



RESEARCH ARTICLE

10.1029/2024MS004462

Neural Data Assimilation for Regime Shift Monitoring of an Idealized AMOC Chaotic Model

Perrine Bauchot^{1,2} , Angélique Drémeau^{1,2}, Florian Sévellec^{3,4}, and Ronan Fablet^{1,4} 

¹Lab-STICC, CNRS ENSTA IMT Atlantique UBO ENIB UBS, Brest, France, ²Institut Polytechnique de Paris (IPP), Palaiseau, France, ³Laboratoire d'Océanographie Physique et Spatiale, Univ Brest CNRS IRD Ifremer, Brest, France, ⁴ODYSSEY Project Team, INRIA CNRS IMT, Brest, France

Key Points:

- Neural data assimilation approaches can significantly improve the monitoring of poorly-observed oceanic multiscale chaotic dynamics
- Regular observation strategies reconstruct better the Atlantic Meridional Overturning Circulation abrupt slowdowns considered as climate extremes during the last glacial period

Correspondence to:

P. Bauchot,
perrine.bauchot@ensta.fr

Citation:

Bauchot, P., Drémeau, A., Sévellec, F., & Fablet, R. (2025). Neural data assimilation for regime shift monitoring of an idealized AMOC chaotic model. *Journal of Advances in Modeling Earth Systems*, 17, e2024MS004462. <https://doi.org/10.1029/2024MS004462>

Received 17 MAY 2024

Accepted 4 FEB 2025

Abstract Data assimilation (DA) reconstructs and forecasts the dynamics of geophysical processes using available observations and physical a priori. Recently, the hybridization of DA and deep learning has opened new perspectives to address model-data interactions. This paper explores its potential contribution to the analysis of a chaotic oceanic phenomenon: the centennial to millennial variability of the North Atlantic ocean circulation during the last glacial period. The implemented neural approach—4DVarNet—yields meaningful improvements over a classical variational DA method in reconstructing regime shifts of the Atlantic Meridional Overturning Circulation (AMOC), especially when fewer observations are available. Interestingly, results exhibit that exploiting explicitly the a priori dynamical model does not lead to better performances compared to a data-driven model. Additionally, we compare four sampling strategies to assess how observation patterns influence the capture of unstable AMOC phases. We highlight the gain of regular over random sampling strategies, with reconstruction errors dropping below 2% for a 100-year sampling period. We find that monitoring the AMOC with regular clusters of three consecutive observation points can reduce errors by a factor of five. Eventually, we assess 4DVarNet's robustness in reconstructing a partially-observed system and in generalizing to different dynamical regimes. We also investigate on the maximum sampling periods that 4DVarNet can assimilate without compromising reconstruction quality. This study, based on an idealized yet complex physical model, suggests that neural approaches trained on observations wisely acquired could improve the monitoring of regime shifts in the context of climate change.

Plain Language Summary This paper presents the benefits of deep learning for the monitoring of the centennial to millennial variability of the ocean circulation in the North Atlantic during the last glacial period. By improving the assimilation of observations and the representation of this complex phenomenon with a neural network, we reduce the error of reconstruction of regime shifts in the North Atlantic circulation by two orders of magnitude, compared with a classical method of assimilation. We also conducted experiments on the impact of the amount of observations and their moment of acquisition. Our results suggest that acquiring clusters of three consecutive observations in a regular manner enables to capture accurately these climate regime shifts. We also show that our method applies to a partially-observed system and generalizes beyond the dynamical regime considered in the training data set. We believe that this study establishes groundwork for a better monitoring of regime shifts in the context of climate change.

1. Introduction

Observing the ocean is a challenge, but one that needs to be met to improve the monitoring of oceanic processes in the context of climate change. For instance, one important oceanic phenomena which regulates the climate by redistributing heat globally is the Atlantic Meridional Overturning Circulation (AMOC). The AMOC drives the transport of warm surface water masses toward the North Pole. Through exchanges with the atmosphere, these water masses become denser and sink when arriving to the Arctic Ocean. They continue their journey southwards as dense cold water masses at depth. This phenomenon is associated with a net heat transport in the North Atlantic. It also traps heat and captures excess carbon emissions from the surface to the deep ocean (Buckley & Marshall, 2016).

Various studies have highlighted the complex nature of the AMOC dynamics, starting with Stommel (1961) who proposed a simple 2-box model to demonstrate the existence of different equilibria depending on the temperature and salinity conditions, confirmed later on by the more complex models of Thual and McWilliams (1992) and

Cessi and Young (1992). More recently, Germe et al. (2022) highlighted with an eddy-resolving ocean model that the unpredictable variability or chaotic nature of the AMOC was noticeable at different spatio-time scales. In particular, paleoceanography studies have evidenced abrupt climate shifts during the last glacial interval, referred to as Dansgaard-Oeschger (DO) events (Dansgaard et al., 1993). These events occurred approximately every $1,470 \pm 500$ years and were characterized by an abrupt slowdown of the AMOC. These paleoclimatic events are nowadays studied to understand how current ocean temperature and salinity changes due to sea ice extent variability could reflect on the AMOC (Sévellec & Fedorov, 2015). Given the impact of the AMOC onto the climate regulation, the AMOC shutdown is regarded as a potential climate tipping point (Ditlevsen & Ditlevsen, 2023), which motivates dedicated research effort both in terms of monitoring, modeling, and predictability issues (McCarthy et al., 2020; Rayner et al., 2011).

To reconstruct the AMOC variability of the last glacial interval, we propose to use a generic methodology in geosciences: data assimilation (DA). Most studies on the AMOC variability during this time period propose a proxy reconstruction based on paleorecords or climate models, which are sometimes compared to each other but rarely combined into one unique reconstruction, as DA would do (Hesse et al., 2011; Kessler et al., 2020; Menviel et al., 2017). Over the last decades, DA has been developed to reconstruct and forecast geophysical dynamics from noisy and partial observations, given some prior knowledge on the underlying dynamics (Carrassi et al., 2017). Forecast is achieved by letting the model run over the forecasting time period and correcting its estimate sequentially whenever observation are acquired. In contrast, reconstruction is performed retrospectively on the entire time window by incorporating every available observation (Evensen, 2009). In paleoclimatology, the first applications of DA date back only to 2000 with Von Storch et al. (2000). Paleoclimatology aims at reconstructing climate signals of the past and on a longer time-scale—from decades to centuries—than typically reconstructed in atmosphere or ocean sciences, which requires bigger computing resources (Fang & Li, 2016). This has motivated a stronger focus on ensemble Kalman approaches (Fang & Li, 2016; Okazaki & Yoshimura, 2017; Perkins & Hakim, 2021).

Recently, neural DA, which bridges deep learning and DA, has attracted a greater attention with potential breakthroughs for the targeted inverse problems thanks to their computational efficiency (Boudier et al., 2023; Brajard et al., 2020; Fablet et al., 2021). Neural DA may leverage neural components in state-of-the-art DA schemes, for instance surrogate dynamical models (Brajard et al., 2020; Peyron et al., 2021), as well as end-to-end differentiable DA schemes that can be optimized for a specific task (Boudier et al., 2023; Febvre et al., 2024). These recent advances appear appealing to monitor ocean processes, which remain usually poorly observed.

Indeed, while the ocean encompasses up to 70% of the Earth's surface, only 10% is considered to have been explored (Kim & Seto, 2022). Nowadays, the observation of the AMOC variability relies mostly on a few moorings measuring the dynamic height and the currents, on hydrographic campaigns and on satellite altimetry (McCarthy et al., 2020). Even if technological progress enabled scientists to develop a consistent monitoring network, these observation points remain sparse compared to the immensity of the phenomenon at study, which evolves on thousand kilometers with a temporal scale reaching decades to centuries while inducing vertical transport at deep sea, a place still hard to reach with observation systems. Through an idealized representation of the AMOC, this study investigates the interplay between the scarcity of the observations and the reconstruction schemes to inform DO events.

As Munk (2000) claimed, future advancements in oceanography can only be achieved by an adequate sampling in space, and also in time. Therefore, in this study, we focus on sampling strategies of time series of a few thousands of years. More specifically, we vary both the observation budget and the sampling patterns and explore how neural DA can improve the monitoring of regime shifts in chaotic climate dynamics under such observing conditions. We benchmark a neural DA scheme and classical state-of-the-art 4D-Var DA scheme (Cummings & Smedstad, 2013). The considered neural DA approach is an end-to-end neural model that learns a representation of the considered dynamical system and jointly solves a variational formulation DA problem (Fablet et al., 2021). This method, called 4DVarNet, has shown a great potential to reconstruct chaotic dynamics like Lorenz-63 and Lorenz-96 systems (Fablet et al., 2021), including application to real sea surface dynamics (Fablet et al., 2024; Febvre et al., 2022). We aim to explore how such neural DA schemes could provide a greater ability of neural DA schemes to retrieve the dynamics of the different phases of the AMOC when dealing with very sparse observations.

In this paper, we delve into the potential of neural inversion schemes for improving the reconstruction of climate regime shifts, such as DO events. We investigate how different observation strategies affect the monitoring

performance of 4DVarNet. The structure of the paper is as follows: we start by introducing the idealized dynamical system representing the AMOC and the DO events. Section 3 presents the considered neural and variational DA schemes. In Section 4, we report a comparative study between three assimilation models, from physics-constrained to purely data-driven schemes. We also explore various sampling strategies to help out the regime shift monitoring of the AMOC. Section 5 assesses the robustness of our neural model in face of more challenging observation conditions.

2. The Idealized Chaotic Model of the AMOC

This experimental study exploits a theoretical representation of the AMOC proposed by Sévellec and Fedorov (2014). This idealized model allows us to analyze the typical centennial to millennial variability of DO events during the last ice age. Inspired by the Howard-Malkus loop (Howard, 1971; Malkus, 1972), this model was validated against $\delta^{18}O$ paleorecords. The set of ordinary differential equations (**M**) for this model reads:

$$\mathbf{M} = \begin{cases} \dot{\omega}(t) = -\lambda\omega(t) - \epsilon\beta S_{NS}(t) \\ \dot{S}_{BT}(t) = (\Omega_0 + \omega(t))S_{NS}(t) - KS_{BT}(t) + \frac{F_0 S_0}{h}, \\ \dot{S}_{NS}(t) = -(\Omega_0 + \omega(t))S_{BT}(t) - KS_{NS}(t) \end{cases} \quad (1)$$

where ω is the time-varying component of the AMOC intensity; S_{BT} and S_{NS} are the vertical and meridional salinity gradients respectively; λ is a linear friction coefficient; ϵ is the buoyancy torque coefficient; β is the haline contraction coefficient; Ω_0 is the constant component of the AMOC intensity set according to mean oceanic temperature gradients and surface winds (such as $\Omega = \Omega_0 + \omega$, where Ω is the total AMOC intensity); K is the linear damping coefficient; F_0 is the freshwater flux intensity; S_0 is a salinity reference and h is the depth of the level of no motion for the baroclinic flow. The notation $\dot{\omega}(t)$ (resp. $\dot{S}_{BT}(t)$, $\dot{S}_{NS}(t)$) refers to the time derivative of ω (resp. $S_{BT}(t)$, $S_{NS}(t)$).

The first equation of the system refers to the momentum balance, while the second and the third equations define the evolution of the bottom-top and North-South salinity gradients, respectively. In order to work with an homogeneous state (say X), we apply the coefficient $\frac{\epsilon\beta}{\lambda}$ to the second and third equations. We obtain the following, time-depending, vector of three components $X(t) \triangleq [x_1(t), x_2(t), x_3(t)]^T = [\omega(t), \frac{\epsilon\beta}{\lambda}S_{BT}(t), \frac{\epsilon\beta}{\lambda}S_{NS}(t)]^T$. Thus, the dynamical system can eventually be written as:

$$\dot{X}(t) = \mathbf{M}(X(t)) = \begin{pmatrix} -\lambda(x_1(t) + x_3(t)) \\ (\omega_0 + x_1(t))x_3(t) - kx_2(t) + f \\ -(\omega_0 + x_1(t))x_2(t) - kx_3(t) \end{pmatrix} \quad (2)$$

where $k = \frac{K\lambda}{\epsilon\beta}$ and $f = \frac{F_0 S_0 \lambda}{h\epsilon\beta}$.

The time integration of the model highlights the different regimes of the system (Figure 1). We compute a global parameter relating to the energy of the system such as $d = \sqrt{x_1^2 + x_2^2 + x_3^2}$, from which we can infer experimentally three categories, equivalent to the ones derived in Sévellec and Fedorov (2014). The ON phase corresponds to a stable circulation with $d < 0.060 \text{ yr}^{-1}$, while the OFF phase indicates a shutdown of the AMOC with an energy exceeding $d > 0.130 \text{ yr}^{-1}$. In between, the regime shift corresponds to the transition between the ON and OFF phases. We note a regularity in this chaotic non-linear system, with a two-stage variability: a first centennial almost harmonic period close to 250 years occurs mainly during the ON phase; a second millennial variability of around $1,470 \pm 500$ years is linked to the temporal scale of the DO events, and corresponds to the periodicity of the AMOC shutdowns (i.e., OFF phase). Further information regarding the parametrization and dynamics of this model can be found in Sévellec and Fedorov (2014).

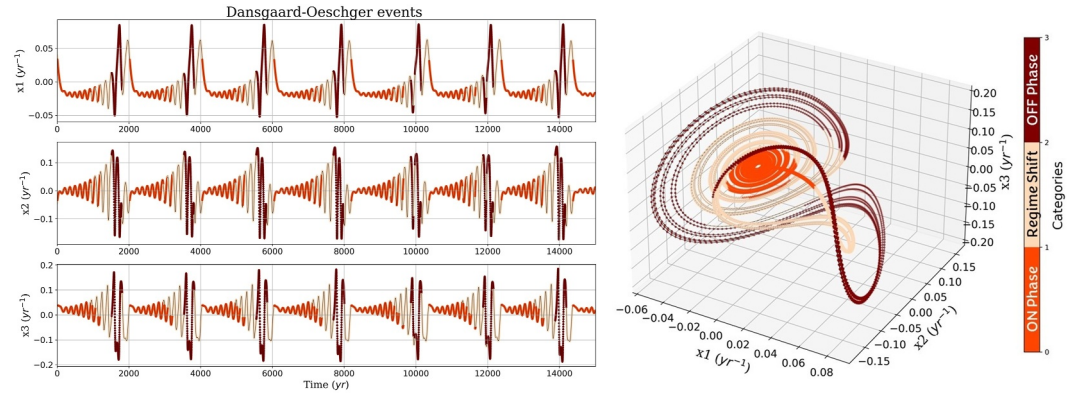


Figure 1. Dansgaard-Oeschger events simulated over 15,000 years. On the left, each line corresponds to the time evolution of the three re-scaled variables (x_1 , x_2 , and x_3). On the right, the model trajectory is shown in the phase-space. The dynamics is split into three categories: the ON Phase, the Phase Shift, and the OFF Phase of the Atlantic Meridional Overturning Circulation.

3. Neural Data Assimilation

The aim of this study is to monitor abrupt changes in a chaotic non-linear dynamical system by integrating it into a DA problem. Data assimilation is widespread in climate sciences due to its ability to solve inverse problems by taking into account both physical models and observations, in order to compute a better estimate of the ground truth (Carrassi et al., 2017; Johnson et al., 2005). Here, we adopt a variational approach of the problem, but we refer the reader to Evensen (2009) for details about the ensemble methods.

Let us consider the following inverse problem:

$$\begin{cases} \dot{\mathbf{X}}(t) = \mathbf{M}(\mathbf{X}(t)) + \mathbf{q}_{\text{dyn}}(t) \\ \mathbf{Y}(t) = \mathbf{H}_t \mathbf{X}(t) + \mathbf{q}_{\text{obs}}(t) \end{cases}, \quad (3)$$

where $\mathbf{X}$ is the state of the system at the considered time and $\mathbf{Y}$ is the vector of observations acquired with a Gaussian noise such as $\mathbf{q}_{\text{obs}} \sim \mathcal{N}(0, \sigma_{\text{obs}}^2 \mathbf{I})$. While $\mathbf{M}$ is the dynamical model containing the physical knowledge, $\mathbf{H}_t$ stands for an observation mask defining the chosen sampling strategy. $\mathbf{q}_{\text{dyn}}$ is the error associated with the model $\mathbf{M}$, which generally includes uncertainties on its parameters and numerical errors linked to its discretization.

The variational 4DVar method produces an estimate $\hat{\mathbf{X}}$ of the true state $\mathbf{X}$ by propagating the information brought by the observations to the rest of the system. This optimization problem is based on the minimization of a variational cost J over a chosen assimilation time window ΔT , such as $\hat{\mathbf{X}} = \text{argmin}_{\mathbf{X}} J(\mathbf{Y}, \mathbf{X}, \phi(\mathbf{X}), \mathbf{H})$ where J is defined as:

$$\begin{aligned} J(\mathbf{Y}, \mathbf{X}, \phi(\mathbf{X}), \mathbf{H}) = & \frac{1}{\Delta T} \left[\alpha_{\text{obs}} \int_{t_0}^{t_0+\Delta T} \|\mathbf{Y}(t) - \mathbf{H}_t \mathbf{X}(t)\|_2^2 dt \right. \\ & + \alpha_{\text{dyn}} \int_{t_0}^{t_0+\Delta T} \|\phi(\mathbf{X})(t) - \mathbf{X}(t)\|_2^2 dt \\ & \left. + \alpha_B \|\mathbf{X}_B - \mathbf{X}(t_0)\|_2^2 \right], \end{aligned} \quad (4)$$

with $\phi(\mathbf{X})(t) = \mathbf{X}(t') + \int_{t'}^t \mathbf{M}(\mathbf{X}(s)) ds$, where t , t' and s are time variables, ds and dt stand for the time unit, and ϕ is the propagator associated with $\mathbf{M}$ defining the dynamics of the AMOC. $\|\cdot\|_2^2$ is a norm such as $\|\mathbf{X}\|_2^2 = d^2$; t_0 is the starting time; $\mathbf{X}_B$ is the background state (i.e., an initial first guess of the optimization); and α_{obs} , α_{dyn} , and α_B

are weights which have to be optimally parameterized. In a Bayesian framework, these weights relate to covariance matrices accounting respectively for the observation error and the model error.

The first term of the variational cost J seeks to reduce the discrepancies between the observations and the estimated state of the system, while the second term favors a small difference between the output of the dynamical model and the estimated state $\mathbf{X}$ on the chosen time window ΔT starting at $t = t_0$. The last term corresponds to a background-dependent regularization of $\mathbf{X}$ in order to constrain the stability of the assimilation by providing an initial condition of the system. To be informative, $\mathbf{X}_B$ has to be a good approximation of the state $\mathbf{X}$ at time t_0 (Zupanski, 1997). In our study, this last term is only implemented in the classical 4D-Var DA method (hereinafter referred to as 4DVar-classic). On Figure 2, the 4DVar-classic workflow is represented by the light blue boxes only.

The minimization of J , usually reached by gradient descent, leads to the analyzed state at iteration $k \in \mathbb{N}^*$:

$$\mathbf{X}_{k+1} = \mathbf{X}_k - \delta \nabla_{\mathbf{X}} J(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H}) \quad (5)$$

where δ is the increment amplitude, and $\nabla_{\mathbf{X}} J(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H})$ stands for the gradient of J evaluated at point $\mathbf{X}_k$. This method is iterated until the variational cost converges. Its convergence also depends on the gradient step δ . It has to be set as a trade-off between a fast calculation and a stable convergence.

As evidenced by a growing literature, DA and deep learning share common theoretical grounds which advocate novel approaches exploring machine learning paradigms in DA problems (Arcucci et al., 2021; Brajard et al., 2020). In this study, we focus on end-to-end neural DA schemes (Boudier et al., 2023; Fablet et al., 2021), and more particularly 4DVarNets, as new means to optimize a DA scheme for given dynamics and observing systems.

3.1. 4DVarNet Architecture

The 4DVarNet architecture reproduces the unfolding of an iterative gradient descent algorithm to minimize the variational cost J as described in Equation 4. We implement a Residual Network (ResNet) architecture composed of convolutive Long Short-Term Memory (LSTM) residual units. Inspired by meta-learning (Vanschoren, 2018), this approach is particularly convenient in the case of the reconstruction of a dynamical system since previous states of the assimilation are stored into memory to help out the learning of the targeted parameters. Here, the parameters to be learned are the weights α_{obs} and α_{dyn} to optimize the minimization of J by speeding up its gradient descent. Therefore, the general architecture of the neural network resembles momentum-based gradient descent (Zhou et al., 2023), enabling faster and more robust convergence of the variational cost J toward an optimal state $\hat{\mathbf{X}}$.

As sketched in Figure 2, the 4DVarNet scheme applies three main steps at each iteration $k \in \mathbb{N}^*$ of the minimization algorithm:

1. We compute the variational cost $J(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H})$ for the observation data $\mathbf{Y}$, the current state $\mathbf{X}_k$, the output of the dynamical model ϕ and the chosen observation mask $\mathbf{H}$;
2. We apply the automatic differentiation according to $\mathbf{X}_k$ to obtain $\nabla_{\mathbf{X}} J(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H})$;
3. The reconstructed state $\mathbf{X}$ is updated such that $\mathbf{X}_{k+1} = \mathbf{X}_k - \delta \nabla_{\mathbf{X}} \tilde{J}(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H})$ where $\nabla_{\mathbf{X}} \tilde{J}(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H})$ is the gradient of $J(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H})$ as the output of a convolutional LSTM block $\mathcal{G}$ defined as $\nabla_{\mathbf{X}} \tilde{J}(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H}) = \mathcal{G}[\nabla_{\mathbf{X}} J(\mathbf{Y}, \mathbf{X}_k, \phi(\mathbf{X}_k), \mathbf{H})]$.

These steps are iterated over a predefined number K of iterations, typically up to a few tens. More details about this end-to-end scheme can be found in Fablet et al. (2021).

3.2. Representation of the Physical Model

The definition of the dynamical model $\mathbf{M}$ is crucial for the computation of the variational cost J , and therefore for the process of DA. To optimize variational DA for our study case, we propose two approaches to compute the physical prior ϕ : a numerical integration of $\mathbf{M}$ and a neural network representing ϕ directly.

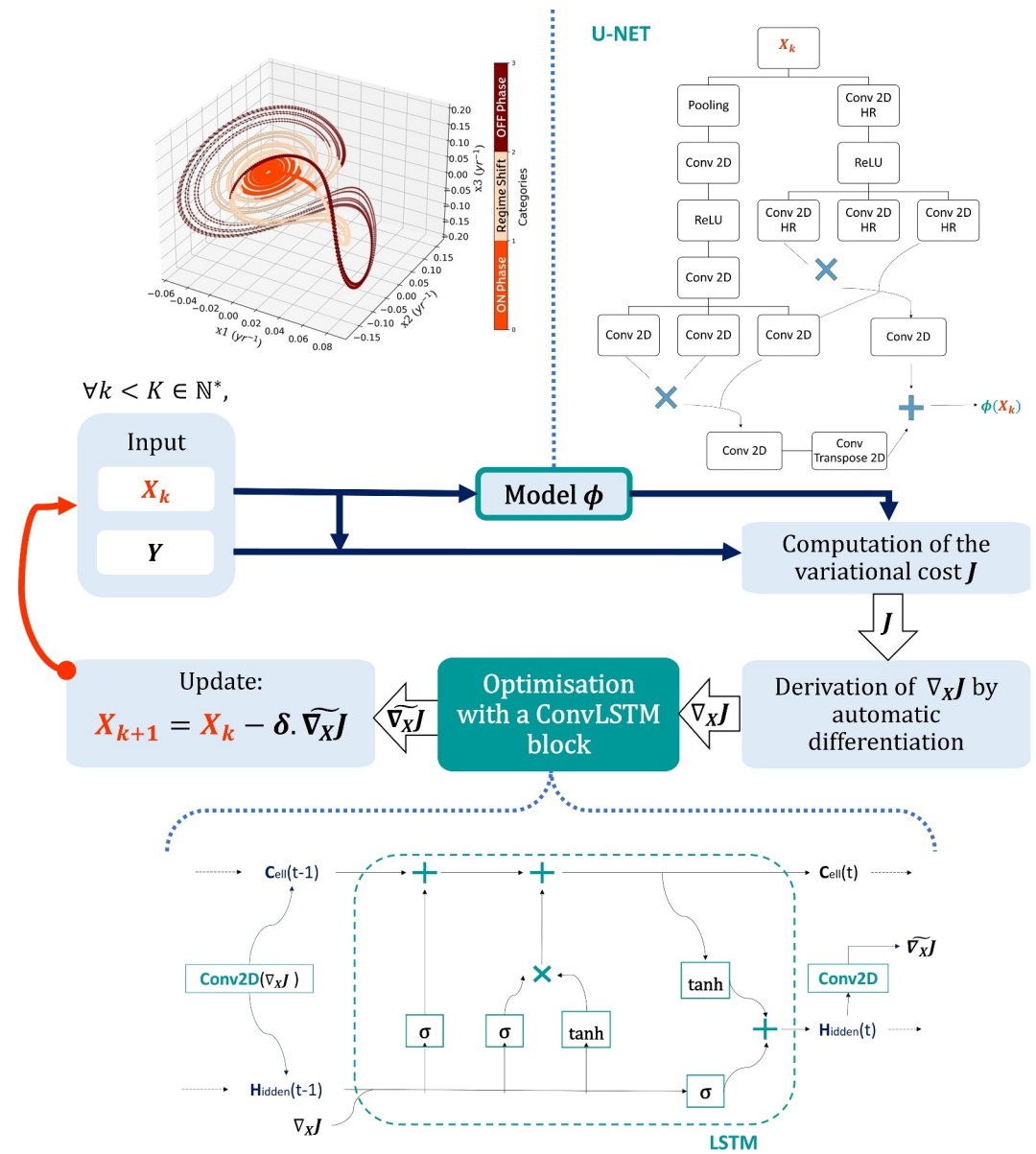


Figure 2. Sketch of the 4D-Var procedure. The model ϕ is computed from the state X in the 4DVar-classic and in the 4DVarNet-ode methods. In the 4DVarNet-unet method, ϕ is a U-Net cell, illustrated in the right upper corner of the Figure. Then, the variational cost J is calculated from the observations, the output of the model and the state X . Subsequently, the variational cost is derived by automatic differentiation. Next, in the 4DVarNet methods, the ConvLSTM network is trained to speed up the gradient descent minimizing the variational cost J . A scheme of the ConvLSTM cell is depicted. Finally, the analyzed state X is updated until the convergence of variational cost. The turquoise color indicates neural network blocks.

ϕ as a numerical integration of M : The numerical integration follows a Runge-Kutta 4 integration scheme and relies on the AMOC equations introduced in Section 2. This approach enables to inform the neural network with physics directly during its training. Physics-Informed Neural Networks (PINNs) is a growing field in machine learning since it provides a physical constraint, which facilitates the interpretability of the final output and of the optimization process (Dabrowski et al., 2023; Raissi et al., 2019).

ϕ as a U-Net operator: The U-Net architecture has been chosen to represent ϕ because of its capacity to learn the multiple scales of the system (Ronneberger et al., 2015). This is a complete data-driven approach, which is useful when the analytical solution of the system M cannot be retrieved. This approach is inspired by recent advances in

data-driven model discovery, made possible by the increasing volumes of data and improvements in computational efficiency over recent decades. These techniques are more and more investigated in geophysics, where the phenomena cannot always be put into equations because of their non-linearities and their chaotic natures (Berg & Nyström, 2019; Rudy et al., 2017). In our case, we know the physical system $\mathbf{M}$ and use it to simulate the data of the training data set, which aims at jointly training the U-Net and the convolutive LSTM block—the solver—, making the optimization of the DA problem powerful.

3.3. Learning Setting

4DVarNet is trained in a supervised way according to the following learning cost:

$$\forall l \in \mathbb{N}^*, \quad L_l = \frac{1}{l} \left[\sum_{t=0}^{(l-1)\Delta t} \alpha_1 \|\hat{\mathbf{X}}(t) - \mathbf{X}^{GT}(t)\|^2 + \alpha_2 (\|\phi(\hat{\mathbf{X}})(t) - \hat{\mathbf{X}}(t)\|^2 + \|\phi(\mathbf{X}^{GT})(t) - \mathbf{X}^{GT}(t)\|^2) + \alpha_3 \|\mathbf{X}^{GT}(t+1) - \mathbf{X}^{GT}(t) - \hat{\mathbf{X}}(t+1) + \hat{\mathbf{X}}(t)\|^2 \right], \quad (6)$$

where l is the number of time steps on which the dynamical system is assimilated, and α_1 , α_2 and α_3 are weights applied to each part of the learning cost. This learning cost is formulated to minimize the mean square error between the ground truth and the reconstructed state $\hat{\mathbf{X}}$, and between the output of the dynamical model ϕ and the state $\mathbf{X}$. A regularization term on the time derivative of $\mathbf{X}$ is added to limit the numerical noise of the reconstructed state $\hat{\mathbf{X}}$.

In our implementation, we consider a 10-iteration 4DVarNet scheme. Our training procedure involves 450 epochs to reach a training convergence. We use an Adam optimizer on batches of size of 128, with a dropout of 20% to avoid overfitting of the data (Kingma & Ba, 2017; Srivastava et al., 2014). The learning rate varies from 1×10^{-3} to 1×10^{-7} throughout the training. The time step of the predicted output is 10 years. The code is implemented in Pytorch and is available at <https://doi.org/10.5281/zenodo.14750068>.

4. Benchmarking 4DVar and Neural DA Schemes

4.1. Experimental Setup

By using the odeint numerical integrator from the scipy package on Python, we carried out 100 simulations of the AMOC system as defined by Equation 2 over a time period of 100,000 years, from which we extracted one training set and one test set (Figure 3). To build our training data set, we extract 50 2,500-year time series out of the first 55,000 years of each simulation. This leads to a training data set composed of 5,000 time series of 2,500 years. 10% of the training data set is reserved for the validation data set, used to evaluate the assimilation model during the learning process. Similarly, the test data set comprises 100 time series of 2,500 years extracted from the last 35,000 years of each simulation. This experimental setting enhances the diversity of situations provided to the learning scheme during the training phase, while guaranteeing the test data set to be independent from the training and validation ones. As DO events occur with a characteristic period of about $1,470 \pm 500$ years, each 2,500-year time series involves on average one DO events. This makes these simulated data sets relevant for the reconstruction of regime shifts in chaotic climate dynamics. We exploit the training and validation data sets during the training phase, and we consider the test data set to apply and benchmark the different DA schemes.

The observation vector $\mathbf{Y}$ is also constructed out of this simulated data set. To satisfy Shannon's criterion for the frequency of DO events, observations are acquired with a minimum sampling frequency of $\frac{1}{120} \text{ yr}^{-1}$. We add a Gaussian measurement noise to the three variables with a variance equivalent to 10% of variance of the centennial oscillation of the system. This percentage applies the same noise ratio across all observed variables. From these observations, we can derive the initial state of the DA, being a linear interpolation of the observations acquired on the DA window.

In our experiments, the three system components are observed simultaneously. We first compare the performance of three DA methods (described below) in order to find the best among them for reconstructing the North Atlantic

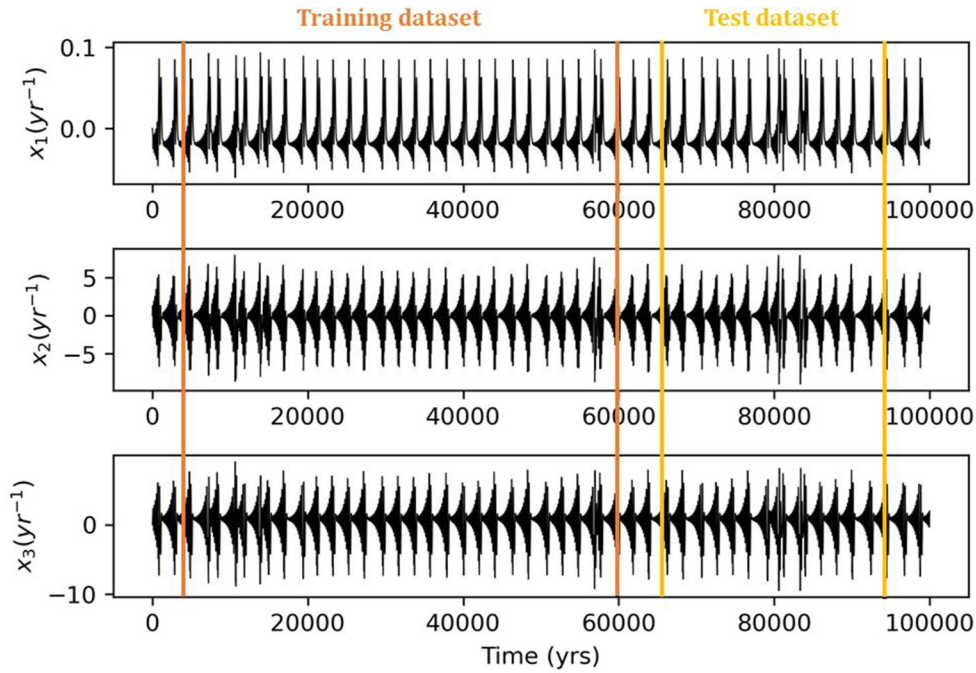


Figure 3. Evolution of the idealized model over one simulation of 100,000 years, 50,2500-year time series were extracted randomly between 5,000 and 60,000 years (in orange); one 2,500-year time series was extracted between 65,000 and 95,000 years to built the test data set (in yellow).

Ocean variability, and more specifically the regime shifts. Then, we analyze the influence of sampling strategies on the resolution of the assimilation problem, and find the best “data assimilation method/sampling strategy” pair to capture DO events.

4.2. Benchmarking Experiments

In this study, we benchmark three different DA methods.

1. **4DVar-classic:** A 4D variational DA method, where J is minimized by a gradient descent, until the convergence criterion is reached. The latter is set by a decay rate of less than 0.05 calculated over the last 10 iterations;
2. **4DVarNet-ode:** A learning DA method informed by physics, where J is optimized through a ResNet and the dynamics ϕ is a numerical integration of the system $\mathbf{M}$;
3. **4DVarNet-unet:** A learning DA method fully data-driven, where J is optimized through a ResNet and the dynamics ϕ is a U-Net operator jointly trained.

To compare these three methods, we first apply a regular sampling strategy with a 50 years sampling period and evaluate the reconstruction performance of each method. To get a initial idea on the performance, we compute the Normalized Mean Square Error (NMSE) for each variable using the following formula:

$$\text{NMSE} = \frac{1}{S\Delta T\sigma_{\text{GT}}^2} \sum_{s=0}^{S=100} \sum_{t=0}^{\Delta T=2500} \|\hat{\mathbf{X}}_{s,t} - \mathbf{X}_{s,t}^{\text{GT}}\|^2 \quad (7)$$

where $\mathbf{X}_s^{\text{GT}}$ is the ground truth and σ_{GT}^2 is its variance, S is the number of simulated trajectories on which the error is computed, and ΔT is the duration of the time series.

We find that the 4DVarNet-ode and 4DVarNet-unet methods allow to divide by a factor 10—and even by a factor from 30 to 80 for x_2 and x_3 —the reconstruction errors compared to a classical DA method (Table 1). This already demonstrates the usefulness of a neural network to improve the optimization problem within a variational DA framework. We also notice that the reconstruction of x_1 is overall better performed than the reconstruction of x_2

Table 1

Normalized Mean Square Error ($\times 10^{-3}$) of the Reconstructed Signal for Each Variable With Each Data Assimilation Method

Method	NMSE x_1	NMSE x_2	NMSE x_3
4DVar-classic	12.63	162.97	164.07
4DVarNet-ode	1.90	4.38	5.75
4DVarNet-unet	0.78	1.88	2.00

Note. Bold values indicate that it is the best performance.

and x_3 . By looking at the temporal evolution of the system, the difference of amplitudes between x_1 and (x_2 , x_3) is obvious, which can explain the difference in NMSE between variables of the system (Figure 4a). This NMSE difference is more important with the 4DVar-classic method, which indicates its difficulty to deal with multi-scale processes.

Qualitative comparison between 4DVarNet-ode and 4DVarNet-unet shows no specific differences in the reconstructed signal (Figure 4). They both fit the ground truth and capture the DO events variations. But, when represented in the phase-space, the differences between the trajectories adopted by 4DVarNet-ode and 4DVarNet-unet are more visible (Figure 4b). We assume

that these differences are linked to the optimization process. The 4DVarNet-unet method performs twice as good as the 4DVarNet-ode method (Table 1). While the 4DVarNet-ode method is guided by physics, potentially constraining trajectory of the AMOC system, 4DVarNet-unet is more efficient at reducing the NMSE, since it optimizes a learning cost minimizing this error as defined by Equation 6. This is aligned with the recent advances in data-driven models discovery, where physical models may not be the most relevant models to perform a particular task: the reconstruction of non-linear chaotic dynamics and, more specifically, to capture DO events. The spectral analysis confirms that 4DVarNet-unet still respects the spectral characteristics of the system (Figure 4c), and in particular the frequency of DO events.

Furthermore, 4DVar-classic has difficulties to capture DO events accurately for x_2 and x_3 (Figure 4). Indeed, the reconstruction of the highest amplitude variations, corresponding to the AMOC slowdown events (i.e., $x_1 > 0.05 \text{ yr}^{-1}$), appears weak and flattened by the 4DVar-classic method. By modifying the variational cost function, it might be possible to steer the assimilation toward a better reconstruction of these climate extremes. However, to produce a consistent comparison of the three DA methods, it was important to choose the same variational cost for each assimilation methods. Only the hyperparameters α_{obs} and α_{dyn} might vary depending on the assimilation method (Table 2). They are optimized for the 4DVarNet-unet and 4DVarNet-ode methods through the training of the neural network, while they are fixed to an optimal value found with a parametrization highly dependent on the experiment setup and assimilated system in the case of the 4DVar-classic method. This is one of the main issues with the 4DVar-classic method, which requires a tedious parametrization adapted to the case under study. In particular, Lorenc and Payne (2007) have shown the limits of a 4DVar-classic method to capture a wide range of scales, as in our case study. Since the regularization of a 4DVar-classic method might be inconvenient due to its high sensitivity, deep learning enables us to directly learn the optimal hyperparameters of our model.

It is worth noting that the variational cost amplitude (J) can significantly vary (Figure 5) depending on the values of the hyperparameters α_{obs} and α_{dyn} (Table 2). The considerable difference between α_{obs} and α_{dyn} in the 4DVar-classic method is due to the change in scale of the first and second terms of the variational cost J in Equation 4.

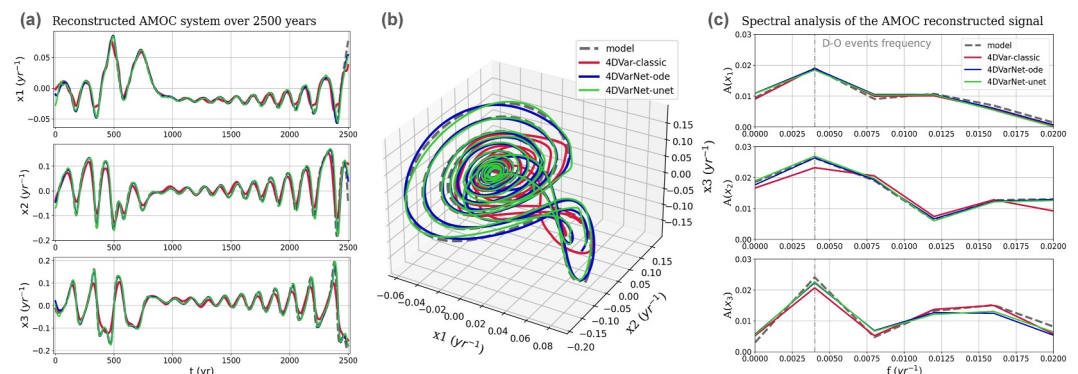


Figure 4. Qualitative comparison of the three Data assimilation (DA) methods (4DVar-classic in red, 4DVarNet-ode in blue and 4DVarNet-unet in green): (a) Atlantic Meridional Overturning Circulation (AMOC) temporal signal reconstructed by the 3 DA methods; (b) Attractor of the AMOC system; (c) Spectral analysis of the reconstructed AMOC signal. The true model trajectory is represented with a dotted black line.

Table 2
Values of the Two Hyperparameters, α_{obs} and α_{dyn} , Used for Each Data Assimilation Method

Data assimilation method	α_{obs}	α_{dyn}
4DVar-classic	1.00	1.00×10^7
4DVarNet-ode	0.41	0.12
4DVarNet-unet	0.29	0.18

Note. Bold values indicate that it is the best performance.

Indeed, observation uncertainties are applied by a diagonal covariance matrix of large amplitude in front of the observation term.

In Figure 5, we plot the minimization paths of the three DA schemes. This plot corresponds to the evolution of the retrieved state in the plane defined by the global NMSE and variational cost with respect to the iterations. As expected, 4DVar-classic iteratively minimizes the variational cost ((4)), which also leads to reducing the NMSE. We can also notice the relatively slow convergence of the fixed-step gradient descent implemented by our 4DVar-classic scheme. After the 10th iteration, 4DVar-classic has barely improved the NMSE and the variational cost remains high. The two 4DVarNet schemes

exhibit different patterns. The first 4 to 5 iterations involve a rapid drop of both the NMSE and variational cost ((4)), reaching a value for the latter in the same range as the 4DVar-classic scheme after convergence. But, the following iterations show a different behavior: the NMSE keeps on decreasing, while the variational cost ((4)) starts increasing. Noticeably, 4DVarNet schemes significantly outperform the 4DVar-classic approach here in terms of NMSE. We may emphasize that the training phase optimizes 4DVarNet schemes with respect to a learning cost involving a NMSE loss ((6)). Overall, this figure suggests that variational cost ((4)) may not provide here a good proxy of the NMSE reconstruction score, such that the 4DVar-classic scheme cannot retrieve a reconstruction as good as the 4DVarNet schemes do. This may also be further illustrated by the comparison of 4DVarNet-unet and 4DVarNet-ode schemes. Whereas 4DVarNet-unet delivers a more accurate reconstruction of the AMOC variability, as reflected by a lower NMSE, 4DVarNet-ode reaches lower values in the variational cost. By construction, the U-Net cell brings an additional flexibility by replacing the physical model used for the assimilation. While it seems that this data-driven model departs from the physics-constrained variational cost, it allows to improve the reconstruction performance without necessarily taking the same optimization path.

4.3. Influence of the Density of Observations

To evaluate the performance of our three DA methods with respect to the observation budget, we computed the error of reconstruction according to the sampling frequencies of observations, going from $\frac{1}{10} \text{ yr}^{-1}$ to $\frac{1}{120} \text{ yr}^{-1}$. This sampling frequency could be compared to the highest frequency of the model $\sim \frac{1}{250} \text{ yr}^{-1}$ and the fastest error growth timescale of ~ 5 years (Sévellec & Fedorov, 2014). As expected, the NMSE increases overall with the sampling period (see Figure 6). Specifically, the 4DVar classic method shows a rapid decline in the accuracy of the reconstruction of x_2 and x_3 when the data become sparse, with an error above 10% once the sampling period surpasses 40 years. We can infer that observations might be lacking to sufficiently constrain this assimilation

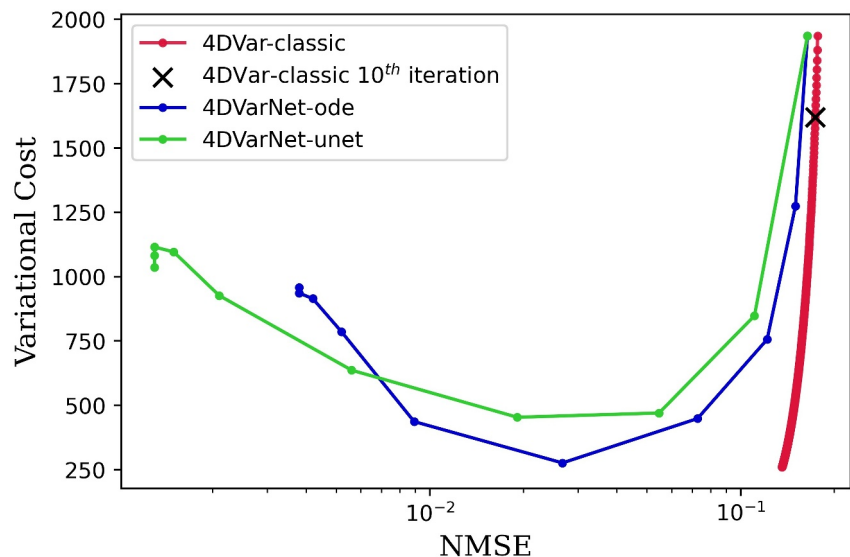


Figure 5. Variational cost (J) evolution according to the Normalized Mean Square Error using 4DVar -classic, -ode, and -net assimilation method, respectively. Each dot represents an iteration of the data assimilation process.

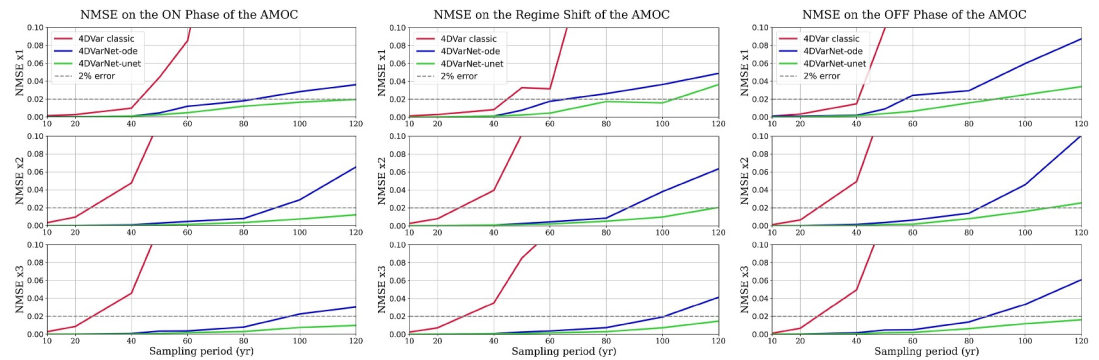


Figure 6. Normalized Mean Square Error evolution for each of the 3 variables (corresponding to each row) according to the sampling period for the 3 data-assimilation methods (corresponding to the 3 color lines) computed on: the ON Phase, the Regime Shift, and the OFF Phase of the Atlantic Meridional Overturning Circulation, respectively.

problem. On the contrary, we observe the excellent performance of the 4DVarNet methods. Even when the sampling period extends beyond 100 years, 4DVarNet consistently yields sufficiently good results, with reconstruction errors below 2% for the 4DVarNet-unet until the minimum tested sampling frequency of $\frac{1}{120} \text{ yr}^{-1}$. Thus, 4DVarNet-unet appears as a promising scheme to provide a high-quality reconstruction of a chaotic model with fewer observations, which is especially sought after in oceanography since observations are often sparse in time and space (Munk, 2000).

We also focus on the ability to reconstruct each phase of the AMOC system. We notice here that DA methods perform better at reconstructing the ON Phase than the OFF Phase. The OFF Phase is characterized by abrupt extreme values and happens on a shorter time span than the ON Phase. As a result, the assimilation of observations during OFF Phases might be more challenging than during the ON Phase. However, the comparison between the 4DVarNet-ode and the 4DVarNet-unet methods also proves the robustness of 4DVarNet-unet in a situation of poor observations. The 4DVarNet-ode method relies on the defined physical system, known to be chaotic and for which the OFF Phase is a deviation from the “normal” state of the ON Phase. While the physical system introduces uncertainties in the modeling of the phase changes of the AMOC, the U-Net have a more stable representation of the system, leading to a better reconstruction, even when observations are sparse and the dynamics unstable. Finally, it is worth noticing that the regime shift is reconstructed with satisfactory results, while it is a transition between the ON and the OFF Phases. By having trustworthy results on the reconstruction of regime shifts, we can consider that a 4DVarNet method could help in the forecast of extreme events happening during the OFF Phases.

4.4. Influence of Sampling Strategies

In this subsection, we test the performance of the 4DVarNet-unet method when the system is only sparsely observed. Here, we investigate how the pattern of observation affects the reconstruction of the DO events, when observations are assimilated with a deep learning approach. According to the previous section, we can choose a sampling period of $T_s = 100$ years and still expect a reconstruction error below 2% with the 4DVarNet-unet method. In this section, we implement four sampling strategies (illustrated in the second column of Table 3):

Table 3
Average Normalized Mean Square Error per Phase of the Atlantic Meridional Overturning Circulation Variability With Each Sampling Strategy

Sampling strategy	NMSE ON Phase	NMSE Regime Shift	NMSE OFF Phase
Regular	3.19×10^{-3}	1.05×10^{-2}	5.40×10^{-2}
Regular cluster	8.47×10^{-4}	2.26×10^{-3}	8.47×10^{-3}
Random	1.10×10^{-2}	6.64×10^{-2}	2.93×10^{-1}
Random cluster	8.07×10^{-3}	4.51×10^{-2}	2.02×10^{-1}

Note. Bold values indicate that it is the best performance.

1. **Regular sampling strategy:** We observe the system with a regular frequency, as in the previous sections. We consider $\mathbf{H}_t = \mathbf{I}_3 \quad \forall t \in [1 : 100 : \Delta T]$;
2. **Regular cluster sampling strategy:** We build clusters composed of three observations acquired with a finer consecutive time steps ($dt = 10$ years), and observed every T_s . Here, $\mathbf{H}_{t-dt} = \mathbf{H}_t = \mathbf{H}_{t+dt} = \mathbf{I}_3 \quad \forall t \in [1 : 100 : \Delta T]$;
3. **Random sampling strategy:** We observe the system at times randomly chosen. The budget of observations is set according to the number of observations acquired by following the regular sampling strategy. We thus consider $\mathbf{H}_t = \mathbf{I}_3$ for t randomly chosen in the ΔT -time window;
4. **Random cluster sampling strategy:** We build clusters of three observations acquired with finer consecutive time steps ($dt = 10$ years), separated by time intervals, which duration varies randomly. Number of clusters is fixed according to the number of clusters acquired by following the regular cluster sampling strategy. Thus, we consider $\mathbf{H}_{t-dt} = \mathbf{H}_t = \mathbf{H}_{t+dt} = \mathbf{I}_3$ for t randomly chosen in the ΔT -time window.

Performances of these sampling strategies for the reconstruction of the AMOC variability are both reported in the third column of Figure 7 and Table 3. Cluster sampling strategies (2 and 4) result in a higher observation budget compared to the two other sampling strategies (1 and 3). These sampling strategies were inspired by possible operational contexts, where deploying acquisition resources is expensive. To make their deployment profitable, we can intend leaving them in place for slightly longer periods of time. Indeed, as detailed in Rayner et al. (2011), the RAPID program deployed an array left in place at 26.5°N for now almost two decades, which enabled to monitor the AMOC continuously and therefore more accurately. This example encourages us to consider the clustering strategies (2 and 4) as more practical scenarios.

We find that random strategies lead to an NMSE that is, on average, 10 times higher than that of regular strategies. Even if this method scan a larger range of states, it does not statistically lead to a better reconstruction. We notice that random strategies fail to capture the pseudo-periodicities of the AMOC system, which undermines their trustworthiness. While incorporating additional trajectories into the learning process might enhance the performance of random strategies, we favor methods requiring less training data for computational efficiency, but also considering that the ocean is only sparsely observed.

The random cluster sampling strategy does not improve the reconstruction of the system compare to the random sampling strategy, whereas it multiplies by 3 the number of observations. On the other hand, if we multiply the number of observations by 3 in the framework of a regular sampling strategy, we reach an NMSE with an order of magnitude less than the NMSE reached by a regular cluster sampling strategy. We deduce that more observations do not necessarily lead to better reconstruction results if observations are inefficiently acquired. It is especially true when comparing the results of the regular sampling strategy in regards of the random cluster sampling strategy. The random cluster sampling strategy has 3 times as much observations as the regular sampling strategy, but produces results with an NMSE twice as high on the ON Phase and almost 4 times as high on the OFF Phase as the NMSE when regular sampling strategy is used.

Since, in an operational context, it might be more practical in some cases to acquire a group of observations on a longer time window, we evaluated the potential benefit of clustering observations rather than increasing the sampling frequency. The random cluster sampling strategy decreases by 26% the error of reconstruction on the ON Phase, and by 31% on the OFF Phase compared to the random sampling strategy. The impact of clustering appears much more efficient on regular sampling strategies. Indeed, the regular cluster sampling strategy reduces the NMSE by 73% on the reconstruction of the ON Phase and by 84% on the reconstruction of the OFF Phase. We can infer that clustering should especially be sought after for the reconstruction—and eventually forecasting—of extreme events difficult to monitor, with a reconstruction error below 1%. We can assume that the regular acquisition of three consecutive samples helps the detection of short-term variations in the system, especially during periods of regime shifts. Consequently, the regular acquisition of observations over a sufficiently long period may facilitate the monitoring of regime shifts such as those induced by DO events, particularly in cases where data are sparse due to low sampling frequency. In Keller (2007) Deutsch, Hall, and Bradford advocated for more frequent observations to enable the detection of early changes in the AMOC, but showed with his co-authors that such an observation system would increase observation costs by a few orders of magnitude. Here, we illustrate that with clustering, we do not necessarily need to increase the acquisition frequency to reconstruct accurately an idealized AMOC dynamics, but only to increase the acquisition time window. From a practical point of view, this can be achieved by installing moorings, such as those that form part of the RAPID network (Rayner

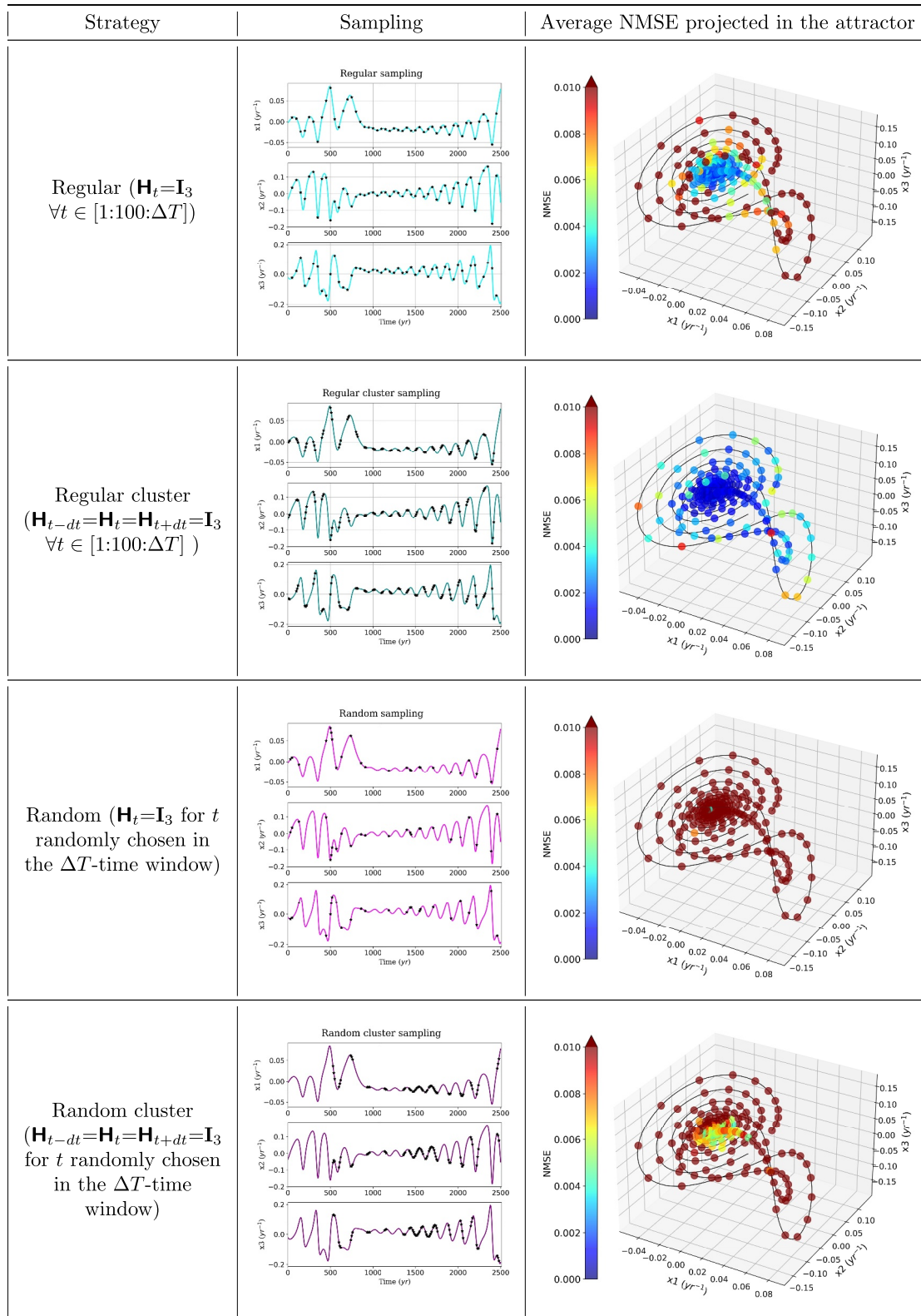


Figure 7. Influence of sampling strategies on the reconstruction of the Atlantic Meridional Overturning Circulation dynamics. The two first columns refers to the sampling strategies at play. The third column displays the Normalized Mean Square Error of the reconstructed system projected on the attractor. Each row corresponds to one sampling strategy.

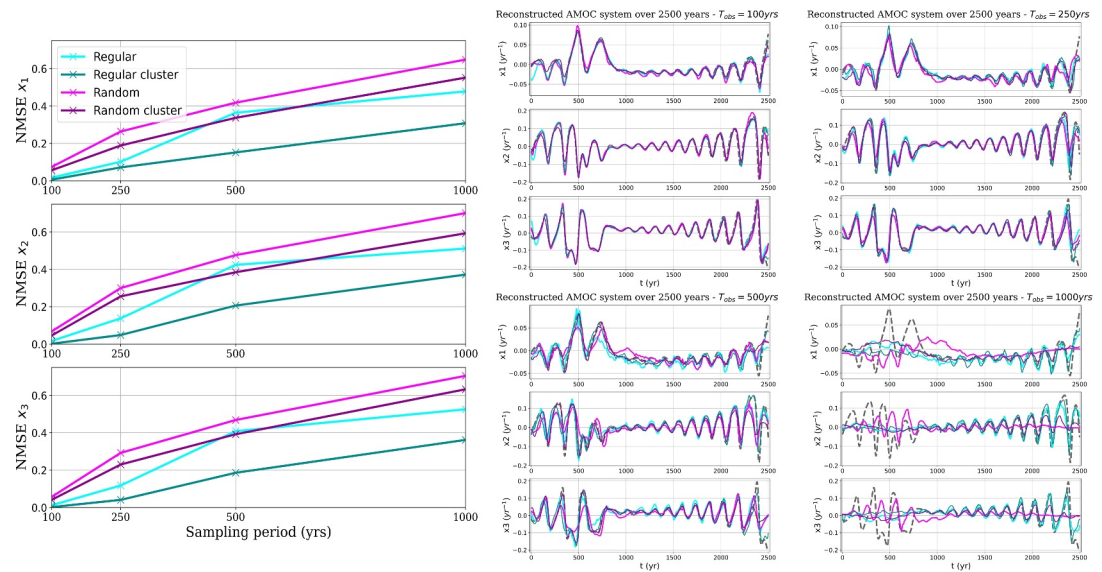


Figure 8. Left panel: Normalized Mean Square Error computed for each variable of the system reconstructed by 4 different sampling strategies on 4 sampling periods: [100 years, 250 years, 500 years, 1,000 years]. Right panel: Reconstruction of the Atlantic Meridional Overturning Circulation variability performed at the 4 sampling periods with the 4 sampling strategies using a 4DVarNet-unet model. The gray dotted line refers to the true trajectory.

et al., 2011). Eventually, improving the reconstruction of regime shifts could improve the reconstruction of extremes by informing scientists on the premises of an OFF Phase, leading to an abrupt climate change, consistently with Early Warning Signal principle (Lenton, 2011). The observation of the phenomena could then be reinforced for the good monitoring of the AMOC at a moment of high instability, and hence help to manage logistical resources in a more sustainable way.

5. Dealing With More Complex DA Settings

The previous Section clearly highlights the potential of neural DA schemes to better exploit sparsely-sampled observations to reconstruct the AMOC. The considered experimental setting remains idealized especially because no model error was included and all state components are observed. This section explores how 4DVarNet-unet could address more challenging and more realistic configurations: namely, extremely sparse observations beyond the Shannon criterion (5.1), partial observations (5.2), imperfectly-known physics (5.3).

5.1. A Very Sparse Observation Case

We wonder how 4DVarNet-unet performs in a case when the observation period does not respect the Shannon criterion of the smaller variability of the AMOC: the centennial oscillations. Therefore, we conduct a new set of experiments by pushing back the sampling period from 100 to 1,000 years following the same experimental setup as in Section 4.

Figure 8 illustrates the error growth for 4 sampling periods: 100, 250, 500 and 1,000 years. As expected, the overall NMSE increases with the sampling period, but at a different rate from one sampling strategy to another. While we emphasized the gain of regular sampling strategies over random sampling strategies in Section 4.4, this benefit becomes even more significant, especially for the regular cluster sampling strategy. We note that the regular strategy yields satisfactory reconstruction results with a sampling period of 250 years but drops off afterward, performing even worse than the random cluster strategy at a sampling period of 500 years. At a sampling period of 1,000 years, the Shannon criterion is no longer met for the D-O events. Even if the NMSE is too high to be confident about the reconstruction provided by 4DVarNet-unet at the 500- and 1,000-year sampling periods, we can still detect the tendencies of the AMOC dynamics with a more qualitative analysis.

It is particularly true for a sampling period of 500 years, where both the DO and the centennial variability are reconstructed, but with less accuracy on the magnitude of the signals. We also note the appearance of numerical

Table 4

Normalized Mean Square Error ($\times 10^{-3}$) of the Reconstructed Signal for Each Variable With Each Data Assimilation Method on a Partially Observed System (Only x_2 and x_3 are Observed)

Method	NMSE x_1	NMSE x_2	NMSE x_3
4DVar-classic	19.47	163.28	162.01
4DVarNet-ode	2.85	12.38	9.39
4DVarNet-unet	0.91	2.58	2.71

Note. Bold values indicate that it is the best performance.

noise, displaying a certain lack of stability from 4DVarNet-unet. At a sampling period of 1,000 years, no sampling strategy is able to capture the entire variability on the considered time window. A phase shift is induced by the lack of observations, and 4DVarNet-unet only manages to realign itself from halfway through the time window, surely after the arrival of an observation. The only exception is the random sampling strategy, which does not manage at all to grasp the AMOC dynamics over the time window. It confirms that both the moment of acquisition and the consecutivity of measurements play an important role in the ability to monitor the dynamics.

5.2. A Partially Observed System

In most ocean DA systems, it is usually unlikely to gather observations on each variable of the system. To test out the robustness of 4DVarNet schemes

when dealing with a partially observed system, we run experiments similar to those reported in Section 4, where we disregard any observation for variable x_1 . For the considered AMOC dynamics, variables x_2 and x_3 refer to salinity gradients, which are realistically observable physical quantities. By contrast, variable x_1 relates to the varying component of the AMOC intensity, which cannot be as easily observed as x_2 and x_3 . In these experiments, we consider a 50-year regular sampling pattern for variables x_2 and x_3 . As reported in Table 4, the reconstruction performance is very similar to the one obtained when the system is fully observed (cf. Table 1). Even if we note a growth of 33% in the average NMSE, this configuration is still better than the performance of 4dvarnet-ode and 4dvar-classic when observing all variables with the same sampling pattern. This result supports the ability of 4DVarNet schemes to inform the full state from partial observations only.

A more qualitative comparison can also be found on Figure 9. We notice that the trajectories reconstructed by the 4DVar-classic method are less accurate in a case of partial observations, especially on the reconstruction of the regime shifts and of the AMOC OFF phases. This conclusion is supported by Figure 10, which also highlights that 4DVarNet methods struggle at reconstructing properly the variability of the AMOC on x_1 , with an NMSE exceeding 10% when the sampling period exceeds 100 years. Conversely, when the system was fully observed, the NMSE on x_1 stayed below 10% for each sampling period tested out. Hence, when x_1 is not observed, the loss in performance can be compensated by observations on x_2 and x_3 acquired at a higher frequency.

Finally, we have a look at the power spectrum of the reconstructed oceanic signals (Figure 9). Similar results to the reconstruction of the fully-observed system are obtained. While the frequency of DO events is well reconstructed by every assimilation methods, the power magnitude of x_1 has slightly been altered by the 4DVarNet methods, whereas 4DVar-classic manages to reconstruct the spectrum almost perfectly. It is however the opposite on the power spectrum of x_2 . It prevents us from deriving any significant conclusion from these results, outside the fact that the provided reconstructions do respect the physical frequencies of our system, even when it is partially observed.

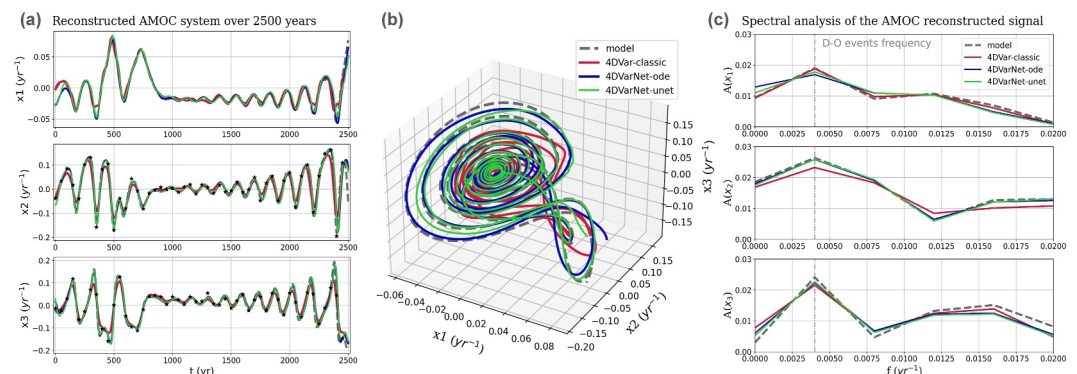


Figure 9. Qualitative comparison of the three Data assimilation (DA) methods on a partially observed system (4DVar-classic in red, 4DVarNet-ode in blue and 4DVarNet-unet in green): (a) Atlantic Meridional Overturning Circulation (AMOC) temporal signal reconstructed by the 3 DA methods; (b) Attractor of the AMOC system; (c) Spectral analysis of the reconstructed AMOC signal. The true model trajectory is plotted with a dotted black line.

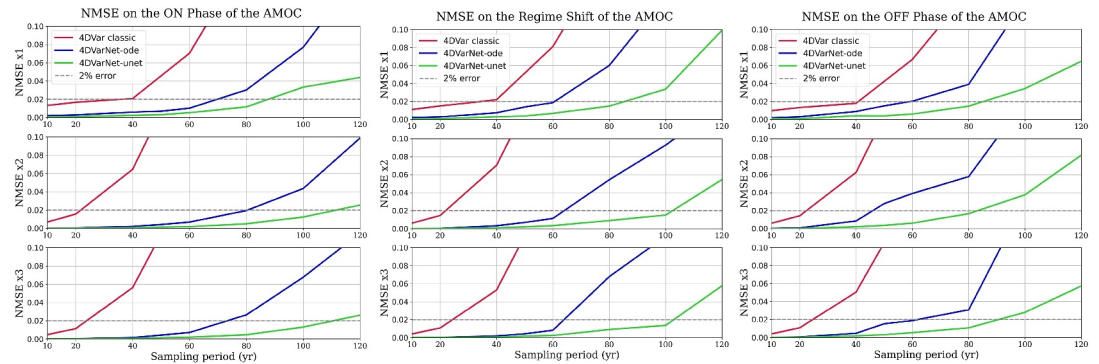


Figure 10. Normalized Mean Square Error evolution for each of the 3 variables (corresponding to each row) according to the sampling period for the 3 data-assimilation methods (corresponding to the 3 color lines) on a partially observed system computed on: (a–c) The ON Phase, the Regime Shift, and the OFF Phase of the Atlantic Meridional Overturning Circulation, respectively.

5.3. Imperfectly-Known Physics

In Section 4, the underlying model ruling the targeted AMOC system is identical to the physical model used as a prior in the assimilation methods, either explicitly through variational cost (4) or implicitly through training and testing phases on data sets issued from the same model configuration. This configuration is often referred to as perfect model approach. In real-world situations, the underlying physics cannot be perfectly known or represented due to unresolved processes (McWilliams, 2007), uncertainties in the forcings (Fox-Kemper et al., 2019), imperfectly-known parametrizations (Le Sommer et al., 2011), etc. Here, we test out this configuration, often referred to as imperfect model approach.

To account for such configurations, we run new experiments where the physics in the training data set differ from the physics of the observed system to be reconstructed. We draw inspiration from Le Bras et al. (2024) and consider nine different model parametrizations (1). All these parametrizations lead to chaotic dynamics and reflect some trade-off between three parameters: the buoyancy torque coefficient ϵ , the steady component of the AMOC intensity Ω_0 and the linear damping coefficient K . In Table 5, we display the parameters' values for each configuration, model 0 being the parametrization considered in Section 4. We consider the 4DVarNet-unet scheme trained on the data set generated for this parametrization for regularly-sampled observations at a frequency of $\frac{1}{50}$ yrs⁻¹ (see Section 4 for the baseline performance of this model). We test this model for ground truth and observation data sets generated from the other nine parametrizations and follow the same evaluation scheme as in Section 4.

To assess the diversity of the 10 different parametrizations, we report in Figure 11 the resulting marginal distributions of a centered and standardized version of variable x_1 . Whereas we can observe a bimodal distribution for parametrization 0, corresponding to the on and off phases of the AMOC, some parametrizations (e.g.,

Table 5
Performances of 4DVarNet-Unet in Face of Different Physical Models, Introducing Model Errors

Simulated model	0	1	2	3	4	5	6	7	8	9
ϵ	0.350	0.350	0.350	0.350	0.175	0.175	0.175	0.175	0.175	0.175
Ω_0 ($\times 10^{-2}$)	-2.50	-1.25	-1.25	-5.00	-2.50	-2.50	-1.25	-5.00	-5.00	-5.00
K ($\times 10^{-4}$)	1	1	2	0.5	1	2	0.5	1	0.5	2
Similarity Index	1.00	0.72	0.70	0.63	0.61	0.62	0.64	0.52	0.54	0.58
NMSE ($\times 10^{-2}$)	0.16	1.90	2.10	48.31	1.44	1.40	2.09	17.61	17.37	16.43

Note. First lines of the table defines each parameterization associated with each model tested out. Model 0 corresponds to the baseline model. Last lines of the table display the Similarity Index of each model compared to the baseline model and the global NMSE computed on the 3 variables of the system.

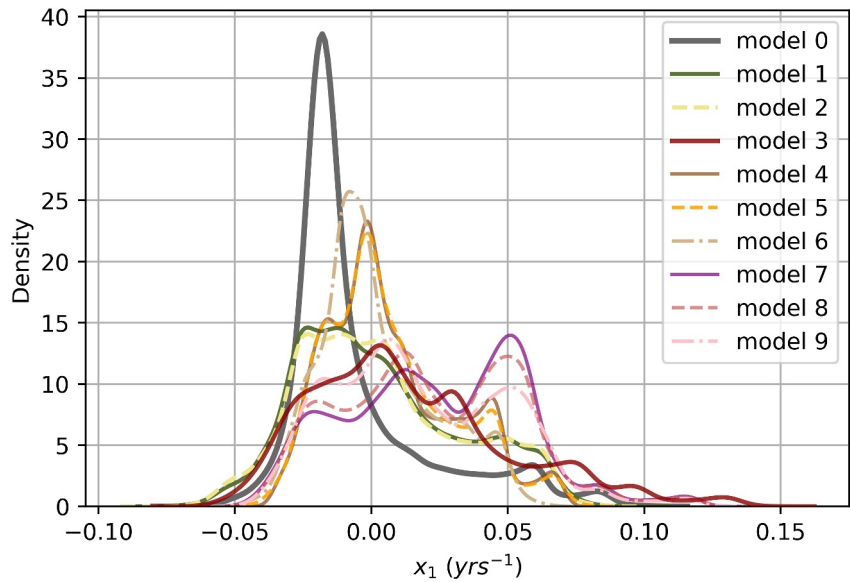


Figure 11. Probability Density Functions of x_1 generated according to the 10 different physical models. Model 0 is the baseline model on which 4DVarNet-unet was trained. Distributions are centered and standardized.

parametrizations 4, 5 and 6) also result in bimodal distribution but depict a more balanced occurrence of the two exhibited phases. By contrast, other parametrizations 1 and 2 rather lead to a unimodal distribution. With a view to evaluating how far a given parametrization is from the baseline model, we introduce the following similarity index (SI):

$$\forall i \in [1, 9], \text{ Similarity Index} = \int_{\min x_1}^{\max x_1} \min(f^0(x_1^0), f^i(x_1^0)) dx_1 \quad (8)$$

where i refers to the index of the parametrization, f^i to the normalized probability density function of variable x_1 for parametrization i . This SI measures the overlap of the normalized probability density function for a given parametrization i and the baseline one ($i = 0$).

Overall, 4DVarNet-unet performs quite well at reconstructing the new observed systems, with a global NMSE around or even below 2% in the best cases. As expected, we notice that the higher the SI, the lower the error of reconstruction. Even when the new model only half resembles the baseline model, the overall error remains below 20% in cases 7, 8 and 9. The only exception lies with the model 3, which is poorly reconstructed despite a rather high SI. Having a look at the distributions on Figure 11, the dynamics of the model 3 appears more widespread and disparate than the other models, which could explain the lower accuracy on its reconstruction. It is especially true when we compare it to the models 4, 5 and 6, which have a close SI to model 3 but benefit from a really good reconstruction of their dynamics. This might be explained by the shapes of their distributions which are very similar to the one of the baseline model, helping out the generalization on these 3 models.

Consequently, we showed the good generalization properties of 4DVarNet-unet on the reconstruction of 9 other various physical systems. This result is encouraging, especially when one of the usual drawback of neural networks is the generalization capacity to other systems. One way to improve the generalization capacity would be to directly train 4DVarNet-unet on a data set uniting trajectories from these 10 models. Nonetheless, even without doing so, we see here that 4DVarNet-unet appears quite robust in face of these different dynamics, with only one exception.

6. Conclusion

The complexity of inverse problems in geophysics is driving the search for new computational methods. Data assimilation has already proved its ability to help solve such problems by optimizing the compromise between

observations and physical knowledge. However, there are still challenges to be met to improve climate monitoring. To achieve finer resolutions and work with low observation rates, the computation efficiency of machine learning opens up new possibilities for processing data, enhancing their full exploitation.

Here, we combined the advantages of DA and the benefits of deep learning to solve an inverse problem of a chaotic non-linear model representing the evolution of the AMOC during the last glacial interval. Two different methods were implemented and compared to a 4D Variational DA method, considered as the state-of-the-art in this study. On the one hand, 4DVarNet-ode is a physics-informed neural network where the physical model is the prior on which the DA problem is relying. On the other hand, 4DVarNet-unet is a fully data-driven method where the prior is encoded by a U-Net cell jointly trained with the rest of the neural network. The power of 4DVarNet lies in optimizing the DA problem by training of a convolutive LSTM neural network, which boosts the gradient descent usually computed to minimize the variational cost. As a result, 4DVarNet performed better than the 4DVar benchmark method, with a reconstruction error divided by 10, and even by 100 for the reconstruction of the varying AMOC intensity (x_1).

We also noticed that 4DVarNet-unet was better than 4DVarNet-ode at reconstructing the AMOC signal when observations are sparse. Despite its fully data-driven training, 4DVarNet-unet is able to capture the physical properties of the AMOC system, including the frequencies and amplitudes characteristic of DO events. This result questions the exploitation of prior knowledge, given as a dynamical model, in DA formulations addressing sparse observations.

This dynamical model puts into equation the forces at play, and more specifically the buoyancy gradient impact on the circulation and the advection of salinity. When only a few noisy observations are available, we can expect that the reconstruction relies further on the specified model, either the true physical system for 4DVarNet-ode or the U-Net cell for 4DVarNet-unet. We may therefore suggest that the U-Net learned a better representation of the sparsely, erroneous observed AMOC variability than the perfect-physical system itself, which goes in line with the recent advances in model-discovery (Zanna & Bolton, 2020). Thus, neural networks provide new means to study processes, for which imperfect or no physical model has been established yet because of their inherent complexity, or even to improve established perfect-models, which have shown their limits of representation in this study. Variational DA is a method historically developed in meteorology, field in which observations are dense and retrieved daily. Even if it also applies as the state-of-the-art in operational oceanography, our study suggests that end-to-end neural approaches could lead to monitoring and forecasting breakthroughs given the scarcity of the available observations for oceanic processes (Cumplings & Smedstad, 2013; Fairbairn et al., 2014; Yaremchuk & Martin, 2014).

The idealized modeling and observation framework considered in this study provides an initial testbed to assess the potential benefits of deep learning for paleo-oceanography. Indeed, 4DVarNet-unet has proved to be robust in the face of challenging observation conditions, such as enhanced data scarcity, only partial observation of the system and situations where the underlying physics is imperfectly known. We tested out the generalization properties of 4DVarNet-unet on 9 different dynamic configurations and showed that it was able to reconstruct them quite accurately. It is encouraging for applying this method up to more complex ocean models. Additionally, we uncovered that 4DVarNet-unet was able to reconstruct satisfactorily the system with sampling frequencies far beyond the Shannon criteria. The full exploitation of the information of acquired data by neural networks might help to push back the limits of sampling theory.

With a sampling period of 100 years, we tested out four different sampling strategies to further improve the reconstruction of these climate extremes. Here, regular sampling methods were more efficient than random sampling methods, since their regularity of observations is suited to reconstruct a signal which evolves according to 2 pseudo-periodicities (one being harmonics). The best reconstruction was reached with a regular cluster sampling strategy, which acquired 3 observations every 100 years on a time window of 30 years. This method captures the variations of the system more accurately, especially during the regime shifts of the AMOC. Admittedly, we could expect that multiplying the number of observations by 3 would lead to better results. However, the random cluster sampling strategy leads to poorer results, with one order of difference on the NMSE compared to the regular cluster sampling strategy, despite the same amount of observations. We conclude that the moment of acquisition matters in climate monitoring, sometimes even more than the amount of observations. Here, the regular cluster sampling strategy can respond to the practical needs of a measuring campaign in

geographical areas that are difficult to reach, like the Northern part of the subpolar North Atlantic for the AMOC monitoring.

For applications to real paleo-oceanic records, future work may better explore the specific features of these data, which generally inform average quantities over decadal to centennial time scales, rather than instantaneous measurements, forming “proxy” records of past oceanic states (Hesse et al., 2011; Kessler et al., 2020; Menviel et al., 2017). As end-to-end neural DA schemes can learn from data the prior on the dynamics, it will provide new means to explore the relevant state-space formulations for such paleo-oceanic studies.

Optimizing the choice of the position and time of acquisition of an observation may certainly have a high impact on our knowledge of a physical phenomenon. As a consequence, the design of sampling strategy has to be adapted according to multiple factors: the region of study, the spatiotemporal scales of the targeted oceanic processes, the sensors and platforms of observations, the cost of acquisition, etc. Adaptive sampling has already proved its benefits to improve sampling strategies in oceanography, by adapting on-line the trajectory of a survey to gather the most useful observations given logistical and physical constraints. Here, the differentiability of end-to-end neural DA solvers could also be of key interest to explore adaptive observation operators stated as neural observation operators and result in the improved monitoring of chaotic non-linear oceanic processes such as the AMOC system considered in this study (Greenhill et al., 2020; Lermusiaux, 2007).

In addition to the sampling scheme, we should also take into account the considered physical variables, as the ocean state is rarely fully observed. Fortunately, 4DVarNet-unet provided reconstructions with an error below 10% in experiments conducted on a partially-observed system, which appeared sufficient to capture the Regime Shifts of the AMOC in our study. Because of global warming, several studies worked recently on the detection of the Early Warning Signals of the increasingly plausible hypothesis of an AMOC collapse, considered as one climate tipping point (Ditlevsen & Ditlevsen, 2023; Jackson & Wood, 2020; Michel et al., 2022). A sudden shutdown of the AMOC would disrupt the climate, particularly in Europe and in the Amazon forest, with an abrupt drop in temperatures or a reversal of the rainfall pattern, as described in (van Westen et al., 2024). An AMOC shutdown would not only require a major adaptation effort from the local populations, but could also lead to cascading tipping points, having an impact on the global climate. Ultimately, a comprehensive and adequate observation network coupled with deep learning techniques could help to anticipate such major climate changes.

Data Availability Statement

The authors provide the source code to generate all the datasets with the various sampling strategies, and to conduct a classical DA and a 4DVarNet DA, both with our physical prior or a neural network. Experiments used in the manuscript are detailed in the code which can be found on the following GitHub public repository: <https://github.com/PerrineBauchot/AMOC-4DVarNet>.

Acknowledgments

This work was supported by CNES (French Space Agency), OceaniX (ANR-19-CHIA-0016) and EU Horizon Europe project EDITO Model Lab (Grant 101 093293). It benefited from HPC and GPU resources from GENCI-IDRIS (Grant 2021-101030) and CPER AIDA GPU cluster supported by The Regional Council of Brittany and FEDER. Florian Sévellec is supported by the ARVOR project funded by the LEFE IMAGO program. We thank the reviewers for their recommendations, comments and corrections, which have enabled us to improve the quality of this study.

References

- Arcucci, R., Zhu, J., Hu, S., & Guo, Y.-K. (2021). Deep data assimilation: Integrating deep learning with data assimilation. *Applied Sciences*, 11(3), 1114. <https://doi.org/10.3390/app11031114>
- Berg, J., & Nyström, K. (2019). Data-driven discovery of PDEs in complex datasets. *Journal of Computational Physics*, 384, 239–252. <https://doi.org/10.1016/j.jcp.2019.01.036>
- Boudier, P., Fillion, A., Gratton, S., Gürol, S., & Zhang, S. (2023). Data assimilation networks. *Journal of Advances in Modeling Earth Systems*, 15(4). <https://doi.org/10.1029/2022MS003353>
- Brajard, J., Carassi, A., Bocquet, M., & Bertino, L. (2020). Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: A case study with the Lorenz 96 model. *Journal of Computational Science*, 44, 101171. <https://doi.org/10.1016/j.jocs.2020.101171>
- Buckley, M. W., & Marshall, J. (2016). Observations, inferences, and mechanisms of the atlantic meridional overturning circulation: A review. *Reviews of Geophysics*, 54(1), 5–63. <https://doi.org/10.1002/2015RG000493>
- Carrasi, A., Bocquet, M., Bertino, L., & Evensen, G. (2017). Data Assimilation in the Geosciences - an overview on methods, issues and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5). <https://doi.org/10.1002/wcc.535>
- Cessi, P., & Young, W. R. (1992). Multiple equilibria in two-dimensional thermohaline circulation. *Journal of Fluid Mechanics*, 241, 291–309. <https://doi.org/10.1017/S00222112092002040>
- Cummings, J. A., & Smedstad, O. M. (2013). Variational data assimilation for the global ocean. In S. K. Park & L. Xu (Eds.), *Data assimilation for atmospheric, oceanic and hydrologic applications* (Vol. II, pp. 303–343). Springer. https://doi.org/10.1007/978-3-642-35088-7_13
- Dabrowski, J. J., Pagendam, D. E., Hilton, J., Sanderson, C., MacKinlay, D., Huston, C., et al. (2023). Bayesian physics informed neural networks for data assimilation and spatio-temporal modelling of wildfires. *Spatial Statistics*, 55, 100746. <https://doi.org/10.1016/j.spasta.2023.100746>
- Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., et al. (1993). Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, 364(6434), 218–220. <https://doi.org/10.1038/364218a0>

- Ditlevsen, P., & Ditlevsen, S. (2023). Warning of a forthcoming collapse of the Atlantic meridional overturning circulation. *Nature Communications*, 14(1), 4254. <https://doi.org/10.1038/s41467-023-39810-w>
- Evensen, G. (2009). *Data assimilation*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-03711-5>
- Fablet, R., Chapron, B., Drumetz, L., Mémén, E., Pannekoucke, O., & Rousseau, F. (2021). Learning variational data assimilation models and solvers. *Journal of Advances in Modeling Earth Systems*, 13(10). <https://doi.org/10.1029/2021MS002572>
- Fablet, R., Chapron, B., Le Sommer, J., & Sévellec, F. (2024). Inversion of sea surface currents from satellite-derived sst-ssh synergies with 4dvarnets. *Journal of Advances in Modeling Earth Systems*, 16(6). <https://doi.org/10.1029/2023MS003609>
- Fairbairn, D., Pring, S. R., Lorenc, A. C., & Roulstone, I. (2014). A comparison of 4DVar with ensemble data assimilation methods. *Quarterly Journal of the Royal Meteorological Society*, 140(678), 281–294. <https://doi.org/10.1002/qj.2135>
- Fang, M., & Li, X. (2016). Paleoclimate data assimilation: Its motivation, progress and prospects. *Science China Earth Sciences*, 59(9), 1817–1826. <https://doi.org/10.1007/s11430-015-5432-6>
- Febvre, Q., Fablet, R., Sommer, J. L., & Ubelmann, C. (2022). Joint calibration and mapping of satellite altimetry data using trainable variational models. In *Icassp 2022 - 2022 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 1536–1540). <https://doi.org/10.1109/ICASSP43922.2022.9746889>
- Febvre, Q., Ubelmann, C., Le Sommer, J., & Fablet, R. (2024). Scale-aware neural calibration for wide swath altimetry observations. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–8. <https://doi.org/10.1109/tgrs.2024.3363503>
- Fox-Kemper, B., Adcroft, A., Böning, C. W., Chassignet, E. P., Curchitser, E., Danabasoglu, G., et al. (2019). Challenges and prospects in ocean circulation models. *Frontiers in Marine Science*, 6. <https://doi.org/10.3389/fmars.2019.00065>
- Germe, A., Hirschi, J. J.-M., Blaker, A. T., & Sinha, B. (2022). Chaotic variability of the atlantic meridional overturning circulation at subannual time scales. *Journal of Physical Oceanography*, 52(5), 929–949. <https://doi.org/10.1175/JPO-D-21-0100.1>
- Greenhill, S., Rana, S., Gupta, S., Vellanki, P., & Venkatesh, S. (2020). Bayesian optimization for adaptive experimental design: A review. *IEEE Access*, 8, 13937–13948. <https://doi.org/10.1109/ACCESS.2020.2966228>
- Hesse, T., Butzin, M., Bickert, T., & Lohmann, G. (2011). A model-data comparison of $\delta^{13}C$ in the glacial Atlantic ocean. *Paleoceanography*, 26(3). <https://doi.org/10.1029/2010PA002085>
- Howard, L. N. (1971). Geophysical fluid dynamics summer school. *WHOI Internal Tech. Rep.*, 71–63, 102–105.
- Jackson, L. C., & Wood, R. A. (2020). Fingerprints for early detection of changes in the amoc. *Journal of Climate*, 33(16), 7027–7044. <https://doi.org/10.1175/JCLI-D-20-0034.1>
- Johnson, C., Nichols, N. K., & Hoskins, B. J. (2005). Very large inverse problems in atmosphere and ocean modelling. *International Journal for Numerical Methods in Fluids*, 47(8–9), 759–771. <https://doi.org/10.1002/fld.869>
- Keller, K., Deutsch, C., Hall, G. M., & Bradford, F. D. (2007). Detection of changes in the north atlantic meridional overturning circulation: Implications for the design of ocean observation systems. *Journal of Climate*, 20(2), 145–157. <https://doi.org/10.1175/JCLI3993.1>
- Kessler, A., Bouttes, N., Roche, D. M., Ninnemann, U. S., Galaasen, E. V., & Tjiputra, J. (2020). Atlantic meridional overturning circulation and $\delta^{13}C$ variability during the last interglacial. *Paleoceanography and Paleoclimatology*, 35(5), e2019PA003818. <https://doi.org/10.1029/2019PA003818>
- Kim, E., & Seto, M. L. (2022). An approach to choose observation systems to observe ocean phenomena. In *2022 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp. 306–313). <https://doi.org/10.1109/CCECE49351.2022.9918203>
- Kingma, D. P., & Ba, J. (2017). Adam: A method for stochastic optimization. *arXiv*. <https://doi.org/10.48550/arXiv.1412.6980>
- Le Bras, P., Sévellec, F., Tandeo, P., Ruiz, J., & Ailliot, P. (2024). Selecting and weighting dynamical models using data-driven approaches. *Nonlinear Processes in Geophysics*, 31(3), 303–317. <https://doi.org/10.5194/npg-31-303-2024>
- Lenton, T. M. (2011). Early warning of climate tipping points. *Nature Climate Change*, 1(4), 201–209. <https://doi.org/10.1038/nclimate1143>
- Lermusiaux, P. F. J. (2007). Adaptive modeling, adaptive data assimilation and adaptive sampling. *Physica D: Nonlinear Phenomena*, 230(1), 172–196. <https://doi.org/10.1016/j.physd.2007.02.014>
- Le Sommer, J., d'Ovidio, F., & Madec, G. (2011). Parameterization of subgrid stirring in eddy resolving ocean models. part 1: Theory and diagnostics. *Ocean Modelling*, 39(1), 154–169. <https://doi.org/10.1016/j.ocemod.2011.03.007>
- Lorenc, A. C., & Payne, T. (2007). 4D-Var and the butterfly effect: Statistical four-dimensional data assimilation for a wide range of scales. *Quarterly Journal of the Royal Meteorological Society*, 133(624), 607–614. <https://doi.org/10.1002/qj.36>
- Malkus, W. V. R. (1972). Non-periodic convection at high and low Prandtl number. *Mem. Societe Royale des Sciences de Liege*, 4, 125–128.
- McCarthy, G. D., Brown, P. J., Flagg, C. N., Goni, G., Houpt, L., Hughes, C. W., et al. (2020). Sustainable observations of the AMOC: Methodology and technology. *Reviews of Geophysics*, 58(1). <https://doi.org/10.1029/2019RG000654>
- McWilliams, J. C. (2007). Irreducible imprecision in atmospheric and oceanic simulations. *Proceedings of the National Academy of Sciences*, 104(21), 8709–8713. <https://doi.org/10.1073/pnas.0702971104>
- Menviel, L., Yu, J., Joos, F., Mouchet, A., Meissner, K. J., & England, M. H. (2017). Poorly ventilated deep ocean at the last glacial maximum inferred from carbon isotopes: A data-model comparison study. *Paleoceanography*, 32(1), 2–17. <https://doi.org/10.1002/2016PA003024>
- Michel, S. L. L., Swingedouw, D., Ortega, P., Gastineau, G., Mignot, J., McCarthy, G., & Khodri, M. (2022). Early warning signal for a tipping point suggested by a millennial atlantic multidecadal variability reconstruction. *Nature Communications*, 13(1), 5176. <https://doi.org/10.1038/s41467-022-32704-3>
- Munk, W. (2000). Chapter 1 Oceanography before, and after, the advent of satellites. In D. Halpern (Ed.), *Chapter 1 Oceanography before, and after, the advent of satellites* (Vol. 63, pp. 1–4). Elsevier Oceanography Series. [https://doi.org/10.1016/S0422-9894\(00\)80002-1](https://doi.org/10.1016/S0422-9894(00)80002-1)
- Okazaki, A., & Yoshimura, K. (2017). Development and evaluation of a system of proxy data assimilation for paleoclimate reconstruction. *Climate of the Past*, 13(4), 379–393. <https://doi.org/10.5194/cp-13-379-2017>
- Perkins, W., & Hakim, G. (2021). Coupled atmosphere–ocean reconstruction of the last millennium using online data assimilation. *Paleoceanography and Paleoclimatology*, 36(5), e2020PA003959. <https://doi.org/10.1029/2020pa003959>
- Peyron, M., Fillion, A., Gürol, S., Marchais, V., Gratton, S., Boudier, P., & Goret, G. (2021). Latent space data assimilation by using deep learning. *Quarterly Journal of the Royal Meteorological Society*, 147(740), 3759–3777. <https://doi.org/10.1002/qj.4153>
- Raissi, M., Perdikaris, P., & Karniadakis, G. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Rayner, D., Hirschi, J. J.-M., Kanzow, T., Johns, W. E., Wright, P. G., Frajka-Williams, E., et al. (2011). Monitoring the atlantic meridional overturning circulation. *Deep Sea Research Part II: Topical Studies in Oceanography*, 12(17–18), 1744–1753. <https://doi.org/10.1016/j.dsr2.2010.10.056>

- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical image computing and computer-assisted Intervention – MICCAI 2015* (pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28
- Rudy, S. H., Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2017). Data-driven discovery of partial differential equations. *Science Advances*, 3(4). <https://doi.org/10.1126/sciadv.1602614>
- Sévellec, F., & Fedorov, A. V. (2014). Millennial variability in an idealized ocean model: Predicting the AMOC regime shifts. *Journal of Climate*, 27(10), 0894–8755. <https://doi.org/10.1175/JCLI-D-13-00450.1>
- Sévellec, F., & Fedorov, A. V. (2015). Unstable AMOC during glacial intervals and millennial variability: The role of mean sea ice extent. *Earth and Planetary Science Letters*, 429, 60–68. <https://doi.org/10.1016/j.epsl.2015.07.022>
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15.
- Stommel, H. (1961). Thermohaline convection with two stable regimes of flow. *Tellus*, 13(2), 224–230. <https://doi.org/10.1111/j.2153-3490.1961.tb00079.x>
- Thual, O., & McWilliams, J. C. (1992). The catastrophe structure of thermohaline convection in a two-dimensional fluid model and a comparison with low-order box models. *Geophysical and Astrophysical Fluid Dynamics*, 64(1–4), 67–95. <https://doi.org/10.1080/03091929208228085>
- Vanschoren, J. (2018). Meta-learning: A survey. *arXiv*.
- van Westen, R. M., Kliphuis, M., & Dijkstra, H. A. (2024). Physics-based early warning signal shows that amoc is on tipping course. *Science Advances*, 10(6), eadk1189. <https://doi.org/10.1126/sciadv.adk1189>
- Von Storch, H., Cubasch, U., Gonzalez-Rouco, J., Jones, J., Voss, R., Widmann, M., & Zorita, E. (2000). *Combining paleoclimatic evidence and gcms by means of Data Assimilation Through Upscaling and Nudging (DATUN)* (Vol. 18, pp. 15–18). GKSS FORSCHUNGSZENTRUM GEESTHACHT GMBH-PUBLICATIONS-E.
- Yaremchuk, M., & Martin, P. (2014). On sensitivity analysis within the 4DVAR framework. *Monthly Weather Review*, 142(2), 774–787. <https://doi.org/10.1175/MWR-D-13-00123.1>
- Zanna, L., & Bolton, T. (2020). Data-driven equation discovery of ocean mesoscale closures. *Geophysical Research Letters*, 47(17), e2020GL088376. <https://doi.org/10.1029/2020GL088376>
- Zhou, Q., Qian, J., Tang, J., & Li, J. (2023). Deep unrolling networks with recurrent momentum acceleration for nonlinear inverse problems. <http://arxiv.org/abs/2307.16120arXiv:2307.16120>
- Zupanski, D. (1997). A general weak constraint applicable to operational 4DVAR data assimilation systems. *Monthly Weather Review*, 125(9), 2274–2292. [https://doi.org/10.1175/1520-0493\(1997\)125<2274:AGWCAT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<2274:AGWCAT>2.0.CO;2)