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

Perrine Bauchot [1], Angélique Drémeau [1], Florian Sévellec [2,3], Ronan Fablet [1,3]

[1] Lab-STICC, CNRS ENSTA Bretagne IMT Atlantique UBO ENIB UBS, Brest, France
[2] Laboratoire d'Océanographie Physique et Spatiale, Univ Brest CNRS IRD Ifremer, Brest, France
[3] ODYSSEY 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 AMOC abrupt slowdowns considered as climate extremes during the last glacial period.

Corresponding author: Perrine Bauchot, `perrine.bauchot@ensta-bretagne.org`

Abstract

Data assimilation (DA) reconstructs and forecasts the dynamics of geophysical processes based on available observations and on physical *a priori*. Recently, the hybridization of DA and deep learning has opened new perspectives to address model-data interactions. In this paper, we investigate its potential contribution to the analysis of a chaotic oceanic phenomenon: an idealized model representing the centennial to millennial variability of the North Atlantic ocean circulation during the last glacial period. The implemented neural approach – 4DVarNet – yields large relative improvements over a classical variational DA method on the reconstruction of the regime shifts of the Atlantic Meridional Overturning Circulation (AMOC). These gains are even more significant when the density of observations decreases. The results also exhibit that the explicit exploitation of the *a priori* dynamical model does not necessarily lead to the best performance compared to a data-driven model. Additionally, we compare four different sampling strategies to assess the impact of the observations on the capture of the unstable phases of the AMOC. We highlight the gain of regular over random sampling strategies, reaching an error of reconstruction below 2% with a sampling period of 100 years. The error on the reconstruction of regime shifts can even be divided by 5 when acquiring clusters of three consecutive observations, sometimes more suited in an operational framework. This study on an idealised, nonetheless 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 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. One important oceanic phenomena which regulates the climate by heat storage, for instance, is the Atlantic Meridional Overturning Circulation (AMOC). The AMOC drives the transport of warm surface water masses towards 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. Various studies have highlighted the chaotic nature of the AMOC dynamics (Buckley & Marshall, 2016; Germe et al., 2022). 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 1470±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, modelling, 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). 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). We can cast data assimilation schemes into two main categories: statistical data assimilation schemes, especially sequential Kalman approaches (Evensen, 2009), and variational data assimilation schemes (Cummings & Smedstad, 2013). Recently, neural data assimilation, which bridges deep learning and data assimilation, has attracted a greater attention with potential breakthroughs for the targeted inverse problems (Boudier et al., 2023; Fablet et al., 2021). 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 idealised 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 advents in oceanography can only be achieved by an adequate sampling in space, but 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. More specifically, Fablet et al. (2021) developed an end-to-end neural model for learning a system's dynamics representation and jointly solving a variational formulation problem. This method, called 4DVarNet, has shown a great potential to reconstruct chaotic dynamics like Lorenz-63 and Lorenz-96 systems (Fablet et al., 2021), and hence it will be used here as a neural DA scheme, which will be benchmarked against a classical state-of-the-art 4D-Var DA scheme (Cummings & Smedstad, 2013). Our results support a much greater ability of neural DA schemes to retrieve the dynamics of the different phases of the AMOC.

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 dynamical system representing the AMOC and the DO events. Section 3 presents the considered neural and variational DA schemes. We experiment various sampling strategies in Section 4 and discuss further our main contributions.

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 ($\mathbf{M}$) for this model reads:

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

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 (such as Ω=Ω_0+ω, 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 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 $\boldsymbol{X}$), we apply the coefficient $\frac{\beta\epsilon}{\lambda}$ to the second and third equations. We obtain the following, time-depending, vector of three components $\boldsymbol{X}(t) \triangleq [x_1(t), x_2(t), x_3(t)]^T = [\omega(t), \frac{\beta\epsilon}{\lambda}S_v(t), \frac{\beta\epsilon}{\lambda}S_{BT}(t)]^T$. Thus, the dynamical system can eventually be written as:

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

The time integration of the model highlights the different regimes of the system (Fig. 1). We compute the energy of the system such as d=$\sqrt{x_1^2 + x_2^2 + x_3^2}$, from which we infer three categories experimentally. The ON phase corresponds to a stable circulation with d<0.060 yr^{-1}, while the OFF phase indicates a shutdown of the AMOC with an energy exceeding d>0.130 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 1470±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 parameterization and dynamics of this model can be found in Sévellec and Fedorov (2014).

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 data assimilation 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 (Johnson et al., 2005; Carrassi et al., 2017).

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:

$$\begin{cases} \dot{\boldsymbol{X}}(t) = & \mathbf{M}(\boldsymbol{X}(t)) \\ \boldsymbol{Y}(t) = & \mathbf{H}_t \boldsymbol{X}(t) + \boldsymbol{q}_{\mathrm{obs}}(t) \end{cases}, \tag{3}$$

where $\boldsymbol{X}$ is the state of the system at the considered time and $\boldsymbol{Y}$ is the vector of observations acquired with a Gaussian noise such as $\boldsymbol{q}_{\mathrm{obs}} \sim \mathcal{N}(0, \sigma^2_{\mathrm{obs}}\mathbf{I})$. While $\mathbf{M}$ is the dynamical model containing the physical knowledge, $\mathbf{H}_t$ stands for an observation mask defining the chosen sampling strategy.

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

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

with $\phi(\boldsymbol{X})(t) = \boldsymbol{X}(t') + \int_{t'}^{t} \mathbf{M}(\boldsymbol{X}(s))ds$, where t, t' and s are time variables, ds and dt stand for the time unit, ϕ is the propagator associated with $\mathbf{M}$, $||.||_2^2$ is a norm such as $||\boldsymbol{X}||_2^2 = d^2$, t_0 is the starting time, $\boldsymbol{X}_B$ 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 parameterised.

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 $\boldsymbol{X}$ on the chosen time window ΔT starting at t=t_0. The last term corresponds to a background-dependent regularization of $\boldsymbol{X}$ in order to constrain the stability of the assimilation by providing an initial condition of the system. To be informative, $\boldsymbol{X}_B$ has to be a good approximation of the state $\boldsymbol{X}$ at time t_0 (Zupanski, 1997). In our study, this last term is only implemented in the classical 4D-Var data assimilation method (hereinafter referred to as 4DVar-classic).

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

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

where δ is the increment amplitude, and $\nabla J(\boldsymbol{Y}, \boldsymbol{X}_k, \phi(\boldsymbol{X}_k), \mathbf{H})$ stands for the gradient of J evaluated at point $\boldsymbol{X}_k$. 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 evidence by a growing literature, data assimilation and deep learning share common theoretical grounds which advocate novel approaches exploring machine learning paradigms in data assimilation problems (Arcucci et al., 2021; Brajard et al., 2020). In this study, we focus on end-to-end neural data assimilation schemes (Fablet et al., 2021; Boudier et al., 2023), and more particularly 4DVarNets, as new means to optimize a data assimilation scheme for given dynamics and observing systems.

3.1 4DVarNet Architecture

The 4DVarNet architecture reproduces the unfolding of an iterative gradient descent algorithm to minimise the variational cost J as described in eq. (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 optimal weights α_{obs} and α_{dyn}. 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 towards an optimal state $\widehat{\boldsymbol{X}}$.

As sketched in Fig. 2, the 4DVarNet scheme applies three main steps at iteration $k \in \mathbb{N}^*$:

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

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 data assimilation. Here, we propose two approaches to compute the physical prior ϕ: a numerical integration of $\mathbf{M}$ and a neural network representing ϕ directly.

ϕ as a numerical integration of $\mathbf{M}$: The numerical integration follows a Runge-Kutta 4 integration scheme and relies on the AMOC equations introduced in Sec.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 optimisation process (Raissi et al., 2019; Dabrowski et al., 2023).

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 $\mathbf{M}$ 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 there 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 dataset, which aims at jointly training the U-Net and the convolutive LSTM block – the solver –, making the optimization of the data assimilation problem powerful.

3.3 Learning setting

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

$$\forall l \in \mathbb{N}^*, \qquad L_l = \frac{1}{l}\Bigg[\sum_{t=0}^{(l-1)dt} \alpha_1 \|\widehat{\boldsymbol{X}}(t) - \boldsymbol{X}(t)\|^2 + \alpha_2\big(\|\phi(\widehat{\boldsymbol{X}})(t) - \widehat{\boldsymbol{X}}(t)\|^2 + \|\phi(\boldsymbol{X})(t) - \boldsymbol{X}(t)\|^2\big) + \alpha_3\|\phi(\boldsymbol{X})(t+1) - \phi(\boldsymbol{X})(t) - \boldsymbol{X}(t+1) + \boldsymbol{X}(t)\|^2\Bigg], \tag{6}$$

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 minimise the mean square error between the ground truth and the reconstructed state $\widehat{\boldsymbol{X}}$, and between the output of the dynamical model ϕ and the state $\boldsymbol{X}$. A regularisation term on the derivative of $\phi(\boldsymbol{X}) - \boldsymbol{X}$ is added to limit the numerical noise of the reconstructed state $\widehat{\boldsymbol{X}}$.

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 optimiser 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 code is implemented in Pytorch and is available at: `https://github.com/PerrineBauchot/AMOC_4DVarNet`.

4 Experiments and Results

4.1 Experimental setup

We carried out 100 simulations of the AMOC system over a time period of 100 000 years, from which we extracted one training set and one test set (Fig. 3). To build our training dataset, we extract 50 2 500-year time series out of the first 60 000 years of each simulation. This leads to a training dataset composed of 5 000 time series of 2 500 years. Similarly, the test dataset 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 dataset to be independent from the training one. As DO events occur with a characteristic period of about 1470±500 years, each 2 500-year time series involves on average one DO events. This makes these simulated datasets relevant for the reconstruction of regime shifts in chaotic climate dynamics.

The observation vector $\boldsymbol{Y}$ is also constructed out of this simulated dataset. To satisfy Shannon's criterion for the frequency of DO events, observations are acquired with a minimum sampling frequency of $\frac{1}{120}$ 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 reconstruction as a linear interpolation of the observations for the assimilation problem.

In our experiments, the three system components are observed simultaneously. We first compare the performance of three data assimilation methods (described below) in order to find the best among them for reconstructing the North Atlantic Ocean variability, and more specifically the regime shifts. Then, we analyse 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 Comparison of three Data Assimilation Methods

In this study, three different data assimilation methods are implemented:

1. **4DVar-classic**: a 4D variational data assimilation method, where J is minimised by a gradient descent;
2. **4DVarNet-ode**: a learning data assimilation method informed by physics, where J is optimised through a ResNet and the dynamics ϕ is a numerical integration of the system M;
3. **4DVarNet-unet**: a learning data assimilation method fully data-driven, where J is optimised 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 year sampling period and evaluate the reconstruction performance of each method. To get a initial idea on the performance, we compute the normalised mean square error for each variable using the following formula:

$$\text{NMSE} = \frac{1}{S\Delta T \sigma_{\text{GT}}^2} \sum_{s=0}^{S=100} \sum_{t=0}^{\Delta T=2500} \|\widehat{\boldsymbol{X}}_s(t) - \boldsymbol{X}_s^{\text{GT}}(t)\|^2 \tag{7}$$

where $\boldsymbol{X}_s^{\text{GT}}$ is the ground truth and σ_{GT}^2 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 50 for x_2 and x_3 – the reconstruction errors compared to a classical data assimilation method (Tab. 1). This already demonstrates the usefulness of a neural network to improve the optimisation problem within a variational data assimilation framework. We also notice that the reconstruction of x_1 is overall better performed than the reconstruction of x_2 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 (Fig. 4(a)). 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 (Fig. 4(a)). 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 (Fig. 4(b)). We assume that these differences are linked to the optimisation process. The 4DVarNet-unet method performs twice as good as 4DVarNet-ode method (Tab. 1). While 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 optimises a learning cost minimising this error as defined by eq. (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 (Fig. 4(c)), and in particular the frequency of DO events.

Furthermore, 4DVar-classic has difficulties to capture DO events accurately for x_2 and x_3 (Fig. 4). Indeed, the reconstruction of the highest amplitude variations, corresponding to the AMOC slowdown events (i.e., x_1>0.05 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 towards a better reconstruction of these climate extremes. However, to produce a consistent comparison of the three data assimilation 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 (Tab. 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 parametrisation 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 parameterization adapted to the case under study. In particular, Lorenc and Payne (2007) has 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 varies (Fig. 5) depending on the values of the hyperparameteres α_{obs} and α_{dyn} (Tab. 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 eq. (4). Indeed, observation uncertainties are applied by a diagonal covariance matrix (of large amplitude).

The NMSE values displayed in Figure 5 are consistent with those computed in Table 1. 4DVarNet-unet optimization starts with a lower NMSE than the 4DVarNet-ode and the 4DVar-classic methods. Therefore, we infer that the physical model leads to a less accurate AMOC variability than the U-Net cell – even at the beginning of the training. While 4DVar-classic and 4DVarNet-unet methods show a monotonously variational cost decrease with the NMSE and also improve quite rapidly, the 4DVarNet-ode optimisation path appears quite different. Indeed, it reaches low values of NMSE quickly but without necessarily being correlated with a smaller variational cost. This suggests that this optimization method is not as well posed as the two others. As a reminder, the 4DVarNet-ode method uses the physical model for the computation of the variational cost, just as in a 4DVar-classic method. The optimisation of this assimilation problem is therefore handled differently by our neural model, compared to classical gradient descent. In fact, what matters in the training of a neural model is the convergence of the learning cost only then to reach the convergence of the variational cost. In the 4DVarNet-ode method, the learning cost has converged but might have reached low NMSE values with a lower reliance on the variational cost, flawed by an ideal physical model probably too chaotic to be useful for the optimisation of the assimilation problem.

4.3 Influence of the Density of Observations

To evaluate the performance of our three data assimilation methods with respect to the observation budget, we computed the error of reconstruction according to the sampling frequencies of observations, going from $\frac{1}{10}$ yr^{-1} to $\frac{1}{120}$ yr^{-1}. This sampling frequency could be compared to the highest frequency of the model $\sim\frac{1}{250}$ yr^{-1} and the fastest error growth timescale of $\sim$5 yr (Sévellec & Fedorov, 2014). As expected, the NMSE increases overall with the sampling period (see

Fig. 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 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 $\frac{1}{120}$ yr^{-1}. Thus, 4DVarNet-unet appears as 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 only punctual in time and space (Munk, 2000).

We also focus on the ability to reconstruct each phase of the AMOC system. We notice here that data assimilation methods perform better at reconstructing the ON Phase than the OFF Phase. The OFF Phase is characterised 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 Tab. 7):

1. **Regular sampling strategy:** We observe the system with a regular frequency, as in the previous sections. We consider $\mathbf{H}_t=\mathbf{I}_3\ \forall t \in [1{:}100{:}\Delta T]$;
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, $\mathbf{H}_{t-dt}=\mathbf{H}_t=\mathbf{H}_{t+dt}=\mathbf{I}_3\ \forall t \in [1{:}100{:}\Delta T]$;
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.

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 2007, Keller, 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 idealised 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 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 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 optimising 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 data assimilation 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 Data Assimilation 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 data assimilation 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 optimising the dataassimilation problem by training of a convolutive LSTM neural network, which boosts the gradient descent usually computed to minimise 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 data assimilation 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 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 learnt 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 data assimilation 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 (Fairbairn et al., 2014; Yaremchuk & Martin, 2014; Cummings & Smedstad, 2013).

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.

Optimising 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... 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 (Lermusiaux, 2007; Greenhill et al., 2020).

In addition to the sampling scheme, we should also take into account the considered physical variables and measurement positions to propose a complete and optimised observation system able to capture Early Warning Signals of a possible AMOC collapse. With the increasingly plausible hypothesis of reaching this tipping point because of global warming, several studies have tackled this task recently (Jackson & Wood, 2020; Michel et al., 2022; Ditlevsen & Ditlevsen, 2023). 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.

Acknowledgments

This work was supported by the OceaniX project funded through the French ANR program. Florian Sévellec is supported by the ARVOR project funded by the LEFE IMAGO program.

6 Open Research

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` .

References

Arcucci, R., Zhu, J., Hu, S., & Guo, Y.-K. (2021, January). Deep Data Assimilation: Integrating Deep Learning with Data Assimilation. *Applied Sciences*, *11*(3). (Number: 3 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/app11031114

Berg, J., & Nyström, K. (2019, May). Data-driven discovery of PDEs in complex datasets. *Journal of Computational Physics*, *384*, 239–252. doi: 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). doi: 10.1029/2022MS003353

Brajard, J., Carassi, A., Bocquet, M., & Bertino, L. (2020, July). 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. (arXiv:2001.01520 [physics, stat]) doi: 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. doi: 10.1002/2015RG000493

Carrassi, A., Bocquet, M., Bertino, L., & Evensen, G. (2017, September). Data Assimilation in the Geosciences - An overview on methods, issues and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, *9*. doi: 10.1002/wcc.535

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). Berlin, Heidelberg: Springer. Retrieved from `https://doi.org/10.1007/978-3-642-35088-7_13` doi: 10.1007/978-3-642-35088-7_13

Dabrowski, J. J., Pagendam, D. E., Hilton, J., Sanderson, C., MacKinlay, D., Huston, C., … Kuhnert, P. (2023). Bayesian physics informed neural networks for data assimilation and spatio-temporal modelling of wildfires. *Spatial Statistics*, *55*, 100746. doi: 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., … Bond, G. (1993, July). Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, *364*(6434), 218–220. (Number: 6434 Publisher: Nature Publishing Group) doi: 10.1038/364218a0

Ditlevsen, P., & Ditlevsen, S. (2023, July). Warning of a forthcoming collapse of the Atlantic meridional overturning circulation. *Nature Communications*, *14*(1), 4254. (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/s41467-023-39810-w

Evensen, G. (2009). *Data Assimilation*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-03711-5

Fablet, R., Chapron, B., Drumetz, L., Mémin, E., Pannekoucke, O., & Rousseau, F. (2021, October). Learning Variational Data Assimilation Models and Solvers. *Journal Of Advances In Modeling Earth Systems*, *13*(10). doi: 10.1029/2021MS002572

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. doi: 10.1002/qj.2135

Germe, A., Hirschi, J. J.-M., Blaker, A. T., & Sinha, B. (2022, May). Chaotic Variability of the Atlantic Meridional Overturning Circulation at Subannual Time Scales. *Journal of Physical Oceanography*, *52*(5), 929–949. (Publisher: American Meteorological Society Section: Journal of Physical Oceanography) doi: 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. (Conference Name: IEEE Access) doi: 10.1109/ACCESS.2020.2966228

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. Retrieved from `https://journals.ametsoc.org/view/journals/clim/33/16/jcliD200034.xml` doi: 10.1175/JCLI-D-20-0034.1

Johnson, C., Nichols, N. K., & Hoskins, B. J. (2005, March). Very large inverse problems in atmosphere and ocean modelling. *International Journal for Numerical Methods in Fluids*, *47*(8-9), 759–771. doi: 10.1002/fld.869

Keller, K., Deutsch, C., Hall, G. M., & Bradford, F. D. (2007, January). Detection of changes in the north atlantic meridional overturning circulation: Implications for the design of ocean observation systems. *Journal Of Climate*, *20*. doi: 10.1175/JCLI3993.1

Kim, E., & Seto, M. L. (2022, September). An Approach to Choose Observation Systems to Observe Ocean Phenomena. In *2022 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp. 306–313). (ISSN: 2576-7046) doi: 10.1109/CCECE49351.2022.9918203

Kingma, D. P., & Ba, J. (2017, January). *Adam: A Method for Stochastic Optimization.* arXiv. Retrieved from `http://arxiv.org/abs/1412.6980` (Published as a conference paper at ICLR 2015) doi: 10.48550/arXiv.1412.6980

Lenton, T. M. (2011, July). Early warning of climate tipping points. *Nature Climate Change*, *1*(4), 201–209. (Number: 4 Publisher: Nature Publishing Group) doi: 10.1038/nclimate1143

Lermusiaux, P. F. J. (2007, June). Adaptive modeling, adaptive data assimilation and adaptive sampling. *Physica D: Nonlinear Phenomena*, *230*(1), 172–196. doi: 10.1016/j.physd.2007.02.014

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. doi: 10.1002/qj.36

Malkus, W. V. R. (1972). Non-periodic convection at high and low Prandtl number. *Mém. Soc. Roy. Sci. Liège*, *4*, 125–128.

McCarthy, G. D., Brown, P. J., Flagg, C. N., Goni, G., Houpert, L., Hughes, C. W., ... Smeed, D. A. (2020). Sustainable Observations of the AMOC: Methodology and Technology. *Reviews of Geophysics*, *58*(1). doi: 10.1029/2019RG000654

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*. doi: https://doi.org/10.1038/s41467-022-32704-3

Munk, W. (2000, January). Chapter 1 Oceanography before, and after, the advent of satellites. In D. Halpern (Ed.), *Elsevier Oceanography Series* (Vol. 63, pp. 1–4). Elsevier. doi: 10.1016/S0422-9894(00)80002-1

Raissi, M., Perdikaris, P., & Karniadakis, G. (2019, February). 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. doi: 10.1016/j.jcp.2018.10.045

Rayner, D., Hirschi, J. J.-M., Kanzow, T., Johns, W. E., Wright, P. G., Frajka-Williams, E., ... Cunningham, S. A. (2011). Monitoring the atlantic meridional overturning circulation. *Deep Sea Research Part II: Topical Studies in Oceanography*, *12*, 1744-1753.

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). Cham: Springer International Publishing. doi: 10.1007/978-3-319-24574-4_28

Rudy, S. H., Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2017, April). Data-driven discovery of partial differential equations. *Science Advances*, *3*(4). (Publisher: American Association for the Advancement of Science) doi: 10.1126/sciadv.1602614

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*.

Sévellec, F., & Fedorov, A. V. (2014, May). Millennial Variability in an Idealized Ocean Model: Predicting the AMOC Regime Shifts. *Journal of Climate*, *27*(10), 0894-8755, 1520-0442. doi: 10.1175/JCLI-D-13-00450.1

Sévellec, F., & Fedorov, A. V. (2015, November). Unstable AMOC during glacial intervals and millennial variability: The role of mean sea ice extent. *Earth and Planetary Science Letters*, *429*, 60–68. doi: 10.1016/j.epsl.2015.07.022

Vanschoren, J. (2018, October). *Meta-Learning: A Survey.* arXiv. Retrieved 2023-04-14, from `http://arxiv.org/abs/1810.03548` (arXiv:1810.03548 [cs, stat])

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. Retrieved from `https://www.science.org/doi/abs/10.1126/sciadv.adk1189` doi: 10.1126/sciadv.adk1189

Yaremchuk, M., & Martin, P. (2014, February). On Sensitivity Analysis within the 4DVAR Framework. *Monthly Weather Review*, *142*(2), 774–787. (Publisher: American Meteorological Society Section: Monthly Weather Review) doi: 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. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL088376` (e2020GL088376 10.1029/2020GL088376) doi: https://doi.org/10.1029/2020GL088376

Zhou, Q., Qian, J., Tang, J., & Li, J. (2023, August). *Deep Unrolling Networks with Recurrent Momentum Acceleration for Nonlinear Inverse Problems.* arXiv. Retrieved 2023-09-05, from `http://arxiv.org/abs/2307.16120` (arXiv:2307.16120 [cs])

Zupanski, D. (1997, September). A General Weak Constraint Applicable to Operational 4DVAR Data Assimilation Systems. *Monthly Weather Review*, *125*(9), 2274–2292. (Publisher: American Meteorological Society Section: Monthly Weather Review) doi: 10.1175/1520-0493(1997)125⟨2274:AGWCAT⟩2.0.CO;2

7 Tables

Method	x_1	x_2	x_3
4DVar-classic	8.45	155	155
4DVarNet-ode	1.90	4.38	5.75
4DVarNet-unet	**0.78**	**1.88**	**2.00**

Table 1. NMSE ($\times 10^{-3}$) of the reconstructed signal for each variable with 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

Table 2. Values of the two hyperparameters, α_{obs} and α_{dyn}, used for each data assimilation method.

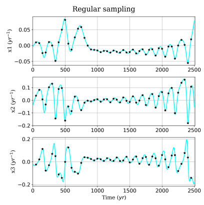
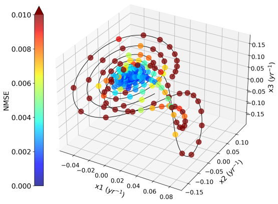
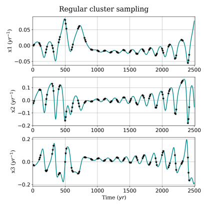
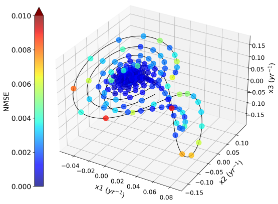
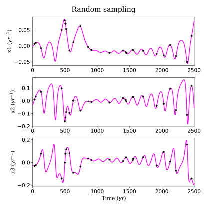
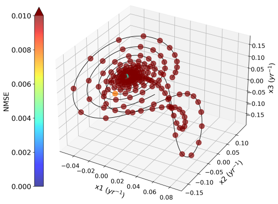
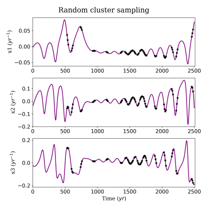
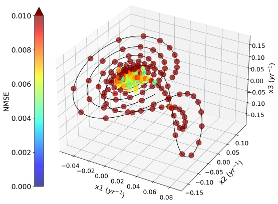
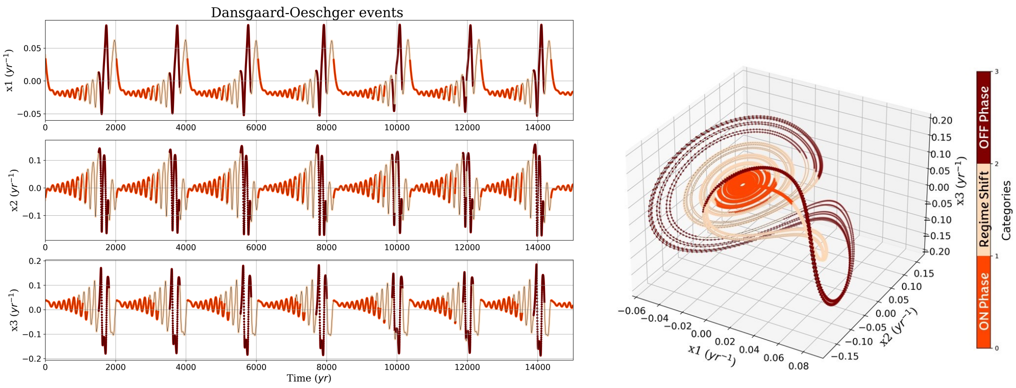
Strategy	Sampling	NMSE	Average NMSE per phase
Regular			ON Phase: 3.19×10^{-3} Regime Shift: 1.05×10^{-2} OFF Phase: 5.40×10^{-2}
Regular cluster			ON Phase: $\mathbf{8.47\times10^{-4}}$ Regime Shift: $\mathbf{2.26\times10^{-3}}$ OFF Phase: $\mathbf{8.47\times10^{-3}}$
Random			ON Phase: 1.10×10^{-2} Regime Shift: 6.64×10^{-2} OFF Phase: 2.93×10^{-1}
Random cluster			ON Phase: 8.07×10^{-3} Regime Shift: 4.51×10^{-2} OFF Phase: 2.02×10^{-1}

Table 3. Influence of sampling strategies on the reconstruction of the AMOC dynamics. The two first columns refers to the sampling strategies at play. The third column displays the Normalised Mean Square Error (NMSE) of the reconstructed system on the attractor. Last column reports the NMSE computed for the ON Phase, the Regime Shift and the OFF Phase of the AMOC. Each row corresponds to one sampling strategy.

8 Figures



(a) (b)

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 AMOC.

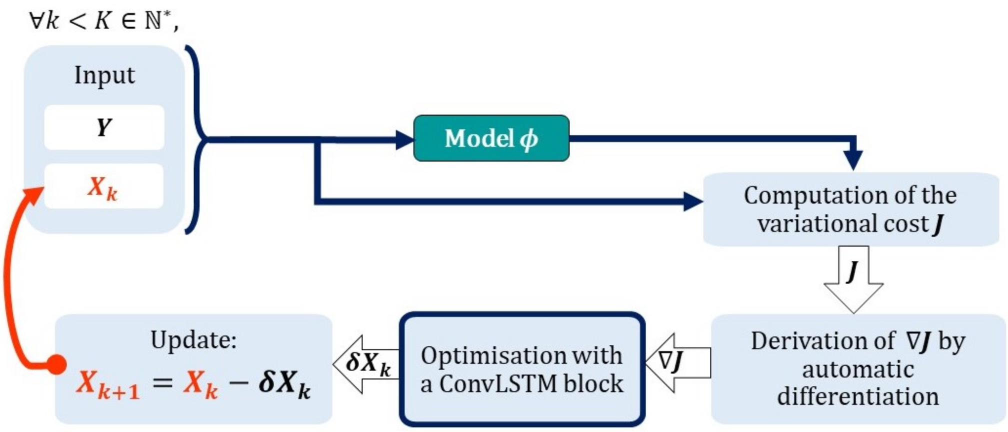


Figure 2. Sketch of the 4DVarNet procedure. The model ϕ is computed from the state X. Then, the variational cost J is calculated from the observations, the output of the model and the state X. Subsequently, the variational cost is derived by automatic differentiation. Next, the neural network is trained to optimise the minimisation of the variational cost J. Finally, the analysed state X is updated until the convergence of the learning and variational costs.

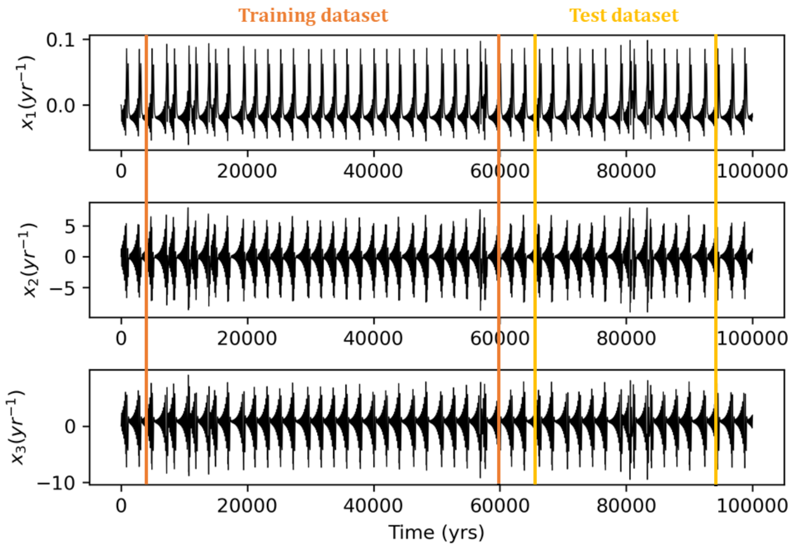
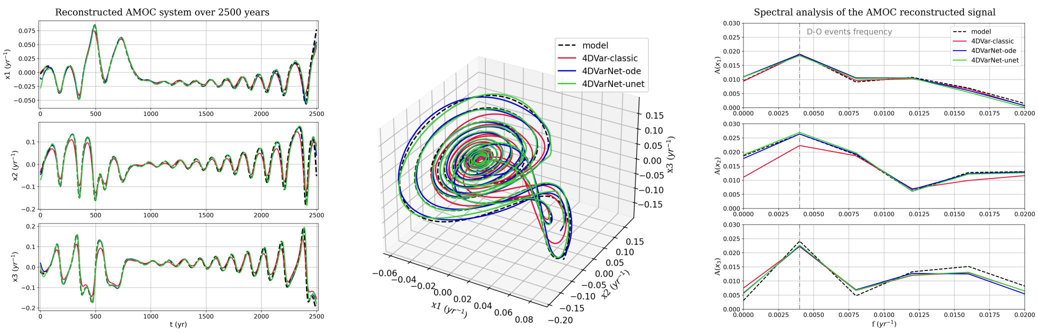


Figure 3. Evolution of the idealized model over one simulation of 10 0000 years, 50 2 500-year time series were extracted randomly between 5 000 and 60 000 yrs (in orange); one 2 500-year time series was extracted between 65 000 and 95 000 yrs to built the test dataset (in yellow).



(a) (b) (c)

Figure 4. Qualitative comparison of the three data assimilation methods (4DVar-classic in orange, 4DVarNet-ode in purple and 4DVarNet-unet in green): (a) AMOC temporal signal reconstructed by the 3 data assimilation methods; (b) Attractor of the AMOC system; (c) Spectral analysis of the reconstructed AMOC signal

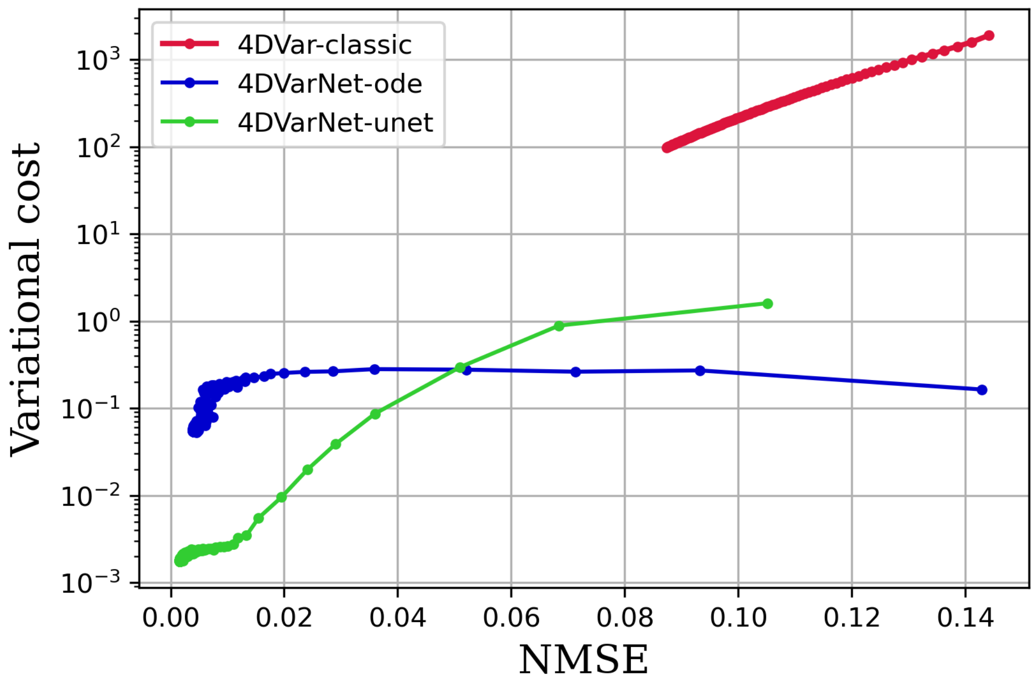
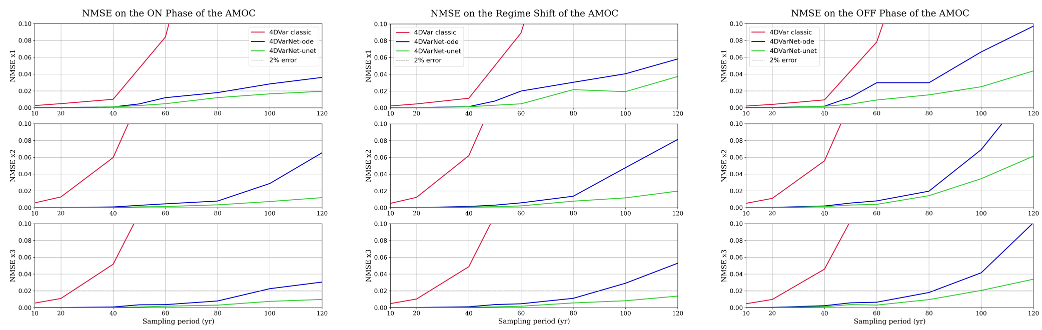


Figure 5. Variational cost (J) evolution according to the NMSE using 4DVar -classic, -ode, and -unet assimilation method, respectively. Each dot represents an iteration of the data assimilation process.



(a) (b) (c)

Figure 6. NMSE evolution for each of the 3 variables (corresponding to each row) acccording to the sampling period for the 3 data-assimilation methods (corresponding to the 3 color lines) computed on: (a-c) the ON Phase, the Regime Shift, and the OFF Phase of the AMOC, respectively.