

Responses to Reviewers' Comments for Manuscript egusphere-2025-6078

**Solving calibration and reanalysis challenges
of ocean biogeochemical dynamics with
neural schemes: a 1D vertical model
case-study.**

Addressed Comments for Publication to

Biogeosciences

by

Jean Littaye, Laurent Memery and Ronan Fablet

Dear editorial support team,

Please find enclosed the revised version of our previous submission entitled “Solving calibration and reanalysis challenges of ocean biogeochemical dynamics with neural schemes: a 1D vertical model case-study.” with manuscript number egosphere-2025-6078. We thank you and the reviewers for the valuable comments to improve the quality of our manuscript. In this revision, we have meticulously addressed the comments made by the reviewers. The following section provides a concise overview of the primary modifications made, in addition to a meticulous, point-by-point response to the observations raised by Dr. Banerjee and Dr. Brajard.

Sincerely,

Jean Littaye, Laurent Memery and Ronan Fablet

Note: To enhance the legibility of this response letter, all the editor’s and reviewers’ comments are typeset in boxes. The respective corrections and modifications appear in the annotated version of the manuscript: deleted sentences are in red and crossed out, added sentences are in blue.

Authors' response to Dr. Brajard

General Comments. My impression is that the comparison between 4dVar and UNET is unfair. 4DVar assumes perfect forcing and imperfect model dynamics, while in the experiments, the model is actually perfect and the forcing is not. On the other hand, UNET is able to correct for the forcing error. As a consequence, it is possible that 4DVar overcorrects parameters and state variables to compensate for the forcing error, which could explain the worse results. I acknowledge that the flexibility of UNET, which does not rely on strong assumptions, is an advantage, but it would be good to include at least one fair comparison, for example, a case with no forcing error (even if forcing uncertainty is central to the work). Another option would be to rewrite the 4DVar to explicitly account for forcing error (essentially replacing the model-error term with a forcing-error term). I understand that this may be too much work for a revised version, but at the very least, the differing assumptions between the algorithms should be more clearly highlighted before the discussion section. The inconvenience of the 4DVar algorithm is acknowledged in the discussion, but it remains difficult to understand the reason for the improvement: is it the optimisation scheme itself (4DVar vs UNET) or the fact that forcing uncertainty is accounted for?

Response: Thank you for your remark.

As you note, the 4DVar approach over-corrects both the parameters and the initial state to compensate for physical uncertainty. This stems from the formulation, which does not explicitly account for noisy forcings. Whereas a classic solution is to specify a model error covariance, this is a very challenging task when dealing with uncertainties in the physical forcings. Due to the propagation of these uncertainties through the dynamical model, there is no simple analytical solution to derive a fixed model error covariance. From a model-based perspective, a more principled approach is to shift to a coupled data assimilation problem regarding the noisy forcings as a noisy observation of the true forcings and aiming at sampling the joint posterior of the forcing variables, the ocean

BGC state and the model parameters. Due to the challenges faced in the literature by coupled data assimilation systems [1, 2], we do not include this configuration as a baseline for comparison. Here, our goal is rather to emphasize the potential impact of noisy forcings, if not relevantly accounted for, on forced data assimilation configurations to address ocean BGC calibration, as it remains the classic approach used in the literature. To deliver a fairer comparison between the DA scheme and the learning-based approach, we have carried out a complementary experiment based on noise-free forcings. In this experimental setup, the 4DVar achieves a better calibration performance than the learning-based method. These findings are in agreement with our previous study for a 0D setup [3]. They suggest the greater sensitivity of the 4DVar scheme to forcings’ uncertainties representative of real ocean physics reanalysis datasets. By contrast, the neural approach learns from the training dataset how to account for noisy forcings, which results in a greater calibration robustness.

An additional baseline experiment has been included in the appendix (see Appendix D). This experiment compares the 4Dvar-based and UNet-based calibration methods for a CTD+Rosette observation setup in the absence of uncertainty in the physical forcing.

Comment 1

L5: The complexity of the biogeochemical processes themselves is also a major source of uncertainty.

Response:

It is accurate to assert that the BGC model error is also a significant source of uncertainty in simulations. In this study, however, we focus on model error arising from poorly constrained parameters. We discuss how the proposed approach could address epistemic uncertainties (e.g., model errors in the equations or parameterisation) in future work.

Comment 2

L6: ocean data assimilation of the physics or the BGC observations?

Response:

Our study addresses BGC model calibration and discusses how physics-related uncertainties hinder the assimilation of BGC observations if not accounted for [4, 5].

To deal with this issue, a principled model-based approach is to leverage a coupled data assimilation scheme to address jointly coupled physical-BGC dynamics. Despite demonstrations in the literature [6, 4], the design of such joint data assimilation schemes is both computationally demanding and challenging in terms of optimization problem [7]. This advocates for the exploration of alternative approaches, such as learning-based ones, which could show a greater robustness to forcings' uncertainties.

Few words have been added L.7-8 to clarify this aspect.

Comment 3

L42–55: Same remark: this could be synthesised in a few lines, reminding the reader that observations are scarce, incomplete at the surface due to clouds, lack high-resolution physics information, and are very limited in the subsurface.

This part has been synthesized with lines 52-59.

Comment 4

L7; L80: The term BGC dynamics is not very clear in this context. While BGC processes are dynamic, the reanalysis reconstructs the state, not the dynamics themselves.

Response:

In this context, the term 'BGC dynamics' is used in a general sense to describe the BGC variable signal, rather than the parameterisation itself.

This term has been modified, L.85.

Comment 5

L110: The G term -> The G factor?

The G term is indeed a factor and has been corrected, L.119.

Comment 6

Eq. 1 and Eq. 2: These are two time-dependent equations, but it is unclear how they are linked. Are NO_3 , NH_4 , etc. defined over the vertical?

Response:

In Eq. 1, the SMS (source minus sink) term is resolved using Eq. 2, where C refers to all tracers (NO_3 , NH_4 , P, Z and D). Combining the two equations for nitrate concentration, for example, gives the following equation:

$$\frac{\partial \text{NO}_3}{\partial t} = \frac{\partial}{\partial z} \left(\mathbf{K}_z \frac{\partial \text{NO}_3}{\partial z} \right) + \mu \text{NH}_4 - G \frac{\text{NO}_3}{\kappa + \text{NO}_3} e^{-\Psi \text{NH}_4 P} + \lambda (\beta - \text{NO}_3) \quad (1)$$

Few lines have been added (L.107-116) to clarify this point.

Comment 7

Eq. 3: how is U linked to $\mathbf{K}_z$?

Response:

$\mathbf{U}$ is a multi-dimensional tensor, composed of the PAR (I_0) and the vertical diffusion coefficient (K_z). This will be specified in the revised version.

This has been specified L.168.

Comment 8

L203: The variables have time and depth dimensions, but in the previous paragraph (L158), $\mathbf{U}$ had no depth dimension.

Response:

In the previous paragraph, the depth dimension was only implied through the definition of the N_z -dimensional vector $\mathbf{U}(t, p)$, which contains the forcing at time t and horizontal location p across all depths. Now, because the observation operator $\mathcal{H}(\cdot)$ returns measurements at specific, possibly non-uniform depths depending on the scenario, it is necessary to represent the depth dimension explicitly.

Few words have been included, L.176-177, to specify this point.

Comment 9

Eq. 5: The cost function is not introduced in the text and appears suddenly.

Response:

The equation is referenced on L.254 to introduce the cost function.

Comment 10

L247: It is difficult to understand that $M^{(i)}$ integrates from time 0 to time i . This could be clarified. Perhaps start with M_θ and then introduce the composition of several M_θ .

A few additional words have been included in lines 265–266 to clarify this point.

Comment 11

L252: Equation is not numbered. There seems to be an inconsistency in the Delta notation. From line 238, Delta appears to be a time value, while in L251 it appears to be an index.

Response:

The symbol Δ refers to an index such that there are d sub-periods, with $T\delta t = d\Delta\delta t$. Therefore, Δ represents the number of time steps δt within a sub-period.

The notation has been clarified on L.255.

Comment 12

Eq. 7: The role of the factor τ_{DA} is unclear. In traditional DA, the weighting between the background term, the model-error term, and the observation term is handled by the covariance matrices B and R .

Response:

Thank you for this remark, the factor τ_{DA} can indeed be handled by the covariance matrix B , in order to keep traditional notation.

The covariance matrix B has been defined, L.284-287, and now includes the factor τ_{DA} .

Comment 13

L268: Why is this needed since the cost is computed in the observation space?

Response:

We acknowledge the term 'exact state' is misleading. The averaged value when the state is not observed refers to the initial condition of the gradient descent.

We removed the term 'exact state' and the sentence (L.292-293) was modified to provide greater clarity.

Comment 14

L299: The phrasing "no regard to optimisation" seems too strong. You could say that it is not fully optimised, since the focus is on demonstrating the potential of the method.

The phrasing has been modified accordingly, L.324-326.

Comment 15

Figure 5: I do not understand the metric for the parameter error. In section 3.4, it is said that the Normalized Mean Square Error is used (a positive value), but I see negative values in panel d.

Response:

Thank you for pointing this out. Indeed, the metric in question is defined as $Err(\theta) = \frac{\theta - \bar{\theta}}{\bar{\theta}}$, where $\bar{\theta}$ denotes the mean value of the parameter. Thus, it is not a normalised mean squared error, but rather a normalised (single) error.

Several lines were incorporated into the methodology section (L.355-357) to introduce the aforementioned metrics.

Comment 16

L253 (353?): compared to a mean value of 0.99 and a minimum at 0.68: please specify that this refers to the UNET.

The approach has been specified, L.400.

Comment 17

L359: This definition could be moved to section 3.4.

The definition has been modified and moved to Section 3.5, L.355-357.

Comment 18

Section 4.2 and 4.3: Please remind the reader which algorithm is used.

Response:

The algorithm that has been used for these sections is the learning-based scheme.

Few words have been added, L.417-418 and L.434 to specify the algorithm.

Comment 19

L427: This third hybrid algorithm arrives suddenly. Why is it not introduced in the methodology section? The justification is unclear except for the improved a posteriori performance.

Response:

The novel approach presented here offers an additional way to evaluate the relevance of the corrected physical forcing produced by the UNet.

A clear description of this method, along with its justification, has been included as a new Section 3.4, before the metrics section.

Comment 20

Figure 10: It is not clear how the standard deviation is computed. The caption says this is one sample at the beginning, and that there is an ensemble of 10 members at the end. Please explain this clearly, or recall in the main text how the ensemble is computed.

Response:

In this context, the term "sample" denotes a 10-member ensemble, relying on a single set of parameters and various forcing realisations according to Eq. 3. The standard deviation is computed from the 10 members.

The caption of this Figure (now designated as Fig. 11) has been modified to clarify this point.

Comment 21

L450: It would be interesting to see improvement over non-observed variables (other variables or future variables), because in this case a cubic-spline interpolation might already give a reasonable result.

Response:

As suggested, an additional configuration has been conducted using the CTD-only observation setup, in which NH_4 and Z are not observed.

The results of this experiment are presented in the appendix. It was decided not to incorporate this in the main text in order to avoid redundancy with the current (CTD+Rosette) configuration, which represents the baseline configuration. The same metrics, that is to say the table and the figure, are included in Appendix F.

Comment 22

L491: I do not see why the forcing could not be added as a control term in the 4DVar loss.

Response:

The sentence is misleading and will be corrected. In fact, it is possible to incorporate the forcing as a control term within the variational cost. Introducing a control term in the 4DVar loss amounts to considering a joint physical-BGC assimilation problem and requires stating a prior on the physical forcing. Such coupled data assimilation schemes have been explored in the literature. While they are computationally-demanding, they also face optimization challenges for operational applications [7]. We may point out that, even for 1D ocean BGC framework as considered in this study, the design of a relevant coupled physical-BGC and/or atmosphere-ocean data assimilation scheme is highly-challenging both in terms of appropriate modelling framework for ocean-atmosphere dynamics and of model error covariances and goes beyond the scope of our study to provide a comparison baseline.

Additional lines have been incorporated, L.547-552, to elucidate this point.

Comment 23

Section 5.2: The problem of accounting for various observation errors is not discussed. One advantage of 4DVar is the ability to handle evolving observing settings with changing observation density and varying error characteristics. Is it possible to add this in the training strategy? How would the algorithms react if observation error changes at inference time?

BGC models and observing systems evolve constantly. The training approach requires generating a large set of simulations for training data. How adaptive is the method if the model or observing system changes? Would retraining be required for every evolution?

Response:

As pointed out by the reviewer, the observation error distribution is assumed to be known *a priori* and constant. This seems a reasonable approximation for a regional case study. The observation setup is also kept fixed for each trained/evaluated NN version. Accordingly, a separate training and evaluation has been carried out for each configuration shown in Figures 7–10. Despite some expected level of generalization, a first strategy to apply the proposed scheme to other observational settings (in terms of observation error and/or sampling patterns) would be to retrain or fine-tune for novel configurations. Numerous works, particularly in image processing [8, 9], have demonstrated that neural approaches can handle a broad variety of noise levels and sampling schemes, provided they are trained appropriately. A natural strategy is therefore to construct training datasets that encompass multiple observing system configurations. This would likely increase the complexity of the neural architectures and require supplying additional inputs describing the observation setup (e.g., noise intensity, sampling patterns, observation modalities). Notably, state-of-the-art UNet architectures, such as those used in our study, can already incorporate such extra conditioning variables, as demonstrated in conditional image generation tasks [10, 11].

This point is treated with additional lines, L.590-601.

Comment 24

Section 5.3: Could you give more details on how the emulator fits in your framework? Would you have an emulator of the physics only? Could you comment on the relative computational cost of physics models versus BGC models?

Response:

Emulators provide an ensemble of ocean state estimates derived from a combination of forcing and parameters, without the necessity of solving any equation. This emerging data-driven approach offers a substantial tool to counter current ocean models that can be computationally prohibitive.

To date, the development of emulators has been primarily focused on the physical components of the ocean. [12, 13] propose realistic representation of the global ocean circulation. These data-driven models represent a novel alternative to conventional ocean physical models. In such cases, biogeochemical models can be forced by emulated physical fields, as demonstrated in [14]. Nevertheless, the primary bottleneck in our study is the emulation of the BGC variables. To illustrate this point, consider that the PISCES model requires 3.4 times the resources of the NEMO model. We are aware of ongoing investigations for the neural emulation of ocean BGC dynamics, currently concentrating on variables such as surface chlorophyll and the air–sea CO₂ partial pressure difference, driven by physical input fields [15, 16]. We anticipate that research on neural approaches for ocean BGC dynamics will progress quickly, in a manner comparable to recent advances in global ocean and atmospheric physics.

These explanations were incorporated into the text, L.650-658.

Authors' response to Dr. Banerjee

General Comments. This work paper addresses calibration and reanalysis challenges in ocean biogeochemical dynamics using neural networks in a 1D vertical model framework. The preprint is generally well-motivated and the topic is timely, especially at the interface of data assimilation, ocean modelling, and biogeochemistry. I find the manuscript promising overall. However, I have a few scientific concerns, mainly regarding the consistency of the DA formulation, the interpretation of the “weak-constrained 4DVar” setup, clarification of the closed-system vs authors’ nitrate nudging argument and whether the 2-year spin-up is sufficient to ensure a repeatable annual cycle and negligible year-to-year drift in the nitrogen inventory.

Response: Thank you for your remark.

Your concerns pertain to the lack of clarity in the descriptions of the experiments and the oversight regarding the assumptions. The ensuing discussion methodically explores and clarifies each of these aspects.

Comment 25

Section 3.1, page 9 line 218-225: “the considered forcings are U^* since exact forcings U are not available” the problem is posed due to noisy forcings and the target operator is meant to coestimate parameters, states, and corrected forcings as well. But in section 3.2 (page 11, lines 265-275), the author states DA uses forcings as error-free (“..the considered forcings is assumed to be error-free..”) and forcing uncertainty is being pushed into the observation-error covariance. This is concerning and a major conceptual mismatch, as the whole point is about calibration under uncertain physical forcing. Authors should clarify this.

Response:

As you mention, the learning-based approach directly accounts for physical uncertainties in its formulation. In this way, it coestimates different variables from noisy forcings. Conversely, the 4D-Var approach operates under the assumption of error-free forcing, as it does not explicitly account for the aforementioned uncertainty. However, explicitly representing these physical errors is a very complex task in the considered state-space formulation due to the non-linear relationship between the physics and the BGC state variables. Conversely, considering a principled model-based approach would lead to a coupled physical-BGC data assimilation framework which is computationally prohibitive in realistic 3D configurations and challenging from an optimization point of view.

The limits of the comparison you are referring to are discussed with additional lines, L.546-552.

Comment 26

Section 3.2, page 10, lines 233–240: “This weakly constrained 4Dvar scheme (Fablet et al., 2021b; Frerix et al., 2021; Tr.molet, 2007) seeks to identify an optimal set of BGC parameters.” and “This error is handled by dividing the studied time series into sub-windows.” : The authors implemented a “weak-constraint 4DVar” via sub-window consistency or a model error penalty ($\tau_{DA} = 10^{-3}$), which is not a standard weak-constraint 4Dvar in the true sense with explicit model error/forcing error control. This approach is okay, but the author needs to justify whether they want to call it weak constraint 4dvar in the “standard sense” or something more appropriate.

Response:

In the definition of the variational cost functions (Equations 5–7), the model error is weighted by the factor τ_{DA} relative to the observation error. We recognize that this choice is uncommon and clarify this point in the revised version.

We restate the considered variational formulation with classic covariance-based terms and revise our reference to classic weak-constrained 4DVar schemes, L.284-287.

Comment 27

Section 2.1, page 4, lines 100–104: “This Neumann condition guarantees a closed model with no outflows...” and on contrary they say in section 2.1, page 5, lines 115–125: “the sinking of detritus leads to the depletion of nitrogen from the system over time... an additional term is introduced...”: The authors state Neumann conditions gives a closed model with no outflows but then they say detritus sinking depletes nitrogen from the system and they need nitrate nudging to compensate. The author needs to clarify this. So, is the system closed or not?

Response:

The system shows no outflow, neither at the surface nor in the bottom layer. The latter statement reflects the diffusive behavior of the BGC components. Nonetheless, the downward flux of detritus transports nitrogen to deeper layers, from which it cannot be returned to the surface during winter mixing. Consequently, a nitrate nudging procedure is applied in winter to compensate for this loss. The expression “closed model” was therefore inappropriate and has been corrected.

The explanation of the transport model has been revised in L.107–111, and a few additional words have been added in L.128–129 to reinforce the rationale for the nudging.

Comment 28

Section 2.2, page 7, lines 143–144: “Each simulation includes a 2-year period of spin-up with a constant nitrogen concentration as initial condition.” The authors should explicitly demonstrate that the chosen spin-up period is sufficient. For example, Bianchi et al. (2023) performed a 650-year spin-up in a 1D nitrogen cycle model and reported that steady state was achieved only after about 100 years under their setup. This does not necessarily mean that the present study also requires such a long spin-up, as the model configuration is different. However, the authors should show whether the final annual cycle is repeatable and whether the vertically integrated nitrogen inventory exhibits negligible year-to-year drift by the end of spin-up. This is important because biogeochemical model skill can depend on spin-up duration and residual dependence on initial conditions, and equilibration times are sensitive to the chosen boundary and/or restoring formulation.

Response:

An experiment was previously conducted in order to assess the influence of this spin-up length on the simulated BGC variables. Results exhibit a negligible impact on the integrated nitrogen after a period of two years.

The aforementioned experiment has been added in the appendix section (Appendix G), in order to justify this 2-year spin-up choice.

Comment 29

Section 3.3, page 12, line 305-310: The authors state the input tensor has size (Nbatch*Nch*NT*NT) where NT=240, but the linkage between the 120-day analysis period, 35 vertical levels and this 240*240 representation is not clear to me and needs more explanation.

Response:

Thank you for pointing out this error. The input tensor has dimensions $(N_{batch} \times (N_{ch} \times (N_z + 1)) \times 2N_T)$, where N_{batch} is the number of samples per batch, $N_{ch} = 7$ is the number of channels, $N_z = 34$ is the number of layers at which tracers are resolved (vertical diffusion is computed at the interfaces, i.e., at $N_z + 1 = 35$ depths), and N_T is the time-window length in days (since the forcings are provided at 12-hour intervals).

The description of the input tensor has been corrected, L.335-338.

Comment 30

Section 2.3, page 7, lines 151–165: The authors represented uncertainty as spatial shifting of forcing, but in a real scenario, reanalysis error includes amplitude bias, timing bias, and vertical profile error due to mixing. The author should discuss these limitations properly, which will strengthen the paper.

Response:

In the present study, the physical errors under consideration are represented as a time-varying spatial shift. This spatio-temporal uncertainty is responsible for the mislocated submesoscale structures observed in contemporary physical reanalyses. Depending on the spatial variability of the input forcing fields, this time-varying spatial shift can introduce errors in terms of both mean biases, amplitude biases, vertical profile changes. The considered error scale represents the finer simulated scale with regard to current small-scale physical reanalyses [17, 18]. A pertinent question to pose would be an evaluation of the impact of larger-scale biases, including amplitude and timing biases, on these calibration approaches. These errors are common in global simulations, such as reported in [19].

This limit has been added to the description of the physical uncertainty in L.178-182. We chose not to include it in the discussion section because it is too removed from the topics currently under consideration.

Comment 31

Section 4.4, paragraph 1: page 20, around lines 423–429: The authors introduced a hybrid method here, which, as a reader, appears quite late. I suggest the exact role of the hybrid relative to the two other methods should be stated earlier in the methods, especially what is and what is not re-estimated in the DA stage.

Response:

The novel approach presented here offers an additional method for evaluating the relevance of the corrected physical forcing produced by the UNet. This additional method is intended to refine the states while maintaining the UNet calibrated parameters constant.

A description of this approach, accompanied by its justification, has been included in the method part, and represents the new Section 3.4.

Comment 32

- (1) Table 1: λ is $0.05\ d^{-1}$, whereas in the model description (Section 2.1, paragraph 2: page 5, around lines 120–123), λ is set to $1\ d^{-1}$ during Feb and 0 otherwise. One of them is wrong and the other is correct, I suppose. Clarifications needed.
- (2) Figure 5 caption: “leaning-based scheme” should be learning-based scheme.
- (3) Fig. 7 caption: “30-day state sampling” appears twice, in both red and brown distributions.
- (4) Section 3.2, page 11, line 273: “the considered forcings are assumed” should be “... forcings are assumed”.

Response:

The value of λ is indeed $0.05\ d^{-1}$ during February and 0 otherwise.

All of these technical points have been addressed and rectified individually.

References

- [1] Stephen G Penny et al. “Coupled data assimilation for integrated earth system analysis and prediction: goals, challenges, and recommendations”. In: (2017). URL: https://repository.library.noaa.gov/view/noaa/28431/noaa_28431_DS1.pdf.
- [2] Laurence A Anderson, Allan R Robinson, and Carlos J Lozano. “Physical and biological modeling in the Gulf Stream region:: I. Data assimilation methodology”. In: *Deep Sea Research Part I: Oceanographic Research Papers* 47.10 (2000), pp. 1787–1827. DOI: 10.1016/S0967-0637(00)00019-4.

- [3] Jean Littaye, Ronan Fablet, and Laurent Memery. “Learning-based calibration of ocean carbon models to tackle physical forcing uncertainties and observation sparsity”. In: *Journal of Advances in Modeling Earth Systems* 17.10 (2025), e2024MS004775. DOI: 10.1029/2024MS004775.
- [4] Hajoon Song et al. “Data assimilation in a coupled physical-biogeochemical model of the California Current System using an incremental lognormal 4-dimensional variational approach: part 2—Joint physical and biological data assimilation twin experiments”. In: *Ocean Modelling* 106 (2016), pp. 146–158. DOI: 10.1016/j.ocemod.2016.09.003.
- [5] Benoit Pasquier et al. “Optimal parameters for the ocean’s nutrient, carbon, and oxygen cycles compensate for circulation biases but replumb the biological pump”. In: *EGUsphere* (2023), pp. 1–38. DOI: 10.5194/bg-20-2985-2023.
- [6] Jann Paul Mattern et al. “Data assimilation of physical and chlorophyll a observations in the California Current System using two biogeochemical models”. In: *Ocean Modelling* 109 (2017), pp. 55–71. DOI: 10.1016/j.ocemod.2016.12.002.
- [7] Blaise Faugeras. “Assimilation variationnelle de données dans un modèle couplé océan-biogéochimie”. Thèse de doctorat dirigée par Blum, Jacques et Verron, Jacques Mathématiques appliquées Grenoble 1 2002. PhD thesis. 2002, 1 volume (199 pages). URL: <http://www.theses.fr/2002GRE10141>.
- [8] Jean Michel Amath Sarr. “Étude de l’augmentation de données pour la robustesse des réseaux de neurones profonds”. Thèse de doctorat dirigée par Cambier, Christophe et Bah, Alassane Informatique Sorbonne université 2023. PhD thesis. 2023. DOI: 10.70675/37ed465dzd027z4358z94a3z33afcee96511. URL: <https://theses.fr/2023SORUS072>.
- [9] Mingkun Tan, Daniel Langenkämper, and Tim W Nattkemper. “The impact of data augmentations on deep learning-based marine object classification in benthic image transects”. In: *Sensors* 22.14 (2022), p. 5383. DOI: 10.3390/s22145383.

- [10] Xingyang Nie et al. “Image-conditional diffusion transformer for underwater image enhancement”. In: *arXiv preprint arXiv:2407.05389* (2024). DOI: 10.48550/arXiv.2407.05389.
- [11] Meisheng Guan et al. “DiffWater: Underwater image enhancement based on conditional denoising diffusion probabilistic model”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 17 (2023), pp. 2319–2335. DOI: 10.1109/JSTARS.2023.3344453.
- [12] Surya Dheeshjith et al. “Samudra: An AI global ocean emulator for climate”. In: *Geophysical Research Letters* 52.10 (2025), e2024GL114318. DOI: 10.1029/2024GL114318.
- [13] Anass El Aouni et al. “GLONET: Mercator’s End-to-End Neural Global Ocean Forecasting System”. In: *Journal of Geophysical Research: Machine Learning and Computation* 2.3 (), e2025JH000686. DOI: <https://doi.org/10.1029/2025JH000686>.
- [14] Said Ouala and Zouhair Lachkar. “Neural-BGC: An Observation-Driven Emulator for Hybrid Physical-Biogeochemical Modeling”. In: (2026). DOI: 10.22541/essoar.15002003/v1.
- [15] Edward Gow-Smith and Roland Séférian. *Coupling of NEMO to a neural network emulator of PISCES*. Tech. rep. Copernicus Meetings, 2026. DOI: 10.5194/egusphere-egu26-12174.
- [16] Nabiz Rahpoe and Raffaele Bernardello. “A Deep Learn Emulator for Ocean Biogeochemical Modelling”. In: *EGU General Assembly Conference Abstracts*. 2025, EGU25–4191. DOI: 10.5194/egusphere-egu25-4191.
- [17] Romain Escudier et al. “A high resolution reanalysis for the Mediterranean Sea”. In: *Frontiers in Earth Science* 9 (2021), p. 702285. DOI: 10.3389/feart.2021.702285.
- [18] Andrea Storto et al. “Ocean reanalyses: recent advances and unsolved challenges”. In: *Frontiers in Marine Science* 6 (2019), p. 418. DOI: 10.3389/fmars.2019.00418.

- [19] Olivier Boucher et al. “Presentation and evaluation of the IPSL-CM6A-LR climate model”. In: *Journal of Advances in Modeling Earth Systems* 12.7 (2020), e2019MS002010. DOI: 10.1029/2019MS002010.