

Learning stochastic representations of geophysical dynamics

Said Ouala, Ronan Fablet, Cédric Herzet, Bertrand Chapron, Ananda Pascual, Fabrice Collard, Lucile Gaultier

▶ **To cite this version:**

Said Ouala, Ronan Fablet, Cédric Herzet, Bertrand Chapron, Ananda Pascual, et al.. Learning stochastic representations of geophysical dynamics. ICASSP 2019 : 44th International Conference on Acoustics, Speech, and Signal Processing, May 2019, Brighton, United Kingdom. 10.1109/ICASSP.2019.8682929 . hal-02005403

HAL Id: hal-02005403

https://imt-atlantique.hal.science/hal-02005403

Submitted on 4 Feb 2019

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.

LEARNING STOCHASTIC REPRESENTATIONS OF GEOPHYSICAL DYNAMICS

Said Ouala[1], *Ronan Fablet*[1], *Cédric Herzet*[1,2], *Bertrand Chapron*[3]
Ananda Pascual[4], *Fabrice Collard*[5], *Lucile Gaultier*[5]

(1) IMT Atlantique; Lab-STICC, Brest, France
(2) INRIA Bretagne-Atlantique, Fluminance, Rennes, France
(3) Ifremer, LOPS, Brest, France
(4) IMEDEA, UIB-CSIC, Esporles, Spain
(5) ODL, Brest, France

ABSTRACT

In the last years, Neural Networks have enriched the state-of-the-art in probabilistic modeling. This is principally due to the advances in deep learning which allow a better understanding of complex systems. However, the stochastic representation of spatio-temporal fields is still an open challenge that may benefit from the recent advances in probabilistic modelization. In this work, we explore neural network to derive a stochastic representation of spatio-temporal dynamical systems based on ensemble forecasting. Trough the implementation of our stochastic model in a classical Kalman filtering scheme, we demonstrate the relevance of the proposed architecture in the reconstruction of geophysical fields with respect to the state-of-the-art approaches.

***Index Terms*—** Probabilistic modeling, Dynamical systems, Neural networks, Kalman filter

1. PROBLEM STATEMENT AND RELATED WORK

The high-resolution monitoring of sea surface geophysical parameters is one of the significant challenges in oceanography. Producing high resolution gridded spatio-temporal products of physical variables such as sea surface temperature, sea surface height and sea surface salinity is of key interest for a variety of scientific fields and applications, from numerical weather prediction to maritime operations [1, 2, 3].

Data assimilation is known as the state-of-the-art techniques used for the reconstruction of such geophysical fields. They are present in two main categories : variational and stochastic data assimilation. While variational data assimilation relies on the minimization of a cost function and requires the computation of the cost function gradient through the adjoint of the dynamical model [4]. Stochastic data assimilation involves Monte Carlo strategies through an ensemble formulation to explore the state space of our physical model [5].

The Key feature of both variational and stochastic data assimilation techniques is the mathematical formulation of the dynamical model describing the temporal evolution of our physical variables. The definition of this dynamical model dramatically limits the use of data assimilation in the case of geophysical fields interpolation. First, analytical modeling techniques may result in complex representations that can not be used in assimilation schemes due to their numerical complexity [6]. Complex dynamics may also involve uncertain parameterizations, whereas simplified parameterizations may only hold for specific dynamical modes [7].

Data-driven strategies in the other hand arise as a relevant alternative to model-driven techniques[8]. several representations have shown a great success in forecasting physical dynamics in several domains [9, 8, 10]. However, when dealing with chaotic systems such as oceanic tracers, modeling the variability in terms of probabilistic representation is still an open challenge due to the complexity of such dynamics.

Neural networks, especially deep models, has become the state-of-the-art in a range of machine learning applications. Among others, neural networks with stochastic layers and generative models have significantly advanced the field of probabilistic modeling for complex processes and signals [11, 12].

In this study, we are specifically interested in the probabilistic modeling of spatio-temporal fields for simulation and assimilation issues. We investigate stochastic data-driven representations of spatio-temporal dynamics based on neural networks. The key idea of our work is to design a Gaussian neural-network representation for the propagation of spatio-temporal dynamics. Inspired by the variational approximation based techniques [11], the proposed approach can be regarded as a neural-network representation of the ensemble forecasting based scheme classically used in Bayesian filters (e.g., particle filters, Kalman filters, Ensemble Kalman fil-

This work was supported by GERONIMO project (ANR-13-JS03-0002), Labex Cominlabs (grant SEACS), Region Bretagne, CNES (grant OSTST-MANATEE), Microsoft (AI EU Ocean awards) and by MESR, FEDER, Région Bretagne, Conseil Général du Finistère, Brest Métropole and Institut Mines Télécom in the framework of the VIGISAT program managed by "Groupement Bretagne Télédétection" (BreTel).

ters) to propagate probabilistic representations of non-linear dynamics [5]. The proposed neural network representation is learnt from data. It can be combined with a classic Kalman recursion to address assimilation issues. Overall, our key contributions are three-fold : i) we propose a novel Gaussian NN-based representation of high-dimensional geophysical dynamics, ii) we derive the associated NN-based Kalman filtering scheme for spatio-temporal interpolation issues, iii) We demonstrate the relevance of these contributions with respect to state-of-the-art approaches [13, 14, 15] for an application to the spatio-temporal interpolation of satellite-derived data.

This paper is organized as follows. Section 2 describes the proposed stochastic neural-network dynamical model and its implementation in a Kalman filtering scheme. Section 3 presents the results of the numerical experiments. We further discuss our contributions in Section 4.

2. LEARNING GAUSSIAN DYNAMICAL MODELS

2.1. Proposed stochastic dynamical model

We aim to model the propagation of a Gaussian approximation (*i.e.*, a mean and covariance) of the true dynamics of the system of interest. Contrary to parametric techniques [16, 17] based on an analytic PDE formulation of the propagation of a Gaussian dynamical model using specific dynamical priors, we aim to benefit from the flexibility and genericity of neural-network architectures to learn our Gaussian dynamical model from a representative training dataset.

Formally, assuming that we are provided with a dataset $X_{0:T}$ of our physical system, our objective is to learn a probabilistic representation of the dynamical model through a parametric model of the conditional distribution of state variable X_t given X_{t-1}. Considering a Gaussian approximation, it comes to the following model :

$$X_t|X_{t-1} \sim \mathcal{N}(X_t = \mathcal{F}(X_{t-1}), \Sigma_t = \mathcal{F}_\Sigma(X_{t-1}, \Sigma_t)) \quad (1)$$

where functions $\mathcal{F}, \mathcal{F}_\Sigma$ are neural networks with parameter vectors $\theta = (\theta_\mu, \theta_\Sigma)$.

Following our previous works on analog data assimilation [18, 15], we consider a patch-based representation as a mean to decompose space-time scales as well as to provide an explicit relationship between the global and local (patch-level) representations. Regarding the mean model $\mathcal{F}$, the proposed architecture proceeds as follows :

— At a given time t, the first layer of the network, which is parameter-free in terms of training, comes to decompose an input field x_t into a collection of N_p $P \times P$ patches $x_{\mathcal{P}_s,t}$, where P is the width and height of each patch and $s \in [1, ..., N_p]$ the patch location in the global field. Each patch is decomposed onto an EOF basis $\mathcal{B}$ according to :

$$z_{\mathcal{P}_s,t} = x_{\mathcal{P}_s,t}\mathcal{B}^T \quad (2)$$

with $z_{\mathcal{P}_s,t}$ the EOF decomposition of the patch $x_{\mathcal{P}_s,t}$. The EOF decomposition matrix $\mathcal{B}$ is trained offline as a preprocessing step ;

— The second layer implements a numerical integration scheme (typically, an Euler or 4th-order Runge-Kutta scheme) using a patch-level dynamical model $\mathcal{F}^{\mathcal{P}_s}$ to predict $z_{\mathcal{P}_s,t+1}$. For patch-level models $\mathcal{F}^{\mathcal{P}_s}$, we consider residual architectures [11] with a bilinear parameterization [19] ;

— The third layer is a reconstruction network $\mathcal{F}_r$. It combines the predicted patches $x_{\mathcal{P}_s,t+1} = z_{\mathcal{P}_s,t+1}\mathcal{B}$ to reconstruct the output field x_{t+1}. This reconstruction network $\mathcal{F}_r$ involves a convolution neural network [20].

The details of the considered parameterizations for the second and third layers are given in Section 3. To train the mean dynamical model $\mathcal{F}$, we apply a two-step procedure. We first learn the local dynamical models $\mathcal{F}^{\mathcal{P}_s}, s \in [1, ..., N_p]$ based on the minimization of the EOF-patch based forecasting error. The reconstruction network $\mathcal{F}_r$ is then optimized using the same criterion over the global field.

Due to dimensionality issues, one cannot expect to model the full covariance of the field. We benefit from the patch-level representation to investigate a block-diagonal approximation of the covariance structure. It comes to consider a diagonal covariance parameterization of the patch-level covariance matrix in the EOF space. Formally, the block-diagonal covariance model is learnt as follows :

— For the t^{th} state X_t in our training set, a patch-based projection is first performed to compute the states in the EOF space $z_{\mathcal{P}_s,t}$ according to equation 2.

— For each $s \in [1, ..., N_p]$, an ensemble $E_{\mathcal{P}_s,t} = \{z^{(1)}_{\mathcal{P}_s,t}, ..., z^{(M)}_{\mathcal{P}_s,t}\}$ of size M is sampled from a Gaussian distribution with mean $z_{\mathcal{P}_s,t}$ and covariance matrix Σ_0 computed as the covariance matrix of our training data in the EOF space.

— From the ensemble data $E_{\mathcal{P}_s,t}$, we build a forecasting ensemble $\mathcal{F}^{\mathcal{P}_s}(E_{\mathcal{P}_s,t})$ by applying to each sample the trained model $\mathcal{F}^{\mathcal{P}_s}$. A neural network covariance model $\mathcal{F}^{\mathcal{P}_s}_{D,\Sigma}$ is then trained to forecast a diagonal covariance matrix in the EOF space $\Sigma^{EOF}_{\mathcal{P}_s,t+1} = \mathcal{F}^{\mathcal{P}_s}_{D,\Sigma}(z_{\mathcal{P}_s,t}, \Sigma_0)$ using the ensemble dataset $\{E_{\mathcal{P}_s,t-1}, \mathcal{F}^{\mathcal{P}_s}(E_{\mathcal{P}_s,t-1})\}$ according to a ML criterion.

— Ones our EOF covariance model learned, the patch-based covariance model $\mathcal{F}^{\mathcal{P}_s}_\Sigma$ is deduced according to :

$$\begin{aligned}\mathcal{F}^{\mathcal{P}_s}_\Sigma(x_{\mathcal{P}_s,t+1}, \Sigma_{\mathcal{P}_s,t+1}) =& \mathcal{B}^t\Psi(\Sigma_{\mathcal{P}_s,t}, \Sigma_0) \\ & \times \mathcal{F}^{\mathcal{P}_s}_{D,\Sigma}(z_{\mathcal{P}_s,t}, \Sigma_0) \cdot \mathcal{B}\end{aligned} \quad (3)$$

where Ψ a scaling function.

— The global covariance model $\mathcal{F}_\Sigma$ is given by the recombination of the N_p local patch-level covariance models $\mathcal{F}^{\mathcal{P}_s}_\Sigma$.

The parametrization of the patch-based covariance model in the EOF-space allows us learning only the variance vector of each patch sins the EOF components are supposed to be orthogonal (resulting in insignificant covariance components in the EOF-space). However, and to illustrate the relevance of the proposed patch-based parametrization of the covariance matrix in the EOF-space, we also investigate a diagonal covariance matrix model in the patch space. according to :

$$\mathcal{F}_{\Sigma}^{\mathcal{P}_s}(x_{\mathcal{P}_s,t+1}, \Sigma_{\mathcal{P}_s,t+1}) = \Psi(\Sigma_{\mathcal{P}_s,t}, \Sigma_0) \cdot \mathcal{F}_{D,,\Sigma}^{\mathcal{P}_s}(x_{\mathcal{P}_s,t}, \Sigma_0) \quad (4)$$

The aim of using an ensemble likelihood optimization criteria is that geophysical dynamical models usually present complex, possibly chaotic, dynamical behaviours. Therefore, using a single forecast to characterize the variability of the dynamical model can result in an underestimation (or an overestimation) of the covariance matrix. The proposed Monte Carlo ML criterion explores the one-step prediction variability resulting in a parametric Gaussian approximation of an ensemble forecasting.

2.2. Application to data assimilation

Given this neural-network representation of the conditional distribution $X_t|X_{t-1}$, we can derive an associated Kalman-based filtering under the assumption that the observation model is linear and Gaussian.

In contrast to classical non linear filtering frameworks illustrated for instance by Extended Kalman filtering and Ensemble Kalman filtering. The proposed Kalman-based filtering technique directly propagates the Gaussian pdf (see equation 5) at a given time step (using the Gaussian transition model). It simply comes to apply the Kalman recursion with the mean and covariance propagated by the trained neural networks :

$$(X_{t+1}^-, \Sigma_{t+1}^-) = (\mathcal{F}(X_t^+), \mathcal{F}_{\Sigma}(X_t^+, \Sigma_t^+)) \quad (5)$$

$$X_{t+1}^+ = X_{t+1}^- + K_{t+1}[Y_{t+1} - H_{t+1}X_{t+1}^-] \quad (6)$$

$$K_{t+1} = \Sigma_{t+1}^- H_{t+1}^t [H_{t+1}\Sigma_{t+1}^- H_{t+1}^t + R_t]^{-1} \quad (7)$$

where H_{t+1} corresponds to some linear observation models. The superscript (-) refers to the forecasting of the mean of the state variable x_{t+1}^- and of its covariance matrix Σ_{t+1}^-, the superscript (+) refers in the other hand to the assimilated mean x_{t+1}^+ and covariance Σ_{t+1}^+. R_t is a Gaussian noise covariance matrix.

As explained above, we consider two different parameterizations of the covariance model $\mathcal{F}_{\Sigma}$. The key aspect of this model is the structure of the patch based covariance operator $\mathcal{F}_{\Sigma}^{\mathcal{P}_s}$. Different assimilation techniques could be envisaged depending on its parametrization :

— Modeling the patch based covariance in the EOF space and projecting it back to the patch space which results in a block-diagonal covariance matrix (illustrated by equation 3) : PB-NNKF-EOF;

— Directly modeling the variance vector of each patch resulting on a diagonal covariance matrix (illustrated by equation 4) : PB-NNKF.

3. NUMERICAL EXPERIMENTS

3.1. Dataset description

As a case-study, we address the spatio-temporal interpolation of satellite-derived sea surface temperature (SST) fields associated with infrared sensors. The time series is delivered by the UK Met Office [13] from January 2008 to December 2015. The spatial resolution of our SST field is $0.05°$ and the temporal resolution $h = 1$ day. The data from 2008 to 2014 were used as training data and we tested our approach on the 2015 data. To perform a quantitative evaluation, we simulated realistic spatio-temporal cloud patterns over the test set using the METOP-AVHRR masks. This sensor is highly sensitive to the cloud cover and results in very high missing data rates.

As case-study region, we select an area off South Africa (from $2.5°E, 38.75°S$ to $32.5°E, 58.75°S$). This region involves complex fine-scale SST dynamics (e.g., fronts, filaments). It makes it relevant for the considered quantitative evaluation.

3.2. Experimental setting

The proposed stochastic dynamical model exploits patch-level representations with non-overlapping 20×20 patches. Each patch $\mathcal{P}_s$ is then projected into an EOF basis learnt from the training data. We keep the first 50 EOF components, which amount to encode 95% of the total variance. For the patch-level NN model $\mathcal{F}^{\mathcal{P}_s}$, we use a bilinear residual neural network architecture as proposed in [10] with 60 linear neurons, 100 bilinear neurons and 10 fully-connected layers with a Relu activation. The reconstruction model $\mathcal{F}_r$ is a convolutional neural network with 3 convolutional layers. The first two layers comprise 64 filters of size 3×3 with a *Relu* activation and the last layer is a linear convolutional layer with one filter. Regarding covariance model $\mathcal{F}_D^{\mathcal{P}_s}$, we consider a diagonal covariance model within each patch. Each element of diagonal involves a 3-layer MLP with 4 neurons and *Relu* activation functions on the hidden layers and a *softplus* activation in the output layer. With a view to evaluating the EOF-based covariance parameterization, we consider both PB-NNKF-EOF and PB-NNKF schemes. A constant scaling function $\Psi() = 1$ in equations 3 and 4 led to the best performances in our experiments.

We perform a quantitative analysis of the interpolation performance of the proposed schemes with respect to an optimal interpolation, the analog data assimilation [15] and the EOF based interpolation method VE-DINEOF. The considered parameter setting is as follows :

— Optimal interpolation (OI) : We use a Gaussian kernel with a spatial correlation length of 100km and a

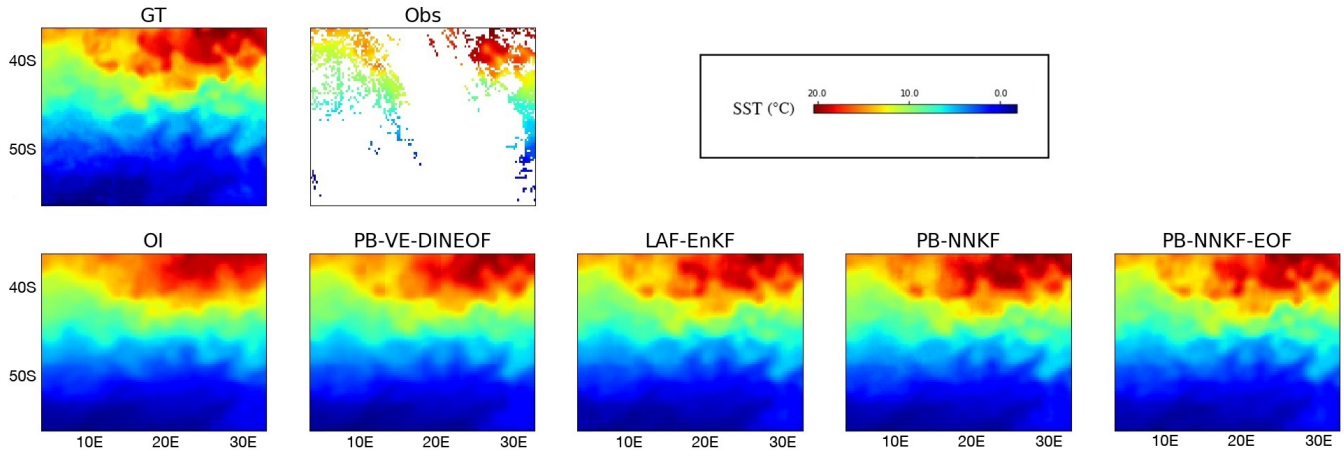


Fig. 1: ***Interpolation of the SST field on July 19 2015 :*** *first row, the reference SST and the observation with missing data (here, 82% of missing data); second row, interpolation results using respectively OI, PB-VE-DINEOF, LAN-ENKF, PB-NN-NNKF, PB-NN-NNKF-EOF.*

temporal resolution length of 3 days. These parameters were empirically tuned for the considered dataset using a cross-validation experiment.

— Analog data assimilation (LAF-EnKF) : We test the local analog data assimilation scheme [15, 21]. Similarly to the proposed architecture, we consider 20×20 patches and 50-dimensional EOF decomposition with an overlapping of 10 pixels. We let the reader refer to [15, 21] for a detailed description of this data-driven approach, which relies on nearest-neighbor regression techniques.

— EOF based reconstruction (PB-VE-DINEOF) : We also compare our approach to the state-of-the-art interpolation scheme based on the projection of our observations with missing data on an EOF basis [14]. The SST field is here decomposed as described in the analog data assimilation application into a collection of 20×20 patches with 10 pixels overlapping. Each patch is then reconstructed using the VE-DINEOF method.

3.3. Results

We report the mean interpolation performance in Tab. 1. Figure 1 illustrates an interpolation example of the tested methods. The proposed NN-based scheme (PB-NNKF-EOF) leads to very significant improvements with respect to the optimal interpolation in terms of RMSE and correlation coefficients, which emphasizes fine-scale structures (e.g., relative improvement of the RMSE above 50% for missing data areas). A clear gain is also exhibited w.r.t. analog data assimilation and PB-VE-DINEOF schemes with a relative gain greater than 20% in terms of RMSE. Although the considered NN-based representation exploits non-overlapping patches, we still come up with significant improvements w.r.t AnDA scheme which involve a 50% overlapping rate between patches. This clearly illustrates the relevance of NN-based representation, which fully embeds the direct and inverse mappings between the SST field and its patch-level representation. Interestingly, Tab.1 also reveals the importance of the EOF-based parameterization of the NN-based covariance model (Equation 3) in the improvement of interpolation results w.r.t. AnDA schemes.

Model		Entire map	Missing data areas
OI	RMSE	0.76	0.75
	Correlation	99.35%	99.37%
PB-VE-DINEOF	RMSE	0.54	0.54
	Correlation	99.68%	99.66%
LAF-EnKF	RMSE	0.43	0.42
	Correlation	99.79%	99.77%
PB-NNKF	RMSE	0.51	0.51
	Correlation	99.75%	99.71%
PB-NNKF-EOF	RMSE	**0.33**	**0.35**
	Correlation	**99.87%**	**99.85%**

Table 1: ***SST interpolation experiment*** *: Mean reconstruction correlation coefficient and RMSE over the SST time series.*

4. CONCLUSION

In this work, we address stochastic data driven representations for spatio-temporal fields identification. we rely on neural networks and ensemble forecasting to derive a neural network Gaussian dynamical model. From our numerical experiments, combining our probabilistic representation with Kalman based data assimilation techniques outperforms classical state-of-the-art techniques in spatio-temporal fields reconstruction.

5. REFERENCES

[1] N. J. Hardman-Mountford, A. J. Richardson, D. C. Boyer, A. Kreiner, and H. J. Boyer, "Relating sardine recruitment in the Northern Benguela to satellite-derived sea surface height using a neural network pattern recognition approach," *Progress in Oceanography*, vol. 59, no. 2, pp. 241–255, oct 2003.

[2] Pierre-Yves Le Traon, "Satellites and operational oceanography," in *Operational Oceanography in the 21st Century*, pp. 29–54. Springer, 2011.

[3] Karina Von Schuckmann, Pierre-Yves Le Traon, Enrique Alvarez-Fanjul, Lars Axell, Magdalena Balmaseda, Lars-Anders Breivik, Robert JW Brewin, Clement Bricaud, Marie Drevillon, and Yann Drillet, "The copernicus marine environment monitoring service ocean state report," *Journal of Operational Oceanography*, vol. 9, no. sup2, pp. s235–s320, 2016.

[4] A. C. Lorenc, S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, and F. W. Saunders, "The Met. Office global three-dimensional variational data assimilation scheme," *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 570, pp. 2991–3012, oct 2000.

[5] Geir Evensen, *Data Assimilation*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[6] Song-You Hong and Jimy Dudhia, "Next-generation numerical weather prediction : Bridging parameterization, explicit clouds, and large eddies," *Bulletin of the American Meteorological Society*, vol. 93, no. 1, pp. ES6–ES9, 2012.

[7] Juan Jose RUIZ, Manuel PULIDO, and Takemasa MIYOSHI, "Estimating model parameters with ensemble-based data assimilation : A review," *Journal of the Meteorological Society of Japan. Ser. II*, vol. 91, no. 2, pp. 79–99, 2013.

[8] Pierre Tandeo, Pierre Ailliot, Bertrand Chapron, Redouane Lguensat, and Ronan Fablet, "The analog data assimilation : application to 20 years of altimetric data," in *International Workshop on Climate Informatics*, Boulder, United States, sep 2015, pp. 1 – 2.

[9] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, Apr. 2016.

[10] Said Ouala, Cedric Herzet, and Ronan Fablet, "Sea surface temperature prediction and reconstruction using patch-level neural network representations," *arXiv :1806.00144 [cs, stat]*, may 2018, arXiv : 1806.00144.

[11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep Residual Learning for Image Recognition," *arXiv :1512.03385 [cs]*, december 2015, arXiv : 1512.03385.

[12] Marco Fraccaro, Søren Kaae Sønderby, Ulrich Paquet, and Ole Winther, "Sequential Neural Models with Stochastic Layers," *arXiv :1605.07571 [cs, stat]*, may 2016, arXiv : 1605.07571.

[13] Craig J. Donlon, Matthew Martin, John Stark, Jonah Roberts-Jones, Emma Fiedler, and Werenfrid Wimmer, "The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system," *Remote Sensing of Environment*, vol. 116, no. Supplement C, pp. 140–158, jan 2012.

[14] Bo Ping, Fenzhen Su, and Yunshan Meng, "An Improved DINEOF Algorithm for Filling Missing Values in Spatio-Temporal Sea Surface Temperature Data," *PLOS ONE*, vol. 11, no. 5, pp. e0155928, may 2016.

[15] R. Fablet, P. H. Viet, and R. Lguensat, "Data-Driven Models for the Spatio-Temporal Interpolation of Satellite-Derived SST Fields," *IEEE Transactions on Computational Imaging*, vol. 3, no. 4, pp. 647–657, dec 2017.

[16] O. Pannekoucke, E. Emili, and O. Thual, "Modelling of local length-scale dynamics and isotropizing deformations," *Quarterly Journal of the Royal Meteorological Society*, vol. 140, no. 681, pp. 1387–1398, 2013.

[17] Olivier Pannekoucke, Sophie Ricci, Sebastien Barthelemy, Richard Ménard, and Olivier Thual, "Parametric Kalman filter for chemical transport models," *Tellus A : Dynamic Meteorology and Oceanography*, vol. 68, no. 1, pp. 31547, dec 2016.

[18] Redouane Lguensat, Phi Huynh Viet, Miao Sun, Ge Chen, Tian Fenglin, Bertrand Chapron, and Ronan FABLET, "Data-driven Interpolation of Sea Level Anomalies using Analog Data Assimilation," oct 2017.

[19] Ronan Fablet, Said Ouala, and Cedric Herzet, "Bilinear residual Neural Network for the identification and forecasting of dynamical systems," *SciRate*, dec 2017.

[20] Yann LeCun, Patrick Haffner, Léon Bottou, and Yoshua Bengio, "Object Recognition with Gradient-Based Learning," in *Shape, Contour and Grouping in Computer Vision*, David A. Forsyth, Joseph L. Mundy, Vito di Gesú, and Roberto Cipolla, Eds., Lecture Notes in Computer Science, pp. 319–345. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[21] Redouane Lguensat, Pierre Tandeo, Pierre Ailliot, Manuel Pulido, and Ronan Fablet, "The Analog Data Assimilation," *Monthly Weather Review*, aug 2017.