
Physically Consistent Sampling For Ocean Model Initialization

Blandine Gorce
Sorbonne Université-CNRS,
LOCEAN Laboratory, Paris

Luther Ollier
ULCO, CNRS-IRD-Wimereux,
LOG Laboratory, Lille

David Kamm
Sorbonne Université-CNRS,
LOCEAN Laboratory, Paris

Etienne Meunier
Inria Paris

Abstract

Accurate simulation of ocean variability and climate response rely on an initialization phase called "spin-up" during which the ocean model reaches equilibrium under an applied forcing. This process comes at a high computational cost that can amount to over 1000 simulation years. Recent advances in deep generative models as climate emulators offer promising acceleration opportunities through their efficiency and ability to capture and generate complex spatio-temporal patterns. However, existing generative models often produce physically inconsistent results due to incomplete representation of underlying physical laws. In this work, we leverage on recent climate emulators to accelerate the initialization phase of ocean models. We introduce physically meaningful constraints on vertical stratification that guide sampling toward physically coherent results. Experiments on idealized ocean simulations demonstrate successful enforcement of vertical stratification.

1 Introduction

Ocean General Circulation Models (OGCMs) are essential climate research tools requiring extended initialization phases where systems reach equilibrium before producing usable simulations. Direct initialization from observations is impossible due to the lack of observational data in hydrography and marine currents, particularly at depth. The standard approach creates a virtual ocean-at-rest initial state and runs the model until equilibrium is reached. Called "spin-up" process, it requires over 1000 simulation years and millions of CPU hours, especially for deep ocean circulation. Since equilibrium states depend heavily on simulation hyper-parameters, each configuration needs separate spin-ups. This initialization phase is crucial for analyzing ocean variability mechanisms and climate responses, but its computational cost severely limits parameter exploration and ensemble simulations needed for uncertainty quantification. Recent deep learning advances have shown promise for climate emulation models [1–4]. However, despite their expressivity and generalization success, these models lack physical understanding and cannot guarantee compliance with the underlying physical laws that traditional OGCMs enforce [5]. Building on preliminary work constraining oceanic temperature and salinity by Meunier et al. [6], we propose a more interpretable yet less constraining approach to generative model conditioning. Using optimization methods for constraint imposition during generative sampling, we introduce density stratification constraints linked to temperature and salinity through a non-linear function. Our contributions include: (1) proposing two constraints based on mean density and density gradient to enforce vertical ocean stratification, (2) applying these constraints during diffusion model inference trained on post-spin-up ocean states, and (3) verifying the generated states are physically coherent while increasing variability compared to previous temperature-salinity conditioning, with potential for reintroduction into the ocean model as post-spin-up states.

2 Method

We learn the distribution of post spin-up ocean state simulations and sample from this learned distribution. As our overarching goal is to minimize the computational cost of the spin-up stage for various simulation conditions, we focused on methods that would allow us to generalize with different constraints and hyper-parameters. The method followed here consists in the following steps (see Figure 1): we first train a model on stable states generated through traditional long-time simulations. We then use improved Langevin sampling methods to guide inference toward physically consistent states. Finally, the results of the different methods are evaluated and compared using key statistical and oceanographic metrics, including ocean stratification error, oceanic pattern fidelity, and stability through reintroduction into the ocean model. This method, on our example ocean dataset, has the potential to cut compute time from $\sim 10^3$ CPU-hours to ~ 48 GPU-hours for training and a few minutes for inferences.

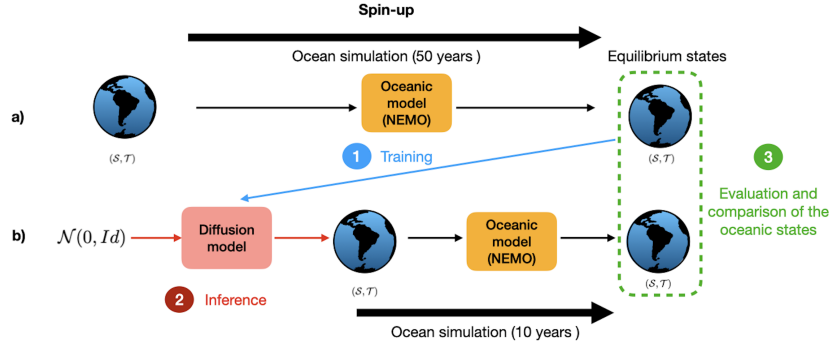


Figure 1: Pipeline a) with traditional spin-up methods, b) with diffusion model inference step.

Dataset: The dataset consists of 1800 3D temperature (T) and salinity (S) fields, extracted from a 50-year simulation of Diabatic Neverworld Ocean (DINO)[7], an idealized Atlantic-only configuration of the model NEMO[8] designed to capture key features of global ocean circulation. More details about the dataset can be found in Appendix A.

Vertical stratification constraint: Our objective is to enforce constraints on physical fields that align with established oceanographic principles. Density is a key variable for characterizing internal ocean dynamics. In the Atlantic ocean, major currents like the Antarctic Bottom Water (AABW) and North Atlantic Deep Water (NADW) are driven by density differentials [9], making preservation of vertical density stratification crucial for oceanographers. More details about density can be found in Appendix D.2.

We aim at generating temperature-salinity fields that respect the overall distribution of ocean density while maintaining the generalization across different dynamical regimes. To this end, we implemented two constraint sets (Eq. 1) to ensure that the evolution of the density with depth remains positive and enforce physically realistic density profiles. The first constrains the mean density per depth level to follow the reference mean density computed on the dataset $(\mu_k)_k$, and the second constrains the density gradient likewise.

$$\mathcal{C}_1(x) = \{x : \|\mu - \frac{1}{N} \sum_{i,j} \rho(x_{i,j})\|_2^2 = 0\} \quad \mathcal{C}_2(x) = \{x : \|\nabla_k \mu - \frac{1}{N} \sum_{i,j} \nabla_k \rho(x_{i,j})\|_2^2 = 0\} \quad (1)$$

The mean density $(\mu_k)_k$ and its gradient were computed on the training dataset. N is the dimension of the surface field. The density ρ is computed from temperature and salinity fields with the Equation Of State (EOS) approximation proposed by Roquet et al. [10] used for the DINO simulation (see D.2 for computation details).

Generative model and constrained sampling: We train a Denoising Diffusion Probabilistic Model (DDPM) [11] on a dataset of physical simulations, following the approach of Lienen et al. [12] for turbulence modeling.

To enforce physical constraints on samples generated by our model without retraining, we formulate a constrained sampling procedure [13] which casts the problem as an optimization task:

$$\underset{q \in \mathcal{P}_2(\mathbb{R}^d)}{\text{minimize}} \ D(q||p) \quad \text{subject to} \ \mathbb{E}_q[C(x)] = 0, \quad (2)$$

where $D(q||p)$ is the Kullback–Leibler divergence between q our sampling distribution subject to constraint C and p the target distribution learned from data. Using a primal–dual decomposition, the optimization is performed through a gradient descent–ascent scheme:

$$\begin{aligned} \text{Primal update:} \quad x_{s+1} &= x_s - \tau_s \nabla_x \mathcal{L}(x_s, \lambda_s) + \sqrt{2\tau_s} \epsilon, \quad \epsilon \sim \mathcal{N}(0, 1), \\ \text{Dual update:} \quad \lambda_{s+1} &= \lambda_s + \eta \nabla_\lambda \mathcal{L}(x_s, \lambda_s). \end{aligned} \quad (3)$$

The resulting potential is defined as $\mathcal{L}(x, \lambda) = \log p(x) + \lambda^\top C(x)$ and τ_s is given by the noise scheduler of the diffusion model. The dual variable λ adaptively increases the effective weight of the constraint in proportion to the violation of $C(x)$ during inference, while the hyperparameter η controls the update rate of λ . Further details on parameter choices are provided in Appendix C.

3 Results

We compare the results of the two constrained sampling and their efficiency to enforce density stratification on the generated fields at two levels: first, we assess after sampling the physical validity and diversity of the generated fields and then we evaluate its potential as a NEMO initialization state.

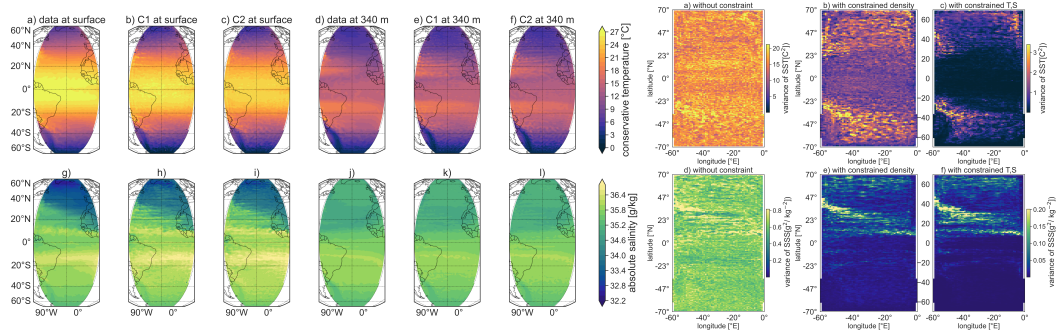


Figure 2: Comparison between training and generated data. Left panel: Temperature and salinity fields from dataset (left) and diffusion model with constraints C1 and C2 showing large-scale oceanic patterns. Right panel: Spatial variability of temperature (top) and salinity (bottom) fields under three conditions: without constraint (left), density-constrained (center), and temperature-salinity constrained (right).

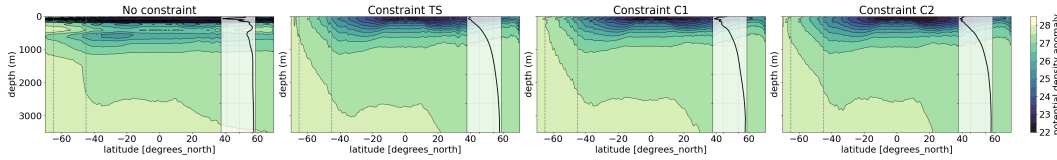


Figure 3: Vertical density stratification. Zonally averaged sections of the potential density fields computed with EOS equation from temperature and salinity fields and density profiles. Left to right: generated without constraint, with temperature and salinity, C1 and C2 constraint.

Quality of the generated fields: Figure 2 compares temperature and salinity fields generated by the diffusion model with original data. The generated fields successfully capture large-scale patterns, including warm, saline waters in the tropics and cooler, fresher waters at higher latitudes, consistent with the surface buoyancy fluxes of the reference simulation used for training. On Figure 3, we confirm that both constraints preserve vertical ocean density structure during inference, particularly near the

surface, unlike the unconstrained case. On the right panel of Figure 2, we notice the constraint on the density gradient allows for more variability in the generation than previous experiments constraining directly on the temperature and salinity vertical profiles as our new constraint is less restrictive.

Table 1: Statistical analysis of water masses and density errors. mean $\pm$ std.

Source	Bottom-Water		Deep-Water		Density Errors
	S	T	S	T	
Data	$35.2 \pm 5.0\text{e-}6$	$4.8 \pm 2.2\text{e-}3$	$35.3 \pm 1.7\text{e-}4$	$2.6 \pm 6.3\text{e-}3$	$0.3 \pm 4.1\text{e-}2$
No constraint	$35.2 \pm 6.4\text{e-}2$	4.7 ± 0.5	$35.3 \pm 8.0\text{e-}2$	2.9 ± 1.0	26.8 ± 9.6
Constraint TS	$35.2 \pm 1.5\text{e-}3$	$4.7 \pm 9.9\text{e-}3$	$35.3 \pm 1.6\text{e-}3$	$2.7 \pm 1.2\text{e-}2$	2.0 ± 0.2
Constraint C1	$35.2 \pm 8.0\text{e-}3$	$4.5 \pm 8.0\text{e-}2$	$35.3 \pm 1.7\text{e-}2$	2.4 ± 0.2	5.8 ± 2.5
Constraint C2	$35.2 \pm 1.2\text{e-}2$	4.5 ± 0.1	$35.3 \pm 1.5\text{e-}2$	2.6 ± 0.2	4.9 ± 1.7

Physical consistency of the generated fields: Accurately reproducing the structure of density layers provides a strong indicator of the physical realism of generated fields. To evaluate this, we compute the characteristics of two key water masses (Bottom-Water and Deep-Water) according to the methodology in Appendix D.1 as well as the density errors in percentage over the ocean volume. The results in Table 1 indicate the constrained generation overall reproduces the main water mass characteristics, except for an offset in the upper layers of the temperature field. Notably, the density errors decreased by an order of magnitude compared to those obtained from the unconstrained generation, which highlights the importance of enforcing the physical constraint to preserve the hydrostatic structure and prevent excessive vertical diffusion in future NEMO simulation. Additionally, in Figure 4, we validate the physical coherence of the generated states through 10 years of NEMO integration. The results show minimal changes over the period of integration both on the zonal average and on the mean density profile computed over the region of interest at low latitudes. We highlight that the generation without constraint underestimates the uplifting of Antarctic dense water around $-60/-70^\circ$, a key aspect in the formation of this water mass, whereas all three constraints are able to resolve this physical property.

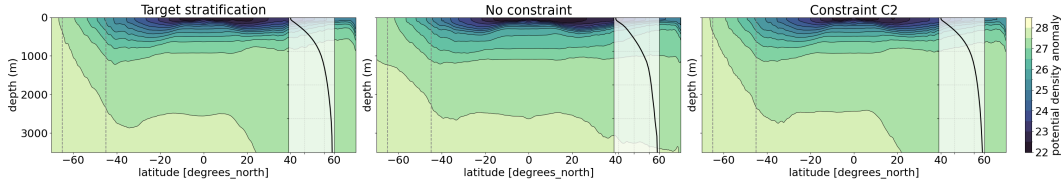


Figure 4: Vertical density stratification after 10 years of NEMO integration with initialization provided by the generative model. Zonally averaged sections of the potential density fields. Left to right : target state computed from long-term NEMO integration, without constraint, and with C2 constraint. All profiles comparison can be found in Appendix E

4 Conclusion

In this work, we introduce a constraint more physically significant than the temperature and salinity constraint imposed in Meunier et al. [6] to enforce vertical stratification on the ocean. In addition, we leverage on established optimization method by Chamon et al. [13] and proposed an implementation to apply it to large scale oceanographic data. Finally, we achieve to preserve the main physical properties, while keeping more variability on the associated temperature and salinity fields. This work enables future advances in physically consistent generative models through more complex constraints and varied sampling methods, offering potential to reduce ocean or even climate model spin-up times. Applications in oceanographic and climate sciences could improve data assimilation and enhance ensemble generation for uncertainty quantification.

References

- [1] Anass El Aouni, Quentin Gaudel, Charly Regnier, Simon Van Gennip, Olivier Le Galloudec, Marie Drevillon, Yann Drillet, and Jean-Michel Lellouche. Glonet: Mercator’s end-to-end neural global ocean forecasting system, 2025. URL <https://arxiv.org/abs/2412.05454>.
- [2] Dmitrii Kochkov, Janni Yuval, Ian Langmore, Peter Norgaard, Jamie Smith, Griffin Mooers, Milan Klöwer, James Lottes, Stephan Rasp, Peter Düben, Sam Hatfield, Peter Battaglia, Alvaro Sanchez-Gonzalez, Matthew Willson, Michael P. Brenner, and Stephan Hoyer. Neural general circulation models for weather and climate. *Nature*, 632(8027):1060–1066, July 2024. ISSN 1476-4687. doi: 10.1038/s41586-024-07744-y. URL <http://dx.doi.org/10.1038/s41586-024-07744-y>.
- [3] Xiang Wang, Renzhi Wang, Ningzi Hu, Pinqiang Wang, Peng Huo, Guihua Wang, Huizan Wang, Senzhang Wang, Junxing Zhu, Jianbo Xu, Jun Yin, Senliang Bao, Ciqiang Luo, Ziqing Zu, Yi Han, Weimin Zhang, Kaijun Ren, Kefeng Deng, and Junqiang Song. Xihe: A data-driven model for global ocean eddy-resolving forecasting, 2024. URL <https://arxiv.org/abs/2402.02995>.
- [4] Oliver Watt-Meyer, Gideon Dresdner, Jeremy McGibbon, Spencer K. Clark, Brian Henn, James Duncan, Noah D. Brenowitz, Karthik Kashinath, Michael S. Pritchard, Boris Bonev, Matthew E. Peters, and Christopher S. Bretherton. Ace: A fast, skillful learned global atmospheric model for climate prediction, 2023. URL <https://arxiv.org/abs/2310.02074>.
- [5] Sam Motamed, Laura Culp, Kevin Swersky, Priyank Jaini, and Robert Geirhos. Do generative video models learn physical principles from watching videos?, 01 2025.
- [6] Etienne Meunier, David Kamm, Guillaume Gachon, Redouane Lguensat, and Julie Deshayes. Learning to generate physical ocean states: Towards hybrid climate modeling, 2025. URL <https://arxiv.org/abs/2502.02499>.
- [7] David Kamm, Julie Deshayes, and Gurvan Madec. Dino: A diabatic model of pole-to-pole ocean dynamics to assess subgrid parameterizations across horizontal scales. *EGUsphere*, 2025: 1–26, 2025. doi: 10.5194/egusphere-2025-1100. URL <https://egusphere.copernicus.org/preprints/2025/egusphere-2025-1100/>.
- [8] Gurvan Madec and NEMO System Team. *NEMO Ocean Engine Reference Manual*, 2024. URL <https://doi.org/10.5281/zenodo.1464816>.
- [9] Céline Heuzé. Antarctic bottom water and north atlantic deep water in cmip6 models. *Ocean Science*, 17(1):59–90, 2021. doi: 10.5194/os-17-59-2021. URL <https://os.copernicus.org/articles/17/59/2021/>.
- [10] Fabien Roquet, Gurvan Madec, Laurent Brodeau, and J. Nycander. Defining a simplified yet “realistic” equation of state for seawater. *Journal of Physical Oceanography*, 45(10):2564 – 2579, 2015. doi: 10.1175/JPO-D-15-0080.1. URL <https://journals.ametsoc.org/view/journals/phoc/45/10/jpo-d-15-0080.1.xml>.
- [11] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models, 2020. URL <https://arxiv.org/abs/2006.11239>.
- [12] Marten Lienen, David Lüdke, Jan Hansen-Palmus, and Stephan Günnemann. From zero to turbulence: Generative modeling for 3d flow simulation, 2024. URL <https://arxiv.org/abs/2306.01776>.
- [13] Luiz F. O. Chamon, Mohammad Reza Karimi, and Anna Korba. Constrained sampling with primal-dual langevin monte carlo, 2025. URL <https://arxiv.org/abs/2411.00568>.
- [14] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015. URL <https://arxiv.org/abs/1505.04597>.
- [15] Patrick von Platen, Suraj Patil, Anton Lozhkov, Pedro Cuenca, Nathan Lambert, Kashif Rasul, Mishig Davaadorj, Dhruv Nair, Sayak Paul, William Berman, Yiyi Xu, Steven Liu, and Thomas Wolf. Diffusers: State-of-the-art diffusion models. <https://github.com/huggingface/diffusers>, 2022.

- [16] Augustus Odena, Vincent Dumoulin, and Chris Olah. Deconvolution and checkerboard artifacts. *Distill*, 2016. doi: 10.23915/distill.00003. URL <http://distill.pub/2016/deconv-checkerboard>.
- [17] Global ocean physics reanalysis. <https://doi.org/10.48670/moi-00021>, 2022.
- [18] L. Vogt, J.-B. Sallée, and C. de Lavergne. Stratification and overturning circulation are intertwined controls on ocean heat uptake efficiency in climate models. *Ocean Science*, 21(3): 1081–1103, 2025. doi: 10.5194/os-21-1081-2025. URL <https://os.copernicus.org/articles/21/1081/2025/>.

A Dataset

The dataset is composed of 1800 1800 snapshots of temperature and salinity fields. Each snapshot represents a 3D ocean state comprising temperature (T) and salinity (S), concatenated to give the entry variable X of the diffusion model. All 3D-fields are of shape ($Z=36$, $W=199$, $H=62$) where the first dimension represents the vertical discretization of the ocean depth in 36 levels. All fields of the training dataset were extracted from 50-years simulation of Diabatic Neverworld Ocean (DINO), a simplified version of the ocean general circulation model NEMO¹ [8] representing the Atlantic ocean. DINO is an idealized, ocean-only model configuration designed to assess eddy parameterizations across a range of horizontal resolutions. DINO parameters allow it to represent key features of global ocean circulation including the Meridional Overturning Circulation (MOC) and the Antarctic Circumpolar Current (ACC) while remaining computationally affordable for high-resolution experiments.

This configuration employs a $1/4^\circ$ horizontal resolution on a Mercator grid, with 36 vertical levels. The model domain extends 60° in longitude and 70° in latitude, covering both hemispheres from the equator to the poles. Vertically, the resolution increases with depth to better capture the internal dynamics of the ocean.

Before training, both temperature and salinity fields are normalized removing the mean and dividing by the standard deviation, both computed from the training set.

B Architecture

A commonly used architecture in diffusion models is the U-Net [14], originally introduced for biomedical image segmentation in 2015. The U-Net follows an encoder–decoder structure, comprising a downsampling path and an upsampling path, with skip connections. In our work², we adopt a U-Net-based architecture provided by Hugging Face’s Diffusers library [15], which offers a flexible and widely adopted PyTorch implementation of diffusion-based generative models. Our implementation builds upon this standard U-Net, with adaptations to match our computational constraints and target resolution. We replaced nearest-neighbor upsampling with bilinear upsampling to dodge checkerboard artifacts in the generated outputs [16] and fixed the downsampling layer sizes to [64, 64, 128, 128] instead of the traditional [64, 128, 256, 512], for computational reasons. For further implementation details, readers are referred to the Diffusers library documentation.

C Training and inference details

The noise scheduler used to fix the $(\alpha_s)_{s=1}^S$ was taken directly from the HuggingFace library ‘squaredcos cap v2’³. AdamW was used for training with an initial learning rate of $1e-4$ and a cosine learning scheduler without warmup steps. Batch size was 8. Training took nearly two days and a half on a single V100 GPU card. We did not perform hyperparameter tuning in this work as the goal was not to find the optimal performance for the diffusion model. We are aware that many aspects of the training can be largely improved. Future work include training with Exponential

¹<https://www.nemo-ocean.eu/>

²<https://github.com/Etienne-Meunier/DINO-Fusion>

³https://huggingface.co/rechido/20231107_latent_interval10_squaredcos_cap_v2

Moving Average (EMA) and investigating the use of Latent Diffusion Models.

For the hydrostatic constraint during the primal-dual inference, the value of λ was initialized at 0 and controlled with the dual updates. η parameter was fixed with the following values: 0.005 for constraint C1 and 0.1 for constraint C2. Parameter η was tested for values between $1e-5$ and 1, and the best values were selected according to the average density profiles fit to the referenced mean density profile.

As we run into boundary issues when using convolutional kernels near solid boundaries, we implement a simple solution for generated fields where the values of cells directly adjacent to the top and bottom walls are copied from their nearest interior cell.

D Metrics

D.1 Bottom-Deep water boxes :

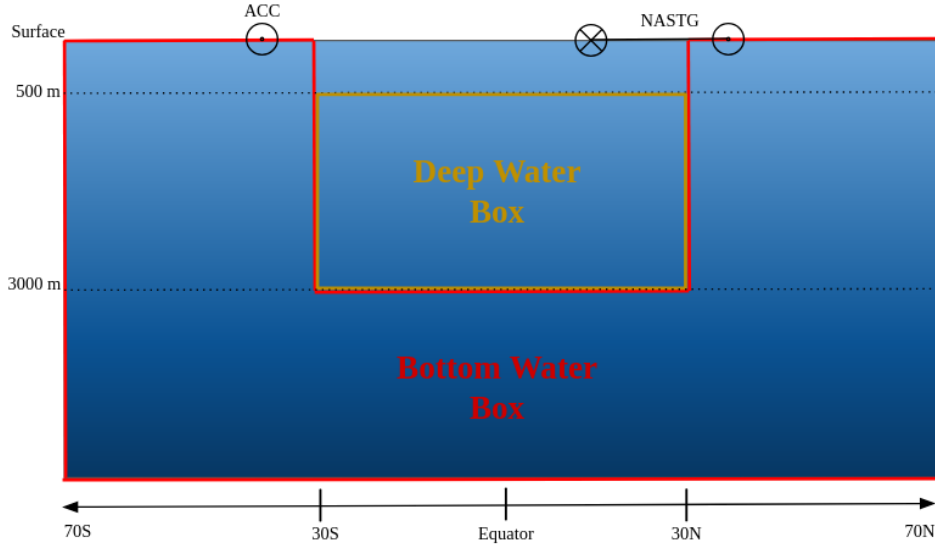


Figure 5: Regions used to calculate the characteristics of Bottom Water and Deep Water. The figure also highlights the strength of the Antarctic Circumpolar Current (ACC) and the North Atlantic Subtropical Gyre (NASTG), both of which play key roles in global ocean circulation.

D.2 Density

Density of seawater is a key parameter in oceanography, used to characterize water masses, their movement, and their properties. It is measured in $kg.m^{-3}$ and varies with both temperature and salinity. Colder water is denser than warmer water (see Figure 6), and saline water is denser than freshwater. The equation we use to compute the density is the simplified Equation Of State (EOS) proposed by Roquet et al. [10]:

$$\rho(T, S, p) = \rho_0 - (a_0 + \frac{1}{2}C_b T_a + T_h p)T_a + b_0 S_a \quad (4)$$

with

- $T_a = T - 10C$
- $S_a = S - 35gkg^{-1}$
- $\rho_0 = 1028$ reference density $kg.m^{-3}$
- $a_0 = 0.1655$ thermal expansion $kg.m^{-3}K^{-1}$
- $b_0 = 0.7655$ haline expansion $kgm^{-3}g^{-1}$

- $C_b = 9.9 \times 10^{-3}$ thermal cabbeling $kgm^{-3}K^{-2}$
- $T_h = 2.4775 \times 10^{-5}$ thermobaricity $kg^{-3}dbar^{-1}K^{-1}$

It defines a simplified thermodynamical framework that considers only the main non-nonlinearties depending on temperature, salinity and depth allowing for efficient computation while maintaining the main stratification and stability properties of the ocean.

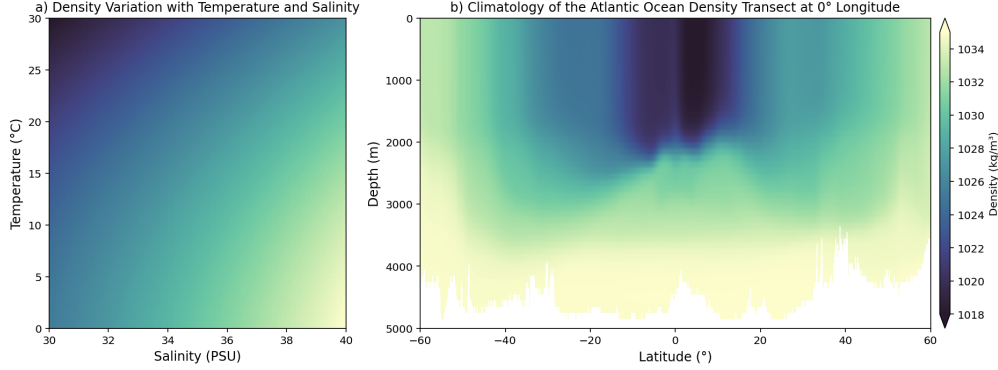


Figure 6: a) Evolution of seawater density as a function of temperature and salinity. b) Vertical section of climatological mean density (in kg/m^3) along the Atlantic Ocean at 0° longitude.

The left panel shows the theoretical dependence of seawater density on temperature ($^\circ C$) and salinity (PSU). In the open ocean, temperature variations are typically larger ($\sim \Delta 15C$) than salinity variations ($\sim \Delta 5PSU$), which means that temperature generally has a greater influence on density than salinity. The right panel presents the vertical climatology density structure of the Atlantic Ocean along 0° longitude computed using the EOS above and the salinity and temperature fields from [17]. The density gradient with depth highlights deep water masses, which are formed by cold and salty waters that sink at high latitudes near the poles. Any deviation in the density vertical structure can lead to important changes in heat uptake and carbon, thereby modulating the ocean's buffering capacity against climate change and significantly impacting long-term climate projections (Vogt et al. [18]).

We introduce a density-based error metric that quantifies the proportion of ocean volume where the vertical density profile violates hydrostatic stability, calculated as follows :

$$DensityError = 100 \times \frac{\sum_{i,j,k} V_{i,j,k} \mathbb{K}(\rho_{i,j,k+1} - \rho_{i,j,k} < 0)}{\sum_{i,j,k} V_{i,j,k}} \quad (5)$$

with $V_{i,j,k}$ the volume of a grid cell, $\rho_{i,j,k}$ the density of a grid cell (i, j, k being, respectively, the longitude, latitude and depth). $\mathbb{K}(\rho_{i,j,k+1} - \rho_{i,j,k} < 0)$ equals 1 when the density difference between two vertical levels is negative (unstable stratification) and 0 otherwise.

D.3 Variability metric

To measure the variability of the generated states, we computed the mean point by point variance of the temperature and salinity fields generated. As illustrated in the right panel of Figure 2, the constraints C1 and C2 allow for more point by point variability showing more diversity in the generated data.

Table 2: Variability metric under different constraints.

	Constraint C1	Constraint C2	No Constraint	Constraint TS
Salinity	0.0105	0.0126	0.2132	0.0095
Temperature	0.4156	0.6326	9.0859	0.3178

E Density profile comparison

We compare the vertical density stratification after 10 years of NEMO integration with initialization provided by the generated constraints. We notice that all three constraint are able to reproduce correctly the uplifting of dense Antarctic waters whereas the unconstrained case produce an uplifting that stops around 1000m. Additionally, we compare the averaged density profiles over the area of interest between the latitudes -65° and -45° that allows us to compare the curve of the dense water uplifting. We can see that all constraint are presenting the expected profile. To differentiate all three constraints a more detailed study of the resulting fields would be required.

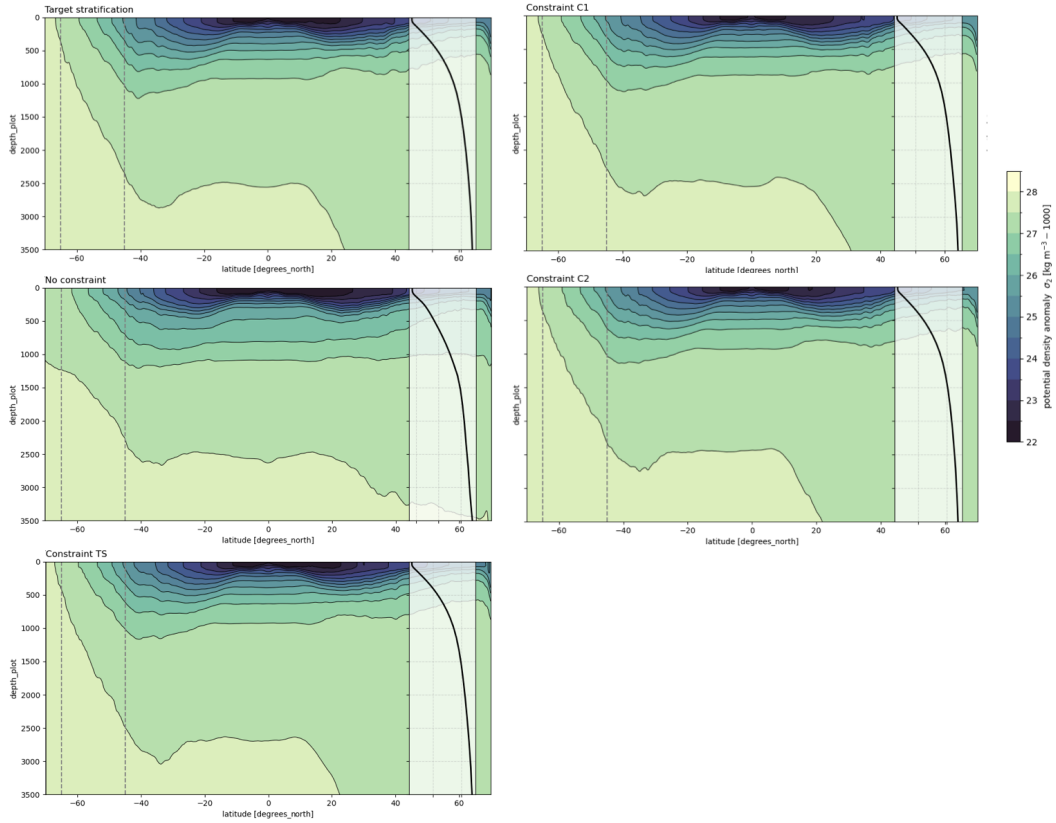


Figure 7: Vertical density stratification after 10 years of NEMO integration with initialization provided by the generative model. Zonally averaged sections of the potential density fields. Left to right, top to bottom: target state computed from long-term NEMO integration, without constraint, with temperature and salinity constraint and with C1 and C2 constraint.