

Fast ocean data assimilation using a neural-network reduced-space regional ocean model of the North Brazil Current

C.A. Quilodrán Casas, N. Sparks and R. Toumi

Space and Atmospheric Physics, Imperial College London, United Kingdom

Abstract

Data assimilation is computationally demanding, typically many times slower than model forecasts. Fast and reliable ocean assimilation methods are attractive for multiple applications such as emergency situations, search and rescue, and oil spills. A novel framework which performs fast data assimilation with sufficient accuracy is proposed for the first time in the open ocean. Speed improvement is achieved by performing the data assimilation in a reduced-space rather than in a full-space. A surface 10 km resolution hindcast of the North Brazil current from the Regional Ocean Modelling System (ROMS) serves as the full-space state. The target variables are sea surface height, sea surface temperature, and surface currents. A dimension reduction of the full-state is made by an Empirical Orthogonal Function analysis which retains most of the explained variance. The dynamics are replicated by a state-of-the-art neural network trained on the truncated principal components of the full-state. An Ensemble Kalman filter assimilates the data in the reduced-space, where the trained neural network produces short-range fore-

Email address: caq13@ic.ac.uk ()

casts from perturbed ensembles. The Ensemble Kalman filter is successful in reducing the root mean squared error by $\sim 45\%$ and increases the correlations between state variables and data. The performance is similar to other full-space data assimilation studies. However, the required computations are two orders of magnitude faster than other full-space data assimilation schemes. The approach could be a useful tool when full regional models are not available to make a forecast.

Keywords: Reduced-space models, Data assimilation, Neural Networks, Kalman filter, Ocean currents

¹ 1. Introduction

² Data assimilation (DA) is a method to conciliate a large variety and
³ volume of observational data with numerical models and can be used to
⁴ find optimal initial conditions [31] for better forecasts by numerical models
⁵ which simulate physical processes. The high-dimensionality of these numer-
⁶ ical models contributes to their high computational-cost. The application
⁷ of a DA scheme in such models increases this cost, because it typically re-
⁸ quires multiple runs of the numerical model making it hard to satisfy the
⁹ requirements of speed and accuracy at the same time [34].

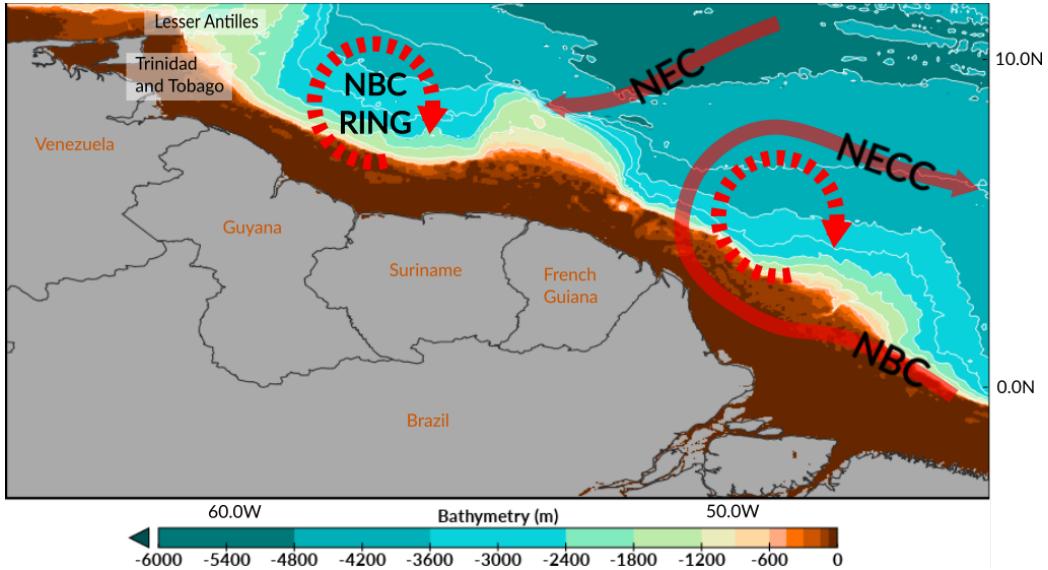


Figure 1: Schematic representation of the North Equatorial Current (NEC), North Equatorial Countercurrent (NECC), North Brazil Current (NBC) and eddy generation (NBC rings) at the western boundary of the tropical Atlantic with warm-water pathways in solid red.

¹⁰ An approach that has gained some attention is to use fast and cheap

11 model substitutes, surrogates or emulators of more complex models to per-
12 form data assimilation [17, 35, 33]. The idea is to reduce the high-dimensionality
13 and/or complexity of the original full-space model and replace it with the
14 model substitute which can be orders of magnitude faster and with an ac-
15 ceptable loss in accuracy. A caveat is that an emulator created for a specific
16 model region or application may not perform well in other regions or appli-
17 cations. Several studies have used emulators to replicate or approximate a
18 full-space model. These applications range from idealised models of ocean
19 eddies [29], to real ocean applications using polynomial chaos expansion [35],
20 rank-reduced statistical emulation [31, 21], and sequential emulator-based
21 assimilation on coastal sediment [33, 32]

22 Other approaches include the replacement of a high dimensional physi-
23 cal model with a lower dimension neural network, trained on reduced-space
24 data. Dimensional reduction of data is performed using Empirical Orthog-
25 onal Functions (EOF) analysis [46]. For the training, it is essential to have
26 a long run of the hindcast model to provide a sufficiently large dataset,
27 which sufficiently captures the physics and dynamics of the original high-
28 dimensional model. Frolov et al. [17] successfully expanded this into a data
29 assimilation framework using data from observations and combining them
30 with an estuary model of the Columbia River. We apply a framework based
31 on Frolov et al. [17], but for the first time in a regional model of the open
32 ocean.

33 The North Brazil Current (NBC) (Fig. 1) was chosen as the testbed for
34 this fast DA scheme, because it is a complex eddy-dominated area. Several
35 studies have been performed on the NBC [24, 15, 13, 10, 16, 25, 4, 40, 23].

36 The NBC is a strong low-latitude western boundary current along the East
37 and North coast of South America carrying water northward supplied by
38 the South Equatorial Current (SEC) from the tip of Brazil near 5°S across
39 the Equator and into the northern hemisphere [10]. Low-latitude western
40 boundary currents are vital in the transport across the Equator and recir-
41 culation of water within zonal equatorial systems [24, 15]. The NBC has a
42 width between 100 and 200 km, with a 500 m depth extension of its flow,
43 and total volume transport of around 50 Sv ($1Sv = 10^6 \text{ m}^3 \text{ s}^{-1}$) across 2°N
44 [13]. During June, July and August (JJA) and September, October and
45 November (SON), the NBC surface flow drifts away from the coast at about
46 6°N and feeds the eastward North Equatorial Countercurrent (NECC). There
47 is observational evidence for seasonal flow variability above the thermocline
48 that connects the NECC, NBC and SEC [6]. The separation process of the
49 NBC is more complex and greatly layered. Flow underneath the thermocline
50 splits apart near 3°N and feeds an eastward thermocline current, while the
51 flow in the thermocline separates just north of the Equator and supplies the
52 Equatorial undercurrent [24].

53 Observations suggest the presence of large long-lived (100 days), anticy-
54 clonic rings that contribute to the mean circulation [24]. These features ap-
55 pear between 4°N and 8°N and are referred to as NBC rings [3]. Detachment
56 of these rings is observed during November through January where the NBC
57 retroflection into the NECC weakens. Figure 1 shows a schematic diagram
58 of these eddies along with the NBC, NECC, and North Equatorial Current.
59 These rings can exceed 450 km in overall diameter with strong azimuthal
60 velocities [16]. Annually, 5 to 6 NBC rings are formed around 50°W and

61 move towards the southeastern Caribbean, near the Lesser Antilles, where
62 they dissolve. In the context of the larger North Atlantic circulation, the
63 transport of the NBC rings has been estimated as 13 - 15 Sv of warm water
64 transport to close the Atlantic Meridional Overturning Circulation [25].

65 **2. The Framework: DA+NN**

66 The framework presented here, DA+NN (Data assimilation and neural
67 network), aims to produce fast DA with sufficiently accurate results. This is
68 achieved by building a reduced-space model that acts as a model substitute
69 to the full-space model. Therefore, the handling of the model and DA occurs
70 in the reduced-space making it faster and less computationally expensive.
71 Reduced dimensionality data is used to train an Artificial Neural Network
72 (NN). To reduce the dimensionality, we use EOFs which have been widely
73 used in atmospheric, ocean and climate sciences [30, 42, 19]. The aim of
74 EOFs is to reduce the number of variables of high-dimensional systems while
75 capturing most of the variance [18]. NNs, used in time series contexts, are
76 skillful tools in extracting patterns from past events and extrapolating these
77 into future time-steps [28]. They approximate any given continuous function
78 with required accuracy and are data-driven. This is particularly helpful when
79 substantial quantities of data are available [20].

80 DA+NN is represented in Fig. 2 and can be outlined in three steps: (a)
81 dimension reduction and NN training, (b) DA, and (c) reconstruction of the
82 full-space.

83 *2.1. ROMS Model and study area*

84 The Regional Ocean Modelling System (ROMS) is used. This is a three-
85 dimensional, free-surface, primitive equations ocean model and it is com-
86 prehensively described by Shchepetkin and McWilliams [44]. An extensive
87 explanation of the following ROMS model setup can be found in Newinger
88 [39] and Newinger and Toumi [40].

89 The domain of the ROMS model (Fig. 1) is enclosed by 2.5°S - 11.5°N and
90 63.5°W - 46°W, and covers about 2000 km × 1500 km including the Ama-
91 zon and Orinoco River mouths, i.e. Northeastern South America including
92 Brazil, French Guyana, Suriname, Guyana and Venezuela, and Trinidad and
93 Tobago along with some islands in the Caribbean. The model has a horizon-
94 tal resolution of 10 km and 50 vertical levels.

95 The open ocean boundary conditions were obtained from the monthly
96 mean GLORYS2v3 model with a 0.25 degree resolution. The National Center
97 for Environment Prediction (NCEP) Climate Forecast System Reanalysis
98 (CFSR) was utilised for the atmospheric boundary conditions.

99 *2.2. Dimension reduction algorithm*

100 The surface layer of the 10-year (2001-01-01 to 2010-12-31) daily hind-
101 cast ROMS simulation of the North Brazil region is utilised as the original
102 full-state (FS), where 2010 serves as a validation year. Therefore, the EOF
103 analysis for training the NN will only consider the first 9 years. The variables
104 extracted from the ROMS model solution were Sea Surface Height (SSH), Sea
105 Surface Temperature (SST), Eastward and Northward horizontal velocities
106 (U and V, respectively) of the surface currents. The FS can provide addi-

107 tional variables, but this study is limited to and targeted at the previously
108 mentioned variables.

109 We used the `eofs` Python package to perform the EOF analysis which
110 is computationally efficient for large output datasets. This package uses
111 a technique based on SVD. A further description of this package and the
112 solving methods can be found in Dawson [9]. The package is available in
113 <http://ajdawson.github.io/eofs/>.

114 Before performing the EOF analysis on FS it is necessary to prepare the
115 surface ocean model state. Each variable is reshaped into a two-dimensional
116 matrix \mathbf{A}_i :

$$\mathbf{A}_i = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix}_i, \quad i = 1, \dots, q \quad (1)$$

117 where a are values of the variable i at the ocean model state grid points, m
118 is the number of data points of the i^{th} variable and n is the number of time
119 steps. The landmask nodes were excluded. The horizontal concatenations of
120 these matrices result in \mathbf{D}^{fs} , as follows:

$$\mathbf{D}^{\text{fs}} = [\mathbf{A}_1, \dots, \mathbf{A}_q] \quad (2)$$

121 where $\mathbf{D}^{\text{fs}} \in \mathbb{R}^{n \times M}$ and $M = m \times q$ is the length of the state vector. \mathbf{D}^{fs}
122 is the original full-state matrix. Because the sub-matrices do not have the
123 same units, it is necessary to standardise them. The dimensionless anomaly
124 matrix $\bar{\mathbf{D}}^{\text{fs}}$ is then defined by:

$$\bar{\mathbf{D}}^{\text{fs}} = [\bar{\mathbf{A}}_1, \dots, \bar{\mathbf{A}}_q] \quad (3)$$

125 where $\bar{\mathbf{A}}_i$ is the standardised data matrix of the i^{th} variable. The standard-
 126 isation is done by removing the column-wise mean and dividing by the i^{th}
 127 variable standard deviation $\sigma_{\mathbf{A}_i}$. This standardisation weights each variable
 128 with equal importance while maintaining the relative magnitude of variability
 129 at each model grid point within a variable.

130 After the EOF analysis is performed, $\bar{\mathbf{D}}^{\text{fs}}$ can be expressed as:

$$\bar{\mathbf{D}}^{\text{fs}} = \mathbf{P}\boldsymbol{\Pi} \quad (4)$$

131 where $\mathbf{P} \in \mathbb{R}^{n \times n}$ contains the time series of the principal components (PCs)
 132 and $\boldsymbol{\Pi} \in \mathbb{R}^{n \times M}$ are the EOFs of $\bar{\mathbf{D}}^{\text{fs}}$.

133 The dimension reduction of the system is then obtained by truncating \mathbf{P}
 134 and $\boldsymbol{\Pi}$. Let N_{eof} be the number of selected EOFs, then the reconstructed
 135 full-state anomaly matrix $\bar{\mathbf{D}}^{\text{rs}}$ is defined by:

$$\begin{aligned} \bar{\mathbf{D}}^{\text{rs}} &= \mathbf{P}_{N_{eof}} \boldsymbol{\Pi}_{N_{eof}} \\ &= [\bar{\mathbf{A}}_1^{\text{rs}}, \dots, \bar{\mathbf{A}}_q^{\text{rs}}] \end{aligned} \quad (5)$$

136 with $\mathbf{P}_{N_{eof}} \in \mathbb{R}^{n \times N_{eof}}$ and $\boldsymbol{\Pi}_{N_{eof}} \in \mathbb{R}^{N_{eof} \times M}$.

137 To reconstruct $\bar{\mathbf{D}}^{\text{rs}}$ back to ocean model state it is necessary to unstan-
 138 dardise the matrix by multiplying each submatrix $\bar{\mathbf{A}}_i^{\text{rs}}$ by $\sigma_{\mathbf{A}_i}$ and adding
 139 the previously calculated column-wise means. Finally, the non-standardised
 140 reduced-dimension matrix \mathbf{D}^{rs} is reshaped to the original grid dimensions and
 141 landmask nodes are placed back at every time step. This will be referred to
 142 as reconstructed full-state (RFS).

143 2.3. Neural Network for short-range forecast

144 At this stage, it is necessary to replicate or emulate the temporal evolution
 145 of the full-space ROMS solution. Our approach consists of training an NN

that will achieve this with an acceptable degree of accuracy. The chosen architecture is an autoregressive feed-forward neural network. This is based on the Multi Layer Perceptron (MLP). An MLP consists of three parts: an input layer, a set of hidden layers, and an output layer (Fig. 2b). Here, we trained the NN to receive $\mathbf{P}_{N_{eof}}$ at time-step t in the input layer. The target of the output layer is to predict $\mathbf{P}_{N_{eof}}$ at time-step $t + 1$ given the information entered in the input layer. The hidden layers are between the input and output layers. The hidden layers consists of nodes. The different nodes of the layers are connected via weights. A single hidden layer can also be used. The NN needs to be properly trained before being used in the DA step. Training the NNs finds optimal hyperparameters which minimise the difference between the target and predicted value in the output layer. These hyperparameters include, but are not limited to, activation functions (linked to the hidden nodes), learning rates, training method, the number of hidden layers and the number of hidden nodes. Different combinations of these hyperparameters determine the prediction ability of the NN. The dataset of $\mathbf{P}_{N_{eof}}$ was divided into subsets for training (70 %), validation (15 %) and testing (15 %). This means that the NN will find the optimal weights that minimise the cost function within the training subset and validate them in the validation subset. During the training and validation stages the target output is known. Our cost function is the mean-squared-error (MSE) between the target and the predicted value. If the cost function can be minimised further, the weights are modified during the validation stage. Since the validation of an NN evaluates whether the chosen weights for training were optimal, this can be considered part of the training method. During the testing stage we

171 obtain predicted values at the output layer with the weights fixed, and the
172 predicted values are not adjusted to the targets accordingly. If the MSE
173 between the predicted values and the target values differ greatly, this means
174 the trained NN is overfitted to the training values and it is not capable
175 of generalising. Thus, it is necessary to modify the hyperparameters and
176 train the NN again. For each combination of hyperparameters, the weights
177 are obtained through the training and validation processes. Overfitting is
178 prevented using early stopping.

179 Different combinations of the hyperparameters can be tested using the
180 function `GridSearch` in Sci-kit learn. The hyperparameters tested include:

- 181 • Learning rate: 10 combinations in the range of 10^{-7} to 10^2 .
- 182 • Activation functions: sigmoid, linear, ReLU [38].
- 183 • Training method: ADAM [27], scalar conjugate gradient (SCG) [37],
184 gradient descent [2].
- 185 • Number of Hidden layers: 1 to 5.
- 186 • Number of Hidden Nodes: 70 combinations in the range of 1 to 1000.

187 The first step including the dimension reduction and the optimal NN
188 training needs to be performed once and then is ready and available to be
189 used in the reduced-space DA.

190 *2.4. Data assimilation*

191 The second step of DA+NN consists of a sequential DA performed in
192 reduced-space. Any appropriate method of DA can be applied at this stage.

193 We chose the Ensemble Kalman Filter (EnKF) [11, 12]. This technique
 194 sequentially assimilates time-distributed observations. The EnKF involves a
 195 forecast step, where a previous state estimate evolves forward to the time
 196 of the observations, and an update step in which the evolved state estimate
 197 is updated with information from the observations. First, we create a 50-
 198 member ensemble ($\mathbf{X}(t)$) for January, 1st 2010, the date at which we start
 199 the EnKF DA. For this purpose we compute the climatological mean using
 200 the time steps of $\mathbf{P}_{N_{eof}}$ related to this day, i.e. January 1st 2001, ..., January
 201 1st, 2009. This mean is perturbed by adding Gaussian noise with zero mean
 202 and normally distributed random numbers with variance 1. This process
 203 creates a perturbed ensemble. Each member of the perturbed ensemble is
 204 then forecasted one time-step forward using the trained NN. This creates
 205 $\mathbf{X}_{\text{forecast}}$. The update of \mathbf{X} is given by:

$$\mathbf{X}_{\text{analysis}} = \mathbf{X}_{\text{forecast}} + \mathbf{K}(\mathbf{y} - \tilde{\mathbf{H}}\mathbf{X}_{\text{forecast}}) \quad (6)$$

206 where $\mathbf{X}_{\text{analysis}}$ is the updated ensemble state of the ensemble $\mathbf{X}_{\text{forecast}}$. The
 207 forecast performs the t to $t + 1$ time stepping. \mathbf{K} is the Kalman gain, \mathbf{y}
 208 is the ensemble of standardised observations and $\tilde{\mathbf{H}}$ is the reduced-space
 209 interpolation matrix. The analysis is the mean of this updated ensemble.

210 The ensemble \mathbf{y} is obtained by adding Gaussian perturbations to the ob-
 211 servational data, with zero mean and data error covariance \mathbf{R} . The reduced-
 212 space interpolation $\tilde{\mathbf{H}}$ is defined as:

$$\tilde{\mathbf{H}} = \mathbf{H}\boldsymbol{\Pi}^T, \tilde{\mathbf{H}} \in \mathbb{R}^{O \times N_{eof}} \quad (7)$$

213 where \mathbf{H} is the full-space observation matrix and O is the number of obser-

214 vations.

215 The Kalman gain, \mathbf{K} , is defined by:

$$\mathbf{K} = \tilde{\mathbf{B}}\tilde{\mathbf{H}} \underbrace{(\tilde{\mathbf{H}}\tilde{\mathbf{B}}\tilde{\mathbf{H}}^T + \mathbf{R})^{-1}}_{\mathbf{W}^{-1}} \quad (8)$$
$$\tilde{\mathbf{B}} = (1 + \alpha)\mathbf{B}$$

216 where the background error covariance matrix \mathbf{B} is calculated directly as the
217 covariance of $\mathbf{X}_{\text{forecast}}$, and $\tilde{\mathbf{B}}$ is the inflated \mathbf{B} using a multiplicative inflation
218 factor of $(1 + \alpha)$, and $\mathbf{W} = \tilde{\mathbf{H}}\tilde{\mathbf{B}}\tilde{\mathbf{H}}^T + \mathbf{R}$. This inflation is done to increase
219 the spread of the background error. The inversion of \mathbf{W} becomes very ex-
220 pensive for high-dimensional systems or for a large dataset of observations,
221 like gridded data. Since this contradicts the goal of performing fast DA, it is
222 necessary to optimise this operation. Considering uncorrelated data errors,
223 \mathbf{R} can be written as a diagonal matrix. The Sherman-Morrison-Woodbury
224 formula is used to decompose \mathbf{W}^{-1} as:

$$\begin{aligned} \mathbf{W}^{-1} &= (\mathbf{R} + \gamma\delta^T)^{-1} \\ &= \mathbf{R}^{-1} - \mathbf{R}^{-1}\gamma(\mathbf{I} + \delta^T\mathbf{R}^{-1}\gamma)^{-1}\delta^T\mathbf{R}^{-1} \\ \gamma &= \tilde{\mathbf{H}}\mathbf{B} \\ \delta &= \tilde{\mathbf{H}}^T \end{aligned} \quad (9)$$

225 This way, the computation of \mathbf{W}^{-1} it is expected to be inexpensive as it
226 depends on the inversion of the diagonal matrix \mathbf{R} and of a square matrix
227 with dimension N_{eof} . In order to build \mathbf{R} , a gridded map of data errors
228 was created using the same landmask of ROMS; and the surface ocean grid
229 points for every variable were replaced by the error value associated with their
230 respective observational products. The data errors were also standardised as

231 the gridded values of the FS. This creates a standardised error across the
 232 grid that creates \mathbf{R} as follows:

$$\mathbf{R} = \begin{bmatrix} diag(\frac{\omega_1}{\sigma_1})^2 & & \\ & \ddots & \\ & & diag(\frac{\omega_{q1}}{\sigma_{q1}})^2 \end{bmatrix} \quad (10)$$

233 where ω is the measurement data error of the observation dataset linked to
 234 the variable to be assimilated, $q_1 \leq q$ is the number of variables to be assimi-
 235 lated, and σ is the aforementioned standard deviation of every variable to be
 236 assimilated. *Diag* indicates diagonal matrices containing the standardised
 237 measurement errors for each variable to be assimilated and the diagonal has
 238 length O . This results in a non-singular square diagonal matrix that can be
 239 used with eq. (9).

240 2.5. Reconstruction of the reduced-space analysis

241 The last step involves the transformation of the sequential reduced-space
 242 analyses back into the reconstructed full-space. The outputs of the EnKF can
 243 only be analysed after the reconstruction. It is important to note that the
 244 reconstruction can be performed at any time, e.g. to assess the dimension
 245 reduction or the performance of the NNs, and not strictly after the data
 246 assimilation is implemented. Thus:

$$\mathbf{X}^{\text{RA}}(t) = \bar{\mathbf{X}}(t)_{\text{analysis}} \boldsymbol{\Pi}_{N_{eof}} \quad (11)$$

247 where, $\mathbf{X}^{\text{RA}}(t)$ is the reconstructed reduced-space analysis at time-step t .
 248 This is subsequently unstandardised and landmask nodes are put back into
 249 place.

250 *2.6. Assimilated data and observational datasets*

251 In this study, we assimilated observations of Sea Surface Height (SSH) and
252 Sea Surface Temperature (SST) together, with further validation using data
253 of SSH, SST, eastward surface currents velocity (U), and northward surface
254 currents velocity (V). The DA commenced on January 1st, 2010. Gridded
255 data from GHRSST (Group for High-Resolution Sea Surface Temperature)
256 was assimilated [26]. GHRSST-L4 is a gap-free gridded satellite product,
257 with complimentary in-situ observations, with a 10 km resolution. To obtain
258 SSH we added the mean dynamical topography from the ROMS climatology
259 to a mean sea level anomaly (MSLA) dataset. The MSLA in delayed time
260 was obtained from the AVISO Ssalto/Duacs database. The latter is a 0.25
261 degree resolution gridded satellite repository. The Ssalto/Duacs altimeter
262 products were produced and distributed by the Copernicus Marine and Envi-
263 ronment Monitoring Service (CMEMS) (<http://www.marine.copernicus.eu>).
264 The Globcurrent (v.2) is a 3-hourly 0.25 degree resolution dataset that in-
265 cludes the Ekman and geostrophic currents at the surface [43], and it is used
266 to validate the ocean surface currents.

267 Before the assimilation, SSH and SST data were interpolated to the
268 ROMS grid resolution, but preserving the data landmarks. The measure-
269 ment data errors considered in this study are 0.04 m for SSH and 0.5°C for
270 SST.

271 *2.7. Validation experiments*

272 In order to address the role of the DA and of the one-day prediction
273 given by the NN within the DA, we perform two additional experiments:
274 1) a projection of the gridded SSH and SST data onto the set of EOFs

275 using a Least Squares (LS) solution, and 2) a persistence experiment. We
 276 would expect DA+NN to retain physical plausibility, and because LS does not
 277 have a physical constraint DA+NN should outperform it. Additionally, the
 278 persistence experiment involves no dynamics and therefore, our DA should
 279 outperform this experiment as well.

280 Before we explain the LS experiment, we need to introduce the concept
 281 of restricted vectors. The restricted vectors are the column vectors of $\Pi_{N_{eof}}$
 282 that are restricted to available data. Since a row vector of $\Pi_{N_{eof}}$ has the same
 283 length as the number of grid points as the ROMS model, we only extract
 284 those column vectors of $\Pi_{N_{eof}}$ where data is available. This is possible as the
 285 observations are interpolated onto the ROMS grid. With these vectors we
 286 construct a matrix of restricted vectors.

287 The LS procedure is defined by the following equation:

$$\mathbf{J} = \|\mathbf{d} - \mathbf{ce}\| \quad (12)$$

288 where $\mathbf{d} \in \mathbb{R}^{n \times O}$ is the matrix of SSH and SST data column vectors in 2010,
 289 $\mathbf{e} = \mathbf{H}\Pi^T \in \mathbb{R}^{N_{eof} \times O}$ is the matrix of restricted vectors to that available data,
 290 and $\mathbf{c} \in \mathbb{R}^{n \times N_{eof}}$ is the matrix of coefficients that minimise $\mathbf{J} \in \mathbb{R}^{n \times O}$. Once
 291 \mathbf{c} is found, we can use it to reconstruct $\Pi_{N_{eof}}$ into the full-space and assess
 292 how the variables are represented in this new full-space reconstruction. This
 293 solution is optimal in that it minimizes the root MSE (RMSE) between the
 294 assimilated observations and analysis but it is likely an overfit and has no
 295 physical constraints, potentially leading to unphysical solutions.

296 The persistence experiment consists of replacing the forward stepping
 297 from the NN with the ensemble persistence and subsequently the data are
 298 assimilated. The drawback of the persistence experiment is that involves

299 no dynamics. At the following time-step, new information is obtained from
300 the observations and the process is repeated. This solution should present a
301 smaller spread of the ensemble during the forecast than the NN forecast and
302 therefore decouple from the truth state, underperforming the DA+NN.

303 The importance of these experiments is to show that DA+NN has a phys-
304 ical constraint (LS experiment) and that the dynamics given by the NN have
305 value (Persistence experiment).

306 3. Results

307 3.1. EOF analysis

308 The concatenated matrices of SSH, SST, U, and V resulted in a state
309 vector of $\sim 10^5$ dimensions at a single time-step. All the landmask gridpoints
310 were ignored in the process, reducing the state vector to almost half of its
311 original dimensionality ($\sim 6 \times 10^4$). The chosen number of EOFs determines
312 the dimensionality of the new system and the explained variance of RFS.
313 Several test cases with different EOFs were performed.

314 Figure 4 shows how the explained variance changes with respect to the
315 number of EOFs and the different combinations of SSH, SST, U, and V that
316 can be included in the model. In this study, the first 100 EOFs (N_{eof} in
317 section 2.2) were retained explaining $\sim 93\%$ of the variance, and all of the
318 presented results are based on this number of modes.

319 3.2. Dimension reduction

320 Figure 3 illustrates a comparison between the RFS and the FS. This is a
321 snapshot of December 1st, 2009 which clearly shows the retroflection of the
322 NBC and formation of NBC rings.

323 Sea surface height (Fig. 3a and 3b) displays a clear representation of
324 the NBC rings aforementioned. Between 48°W - 52°W the FS is not dis-
325 tinguishable from the RFS. The NBC ring positioned around 56°W shows
326 some difference between the RFS and the FS figures. In the FS it can be
327 seen how the NBC ring is about to move in the NW direction, meanwhile in
328 the RFS the eddy is already detached and drifting freely. However, the main
329 features of the NBC rings can be observed in both model solutions, thus the
330 mesoscale features are well represented in both. The RFS and the FS exhibit
331 the same SSH anomalies around the mouth of the Amazon River.

332 The FS snapshots of SST (Fig. 3c) show distinctive swirls of temperature
333 anomalies around the NBC rings, while in the RFS (Fig. 3d) these features
334 are hardly noticeable. Nevertheless, both the FS and the RFS show lower
335 temperatures near the retroflection of the NBC, along the coast of Suriname
336 and the west of Trinidad.

337 The currents (Fig. 3e and 3f) show a NBC ring moving SE-NW along
338 the coastline. The formation of the NBC ring, between 48°W - 52°W , can be
339 seen in both model solutions. While the RFS illustrates these rings clearly,
340 the FS presents more detail of the mesoscale features. For example, in the
341 retroflection of the NBC, near 48°W , the NBC ring advecting NW is already
342 detached from the NBC in the RFS, while it is still attached in the FS. The
343 FS generates the high azimuthal velocity of the full eddy near 56°W , while
344 the RFS shows this feature only over the southern and western region of the
345 ring.

346 *3.3. Seasonal correlations*

347 Between RFS and FS, the seasonal correlations of SSH, SST, U and V are
348 high with $r > 0.99$ and with negligible mean RMSE. Winter, Spring, Summer
349 and Autumn are represented by DJF(December, January, February), MAM
350 (March, April, May), JJA and SON, respectively.

351 For SSH the lowest correlations are observed during JJA in the West
352 of the Caribbean islands. The r values along the coast, the mouth of the
353 Amazon River and boundaries have the lowest values compared to the rest
354 of the grid (not shown).

355 Seasonal correlation coefficients of the horizontal velocities, U and V,
356 show similar behaviour. Like the SSH, both U and V have their lowest
357 correlations in the NW part of the grid and along the coast. The seasonal
358 values of r are larger in the open ocean than in the rest of the domain (not
359 shown). The region affected by Amazon discharge and the NW movement of
360 the NBC show an area of high correlation between the RFS and the FS, but
361 lower than in the open ocean. This is larger during DJF and MAM.

362 The low correlations along the coast and the NW areas also appear in SST.
363 The input of the Amazon River in JJA and SON induces large variability
364 in the SST. The area with higher correlations is the open ocean and the
365 North boundary of the domain (not shown). JJA exhibits the lowest seasonal
366 correlations for SSH, SST, U and V.

367 *3.4. Neural network training*

368 In our NN training, 30000 combinations of hyperparameters were tested,
369 which took approximately 50 hours using 12 cores of Intel Xeon (2.5 GHz)
370 processors. The best result was found for the NN with a single hidden layer,

371 151 hidden nodes with linear activation functions, a learning rate of 0.001,
372 and using the SCG as a training method (Fig. 2).

373 Figure 5 shows the statistics of the training of the NN. The correlation
374 coefficients shown are with respect to $\mathbf{P}_{N_{EOF}}$, which is used to train the NN.
375 As shown, the correlations are 0.982, 0.973, 0.980 for training, validation and
376 testing, respectively. The RMSE shown are of the one-day forecast of the NN
377 of RFS at $t - 1$ versus RFS at time-step t . Therefore, these values are based
378 on the reconstruction to physical values, and they are not the RMSE of the
379 prediction values and targeted values of $\mathbf{P}_{N_{EOF}}$. The RMSE figure shows the
380 different RMSE in m, $^{\circ}\text{C}$, m s^{-1} for SSH, SST, and U and V respectively.
381 While SSH does not vary greatly, there is a seasonality in SST, U and V. For
382 U and V, the worst performance is during the JJA. This may be due to the
383 less well-defined surface currents during JJA and the fact that the reduced-
384 space cannot reproduce the finer detail surface currents. The same is true
385 for SST and, to a lesser extent, for SSH as it is a more tractable variable.

386 *3.5. Analysis assessment*

387 Daily data for 2010 were assimilated. We used an inflation factor of
388 $\alpha = 3$ to increase the spread of the ensemble of the background covariance
389 error in both the NN and persistence experiment. The mean of the diagonal
390 of $\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T$, using the NN, degrades from 0.49, at the first iteration of EnKF,
391 and converges over time to a value within 0.16 and 0.19. The mean of the
392 diagonal of \mathbf{R} is 0.24. These unitless numbers indicate us that the model
393 covariance error and the observations covariance error have the same order
394 of magnitude, and that the model does not detach from the truth. Figures
395 6 and 7 show the results of the DA, where the coloured lines represent the

396 RMSE and correlation coefficients of the RFS and analysis with respect to
 397 the validation datasets of SSH from AVISO, SST from GHRSST, and U and
 398 V from Globcurrent. The RMSE and correlation time series of the FS (not
 399 shown) and the RFS are very similar. The annual mean RMSE between
 400 RFS and the gridded datasets used for their validation are: 0.09 m for SSH
 401 ($r = 0.52$), 0.73 °C for SST ($r = 0.56$), 0.54 m s⁻¹ for U ($r = 0.21$), and 0.62
 402 m s⁻¹ ($r = 0.12$) for V. It is clear that DA+NN does decrease considerably
 403 the mean RMSE across the time series. The 365-day mean RMSEs after the
 404 EnKF using the NN are 0.05 m ($r = 0.84$), 0.35 °C ($r = 0.74$), 0.46 m s⁻¹ (r
 405 = 0.36), and 0.53 m s⁻¹ ($r = 0.35$) for SSH, SST, U, and V, respectively.

406 An additional metric to assess the improvement of the fit of SSH and SST
 407 to the observations is the unitless measure J_{fit} [8]:

$$J_{fit} = \frac{1}{m} \sum_{i=1}^m \left| \frac{x_i^{model} - y_i^{obs}}{e^{obs}} \right| \quad (13)$$

408 where m is the number of data points, x_i^{model} and y_i^{obs} are the model solution
 409 at the location of the i th observation and the i th observation, respectively,
 410 and e^{obs} is the observation error of the assimilated variable. J_{fit} determines
 411 if the fit of the observations to the model is within one standard deviation
 412 of the observation error. If J_{fit} is less than 1, then the analysis residual is
 413 within the observation error [41]. The annual mean J_{fit} of SSH and SST in
 414 the RFS is 1.38 and 1.09, respectively, while is only 0.70 and 0.53 when SSH
 415 and SST data are assimilated, respectively, using the NN. This indicates that
 416 the analysis error is of the same magnitude as their respective observation
 417 errors.

418 The LS solution of the SSH, SST, U and V corresponds to the recon-

419 structure into the full-space of the $\Pi_{N_{eof}}$ matrix using \mathbf{c} . The LS experiment
 420 yields an annual RMSE of 0.04 m ($r = 0.89$), 0.28 °C ($r = 0.84$), 0.64 m
 421 s^{-1} ($r = 0.27$), and 0.68 m s^{-1} ($r = 0.26$) for SSH, SST, U, and V, respec-
 422 tively which beats the DA+NN solution in SST and SSH as expected. The
 423 persistence underperforms DA+NN with an annual RMSE of 0.08 m ($r =$
 424 0.65), 0.71 °C ($r = 0.51$), 0.76 m s^{-1} ($r = 0.20$), and 0.76 m s^{-1} ($r = 0.07$)
 425 for SSH, SST, U and V, respectively. The statistics of both DA experiments,
 426 including DA+NN and persistence, are summarised in table 1.

Table 1: DA annual means of RMSE, correlation coefficient and J_{fit} for the LS, DA+NN and Persistence (P) experiments. The correlation coefficient r and J_{fit} are unitless and only shown for the assimilated variables and the DA experiments.

Variable	RMSE			r (-)			J_{fit} (-)		
	LS	DA+NN	P	LS	DA+NN	P	LS	DA+NN	P
SSH (m)	0.04	0.05	0.08	0.89	0.84	0.65	-	0.70	1.38
SST (°C)	0.28	0.35	0.71	0.84	0.74	0.51	-	0.53	1.09
U (m s^{-1})	0.64	0.46	0.76	0.27	0.36	0.20	-	-	-
V (m s^{-1})	0.68	0.53	0.77	0.26	0.35	0.07	-	-	-

427 We introduce G_R as a unitless ratio, between the Root Mean Square
 428 (RMS) of the geostrophic and the total currents, that provides some measure
 429 of the geostrophicity of the flow. The geostrophic currents were obtained from
 430 the SSH gradients. Here:

$$RMS = \sqrt{\frac{1}{m} \sum_{i=1}^m (x_U)_i^2 + \frac{1}{m} \sum_{i=1}^m (x_V)_i^2} \quad (14)$$

$$G_R = \frac{RMS_{geo}}{RMS_{total}}$$

431 where x_U and x_V are the values of the ocean grid points of the surface currents
 432 velocities. We obtain the joint RMS of the U and V components from the
 433 geostrophic (RMS_{geo}) and total currents (RMS_{total}), and we calculate the
 434 annual G_R for LS, DA+NN, and the Globcurrent observations. The velocities
 435 of the geostrophic currents were obtained from the obtained SSH gradients:

$$u_g = -\frac{g}{f} \frac{\partial P}{\partial y}, \quad v_g = -\frac{g}{f} \frac{\partial P}{\partial x} \quad (15)$$

$$f = 2\Omega \sin\phi, \quad \Omega = \frac{2\pi}{day} \quad (16)$$

436 where P is the horizontal pressure field from the SSH, g is the gravity con-
 437 stant, f is the Coriolis parameter, ϕ is the latitude, and Ω is the angular
 438 velocity of the Earth. The RMS did not include the latitudinal area com-
 439 prising 2.5°S to 2.5°N in the vicinity of the Equator. This is due to the un-
 440 realistically high geostrophic currents that can occur near the Equator since
 441 the Coriolis parameter f is too small in this region. The annual RMS_{geo}
 442 ($m^{-2} s^{-2}$) means are: 0.44 ($\sigma = 0.03$), 0.44 ($\sigma = 0.03$), and 0.45 ($\sigma = 0.04$)
 443 for LS, DA+NN, and Observations, respectively. The annual G_R means are:
 444 0.73 ($\sigma = 0.12$), 1.01 ($\sigma = 0.09$), and 1.03 ($\sigma = 0.05$) for LS, DA+NN, and
 445 Observations, respectively.

446 The following results correspond only to DA+NN. Figure 8 shows the
 447 annual spatial mean RMSE, with respect to the validation datasets of SST
 448 and SSH for the RFS and analysis. After the assimilation, the differences in
 449

450 the advection route of the NBC and NBC rings have been clearly minimised
451 but not so evidently in the coastal regions or the area near the Lesser Antilles
452 (Fig. 8a and 8b). The assimilation of SST is better in those areas, but they
453 still are the most challenging regions (Fig. 8c and 8d). DA improves the
454 annual mean correlation coefficient from 0.52 to 0.80 in SSH when compared
455 to AVISO, and from 0.56 to 0.72 in SST when compared to GHRSST.

456 The improvements of the time and spaced averaged RMSE for DA+NN
457 are 0.04 m (44 %) and 0.38 °C (52 %) for SSH and SST, respectively. It is
458 encouraging to note that the time averaged RMSE of the assimilated fields
459 are similar to those of global reanalysis products like Hybrid Coordinate
460 Ocean Model (HYCOM), which include DA in their methodology [14, 7].

461 Figure 9 is a snapshot of January 11th 2010 comparing observations, the
462 analysis, and RFS. This day was chosen as it portrays the main features of
463 the NBC and how DA+NN improves the representation of the fields in the
464 analysis. The velocity fields of total and geostrophic currents are represented
465 by the vector magnitude of their respective eastward and northward compo-
466 nents. The spatial representation of SSH shows how the analysis closely
467 resembles the altimetry dataset, where the main features like the NBC ring
468 can be easily recognised along with the retroflection of the NBC in the east-
469 ern part of the grid. The SST from RFS appears relatively uniform and
470 featureless before DA but the analysis succeeds in capturing enough infor-
471 mation from the observations in order to match its spatial patterns. The
472 response and modification of U and V is the result of the assimilation of only
473 SSH and SST within the reduced-space. The retroflection of the NBC and
474 its rings appear in the same location for the geostrophic currents and total

475 surface currents. This means that the SSH is closely correlated with the
476 currents since the geostrophic currents are extracted from the SSH gradient.
477 The NBC is captured and well represented by the currents of the assimilated
478 system when compared to Globcurrents. The main difference between the
479 variables in the assimilated system and their validation datasets are in the
480 Amazon River region of influence. In general, the variables present great
481 variability in the mouth of the Amazon River. After the analysis, the rep-
482 resentation of the dynamics of the system is physically plausible showing
483 features like the retroflection of the NBC and the signal of the NBC ring in
484 the NW of the grid across the four variables in the analysis.

485 *3.6. Speed of DA+NN*

486 A critical aspect of DA+NN is the gain in speed. The EOF analysis on FS
487 is performed once in 20 cores of a Intel Xeon E5-4650v2 (2.4 GHz) processor
488 and takes ~ 300 seconds. Typically, the assimilation of SSH and SST takes
489 ~ 20 seconds per model day, including the forward stepping with the NN and
490 daily EnKF DA. This means that a year-long simulation takes between 2-3
491 hours.

492 Other sensitivity runs showed that the assimilation of one variable takes
493 ~ 10 seconds, while the joint assimilation of four variables takes up to 60
494 seconds per model day. On a daily time-step basis, the fast DA framework
495 presented here is computationally more efficient outperforming the full-space
496 ROMS run of this region by 2 orders of magnitude. While the ROMS sim-
497 ulation runs on 256 cores of Intel Xeon (2.5 GHz) processors, the fast DA
498 runs on a single core of an Intel Xeon E5-4650v2 (2.4 GHz) processor.

499 **4. Discussion**

500 The results can be discussed in terms of the effectiveness of the dimension
501 reduction algorithm, the performance of the data assimilation using the
502 EnKF and how the characteristics of the NBC region are represented in this
503 scheme, and finally the execution speed and robustness of DA+NN.

504 The DA diminishes the RMSE with respect to observational data. This is
505 most clear when we observe SSH and SST since these are the state variables
506 directly constrained by data. The reconstructed DA fields also presents a
507 correlation across variables. This suggests that the reduced-space information
508 from the NN and DA solution has value. There are three distinguishable
509 problematic sections in the study area when we perform DA: The Lesser An-
510 tilles islands, the coastline of Northern South America and the outlet of the
511 Amazon River into the Atlantic Ocean.

512 The islands located in the Lesser Antilles are a problematic area since
513 these serve as a natural barrier to the NBC and NBC rings, where the latter
514 collide and dissipate. The freshwater discharge affects the area near the
515 mouth of the Amazon River.

516 Large misfits around these regions can occur because:

- 517 • The ROMS does not reproduce the full spectrum of data variability in
518 these problematic regions. In the Amazon River, due to the utilisation
519 of climatological river runoff and in the Lesser Antilles due to excessive
520 smoothing of the bottom topography around the islands. As a result,
521 even the full set of EOFs is not able to represent the data variability
522 in these regions.

- 523 ● The 100 EOFs selected for this study are not sufficient to represent the
524 data variability in these regions.

525 It is important to assess how much of the variability is explained when
526 the dimension reduction is applied. The dimension of the state vector during
527 the DA problem is 100, and it is much smaller than the available number
528 of data at every time step ($\sim 10^4$). Therefore, the reduced space vectors
529 cannot reproduce the entire spatial variability of the data. Among the vari-
530 ables, altimetry sea level is a more stable variable, with less noise, and strong,
531 large-scale features like the formation of the NBC rings, which retains more
532 explained variability when the dimension reduction is applied. This is fol-
533 lowed by the SST.

534 Despite being part of the reduced-space state vector and the NN dynam-
535 ics, U and V data were not assimilated in this study. These variables are
536 only affected by the changes in the reduced-space during the DA of SSH and
537 SST. The simulation of the currents only improves marginally. The average
538 RMSE and spatial r improve compared to the Globcurrent dataset, but the
539 same problems in the coastlines, mouth of the Amazon River and the Lesser
540 Antilles appear. Globcurrent products are partially derived from altimetry
541 [43]. The impact of assimilating SSH and SST on U and V is in concordance
542 with previous studies [41], where the currents are improved by the assimila-
543 tion of SSH, but the influence of assimilating of SST is less. The geostrophic
544 currents obtained from the SSH analysis present similar spatial patterns to
545 the analysis surface currents.

546 The result of the application of the EnKF is dependent on the construc-
547 tion of the error covariance matrix \mathbf{R} . This has been discussed previously

in Stewart et al. [45]. If information of detailed errors of the gridded uncorrelated observational datasets were available, these could be easily applied to DA+NN to form the diagonal of \mathbf{R} . Here, we assume that the error associated with the observations is uncorrelated and that is the same for the entire grid. This works as a simplification, meaning that \mathbf{R} might be under or overestimating the error covariance with more or less weight, respectively, compared to the error covariance of the model. Moreover, recent studies have shown that introducing observation error correlations improve DA results [47, 36]. However, by assuming 0 observation error covariance we get a diagonal \mathbf{R} which permits the use of the Sherman-Morrison-Woodbury formula (eq. (9)). Since the goal of this framework of a reduced dimensionality model is to perform fast data assimilation we need to avoid a bottleneck when calculating the inversion in the Kalman gain and hence this assumption is important.

In this framework the background error covariance matrix \mathbf{B} in the reduced-space can be explicitly calculated. This allow us to easily apply a multiplicative inflation to increase the spread of the ensemble of the NN forecast within the EnKF. Explicit calculation of \mathbf{B} in high-dimensional systems can restrict the resolution of the DA. The ensemble method also helps to provide a better simulation of the forecasts errors [1], and in this case is computationally cheap to generate. This is only valid in the reduced-space as a correct representation of \mathbf{B} in the full-space will be computationally expensive and would defeat the purpose of having a fast DA framework.

The reduced-space allows us to perform a fast DA. The comparison between the observations and the model state occurs in the observational space,

573 thus more observations will slow down the procedure. This is a cheaper task
574 than working with high-dimensionality systems and therefore powerful pro-
575 cessing power (supercomputer) is not necessary. However, certain features
576 may not appear in the reduced-space which are present in the full-space.
577 More EOFs could be retained at the cost of reducing the speed of DA+NN
578 because of the larger matrices in the RS.

579 The DA also depends on the short-range forecast to create the forecast
580 estimates. A comprehensive search for the optimal hyperparameters of the
581 NN was performed. Any other configurations will change this forecast. If
582 another number of EOFs were to be considered, it would produce a new
583 reduced-space which would also change the configuration of the NN, modify
584 parameters such as the number of hidden layers and affect the time necessary
585 to train a new system. The length of the ROMS data also alters the training
586 and configuration of the NN. If there is not enough data, the NN would not
587 have enough information to effectively replicate the physics and dynamics of
588 the full space system. There is therefore a great deal of effort required to
589 establish the optimal reduced-space DA model. However, once this effort has
590 been undertaken the assimilation is fast.

591 For a daily time-step basis, DA+NN presented here outperforms the full-
592 space ROMS run of this region by 2 orders of magnitude in computational
593 efficiency. In terms of computing time, Frolov et al. [17] also reports a speed-
594 up of two orders of magnitude when comparing the DA on the full-space
595 estuary model of the Columbia River to the one performed on the model
596 surrogate.

597 The robustness of the application of DA+NN can be assessed by com-

598 paring the results to Frolov et al. [17] and by comparing the results of the
599 data assimilation with other models performing DA in the same region. The
600 RMSE results of the assimilation of altimetry and temperature are similar
601 to the ones presented by Frolov et al. [17] for an estuary: between 0.04 to
602 0.07 m in SSH and 0.4 to 1.4 °C for SST. HYCOM has a 10 km resolution,
603 the same as the ROMS model solution employed in this study. A study of
604 the NBC area includes the DA of sea level anomaly, temperature and salin-
605 ity into HYCOM around the western tropical and south Atlantic [5]. The
606 reported RMSE of SSH and SST validation datasets range from 0.05 to 0.15
607 m, and 0.5 to 3.5 °C. Our fast DA framework has a similar magnitude of
608 errors compared to these full-space systems.

609 *4.1. Least squares and persistence experiments*

610 The LS solution sets up a reference for the comparison of the DA+NN
611 solution. As aforementioned, the EOFs cannot reproduce the entire data
612 variability. An idea of the limitations of the variability that can be repro-
613 duced by the EOFs is given by analysing the LS experiment. The LS solution
614 presents the largest errors in the area around the Amazon River mouth and
615 coastlines (not shown), which means that the selected EOFs do not repre-
616 sent properly the data in these regions. It also outperforms the DA solution
617 in RMSE and correlation with SSH and SST data because it provides an
618 optimal solution in terms of minimising the errors of assimilated quantities.
619 Nonetheless, if we reconstruct using \mathbf{D}^{rs} and the \mathbf{c} coefficients found with
620 the least squares problem, then the overfit of the LS solution to SSH and
621 SST data is more clear. This is supported by the RMSE and r time series of
622 U and V in figures 6 and 7. Here, the DA using the NN outperforms the LS

623 solution.

624 The G_R is a measure of how geostrophic a flow is. The annual means
625 of G_R show that the DA solution is closer to the Observations, where the
626 G_R of the Observations is 2 % larger than the G_R of the DA solution. When
627 compared to the LS, the observations present a 41 % larger G_R annual mean.
628 This suggests that the DA solution preserves this ratio much better than the
629 LS solution. This demonstrates that the model surrogate in the DA method
630 provides a crucial physical constraint on the DA process that the LS solution
631 do not have.

632 The persistence experiment allows us to assess the role of the NN in the
633 forecast step within the DA. Since the persistence method does not increase
634 the spread of the ensemble beyond the inflation, we would expect the NN to
635 grow the spread of the ensemble further in the one-day forecast. As previously
636 discussed, the mean of the diagonal of $\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^T$, using the NN, converges and
637 oscillates to a value in the same order of magnitude as the diagonal mean
638 of \mathbf{R} . A bigger spread of the ensemble means that the EnKF synchronises
639 better with the true system trajectory [22, 48]. From table 1 it is clear that
640 DA+NN results in lower RMSE and higher correlations with the validation
641 data, and a J_{fit} less than 1. Thus, the NN outperforms the persistence
642 experiment, and yields better results after the data assimilation, since it can
643 better simulate the dynamics of the full-state.

644 Finally, having an NN for the forecasting task is useful because it is
645 simple to train and there is an increasing availability of state-of-the-art and
646 optimised libraries ready to be used.

647 **Conclusion**

648 We present the first application in the open ocean of a novel fast DA,
649 DA+NN, using a reduced-space. This framework proved to be much faster
650 than full-space DA schemes by two orders of magnitude, with similar accu-
651 racy to similar global reanalyses. The dimension reduction managed to cap-
652 ture most of the explained variance. Since the NN is data-driven, a longer
653 training dataset will be beneficial as it might include extreme events that are
654 not present in the current training dataset.

655 Only SSH and SST data were assimilated. There was a clear improve-
656 ment when these variables were validated against satellite gridded data. Al-
657 though U and V were not assimilated, the reconstructed reduced-space anal-
658 yses present a higher correlation coefficient and lower time-averaged RMSE
659 when compared to the Globcurrent synthetic velocity product.

660 Finally, DA+NN is not limited to outputs from ROMS nor to the four
661 variables studied here, and therefore any global analysis product like HY-
662 COM or CMEMS could be used as the underlying full-state. This framework
663 can be applied to other regional ocean areas around the globe making it the
664 first step to fast forecasting applicable in real emergency situations where
665 new data is quickly available and an immediate prediction of, for example,
666 currents, is necessary. It could be a powerful complimentary tool to full
667 models when they are not readily available.

668 **Acknowledgements**

669 This research was supported by BECAS CHILE, governmental Chilean
670 scholarship from the Comisión Nacional de Investigación Científica y Tec-

671 nológica (CONICYT).

672 **References**

- 673 [1] Bannister, R. N., 2008. A review of forecast error covariance statistics
674 in atmospheric variational data assimilation. ii: Modelling the forecast
675 error covariance statistics. Quarterly Journal of the Royal Meteorological
676 Society 134 (637), 1971–1996.
- 677 [2] Bengio, Y., Simard, P., Frasconi, P., 1994. Learning long-term depen-
678 dencies with gradient descent is difficult. IEEE transactions on neural
679 networks 5 (2), 157–166.
- 680 [3] Bruce, J., Kerling, J., Beatty, W., 1985. On the North Brazilian eddy
681 field. Progress in oceanography 14, 57–63.
- 682 [4] Cochrane, J. D., Kelly Jr, F. J., Olling, C. R., 1979. Subthermocline
683 countercurrents in the western equatorial Atlantic Ocean. Journal of
684 Physical Oceanography 9 (4), 724–738.
- 685 [5] Costa, F., Tanajura, C., 2015. Assimilation of sea-level anomalies and
686 argo data into hycom and its impact on the 24 hour forecasts in the west-
687 ern tropical and south atlantic. Journal of Operational Oceanography
688 8 (1), 52–62.
- 689 [6] Csanady, G., 1985. A zero potential vorticity model of the North Brazil-
690 ian Coastal Current. Journal of Marine Research 43 (3), 553–579.

- 691 [7] Cummings, J. A., 2005. Operational multivariate ocean data assimila-
692 tion. *Quarterly Journal of the Royal Meteorological Society* 131 (613),
693 3583–3604.
- 694 [8] da Rocha Fragoso, M., de Carvalho, G. V., Soares, F. L. M., Faller,
695 D. G., de Freitas Assad, L. P., Toste, R., Sancho, L. M. B., Passos,
696 E. N., Böck, C. S., Reis, B., et al., 2016. A 4d-variational ocean data
697 assimilation application for santos basin, brazil. *Ocean Dynamics* 66 (3),
698 419–434.
- 699 [9] Dawson, A., 2016. eof: A library for eof analysis of meteorological,
700 oceanographic, and climate data. *Journal of Open Research Software*
701 4 (1).
- 702 [10] Didden, N., Schott, F., 1993. Eddies in the North Brazil Current
703 retroflection region observed by Geosat altimetry. *Journal of Geophysical
704 Research: Oceans (1978–2012)* 98 (C11), 20121–20131.
- 705 [11] Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-
706 geostrophic model using monte carlo methods to forecast error statistics.
707 *Journal of Geophysical Research: Oceans* 99 (C5), 10143–10162.
- 708 [12] Evensen, G., 2003. The ensemble kalman filter: Theoretical formulation
709 and practical implementation. *Ocean dynamics* 53 (4), 343–367.
- 710 [13] Flagg, C. N., Gordon, R. L., McDowell, S., 1986. Hydrographic and
711 current observations on the continental slope and shelf of the western
712 equatorial Atlantic. *Journal of Physical Oceanography* 16 (8), 1412–
713 1429.

- 714 [14] Fox, D., Teague, W., Barron, C., Carnes, M., Lee, C., 2002. The modular
715 ocean data assimilation system (MODAS). *Journal of Atmospheric and*
716 *Oceanic Technology* 19 (2), 240–252.
- 717 [15] Fratantoni, D. M., Johns, W. E., Townsend, T. L., 1995. Rings of the
718 North Brazil Current: Their structure and behavior inferred from ob-
719 servations and a numerical simulation. *Journal of Geophysical Research:*
720 *Oceans* (1978–2012) 100 (C6), 10633–10654.
- 721 [16] Fratantoni, D. M., Richardson, P. L., 2006. The Evolution and Demise of
722 North Brazil Current Rings*. *Journal of Physical Oceanography* 36 (7),
723 1241–1264.
- 724 [17] Frolov, S., Baptista, A. M., Leen, T. K., Lu, Z., van der Merwe, R.,
725 2009. Fast data assimilation using a nonlinear Kalman filter and a model
726 surrogate: An application to the Columbia River estuary. *Dynamics of*
727 *Atmospheres and Oceans* 48 (1), 16–45.
- 728 [18] Hannachi, A., 2004. A primer for EOF analysis of climate data. Depart-
729 ment of Meteorology, University of Reading, 1–33.
- 730 [19] Hannachi, A., Jolliffe, I., Stephenson, D., 2007. Empirical orthogonal
731 functions and related techniques in atmospheric science: A review. *Interna-*
732 *tional Journal of Climatology* 27 (9), 1119–1152.
- 733 [20] Hippert, H. S., Pedreira, C. E., Souza, R. C., 2001. Neural networks for
734 short-term load forecasting: A review and evaluation. *Power Systems,*
735 *IEEE Transactions on* 16 (1), 44–55.

- 736 [21] Hooten, M. B., Leeds, W. B., Fiechter, J., Wikle, C. K., 2011. Assessing
737 first-order emulator inference for physical parameters in nonlinear mech-
738 anistic models. *Journal of Agricultural, Biological, and Environmental*
739 *Statistics* 16 (4), 475–494.
- 740 [22] Hunt, B. R., Kostelich, E. J., Szunyogh, I., 2007. Efficient data assimila-
741 tion for spatiotemporal chaos: A local ensemble transform kalman filter.
742 *Physica D: Nonlinear Phenomena* 230 (1-2), 112–126.
- 743 [23] Jochum, M., Malanotte-Rizzoli, P., 2003. On the generation of north
744 brazil current rings. *Journal of Marine research* 61 (2), 147–173.
- 745 [24] Johns, W. E., Lee, T. N., Schott, F. A., Zantopp, R. J., Evans, R. H.,
746 1990. The North Brazil Current retroflection: Seasonal structure and
747 eddy variability. *Journal of Geophysical Research: Oceans (1978–2012)*
748 95 (C12), 22103–22120.
- 749 [25] Johns, W. E., Zantopp, R. J., Goni, G. J., 2003. Cross-gyre transport by
750 North Brazil Current rings. *Elsevier Oceanography Series* 68, 411–441.
- 751 [26] JPL MUR MEaSUREs Project GHRSST, 2010. GHRSST
752 Level 4 MUR Global Foundation Sea Surface Tempera-
753 ture Analysis. Ver. 2. PO.DAAC, CA, USA. Available from:
754 <http://dx.doi.org/10.5067/GHGMR-4FJ01> [Accessed June 2017].
- 755 [27] Kingma, D. P., Ba, J., 2014. Adam: A method for stochastic optimiza-
756 tion. *arXiv preprint arXiv:1412.6980*.
- 757 [28] Knutti, R., Stocker, T., Joos, F., Plattner, G.-K., 2003. Probabilistic

- 758 climate change projections using neural networks. Climate Dynamics
759 21 (3-4), 257–272.
- 760 [29] Kondrashov, D., Berloff, P., 2015. Stochastic modeling of decadal vari-
761 ability in ocean gyres. Geophysical Research Letters 42 (5), 1543–1553.
- 762 [30] Kutzbach, J. E., 1967. Empirical eigenvectors of sea-level pressure, sur-
763 face temperature and precipitation complexes over North America. Jour-
764 nal of Applied Meteorology 6 (5), 791–802.
- 765 [31] Leeds, W. B., Wikle, C. K., Fiechter, J., 2014. Emulator-assisted
766 reduced-rank ecological data assimilation for nonlinear multivariate dy-
767 namical spatio-temporal processes. Statistical Methodology 17, 126–138.
- 768 [32] Margvelashvili, N., Andrewartha, J., Herzfeld, M., Robson, B. J.,
769 Brando, V. E., 2013. Satellite data assimilation and estimation of a
770 3D coastal sediment transport model using error-subspace emulators.
771 Environmental modelling & software 40, 191–201.
- 772 [33] Margvelashvili, N., Campbell, E., 2012. Sequential data assimilation
773 in fine-resolution models using error-subspace emulators: Theory and
774 preliminary evaluation. Journal of Marine Systems 90 (1), 13–22.
- 775 [34] Margvelashvili, N. Y., Herzfeld, M., Rizwi, F., Mongin, M., Baird,
776 M. E., Jones, E., Schaffelke, B., King, E., Schroeder, T., 2016. Emulator-
777 assisted data assimilation in complex models. Ocean Dynamics 66 (9),
778 1109–1124.
- 779 URL <http://dx.doi.org/10.1007/s10236-016-0973-8>

- 780 [35] Mattern, J. P., Fennel, K., Dowd, M., 2012. Estimating time-dependent
781 parameters for a biological ocean model using an emulator approach.
782 Journal of Marine Systems 96, 32–47.
- 783 [36] Miyoshi, T., Kalnay, E., Li, H., 2013. Estimating and including
784 observation-error correlations in data assimilation. Inverse Problems in
785 Science and Engineering 21 (3), 387–398.
- 786 [37] Møller, M. F., 1993. A scaled conjugate gradient algorithm for fast su-
787 pervised learning. Neural networks 6 (4), 525–533.
- 788 [38] Nair, V., Hinton, G. E., 2010. Rectified linear units improve restricted
789 boltzmann machines. In: Proceedings of the 27th international confer-
790 ence on machine learning (ICML-10). pp. 807–814.
- 791 [39] Newinger, C., 2015. The barrier layer and ocean colour in the Ama-
792 zon and Orinoco plume: Competing for the oceanic control on tropical
793 cyclone intensity. Ph.D. thesis, Imperial College London.
- 794 [40] Newinger, C., Toumi, R., 2015. Potential impact of the colored Amazon
795 and Orinoco plume on tropical cyclone intensity. Journal of Geophysical
796 Research: Oceans 120 (2), 1296–1317.
- 797 [41] Phillipson, L., Toumi, R., 2017. Impact of assimilating OSCAR, alti-
798 metry, and limited drifter data on ocean current forecasting in the Angola
799 Basin. Ocean Modelling.
- 800 [42] Preisendorfer, R. W., Mobley, C. D., 1988. Principal component analysis
801 in meteorology and oceanography. Vol. 425. Elsevier Amsterdam.

- 802 [43] Rio, M., Mulet, S., Picot, N., 2014. Beyond GOCE for the ocean circulation
803 estimate: Synergetic use of altimetry, gravimetry, and in situ data provides new insight into geostrophic and Ekman currents. *Geophys.*
804 *Res. Lett.* 41(24), 8918–8925.
- 806 [44] Shchepetkin, A. F., McWilliams, J. C., 2005. The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Modelling* 9 (4), 347–404.
- 809 [45] Stewart, L. M., Dance, S. L., Nichols, N. K., 2007. Correlated observation errors in data assimilation. *Int. J. Numer. Meth. Fluids* (56),
810 1521–1527.
- 812 [46] van der Merwe, R., Leen, T. K., Lu, Z., Frolov, S., Baptista, A. M.,
813 2007. Fast neural network surrogates for very high dimensional physics-based models in computational oceanography. *Neural Networks* 20 (4),
814 462–478.
- 816 [47] Waller, J. A., Dance, S. L., Nichols, N. K., 2016. Theoretical insight
817 into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. *Quarterly Journal*
818 *of the Royal Meteorological Society* 142 (694), 418–431.
- 820 [48] Whitaker, J. S., Hamill, T. M., 2002. Ensemble data assimilation without perturbed observations. *Monthly Weather Review* 130 (7), 1913–
821 1924.

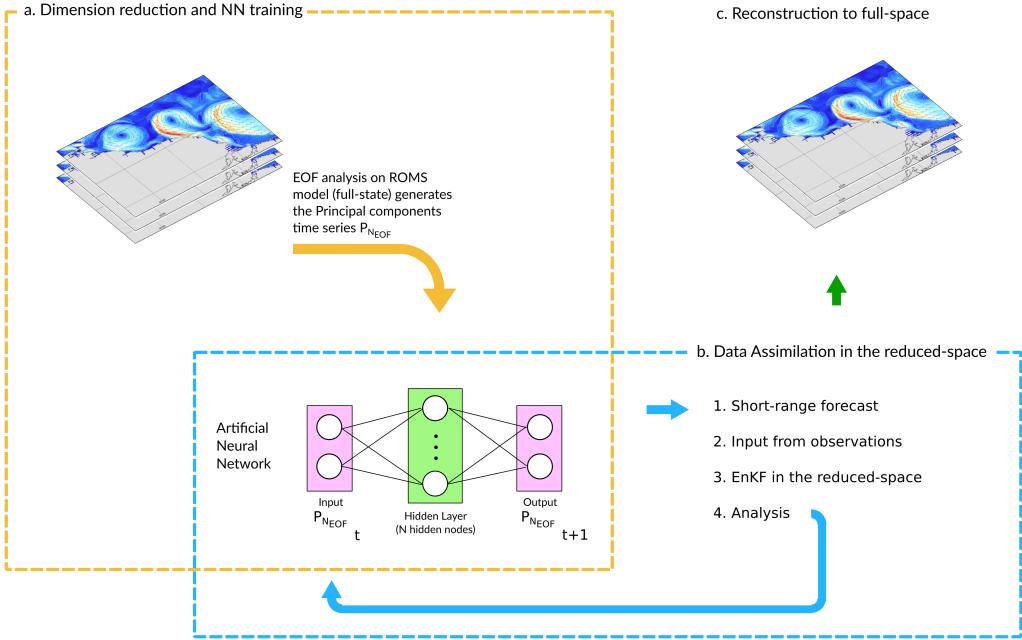


Figure 2: DA+NN framework diagram to perform data assimilation in a reduced-space. **a.** Dimension reduction and NN training, **b.** Data assimilation, and **c.** Analysis reconstruction to full-space. The first step includes the EOF analysis to reduce the dimensionality of the surface layer of a ROMS model solution and obtain the reconstructed full-state (RFS) at every time step. Via an EOF analysis we obtain the truncated time series of the Principal Components (P_{EOF}) which are used to train an NN with a hidden layer and N hidden nodes, depending on the input used. The NN is configured so the $t + 1$ timestep can be predicted from the t timestep of P_{EOF} . The second step is a sequential Ensemble Kalman filter (EnKF). This includes a short-range forecast from the NN, followed by the input of the observational products into the system. This creates a new updated state (analysis) which is used as a new input into the NN from which we obtain a new short-range forecast. Lastly, after the DA, the updated reduced-space analysis is reconstructed into the full-space.

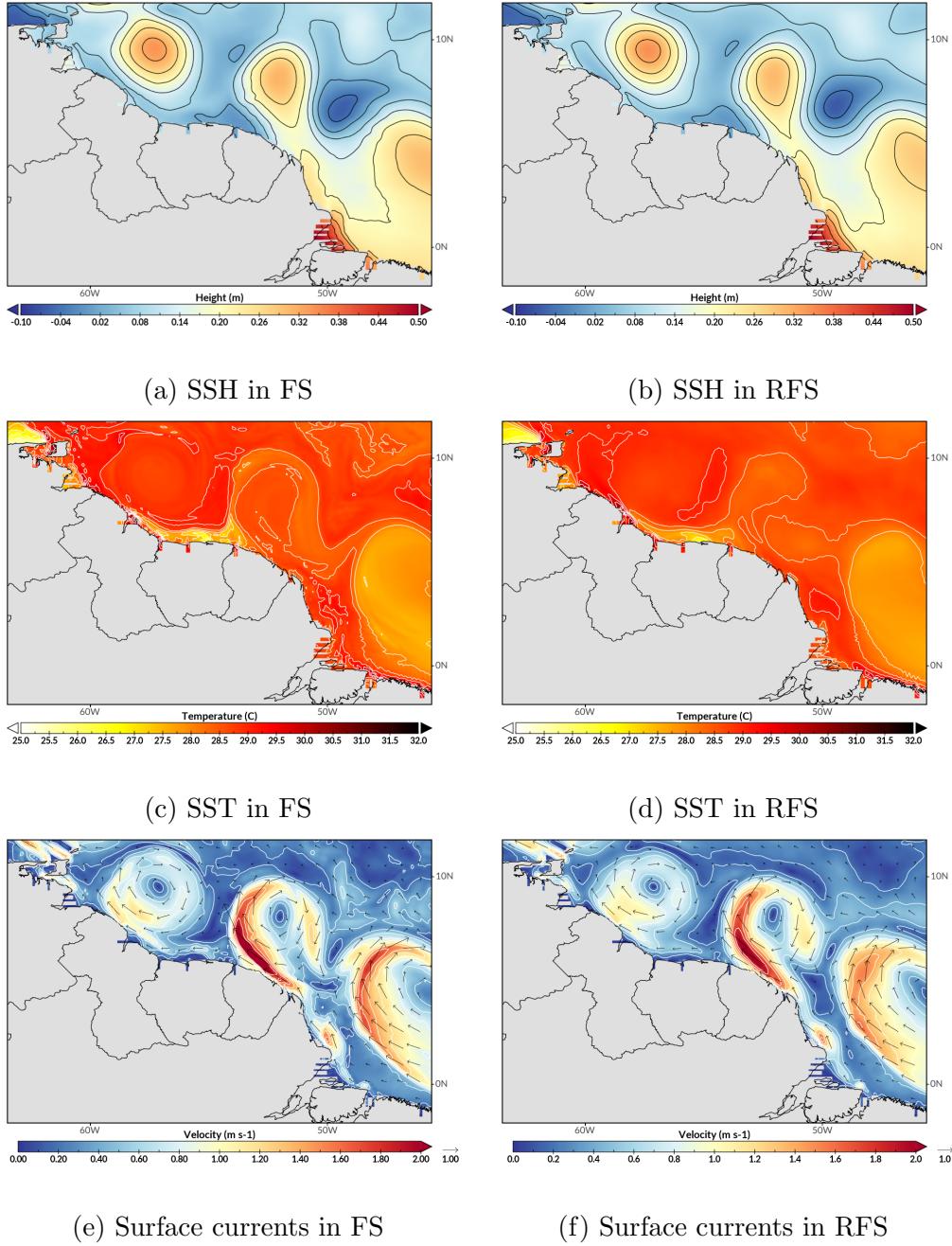


Figure 3: Full-space state (FS) and reconstructed full-state (RFS). Snapshot of December 1st, 2009. This day shows the genesis and advection of the North Brazil Current rings travelling northwestward and the retroreflection of the NBC to supply the North Equatorial Counter Current

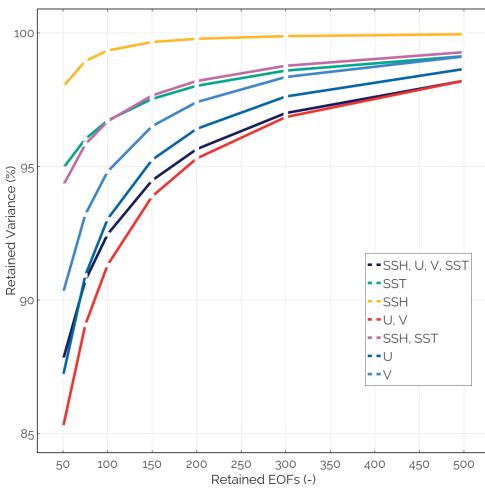
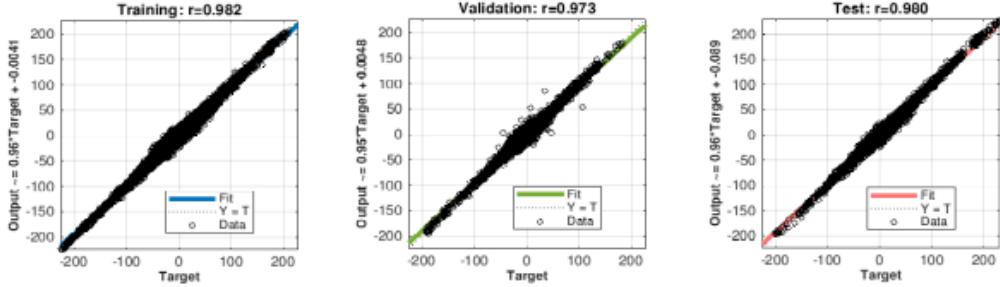
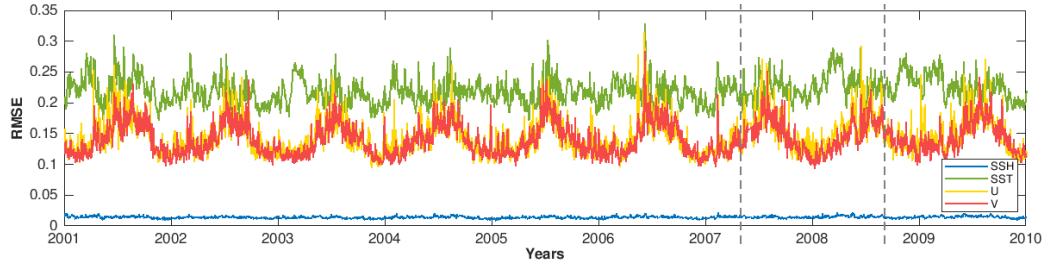


Figure 4: Explained variance versus number of EOFs retained. Different dashed lines represent the explained variability with respect to 50, 75, 100, 150, 200, 300 and 500 EOFs for different combinations of variables used in the model: SSH, SST, U, and V.



(a) Correlations between prediction of output layer and target values



(b) RMSE between one-day prediction by the NN of RFS at $t - 1$ and RFS at time-step t

Figure 5: Statistics of the NN training. **a.** Correlation coefficient between the prediction of the output layer and the target values for training, validation and testing. **b.** Spatial RMSE of the reconstructed one-day prediction by the NN of RFS at $t - 1$ compared to the RFS at time step t . Blue, green, yellow and red solid lines are SSH (m), SST ($^{\circ}$ C), U and V ($m s^{-1}$). The first section is the training dataset, the second section is the validation dataset, and the third section is the testing dataset.

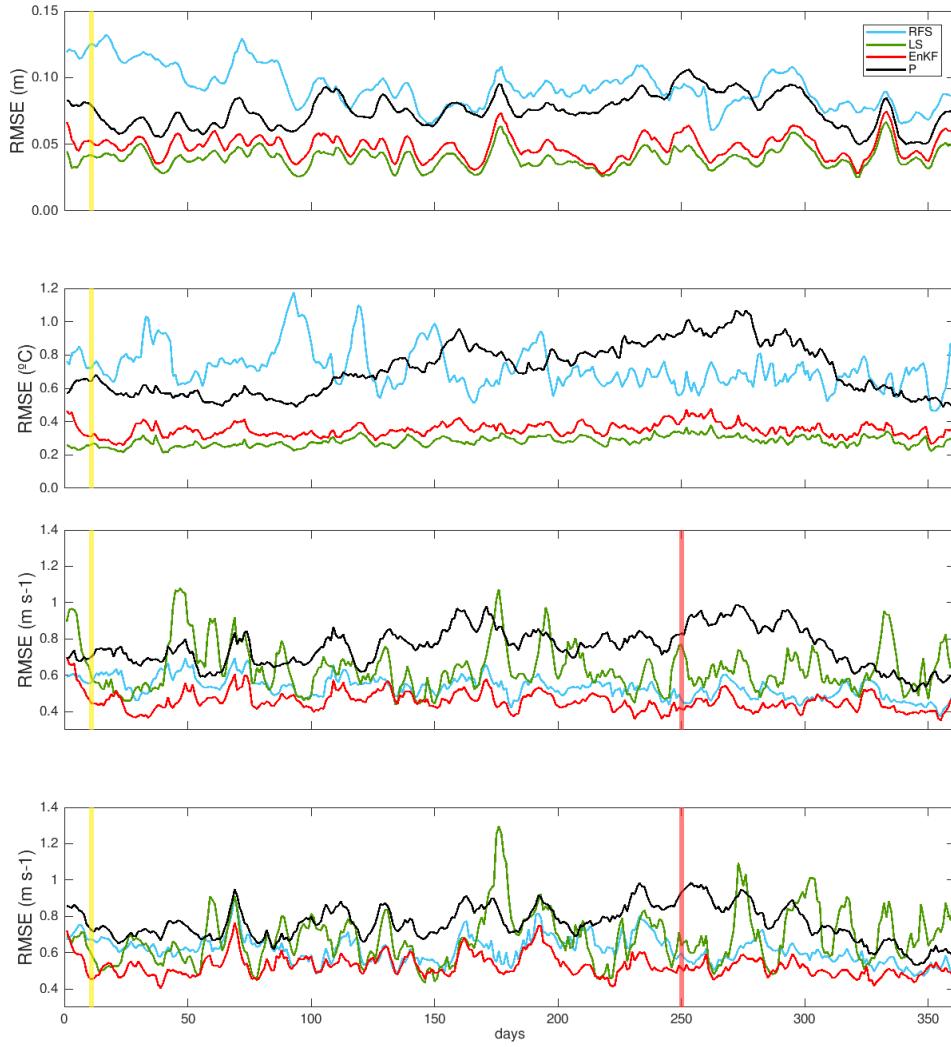


Figure 6: RMSE time series (from January 1st 2010) of the validation with respect to GHRSST, AVISO altimetry, and Globcurrent datasets to validate (from top to bottom): SST, SSH, U, and V. Blue corresponds to the reconstructed full-space (RFS), green is the Least Squares (LS) solution, red represents the ensemble-mean (DA+NN), and black is the analysis using the EnKF and persistence (P). Yellow and pink shades are centered on the snapshot days in figures 9 and 10.



Figure 7: Correlation coefficient (r) time series (from January 1st 2010) of the validation with respect to GHRSST, AVISO altimetry, and Globcurrent datasets to validate (from top to bottom): SST, SSH, U, and V. Blue corresponds to the reconstructed full-space (RFS), green is the Least Squares (LS) solution, red represents the ensemble-mean (DA+NN), and black is the analysis using the EnKF and persistence (P). Yellow and pink shades are centered on the snapshot days in figures 9 and 10.

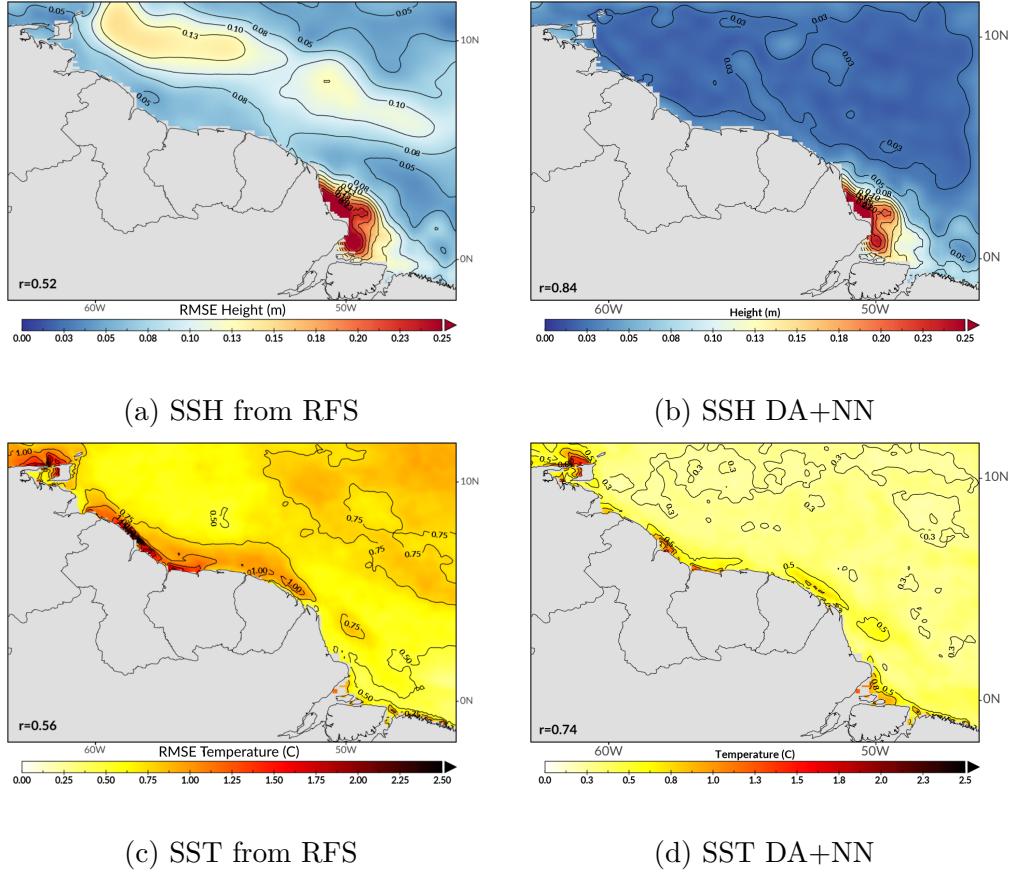


Figure 8: 365-day average spatial RMSE for SSH and SST. Left figures represent the average spatial RMSE of the reduced-space model solution (RFS) during 2010, while right figures represent the average spatial RMSE after the data assimilation with an ensemble Kalman filter (EnKF). The correlation coefficients (r) with respect to the validation datasets AVISO and GHRSST for SSH and SST, respectively, are depicted in the bottom left corner of each subfigure.

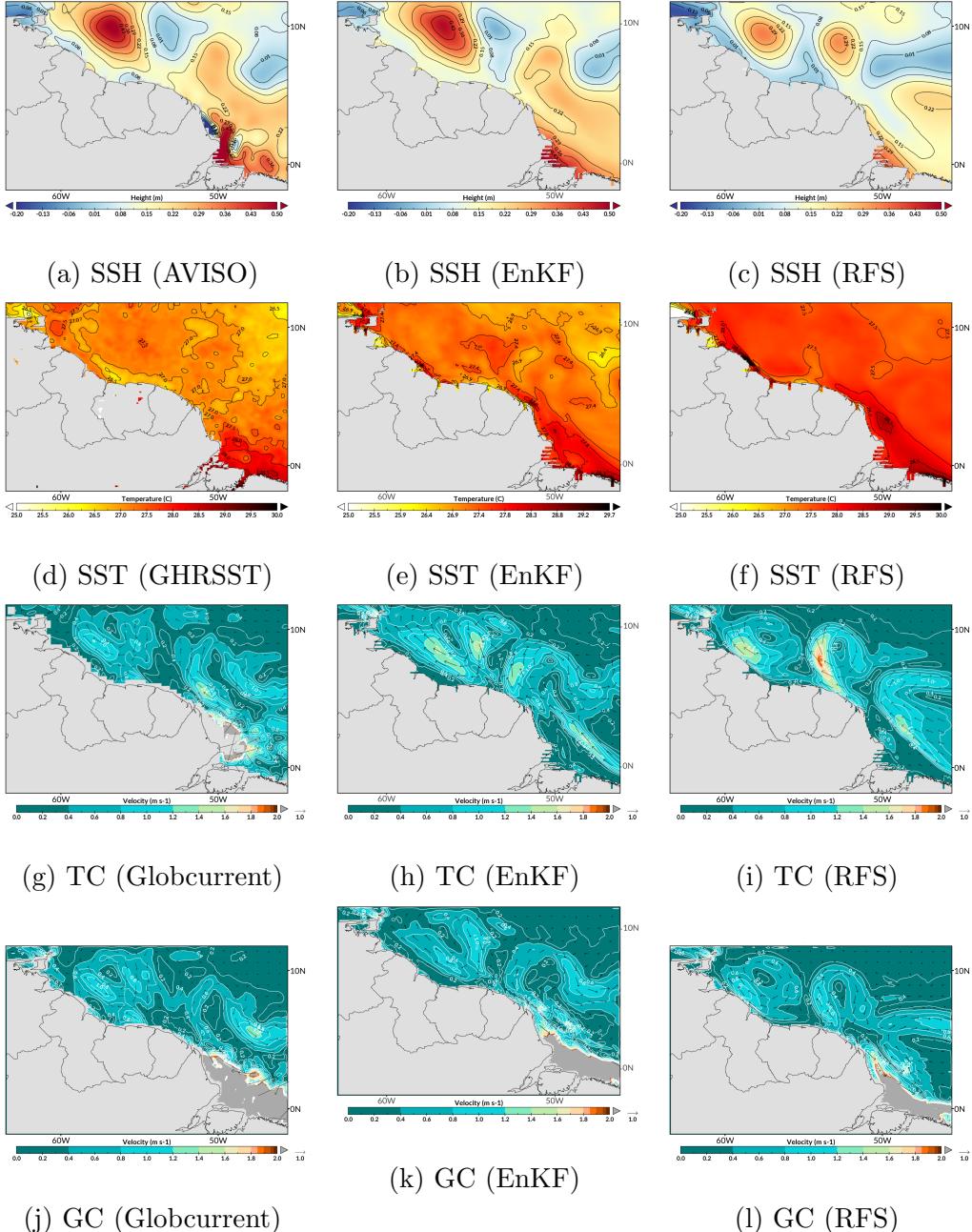


Figure 9: Validation datasets (left column), ensemble-mean of the assimilated reduced-space using EnKF (middle column), and reconstructed full-space model (RFS). Snapshot of January 11th, 2010 (yellow shade in Fig. 6 and 7). This day shows the comparison between the validation datasets and the ensemble-mean of the EnKF on the reduced-space. The validation datasets are SSH from AVISO (0.25-degree resolution), SST from GHRSST (10 km resolution), total currents (TC) from Globcurrent (0.25-degree resolution) and geostrophic currents (GC) from Globcurrent (10km resolution). The surface currents represent the vector magnitude and direction of U and V. Light gray areas correspond to the landmask or where Globcurrent does not have data; dark gray correspond to areas