SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

Reporting year	2025
Project Title:	Enhancing regional and global ocean data assimilation
Computer Project Account:	spitstor
Principal Investigator(s):	Andrea Storto
Affiliation:	CNR ISMAR
Name of ECMWF scientist(s) collaborating to the project (if applicable)	NA
Start date of the project:	1/JAN/2025
Expected end date:	31/DEC/2027

Computer resources allocated/used for the current year and the previous one

(if applicable)
Please answer for all project resources

		Previous year		Current year	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)			9M	9M
Data storage capacity	(Gbytes)				

Summary of project objectives (10 lines max)

The overarching goal of this project, which is in full continuity with the previous Special Project run by the P.I. (spitstor-2019, spitstor-2022) is to allow us to experiment several new ensemble generation and assimilation techniques, which are well aligned with the priorities outlined by the ocean data assimilation community and may also contribute to advancing the ocean data assimilation science at European level. The proposal is articulated in different tasks: 1. Enhancing and updating stochastic physics schemes for NEMO, extending the STOPACK package. 2. Coupled data assimilation experiments with the MESMAR system, formed by NEMO, WRF and HD models and covering the Mediterranean region. 3. Data assimilation and bias correction for historical reanalyses, focussing on use of hybrid covariances and climatological and flow-dependent large-scale mode bias correction methods.

Summary of problems encountered (10 lines max)

None

Summary of plans for the continuation of the project (10 lines max)

In the continuation of the project, we will focus on:

1. Merging the stochastic physics code into NEMO5, testing the new infrastructure that includes several stochastic physics schemes, in the framework of the NEMO Working Group on Machine Learning and Uncertainty Quantification.
2. Further developments of sea level observation operator complemented by pre-trained neural network model for shallow water regions.
3. Tests with climatological bias corrections versus flow-dependent corrections, and sensitivity tests with hybrid ensemble-variational covariances in the CIGAR reanalysis system.

List of publications/reports from the project with complete references

Storto, A., Frolov, S., Slivinski, L., and Yang, C.: Correction of Air-Sea Heat Fluxes in the NEMO Ocean General Circulation Model Using Neural Networks, Geosci. Model Dev. Discuss. [preprint], https://doi.org/10.5194/gmd-2024-185, Accepted, 2025.

Summary of results

If submitted **during the first project year**, please summarise the results achieved during the period from the project start to June of the current year. A few paragraphs might be sufficient. If submitted **during the second project year**, this summary should be more detailed and cover the period from the project start. The length, at most 8 pages, should reflect the complexity of the project. Alternatively, it could be replaced by a short summary plus an existing scientific report on the project attached to this document. If submitted **during the third project year**, please summarise the results achieved during the period from July of the previous year to June of the current year. A few paragraphs might be sufficient.

1. Enhancing and updating stochastic physics schemes for NEMO

Within this research line, we have recently updated the stochastic physics package called STOPACK (Storto and Andriopoulos, 2021) - being used by ECMWF, UKMO, and ECCC- through several tests performed within the ORCA1 and ORCA025 configurations used by CNR ISMAR. Sensitivity tests were run to assess the impact of the new random field generation algorithm, based on Perlin noise (Perlin, 1985). For instance, tests have shown (Figure 1) that the use of Perlin noise (SPP-SI-Perlin) enable independent random generation for each SPP parameter (previously only one, as in SPP-SI-old and SPP-SI-Perlin-1fld), and efficient setup of large decorrelation scales for SPP compared to the old formulation SPP-SI-Old, which both contribute to increase the ensemble spread compared to the old SPP formulation. This is visible in both the SSH ensemble spread timeseries, and sea-ice thickness ensemble spread increase maps.

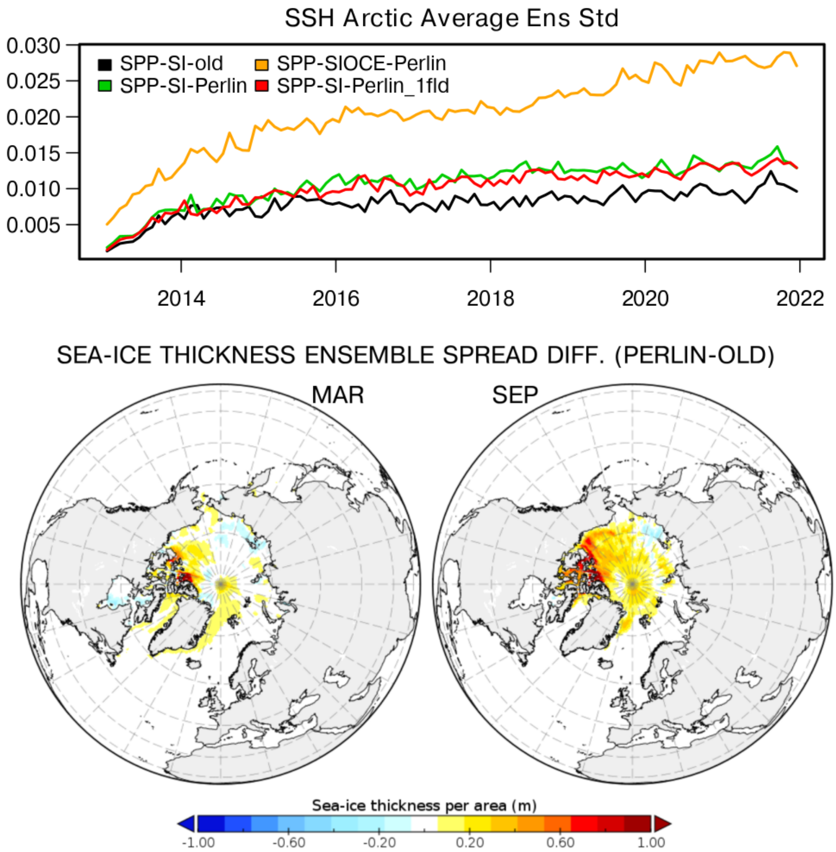

Figure 1. Top panel: ensemble spread of SSH in the Arctic for different experiments as explained in the text; bottom panel: Sea-ice thickness ensemble spread differences between the experiment with Perlin noise and the old filter.

2. Coupled data assimilation experiments

Tests were performed to identify the best configuration for altimetry assimilation within the MESMAR coupled regional reanalysis system (formed by NEMO, WRF and HD). We have tested the best-performing vertical limit of the dynamic height operator for use in the Mediterranean region through a set of experiments, whose results are summarized in Figure 2, and highlight that the best-scoring bottom level for the dynamic height integration in the Mediterranean region is 800m, in terms of correlation and explained variance compared to the fully non-linear increments simulated by the NEMO model.

Additionally, we started preliminary tests to replace the dynamic height operator with a neural network sea level operator, trained on the full-physics NEMO model increments, with the aim to overcome the limitations of the dynamic height operator in shallow waters. Preliminary results are shown in Figure 3, where we show the correlation of the dynamic height operator versus the fully non-linear model, and the increase of correlation when using the neural network observation operator, resulting in much better temporal co-variability, especially in shallow water regions such as the Adriatic Sea. This confirms the potential for the dynamic height replacements in such regions.

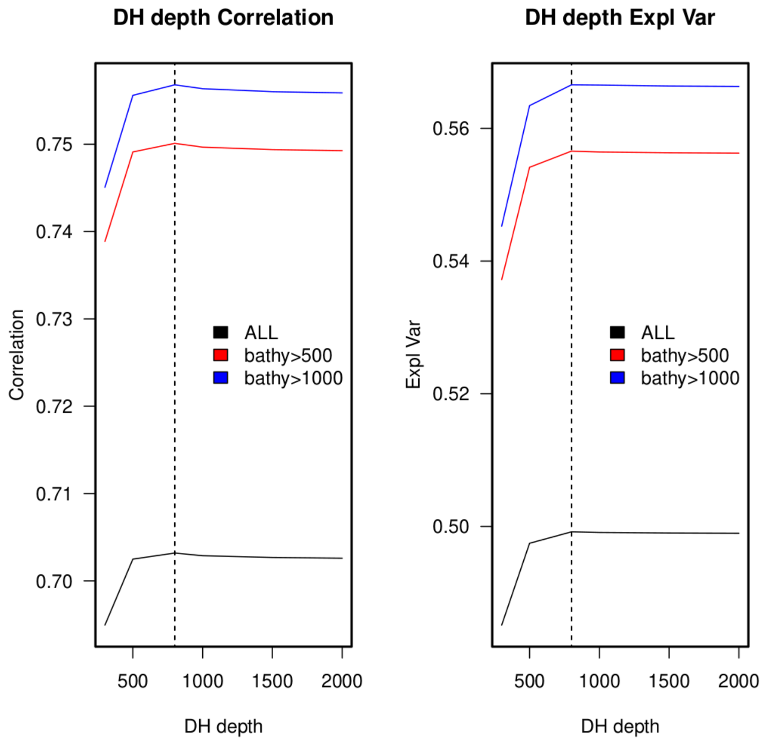

Figure 2. Optimal bottom depth for the vertical integration in the dynamic height balance operator for the Mediterranean Sea.

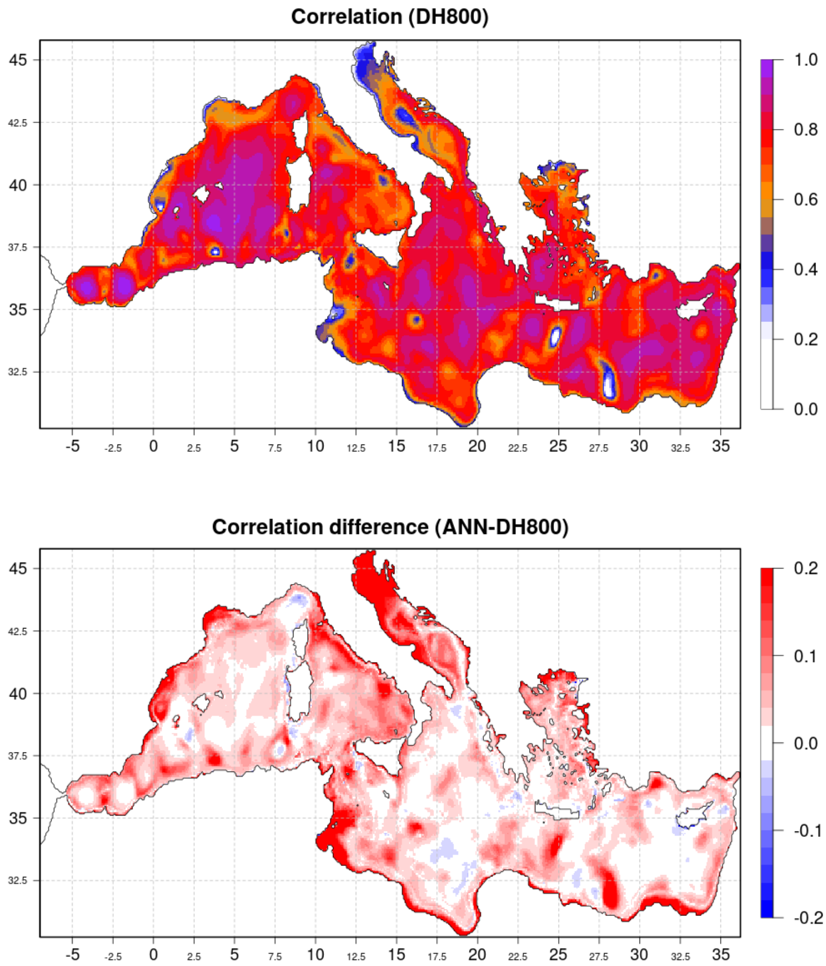

Figure 3. Correlation of the dynamic height operator (DH800) in the top panel, and correlation increase when an artificial neural network (ANN) operator replaces DH800.

3. Data assimilation and bias correction for historical reanalyses

The last task covers activities aimed at bias-correcting historical reanalyses. First, we run several tests leading to the scientific publication (Storto et al., 2025), now accepted, where a pre-trained neural network (NN) model was used to correct biases at the sea surface. The NN model uses ocean and atmospheric state predictors to infer corrections of the non-solar component of the air-sea heat flux, which help mitigate biases in the upper ocean temperature. Within these initial months, we have re-run experiments aimed at assessing the impact of the scheme within forecast experiments.
Secondly, we focused on assessing the impact of enhanced river discharge data, computed from gravimetry data and a water-mass-balance approach developed within the CMEMS Service Innovation Project WAMBOR in collaboration with Magellium. Figure 4 shows the reduction of sea surface salinity Mean Absolute Error (MAE), when the reference NEMO runoff is substituted by that developed in the North Atlantic by WAMBOR (ALL-REF), when the JRA55-do reanalysis is used (J5D-REF) or when only the Amazon river is substituted (AMZ-REF). Furthermore, the mean difference (rightmost panel) indicate that overall the salinity increases with the use of the gravimetry-based dataset.

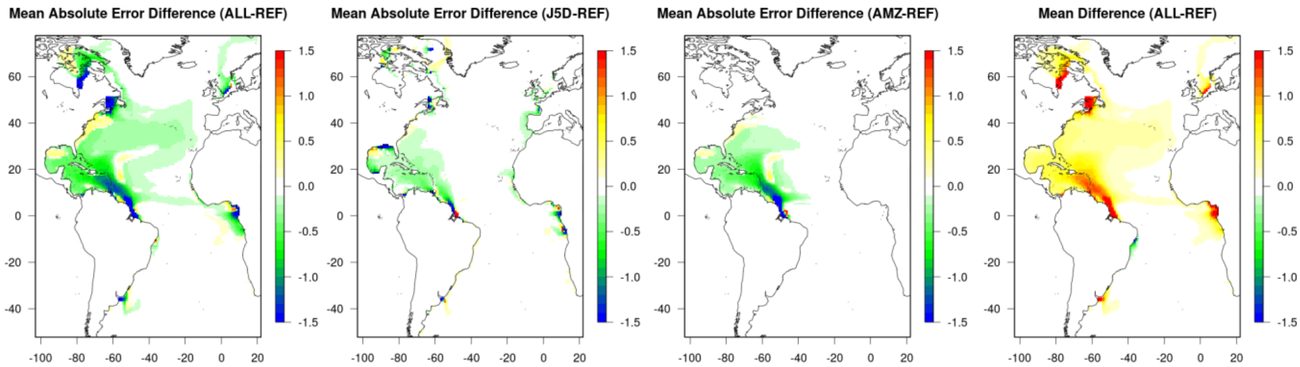

Figure 4. Difference of sea surface salinity mean absolute error (MAE) between the experiments described in the text (ALL, J5D, AMZ) and REF, in panels a, b and c. The MAE is computed versus EN4 objective analyses of salinity (Good et al., 2013). Negative values indicate a reduction of MAE compared to REF. Panel d: time-mean difference of SSS between ALL and REF. The period is 2003-2020.

References

Storto, A., and P. Andriopoulos, 2021: A new stochastic ocean physics package and its application to hybrid-covariance data assimilation. Quarterly Journal of the Royal Meteorological Society, 147, 1691–1725, https://doi.org/10.1002/qj.3990.

Perlin, K., 1985. An image synthesizer. SIGGRAPH Comput. Graph. 19, 3 (Jul. 1985), 287–296. https://doi.org/10.1145/325165.325247