian prior $\mathbf{dx}_k^f | \mathbf{dx}_{k-1}^a \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, where the mean $\boldsymbol{\mu}_k$
 116 and the covariance matrix $\boldsymbol{\Sigma}_k$ are estimated using the so-called locally linear model [1], i.e. a weighted
 117 linear regression between the K nearest analogs and their successors.

118 In the experiments, the diagonal of the observation error matrix $\mathbf{R}_k = Cov(\boldsymbol{\epsilon}_k)$ is not assumed constant
 119 but its values increase according to a parametric function of the hourly time lag between the observation
 120 and the day to estimate:

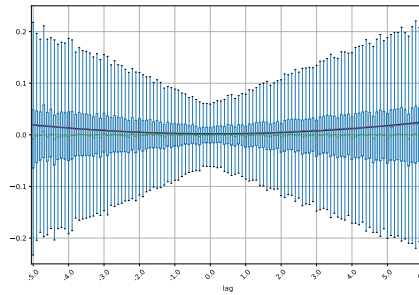


Figure 3. Variance of the observation error $\boldsymbol{\epsilon}_k$ as a function of the hourly lag between the observation and the day to estimate

As in [4], a patch-based version of AnDA coupled with an EOF-based representation of the individual patches is used. The anomaly field dx 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 [5].

3.2. VE-DINEOF

VE-DINEOF is a state-of-the-art interpolation approach [13] 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 [14] to deal with image sequences involving potentially large missing data rates. In this framework, an energy-based representation U_θ to minimize is introduced :

$$U_\psi(dx) = \|dx - \psi(dx)\|^2 \quad (3)$$

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

$$\widehat{dx}_k = I_{U_\psi}(dy_k(\Omega_k)) = \arg \min_{dx} U_\psi(dx) \quad (4)$$

such that $I_{U_\psi}(dy_k(\Omega_k)) = dx_k$ 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:

$$\widehat{\theta} = \arg \min_{\theta} \sum_k \left\| dy_k(\Omega_k) - I_{U_\psi}(dy_k(\Omega_k)) \right\|_{\Omega_k}^2 \quad (5)$$

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 dx 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

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

- 155 2. GE-NN: Second, NN-based Gibbs-Energy (GENN) representations where $d\mathbf{x}_s$, the anomaly
 156 observed at location $\mathbf{s} \in \mathcal{D}$, is supposed to be explained by the potential function $\psi(d\mathbf{x}_{\delta s})$ with
 157 δs a predefined neighbourhood of site \mathbf{s} , thus relating this representation to Markovian priors
 158 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}
 159 ensures to provide a good state space reconstruction. Regarding the architecture involved,
 160 the following scheme is used: an initial 4×4 average pooling, a Conv2D layer with 40 filters,
 161 11×11 kernel, ReLU activation and a zero-weight constraint on the center of the convolution
 162 window, a 1×1 Conv2D layer with 40 filters, a ResNet composed of an initial mapping to an initial
 163 $200 \times 200 \times (5 \times 40)$ space with a Conv2D+ReLU layer and a linear 1×1 Conv2D+ReLU layer with 40
 164 filters. Last, a final 4×4 Conv2DTranspose layer with a linear activation for an upsampling to the
 165 input shape is considered. GE-NN involves 10 residual units for a total of $\approx 450,000$ parameters.

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

173 3.3.2. Fixed-point solver

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

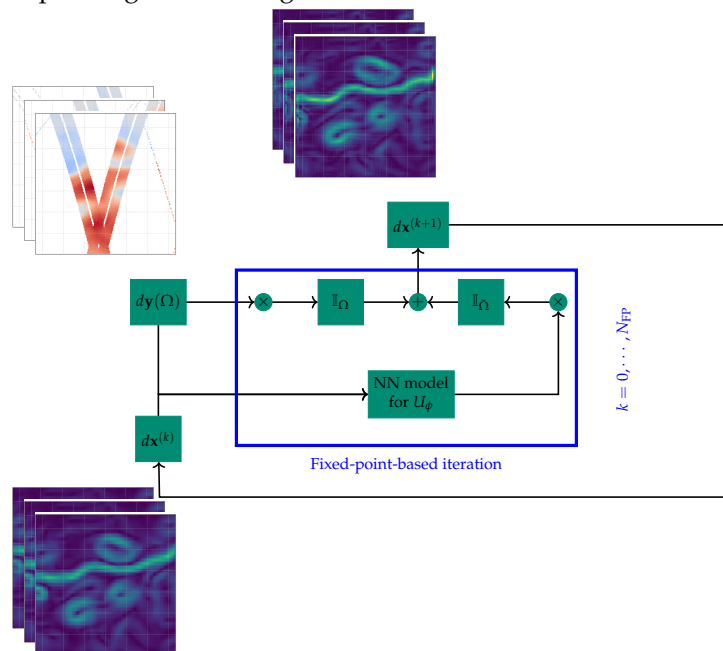


Figure 4. Sketch of the iterative fixed-point algorithm

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

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

177 It is parameter-free and easily implemented as a NN in a joint solution with the
 178 NN-parametrization of U_θ for the interpolation problem. The two NN-architectures are then referred
 179 as FP-ConvAE and FP-GENN. Let note that additional improvements are expected when using
 180 an iterative gradient-based formulation of the solver, where the gradient of U_ψ is replaced by a
 181 ConvNet or LSTM unit $G(\mathbf{x} - \psi(\mathbf{x}))$, thus enabling to solve jointly for the parametrization of ψ and
 182 G . Complementary results on SST datasets regarding this point can be found in [14]. Let precise
 183 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
 184 time window $dT = 5$, centered on time t_k , leading to image time series of length 11. Last, the
 185 above-mentioned works are generalized to establish a connection between 4DVAR variational data
 186 assimilation and joint learning of models and solvers in [15].

187 4. Evaluation

188 4.1. Experimental/benchmarking setup

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

195
 196 Regarding the metrics used in the intercomparison exercise, daily normalized RMSE (nRMSE) time
 197 series are first provided: they give a quick overview of the potential gain obtained with the data-driven
 198 interpolators. Additional correlation and variances scores are also computed, then all displayed
 199 together with the RMSE as Taylor diagrams. We also provide three other indicators, namely the
 200 global reconstruction score (R-score) for the known SSH field areas (Ω), the interpolation performance
 201 (I-score) for the missing data areas ($\overline{\Omega}$), and the reconstruction performance of the trained NN-based
 202 representation of the SSH dynamics for FP-ConvAE and FP-GENN when applied to gap-free SSH
 203 fields (AE-score). Last, signal-to-noise ratios are also computed in the spectral domain, in particular to
 204 assess up to which spatial scale the different interpolators are able to reproduce the ground truth.

205 4.2. GULFSTREAM

206 We first have to discuss the time window parameter d related to the aggregation of along-track
 207 data over a specific day t_k , see Section 2.2. A same value of this parameter may not be optimal
 208 for all the interpolators: AnDA exhibits a better performance when considering only along-track
 209 nadir data of the day ($d = 0$), thus contradicting the previous optimal results of $d = 5$ found by
 210 [4] over the Mediterranean sea, which may indicate AnDA responds differently to the along-track
 211 aggregation strategy depending on the energetic dynamical regime of the region. On the other hand,
 212 both FP-ConvAE and FP-GENN interpolators performs better (not shown here) by aggregating nadir
 213 data over a 5-day time window. As a consequence, the results presented in what follows will use value
 214 of $d = 0$ for AnDA and VE-DINEOF and $d = 5$ for FP-ConvAE and FP-GENN.

215 Next, to evaluate the behaviour of the different interpolators on both along-track nadir samplings

216 and their fusion with wide-swath SWOT datasets and make the comparison possible, we have to
 217 preliminary define if the NN-based interpolator will be used under a supervised or unsupervised
 218 learning strategy. Figure 5 depicts how the FP-GENN interpolator performs using nadir data (a) or
 219 their joint use with SWOT (b), according to the input and target data used for the training. Six possible
 220 configurations have been tested. Two supervised versions using the gap-free NATL60 simulations
 221 as target, and either the pseudo-observations or the gap-free maps as input, respectively denoted as
 222 FP-GENN-MNM and FP-GENN-NMNM. A fully unsupervised FP-GENN-MM version is also used in
 223 which both input and target are only made of the pseudo-observations. These three configurations
 224 are also tested when adding the DUACS OI product as a covariate for input data, because we think
 225 that this may give a prior information about how the anomaly field dx is distributed. Within this
 226 part-GULFSTREAM domain, we clearly see the best performance is obtained by the unsupervised
 227 configuration of FP-GENN: it is a keypoint result because the learning network abilities seems to be
 228 better when it is fully data-driven, meaning that it benefits from its knowledge of the spatio-temporal
 229 location and occurrence of the data, which is a fairly new avenue for data assimilation related problems.
 230 The use of the OI as a covariate improve the FP-GENN behaviour but not systematically.

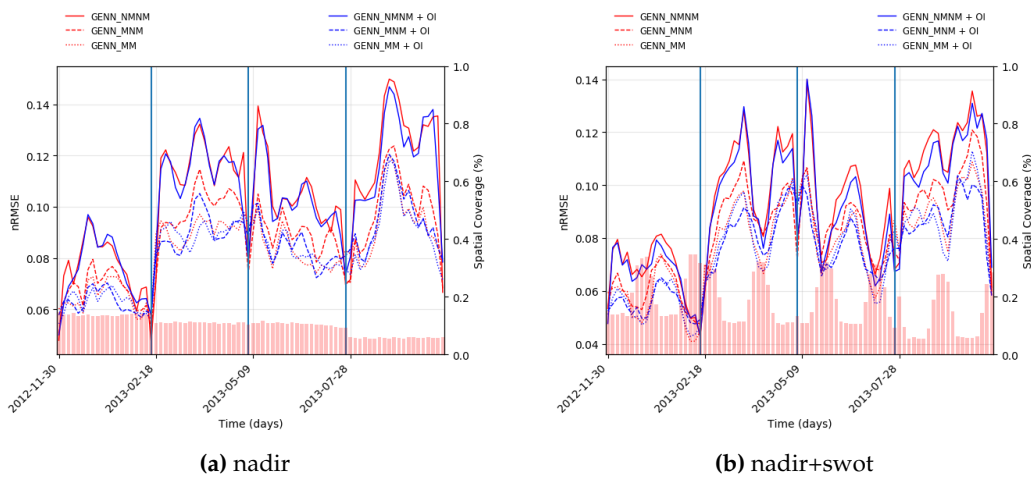


Figure 5. Daily spatial nRMSE computed on the 80-days non-continuous validation period for the six supervised/unsupervised FP-GENN configurations. The spatial coverage of 0-days accumulated along-track nadir (a) expanded with wide-swath SWOT data (b) is provided by the red-colored barplot

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

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

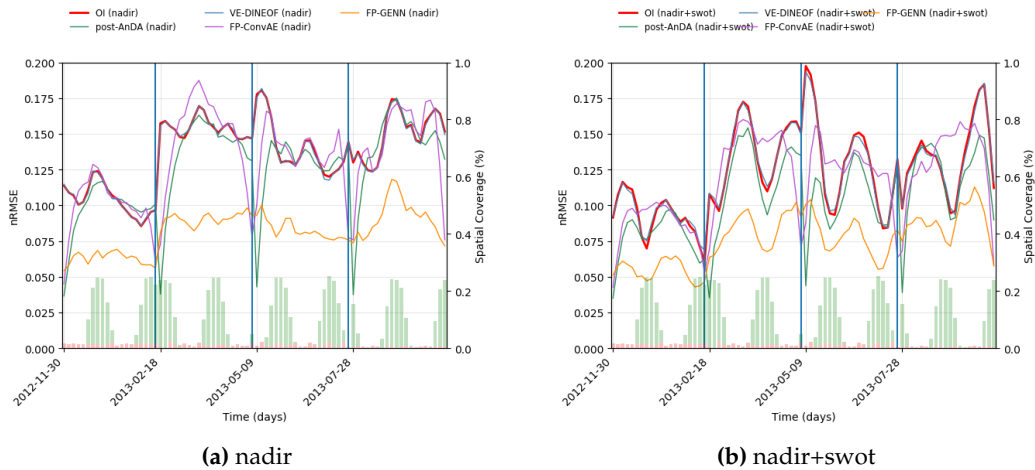


Figure 6. Daily spatial nRMSE computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN. The spatial coverage of 0-days accumulated along-track nadir and wide-swath SWOT data are respectively provided by the red and green-colored barplots

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

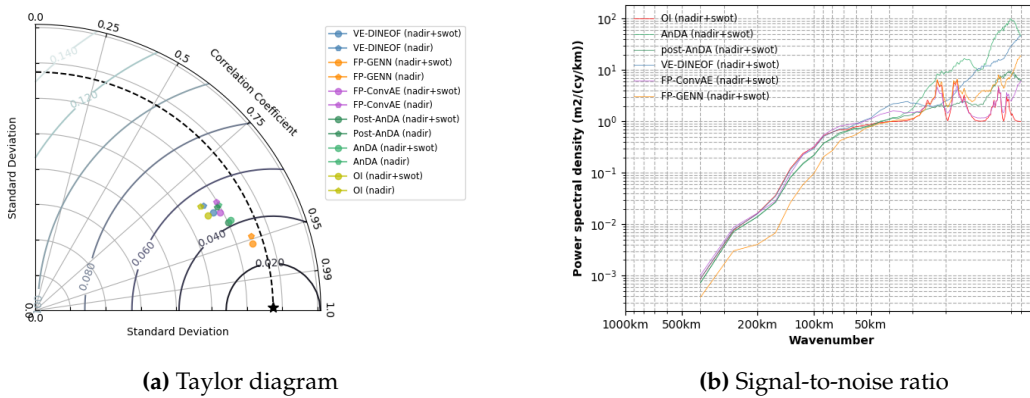


Figure 7. Taylor diagram and Signal-to-noise ratio computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN computed for both nadir use only and joint assimilation/learning with wide-swath SWOT data

252 In Table 1, R/I/AE-scores are applied to both SSH (after application of a retrieving high-pass filter
 253 to keep only the small scales information) and its gradient (module). Regarding the R-scores, AnDA
 254 and VE-DINEOF are often the best way to keep track of the known areas, which is not surprising since
 255 these two methods makes an explicit use of the observational altimetric data into their mapping process.
 256 When looking at the I-scores, where no data is available, FP-GENN now clearly stands out from the
 257 other interpolators, which motivate its future use for irregularly-sampled data with large missing data
 258 rates. In addition, because its reconstruction scores remain overall satisfactory, in particular when
 259 considering the joint learning on nadir and SWOT data, these results are supplementary arguments on
 260 account of this markovian-related NN-based formulation.

	Model type	R-score	I-score	AE-score		Model type	R-score	I-score	AE-score
nadir	OI	87.32	72.17	–	nadir	∇_{OI}	78.03	75.97	–
	AnDA	94.85	77.91	–		∇_{AnDA}	85.56	79.14	–
	VE-DINEOF	96.11	72.72	–		$\nabla_{\text{VE-DINEOF}}$	82.69	75.61	–
	FP-ConvAE	87.82	76.32	82.85		$\nabla_{\text{FP-ConvAE}}$	77.80	76.81	75.89
	FP-GENN	91.78	84.56	93.15		$\nabla_{\text{FP-GENN}}$	81.05	80.56	84.24
nadir + SWOT	OI	93.25	74.25	–	nadir + SWOT	∇_{OI}	73.83	75.78	–
	AnDA	96.05	83.55	–		∇_{AnDA}	89.89	82.88	–
	VE-DINEOF	97.13	75.28	–		$\nabla_{\text{VE-DINEOF}}$	88.19	76.69	–
	FP-ConvAE	80.63	77.51	83.26		$\nabla_{\text{FP-ConvAE}}$	76.20	76.49	75.84
	FP-GENN	96.49	90.13	95.58		$\nabla_{\text{FP-GENN}}$	86.96	85.33	88.23

Table 1. 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

261 Last, when computing the radially averaged power spectra as a spatial domain averaged over the
 262 80-days validation period and the associated signal-to-noise ratio for joint use of along-track nadir
 263 with SWOT data (Figure 7b), we observe that AnDA and FP-GENN lead to a better constraint of the
 264 SSH spectrum compared to the actual OI capabilities. In particular, these two methods produce a
 265 spectrum closer to the ground truth real spectrum, by catching up the submesoscale range up to 50km
 266 if along-track nadir are only used (not shown here), a result improved up to 40km for AnDA and
 267 30km for FP-GENN when considering additional SWOT data. Let note on Figure 7b the importance
 268 of the patch-based AnDA post-processing on its performance which clearly appears on the spectra:
 269 its overestimation by the blocky patch-based AnDA rough outputs is partly mitigated thanks to the
 270 smoothing produced by averaging the patches overlapping areas. This result may certainly be further
 271 improved, for instance by training a CNN rather than using a simple average-based smoothing.

272
 273 To further enhance the visualization of the improvements brought by the different interpolators,
 274 Figures 8 and Figure 9 depict the velocity ground truth as well as its global reconstruction based on OI,
 275 (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN with both single along-track nadir data and
 276 joint use with wide-swath pseudo-observations on August 4, 2013. In Appendix A, complementary
 277 figures are provided for the SSH on the same day. To support what has already been said through the
 278 performance analysis previously discussed, FP-GENN using 5-days accumulated nadir observations
 279 appears closer to the ground truth SSH field than the reconstruction obtained with FP-ConvAE
 280 using a similar solver but a simple auto-encoder representation of the dynamics. The latter clearly
 281 oversmooths the true field and also exhibits some unnecessary artefacts on the SSH gradient thus
 282 explaining the noisy-related small scale energies on the spectra. The same artefacts appears on the
 283 VE-DINEOF mapping which exhibits discontinuities between the known wide-swath-informed areas
 284 and the filled missing data. Last, AnDA also behaves well, especially because the wide-swath SWOT
 285 data coverage on this specific day is important, getting its performance closer to FP-GENN than the
 286 day without the 2D-SWOT information.

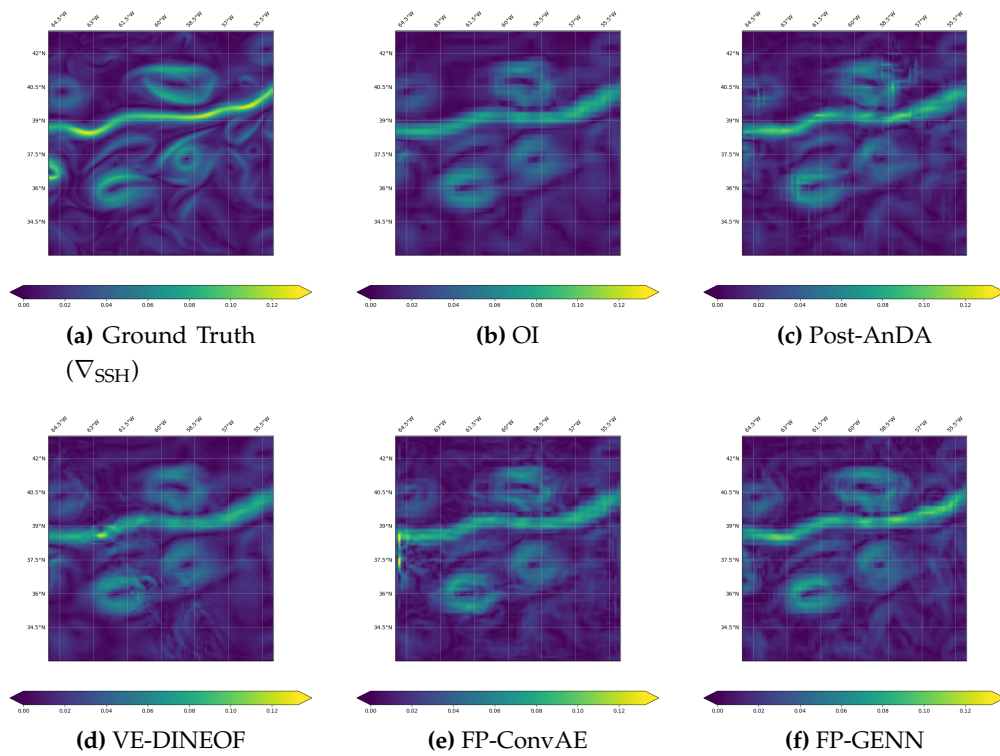


Figure 8. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

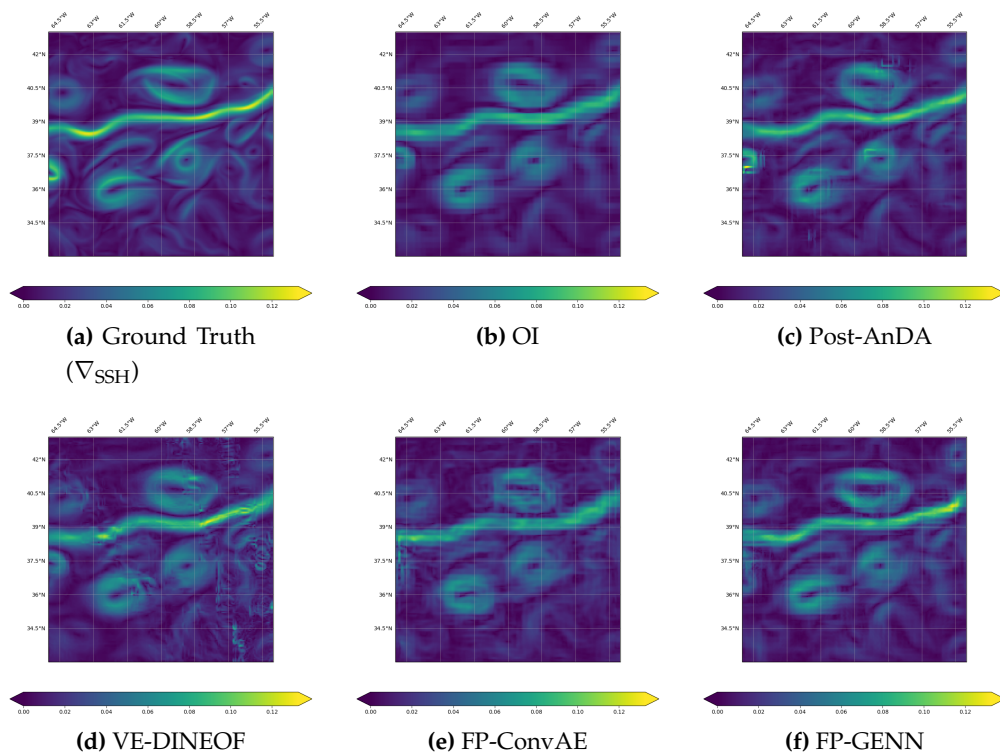


Figure 9. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

287 4.3. OSMOSIS

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

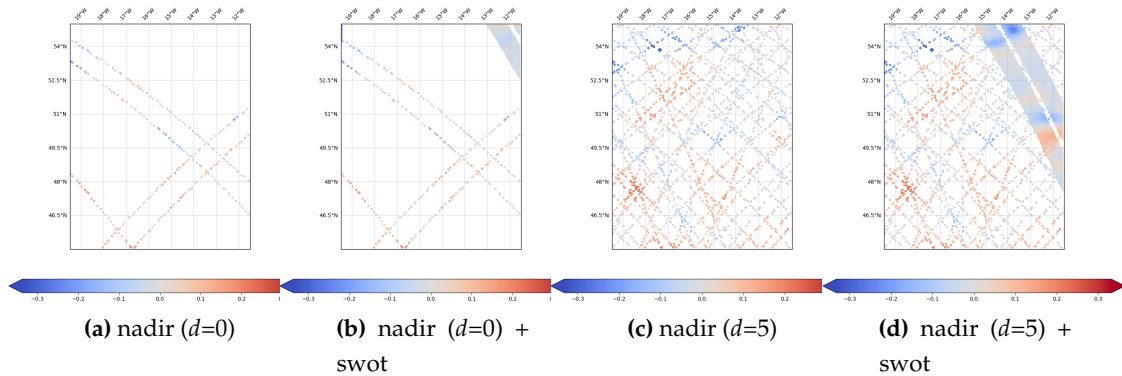


Figure 10. 0 and 5-days accumulated along-track nadir and wide-swath pseudo-observations on August 4, 2013 (a,b) and August 5, 2013 (c,d)

291 The daily nRMSE as a function of the along-track nadir time window parameter d (not shown
 292 here) leads to the same GULFSTREAM-related optimal values, namely ANDA behaves best when
 293 considering only the data retained to the targetted day t_k and both FP-ConvAE and FP-GENN
 294 performs better with $d = 5$.

295 Regarding the GENN configuration, the fully unsupervised FP-GENN-MM + OI configuration, the
 296 one using only the observations as both target and input (with OI as additional covariate) does not
 297 seem to perform well on the OSMOSIS domain, while it was the best option in the GULFSTREAM
 298 region.

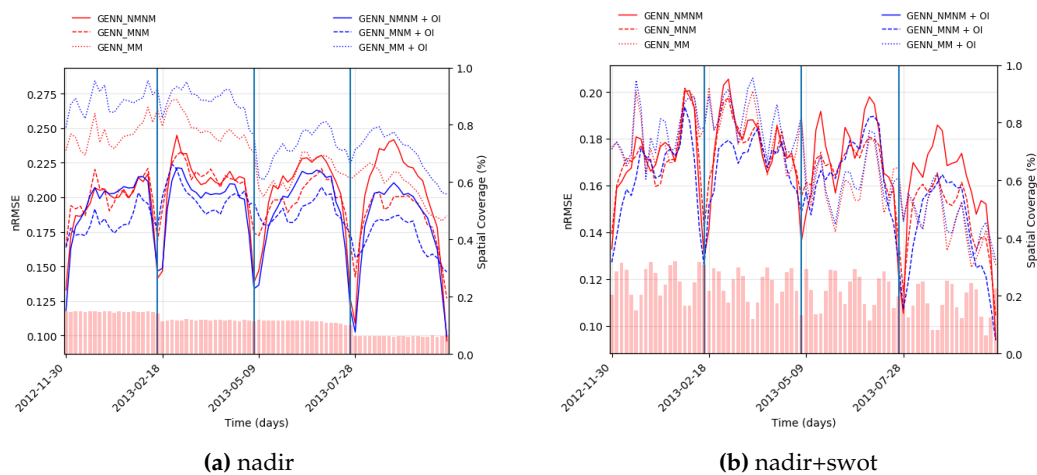


Figure 11. Daily spatial nRMSE computed on the 80-days non-continuous validation period for the six supervised/unsupervised FP-GENN configurations. The spatial coverage of 0-days accumulated along-track nadir (a) expanded with wide-swath SWOT data (b) is provided by the red-colored barplot

299 It is especially noticeable on the 20-days long time series, see Figure 11. However, this result
 300 should be qualified because when replicating the same preliminary work to find the best FP-GENN
 301 configuration but with no observation errors, see Figure A12 in Appendix B, the unsupervised
 302 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 2. Still, FP-GENN performs in a very similar way and the single use of nadir data is largely inferior to FP-GENN-MNM + OI.

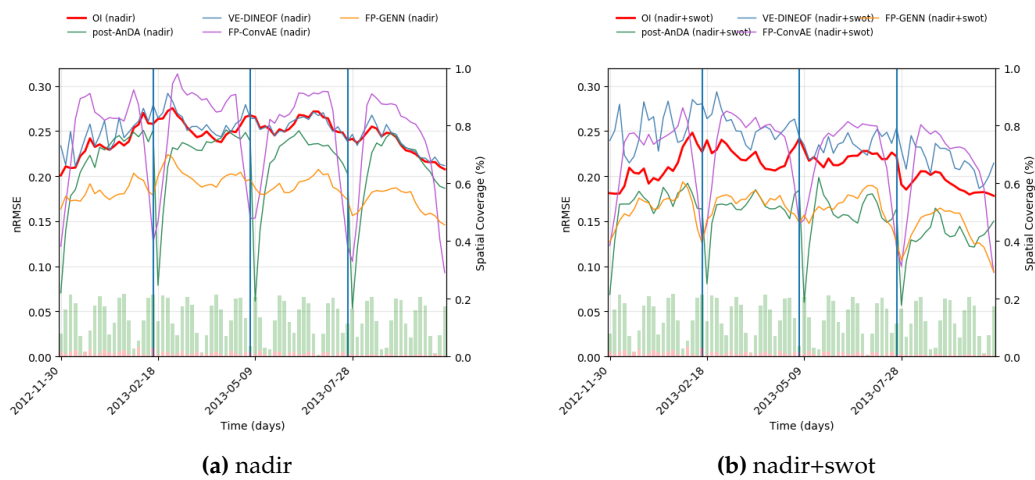


Figure 12. Daily spatial nRMSE computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN. The spatial coverage of 0-days accumulated along-track nadir and wide-swath SWOT data are respectively provided by the red and green-colored barplots

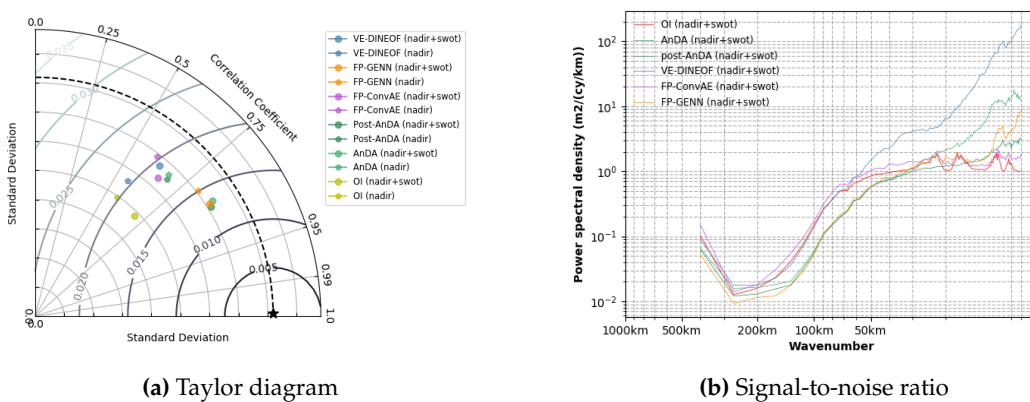


Figure 13. Taylor diagram and Signal-to-noise ratio computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN computed for both nadir use only and joint assimilation/learning with wide-swath SWOT data

315 On the spectral analysis in Figure 13b, the signal-to-noise ratio of FP-GENN and AnDA indicates
 316 a capability to retrieve spatial scales up to 40km, while the OI clearly only catches again the spatial
 317 scales over 100km. Again, let remain that when no observational errors are introduced, see Figure
 318 A14b in Appendix B, the fully unsupervised configuration of FP-GENN still behaves better. The single
 319 use of along-track nadir data clearly downgrades the performance of interpolations even if the gain
 320 remains significant for FP-GENN.

	Model type	R-score	I-score	AE-score		Model type	R-score	I-score	AE-score
nadir	OI	42.05	32.11	–	nadir	∇_{OI}	48.83	47.57	–
	AnDA	58.85	47.02	–		∇_{AnDA}	58.78	55.17	–
	VE-DINEOF	26.29	30.61	–		$\nabla_{VE-DINEOF}$	33.11	35.28	–
	FP-ConvAE	37.20	31.67	47.77		$\nabla_{FP-ConvAE}$	32.15	35.87	41.24
	FP-GENN	67.94	62.52	80.40		$\nabla_{FP-GENN}$	50.53	52.12	60.41
nadir + SWOT	OI	54.21	47.75	–	nadir + SWOT	∇_{OI}	36.83	47.30	–
	AnDA	81.15	70.91	–		∇_{AnDA}	72.35	67.59	–
	VE-DINEOF	69.08	32.98	–		$\nabla_{VE-DINEOF}$	22.08	24.90	–
	FP-ConvAE	45.15	42.70	47.93		$\nabla_{FP-ConvAE}$	38.22	43.13	42.03
	FP-GENN	77.16	69.56	83.08		$\nabla_{FP-GENN}$	56.29	59.21	67.69

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

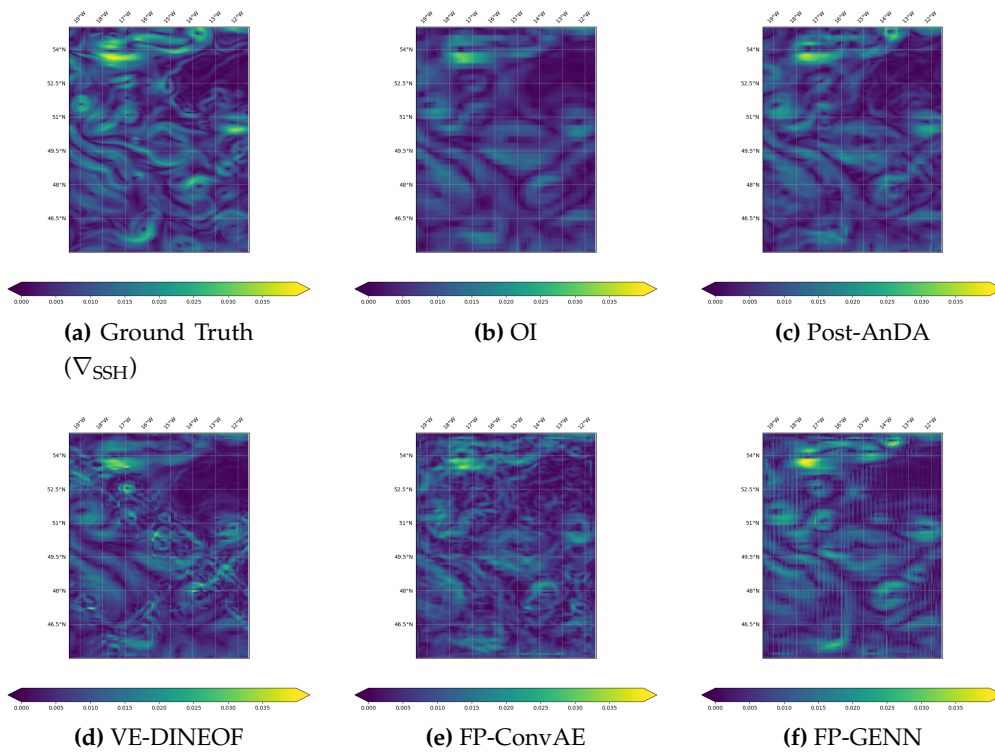


Figure 14. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

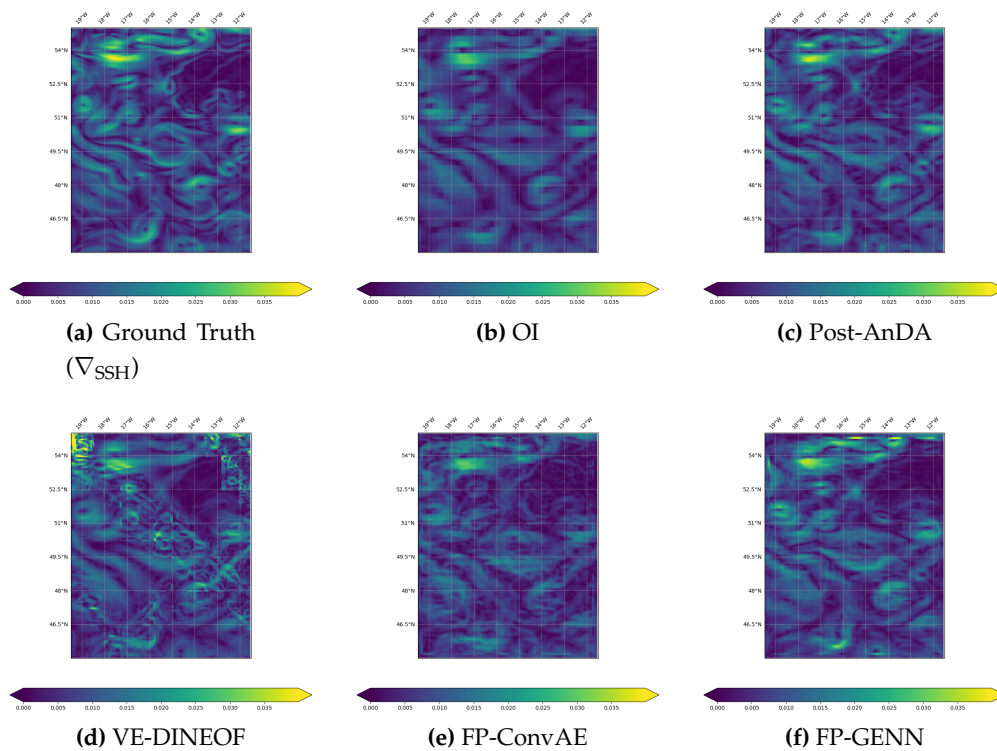


Figure 15. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

321 5. Discussion

322 In this study focusing on how data-driven and learning-based algorithms may help to improve
 323 the reconstruction performance of altimetric fields generally given by a state-of-the-art optimal
 324 interpolation (OI) baseline, here provided by the DUACS processing chain, we used two small
 325 areas with different energetic dynamics: the $10^\circ \times 10^\circ$ GULFSTREAM domain mainly driven by
 326 mesoscale processes and the $8^\circ \times 10^\circ$ OSMOSIS domain, less energetic but labelled with more
 327 small scale structures. Based on the NATL60 numerical simulations [6], some experiments were
 328 designed in which pseudo observational along-track nadir and wide-swath SWOT realistic datasets
 329 are generated. Because the DUACS OI [12] of these pseudo-observations is used as the reference, all
 330 the investigated methods are applied in a multi-scale decomposition framework where the anomaly dx
 331 is seen as the difference between the original field x and the large-scale component \bar{x} provided by the OI.

332
 333 Knowing the underlying reality, it was possible to precisely assess the reconstruction abilities of
 334 both AnDA and DINEOF data-driven methodologies, already consolidated with numerous experiences
 335 and methodological developments reported in the literature [1,4,5,13]. As a new competitive
 336 learning-based approach, we proposed to apply specifically interpolation-designed neural networks
 337 involving a joint interpolation and representation learning for irregularly-sampled satellite-derived
 338 geophysical fields [14]. As a short synthesis of these evaluations reported in Sections 4.2 and 4.3, some
 339 key points can be retrieved:

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

- 345 • A significant contribution from the 2D spatial information provided by the additional SWOT
346 sampling to improve the reconstruction of altimetric fields with a relative gain up to 30% on the SSH
347 daily mean squared error, when comparing to the single use of along-track nadir 1D information.
348 Within this combined use of the two datasets, the spectral analysis indicates the new capability to
349 reconstruct spatial scales up to 40km which is an important improvement compared to the scales
350 that OI is handling by now; on the other hand, the temporal sampling being less important than
351 nadir tracks, in particular on the GULFSTREAM domain where periods of several days without
352 any SWOT information appears, the reconstruction on these specific periods is sometimes better
353 when learning only with along-track nadir as inputs: we believe that a longer training period (not
354 available here) should improve the behaviour of the NN on this specific issue;
- 355 • The possibility of neural network methods to learn from the single observations, without requiring
356 any numerical simulations, which is of particular interest on low latitude areas where the Rossby
357 radius of deformation is large, thus requesting an important catalog to efficiently retrieve the SSH
358 dynamics over the year.

359 As it stands, the results obtained are very encouraging: FP-GENN is a "plug-and-play" algorithm
360 whose conceptual use easily enables its implementation on new datasets. Many perspectives have to
361 be considered in the short and medium terms.

362 The configuration of FP-GENN used here aims at minimizing the difference between the true anomaly
363 state of the system dx and its representation $\psi(dx)$ through energy form $\|dx - \psi(dx)\|^2$. Alternate
364 energy forms have to be investigated, considering extremes or more generally the whole pdf. In
365 addition, the fixed-point solver used in the joint interpolation approach with GENN never goes too
366 far from the observations, even though they are noisy, which can be an issue in the case of a strong
367 noise including spatial and/or temporal correlations, which was already seen when using SWOT data
368 without any preprocessing (not shown here).

369 From a methodological point of view, the next developments are expected in the coming related works
370 to increase the gain already observed with FP-GENN:

- 371 • use a joint learning of the dynamical representation ψ and the solver Γ minimizing its reconstruction
372 error. A significant gain on the reconstruction performance is expected according to preliminary
373 results obtained with toy models [15];
- 374 • a stochastic extension of GENN for including in the NN-based framework an estimation of the
375 uncertainties, thus enabling this new reconstruction method to fully compete with the other
376 interpolators in a "data assimilation" context, with a possible link with Gaussian Processes and the
377 related Stochastic PDE formalism [16,17].

378 Besides methodological aspects, new applications are also promising. If we focused here on small
379 North-Atlantic subdomains, the transfer of the NN-based interpolators to an operational process chain
380 will be to reproduce a similar work on the whole basin where the computational constraints in this
381 learning-based setting with large number of parameters is still a challenge. Using a Deep Learning
382 multi-GPU framework and build a pre-operational demonstrator should be of great interest in the
383 community, as are other SWOT use cases, e.g. using a pre-learning on SWOT data to produce a new
384 interpolation of historical along-track nadir datasets, or taking advantage of the SWOT fast-sampling
385 phase data as inputs for learning prior to its use with SWOT upcoming "operational" data. Last,
386 because the 2D information brought by SWOT showed a significant gain in the reconstruction, a
387 natural extension of this work would be to consider pseudo-observations SKIM datasets [18], whose
388 swath width is more than twice larger (110km vs 270km), and also to propose multivariate analyses
389 including complementary datasets (SST/SSS), already existing in other data-driven schemes like AnDA
390 with an easy extension as additional channels in a neural networks framework.

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

393 **Author Contributions:** "Conceptualization, Ronan Fablet and Maxime Beauchamp; methodology, Ronan Fablet
394 and Maxime Beauchamp; software, Ronan Fablet and Maxime Beauchamp; validation, Maxime Beauchamp;

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

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

399 References

400

- 401 1. Lguensat, R.; Tandeo, P.; Aillot, P.; Fablet, R. The Analog Data Assimilation. *Monthly Weather Review* **2017**.
- 402 2. Lguensat, R.; Huynh Viet, P.; Sun, M.; Chen, G.; Fenglin, T.; Chapron, B.; Fablet, R. Data-driven
403 Interpolation of Sea Level Anomalies using Analog Data Assimilation **2017**.
- 404 3. Fablet, R.; Viet, P.H.; Lguensat, R. Data-Driven Models for the Spatio-Temporal Interpolation
405 of Satellite-Derived SST Fields. *IEEE Trans. on Computational Imaging* **2017**, *3*, 647–657.
406 doi:10.1109/TCI.2017.2749184.
- 407 4. Lopez-Radcenco, M.; Pascual, A.; Gomez-Navarro, L.; Aissa-El-Bey, A.; Chapron, B.; Fablet, R. Analog
408 Data Assimilation of Along-Track Nadir and Wide-Swath SWOT Altimetry Observations in the Western
409 Mediterranean Sea. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2019**, pp.
410 1–11. doi:10.1109/JSTARS.2019.2903941.
- 411 5. Oualla, S.; Fablet, R.; Herzet, C.; Chapron, B.; Pascual, A.; Collard, F.; Gaultier, L. Neural Network Based
412 Kalman Filters for the Spatio-Temporal Interpolation of Satellite-Derived Sea Surface Temperature. *Remote*
413 *Sensing* **2018**, *10*, 1864. doi:10.3390/rs10121864.
- 414 6. Molines, J.M. meom-configurations/NATL60-CJM165: NATL60 code used for CJM165 experiment, 2018.
415 doi:10.5281/zenodo.1210116.
- 416 7. Ballarotta, M.; Ubelmann, C.; Pujol, M.I.; Taburet, G.; Fournier, F.; Legeais, J.F.; Faugère, Y.; Delepouille,
417 A.; Chelton, D.; Dibarboure, G.; Picot, N. On the resolutions of ocean altimetry maps. *Ocean Science* **2019**,
418 *15*, 1091–1109. doi:10.5194/os-15-1091-2019.
- 419 8. Gaultier, L.; Ubelmann, C.; Fu, L.L. The Challenge of Using Future SWOT Data for
420 Oceanic Field Reconstruction. *Journal of Atmospheric and Oceanic Technology* **2015**, *33*, 119–126.
421 doi:10.1175/JTECH-D-15-0160.1.
- 422 9. Esteban-Fernandez, D. SWOT project mission performance and error budget document. Technical report,
423 JPL, NASA, 2014.
- 424 10. Gaultier, L.; Ubelmann, C. SWOT Simulator Documentation. Technical report, JPL, NASA, 2010.
- 425 11. Metref, S.; Cosme, E.; Le Guillou, F.; Le Sommer, J.; Brankart, J.M.; Verron, J. Wide-Swath Altimetric
426 Satellite Data Assimilation With Correlated-Error Reduction. *Front. Mar. Sci.* **2020**, *6*:822. doi:doi:
427 10.3389/fmars.2019.00822.
- 428 12. Taburet, G.; Sanchez-Roman, A.; Ballarotta, M.; Pujol, M.I.; Legeais, J.F.; Fournier, F.; Faugere, Y.;
429 Dibarboure, G. DUACS DT2018: 25 years of reprocessed sea level altimetry products. *15*, 1207–1224.
430 Publisher: Copernicus GmbH, doi:https://doi.org/10.5194/os-15-1207-2019.
- 431 13. Ping, B.; Su, F.; Meng, Y. An Improved DINEOF Algorithm for Filling Missing Values in Spatio-Temporal
432 Sea Surface Temperature Data. *PLOS ONE* **2016**, *11*, e0155928. doi:10.1371/journal.pone.0155928.
- 433 14. Fablet, R.; Drumetz, L.; Rousseau, F.; Beauchamp, M. Joint interpolation and representation learning for
434 irregularly-sampled satellite-derived geophysical fields.
- 435 15. Fablet, R.; Drumetz, L.; Rousseau, F. Joint learning of variational representations and solvers for inverse
436 problems with partially-observed data, 2020, [arXiv:cs.LG/2006.03653].
- 437 16. Lindgren, F.; Rue, H.; Lindström, J. An explicit link between Gaussian fields
438 and Gaussian Markov random fields: the stochastic partial differential equation
439 approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **2011**,
440 *73*, 423–498, [https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2011.00777.x].
441 doi:10.1111/j.1467-9868.2011.00777.x.
- 442 17. Sidén, P.; Lindsten, F. Deep Gaussian Markov random fields, 2020, [arXiv:stat.ML/2002.07467].
- 443 18. Arduin, F.; Brandt, P.; Gaultier, L.; Donlon, C.; Battaglia, A.; Boy, F.; Casal, T.; Chapron, B.; Collard, F.;
444 Cravatte, S.; Delouis, J.M.; De Witte, E.; Dibarboure, G.; Engen, G.; Johnsen, H.; Lique, C.; Lopez-Dekker,

445 P.; Maes, C.; Martin, A.; Marié, L.; Menemenlis, D.; Nouguier, F.; Peureux, C.; Rampal, P.; Ressler, G.; Rio,
446 M.H.; Rommen, B.; Shutler, J.D.; Suess, M.; Tsamados, M.; Ubelmann, C.; Van Sebille, E.; Van Den Oever,
447 M.; Stammer, D. SKIM, a Candidate Satellite Mission Exploring Global Ocean Currents and Waves, 2019.
448 doi:<https://doi.org/10.3389/fmars.2019.00209>.

449 **Appendix A. Complementary figures for SSH interpolations**

450 *Appendix A.1. GULFSTREAM*

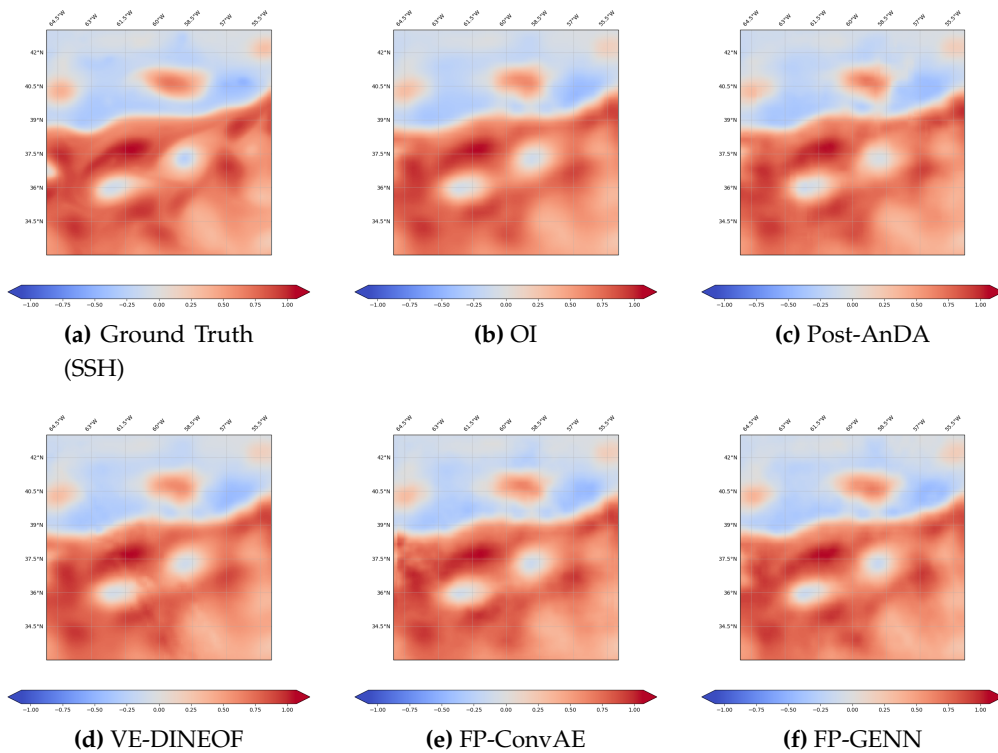


Figure A1. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

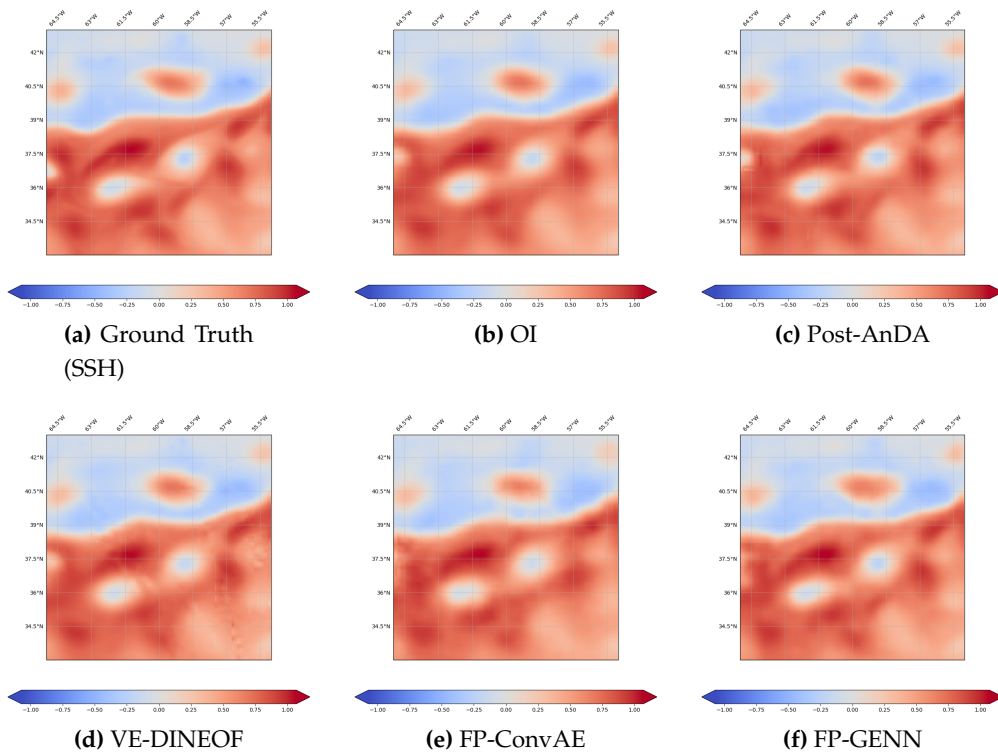


Figure A2. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

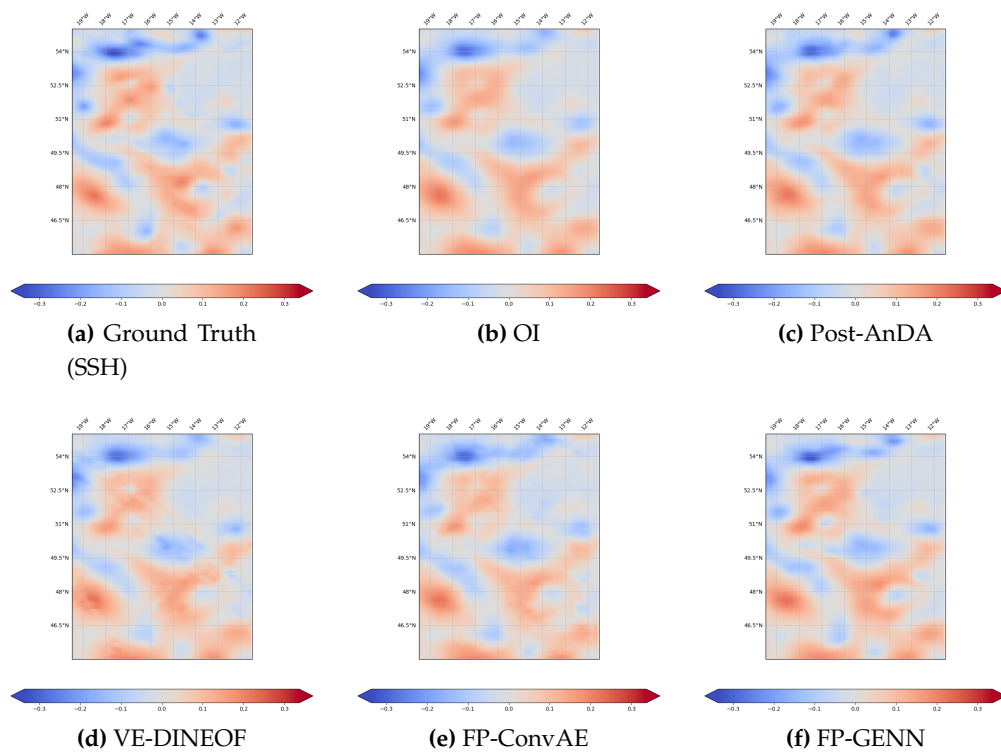
451 *Appendix A.2. OSMOSIS*

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

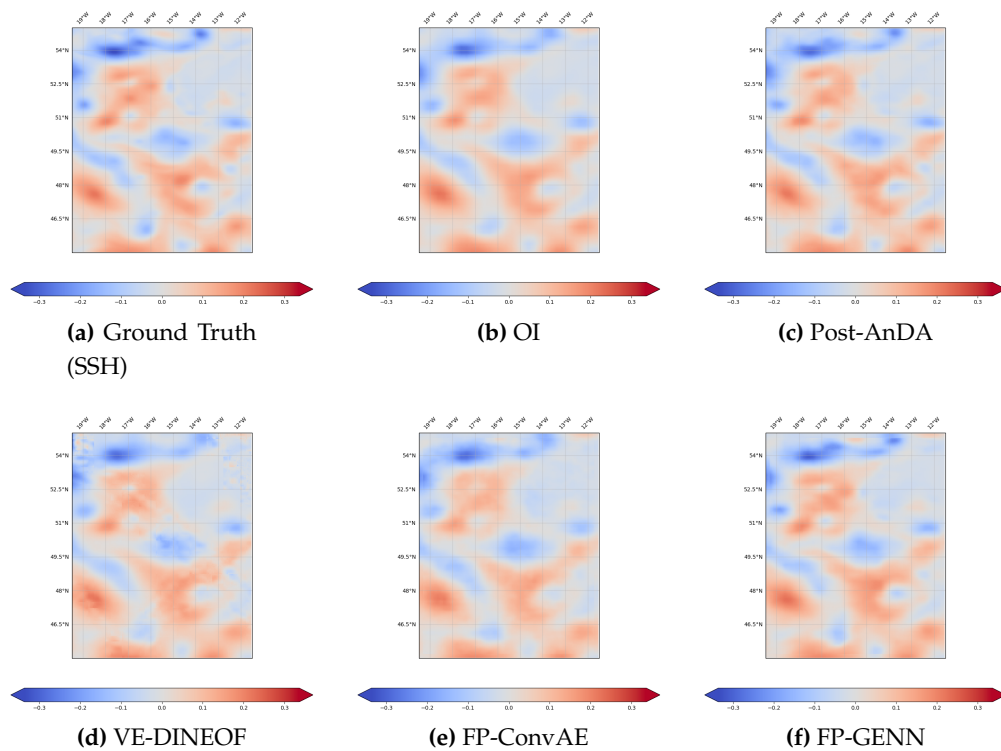


Figure A4. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

452 **Appendix B. OSSE without observation errors**

453 *Appendix B.1. GULFSTREAM*

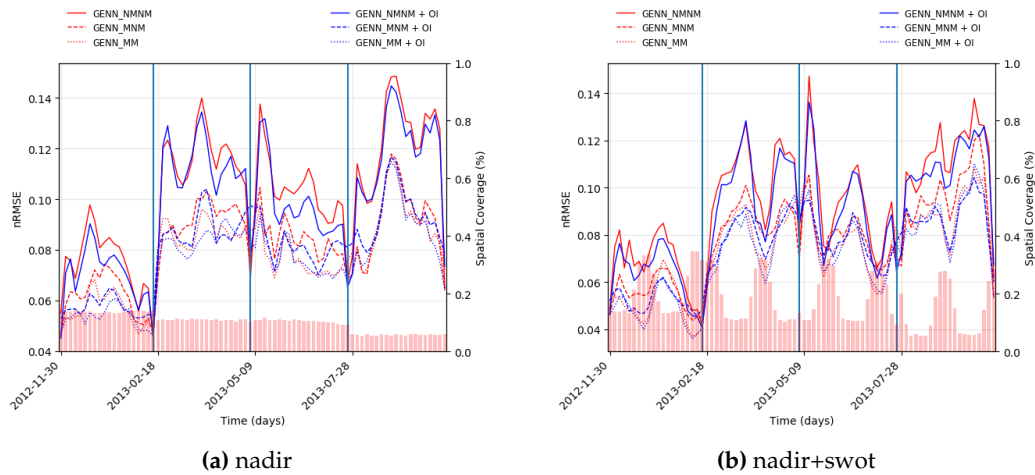


Figure A5. Daily spatial nRMSE computed on the 80-days non-continuous validation period for the six supervised/unsupervised FP-GENN configurations. The spatial coverage of 0-days accumulated along-track nadir (a) expanded with wide-swath SWOT data (b) is provided by the red-colored barplot

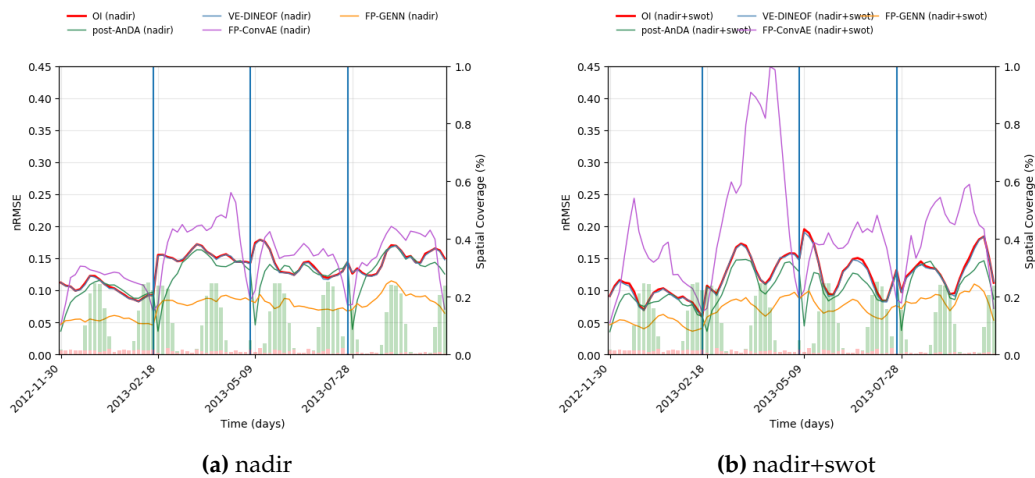


Figure A6. Daily spatial nRMSE computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN. The spatial coverage of 0-days accumulated along-track nadir and wide-swath SWOT data are respectively provided by the red and green-colored barplots

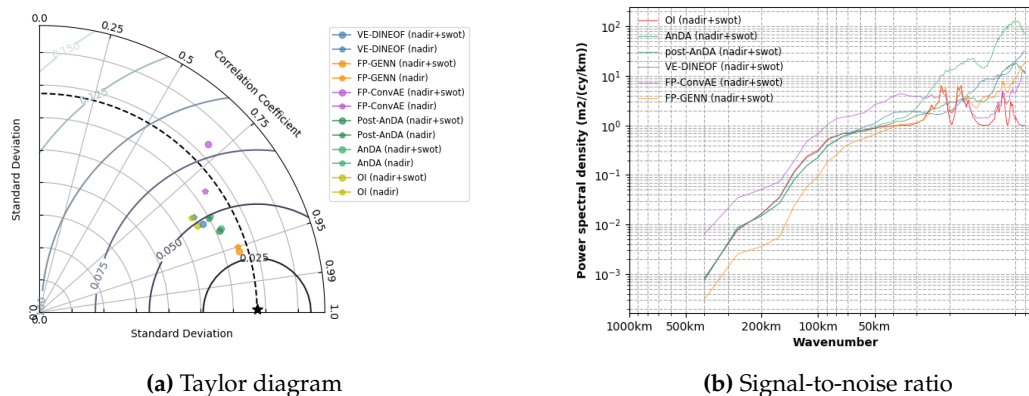


Figure A7. Taylor diagram and Signal-to-noise ratio computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN computed for both nadir use only and joint assimilation/learning with wide-swath SWOT data

	Model type	R-score	I-score	AE-score		Model type	R-score	I-score	AE-score
nadir	OI	86.53	72.25	–	nadir	∇_{OI}	76.14	72.41	–
	AnDA	90.56	76.81	–		∇_{AnDA}	81.81	76.15	–
	VE-DINEOF	91.33	72.58	–		$\nabla_{VE-DINEOF}$	80.09	72.07	–
	FP-ConvAE	69.46	63.82	79.86		$\nabla_{FP-ConvAE}$	58.30	59.79	70.14
	FP-GENN	95.15	91.28	96.32		$\nabla_{FP-GENN}$	84.75	84.63	88.05
nadir + SWOT	OI	91.76	75.30	–	nadir + SWOT	∇_{OI}	71.41	72.31	–
	AnDA	91.72	82.43	–		∇_{AnDA}	85.85	79.80	–
	VE-DINEOF	92.47	76.00	–		$\nabla_{VE-DINEOF}$	84.73	73.36	–
	FP-ConvAE	42.78	34.96	79.93		$\nabla_{FP-ConvAE}$	31.78	36.48	69.72
	FP-GENN	97.31	91.45	96.87		$\nabla_{FP-GENN}$	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

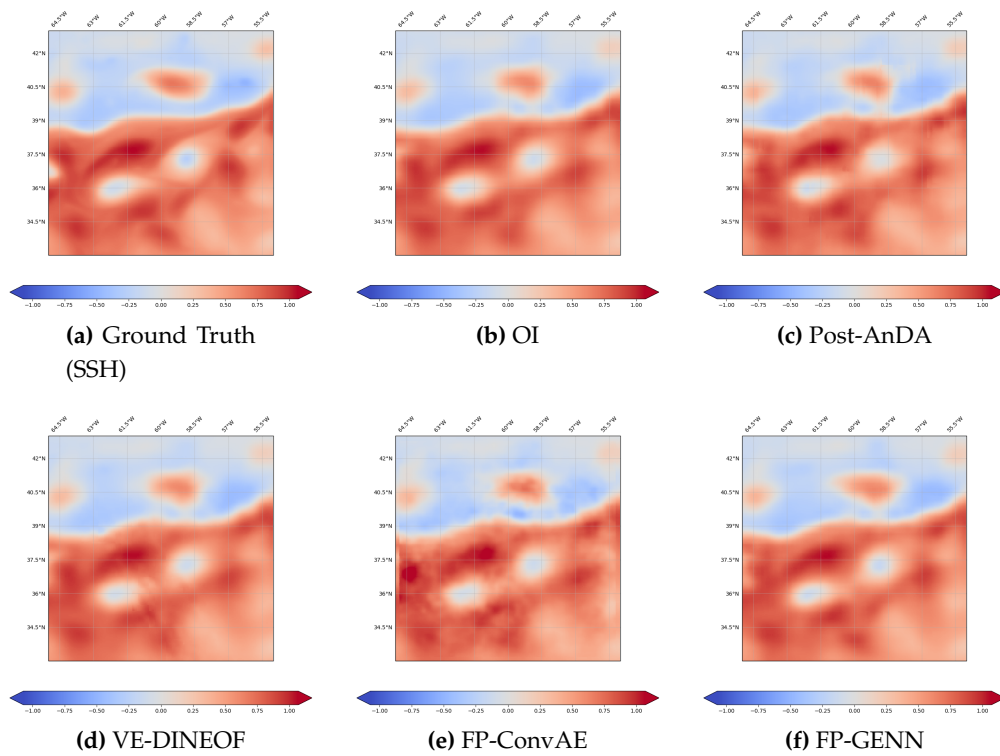


Figure A8. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

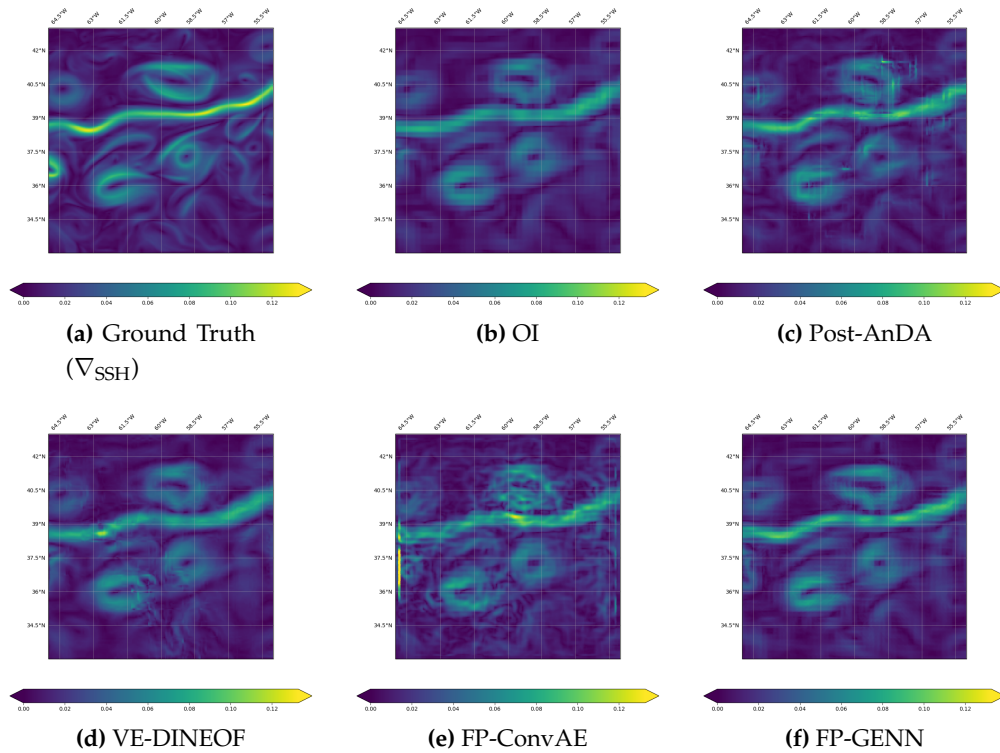


Figure A9. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

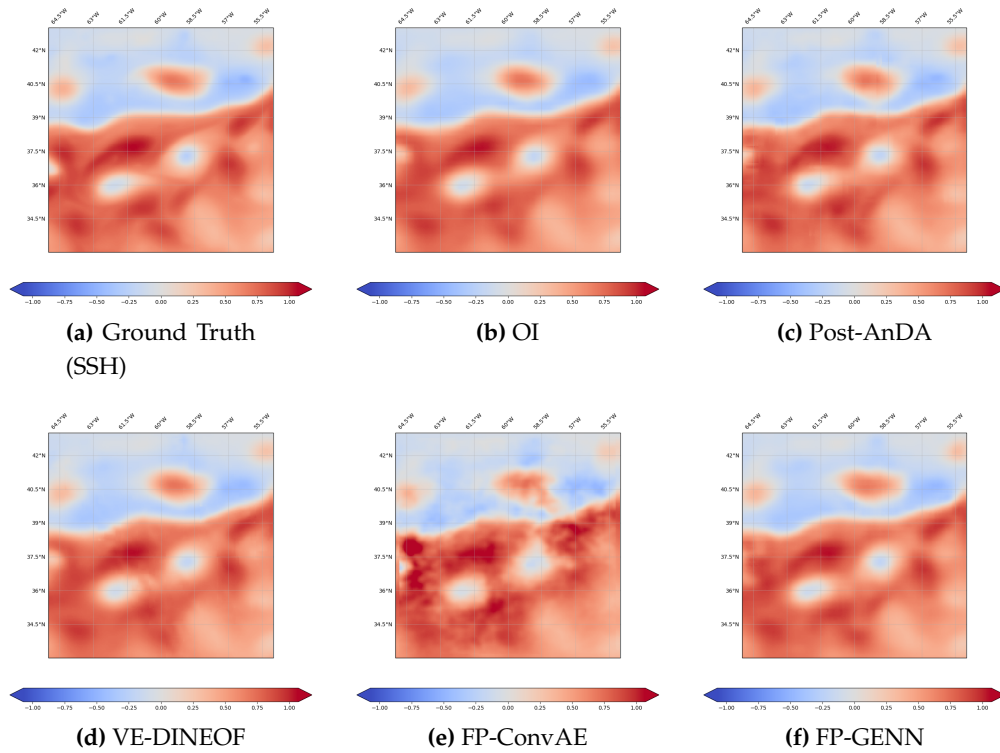


Figure A10. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

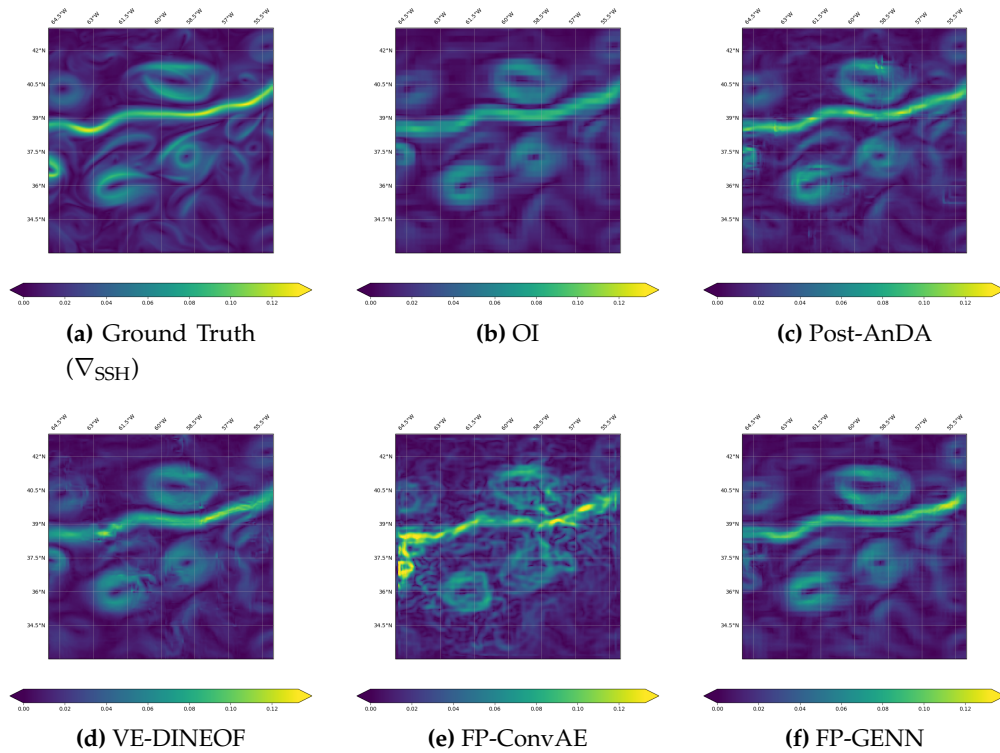


Figure A11. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

454 *Appendix B.2. OSMOSIS*

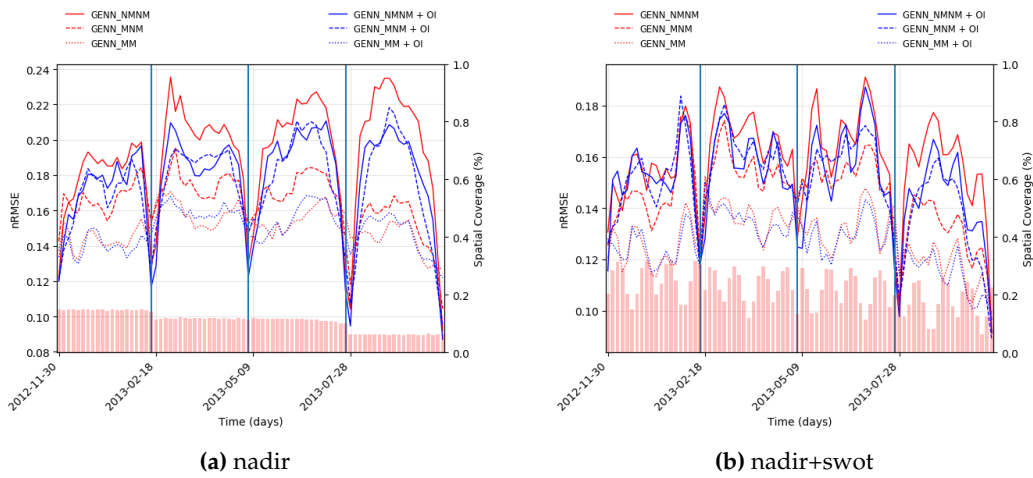


Figure A12. Daily spatial nRMSE computed on the 80-days non-continuous validation period for the six supervised/unsupervised FP-GENN configurations. The spatial coverage of 0-days accumulated along-track nadir (a) expanded with wide-swath SWOT data (b) is provided by the red-colored barplot

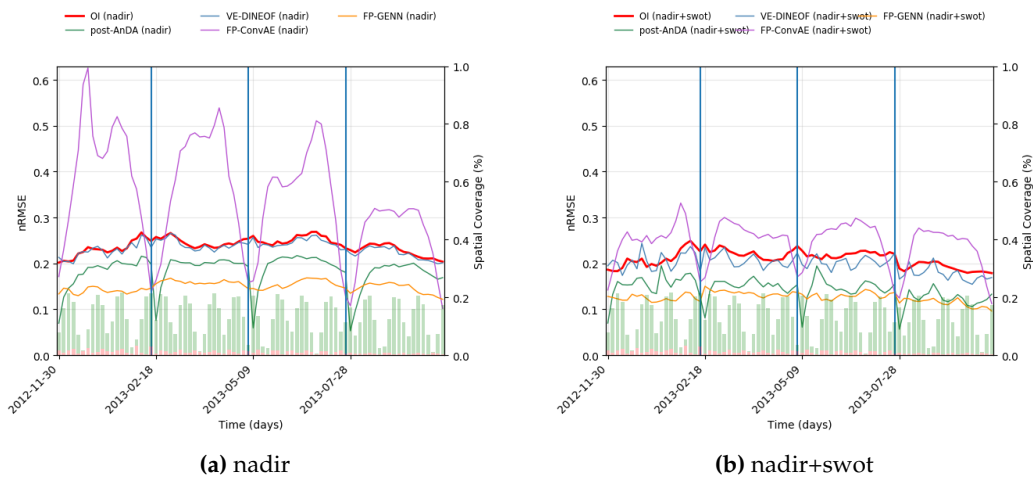


Figure A13. Daily spatial nRMSE computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN. The spatial coverage of 0-days accumulated along-track nadir and wide-swath SWOT data are respectively provided by the red and green-colored barplots

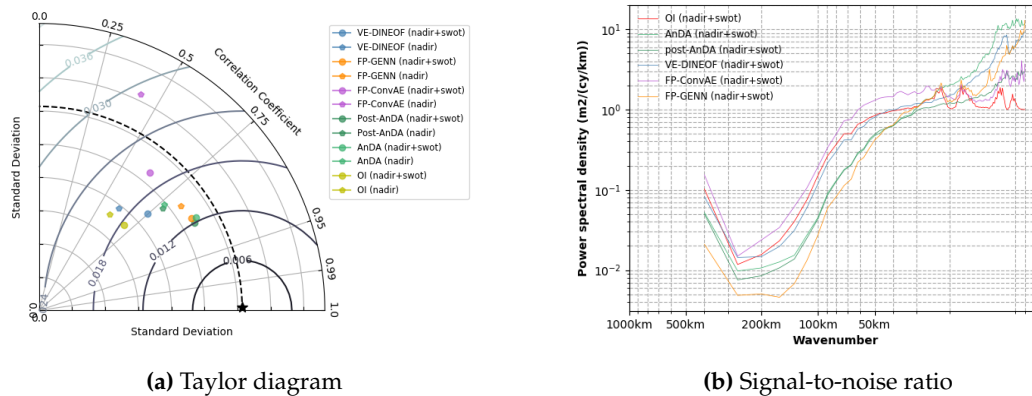


Figure A14. Taylor diagram and Signal-to-noise ratio computed on the 80-days non-continuous validation period for OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN computed for both nadir use only and joint assimilation/learning with wide-swath SWOT data

	Model type	R-score	I-score	AE-score		Model type	R-score	I-score	AE-score
nadir	OI	44.63	34.93	–	nadir	∇_{OI}	49.53	48.20	–
	AnDA	76.60	59.42	–		∇_{AnDA}	64.56	59.88	–
	VE-DINEOF	77.17	37.66	–		$\nabla_{\text{VE-DINEOF}}$	58.71	45.61	–
	FP-ConvAE	28.39	17.00	42.94		$\nabla_{\text{FP-ConvAE}}$	22.47	19.12	36.66
	FP-GENN	84.35	76.17	86.30		$\nabla_{\text{FP-GENN}}$	62.47	61.64	64.88
nadir + SWOT	OI	54.31	47.87	–	nadir + SWOT	∇_{OI}	37.55	47.93	–
	AnDA	83.07	74.95	–		∇_{AnDA}	75.13	70.22	–
	VE-DINEOF	83.47	51.50	–		$\nabla_{\text{VE-DINEOF}}$	79.31	49.32	–
	FP-ConvAE	36.80	33.37	47.56		$\nabla_{\text{FP-ConvAE}}$	30.85	35.06	39.06
	FP-GENN	90.67	81.35	88.04		$\nabla_{\text{FP-GENN}}$	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

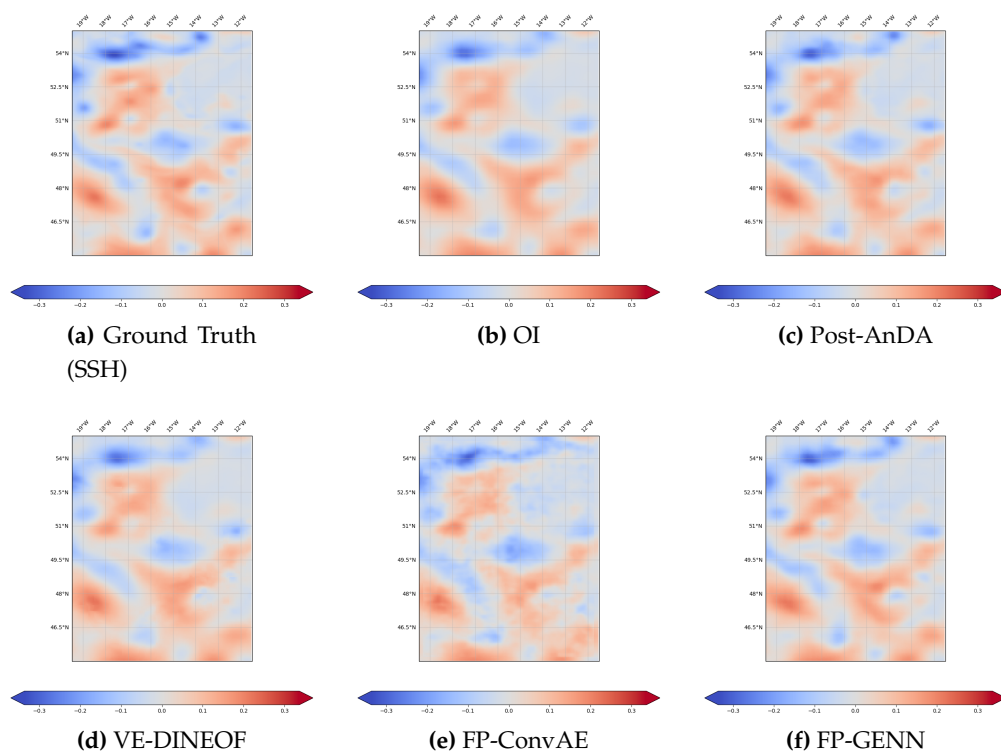


Figure A15. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

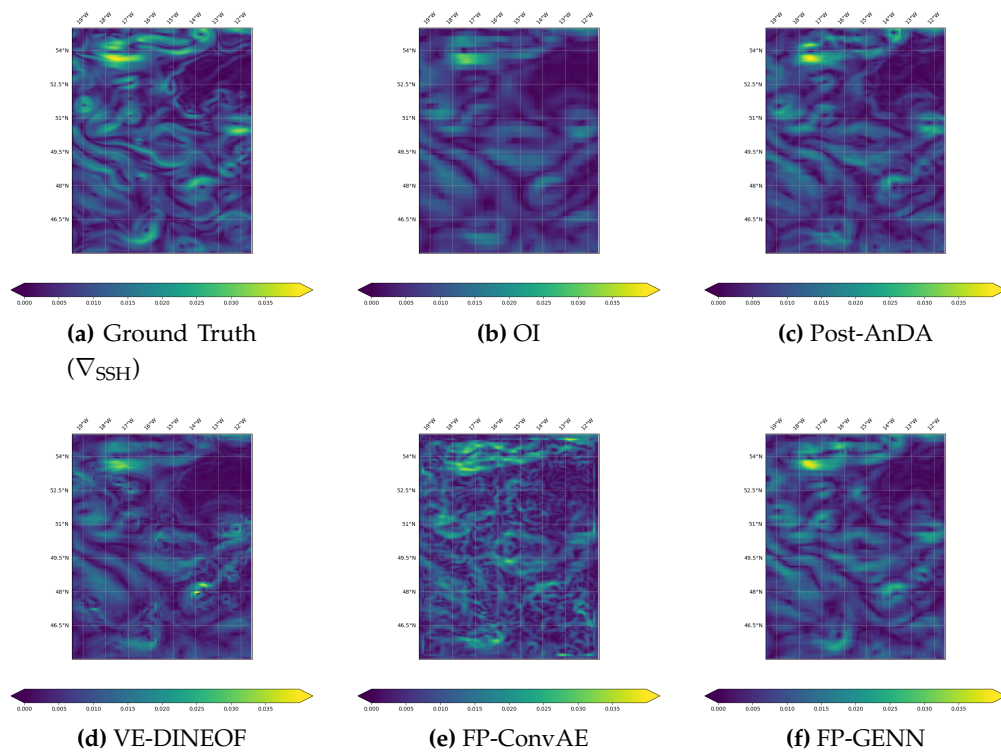


Figure A16. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN using along-track nadir data only

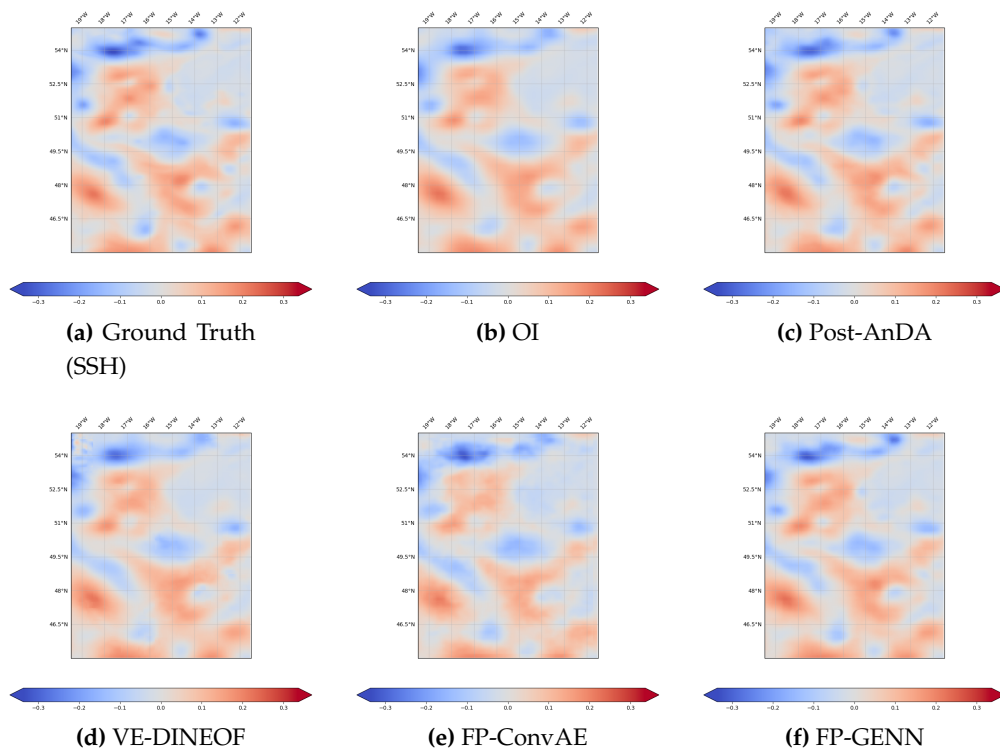


Figure A17. Global SSH field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

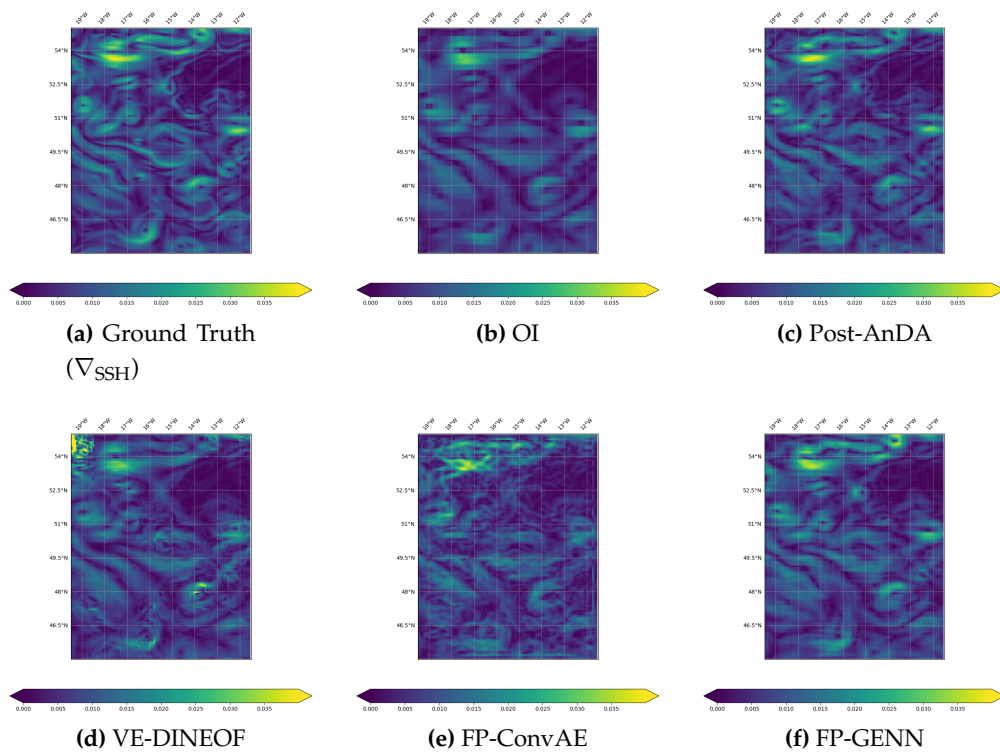


Figure A18. Global SSH gradient field reconstruction obtained by OI, (post-)AnDA, VE-DINEOF, FP-ConvAE and FP-GENN for a joint assimilation/learning of along-track nadir with wide-swath SWOT data

455 © 2020 by the authors. Submitted to *Remote Sens.* for possible open access publication
456 under the terms and conditions of the Creative Commons Attribution (CC BY) license
457 (<http://creativecommons.org/licenses/by/4.0/>).