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	2023
Project Title:	Enhancing regional ocean data assimilation in high and mid latitude European seas
Computer Project Account:	spitstor
Principal Investigator(s):	Andrea Storto
Affiliation:	CNR ISMAR
Name of ECMWF scientist(s) collaborating to the project (if applicable)	…………………………………………………………….…… …………………………………………………………….……
Start date of the project:	01/01/2022
Expected end date:	31/12/2024

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)	5 800 000	7 000 000	5 800 000	4 260 000*
Data storage capacity	(Gbytes)	10 TB	10 TB	20 B	21 B

* Expected to finish by the end of the year.

Summary of project objectives (10 lines max)

The project develops along with four main research topics that are summarized below: i) Stochastic physics formulation and experiments, mostly using a North Atlantic-Arctic-Mediterranean configuration of NEMO at multiple resolutions (1/4°, 1/12°, and 1/36°); ii) Altimetry data assimilation at high latitudes, focussing on the use of high sampling rate (5Hz) altimetry datasets; iii) coupled data assimilation algorithms, using as a target configuration a regional Earth system model; iv) data assimilation tests for historical reconstruction (centennial ocean reanalyses). Note, the latest point is slightly modified from the original plan. Although the topics are quite diverse, the project serves the purpose to support several ocean data assimilation research activities performed at CNR ISMAR.

Summary of problems encountered (10 lines max)

No specific problems were encountered. The new Atos machine runs well as expected. We have underestimated the initial computational resources and plan to ask for an extension of this through a re-modulation of the allocation of resources.

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

In the continuation of the project, we plan to investigate and prioritize the following scientific topics:

- Perform a long run of the weakly coupled data assimilation system for detailed assessment and as a reference for strongly coupled data assimilation experiments. The testing period covers the period 1998-2022, using the MESMAR system (see next section) with coupled assimilation.
- Perform sensitivity tests and long runs with the historical reconstruction systems, where the ocean reanalysis is forced by the 20CRv3 atmospheric reanalysis, and an ensemble of observations is used to characterize the uncertainty.
- Perform data assimilation sensitivity tests using the CREG model at both eddy-permitting and eddy-resolving resolution.

List of publications/reports from the project with complete references

1. Storto A and Yang C (2023) Stochastic schemes for the perturbation of the atmospheric boundary conditions in ocean general circulation models. Front. Mar. Sci. 10:1155803. doi: 10.3389/fmars.2023.1155803
2. Storto, A., Hesham Essa, Y., de Toma, V., Anav, A., Sannino, G., Santoleri, R., and Yang, C.: MESMAR v1: A new regional coupled climate model for downscaling, predictability, and data assimilation studies in the Mediterranean region, Geosci. Model Dev. Discuss. [preprint], https://doi.org/10.5194/gmd-2023-77, in review, 2023.
3. Storto, Yang, 2023: Acceleration of the ocean warming from 1961 to 2022 unveiled by large-ensemble reanalyses. In review.
4. Storto, et al., 2023: Assimilation of high-sampling rate altimetry for sea level studies in the Nordic Seas and Arctic Ocean. OceanPredict Data Assimilation Workshop, Rome 9-11 May 2023. Available at: https://oceanpredict.org/docs/Documents/Task%20Teams/DA-TT/Meetings/Rome-May-2023/Presentations/3.2_20230510_1105_Storto.pdf
5. Storto, et al., 2023: Reconstructing historical ocean heat content from reanalyses: an uncertainty assessment. OceanPredict Data Assimilation Workshop, Rome 9-11 May 2023. Available at: https://oceanpredict.org/docs/Documents/Task%20Teams/DA-TT/Meetings/Rome-May-2023/Presentations/6.3_20230511_1145_Storto.pdf

In all these communications, ECMWF HPC resources were acknowledged in the proper section.

Summary of results

Stochastic physics

New developments of stochastic physics have been extensively tested on the ECMWF HPC system, using a NEMO configuration that includes the North Atlantic and the Arctic oceans and the Mediterranean Sea, at ¼° of spatial resolution and 75 vertical levels. The configuration is called CREG025 and is used as a testbed for ensemble experiments.

The developments build upon the stochastic physics package STOPACK (https://doi.org/10.1002/qj.3990). The package includes three schemes applied simultaneously: stochastically perturbed parametrization tendencies (SPPT), stochastically perturbed parameters (SPP), and stochastic kinetic energy backscatter (SKEB) schemes. The three schemes allow for different temporal and spatial perturbation scales, all implemented within NEMO (initially in NEMO 3.6 and recently ported to NEMO4).

During the recent developments, we formulate and revise different approaches to perturb the air-sea fluxes used within the atmospheric boundary conditions. In particular, perturbation of the fluxes is performed either through i) stochastic modulation of the air-sea transfer coefficients (SPP); ii) stochastic modulation of the air-sea flux tendencies (SPPT); iii) coarse-graining of stochastic sub-grid computation of the fluxes (SCBF); or iv) multiple bulk formulas (MAES). A series of 22-year 4-member ensemble experiments with different stochastic schemes are performed and analyzed for the period 2000-2021, and results are compared in terms of the ensemble mean and, when applicable, ensemble spread of the principal oceanic diagnostics. Results indicate that the schemes, in general, can significantly improve some verification skill scores (e.g. against drifter current speed, SST analyses, and hydrographic profiles) and, in some cases, enhance the mesoscale activity and weaken the large-scale circulation. The response, however, is different depending on the specific scheme, whose choice thus depends on the target application, as detailed in the paper. These findings foster the adoption of these schemes in both extended-range operational ocean forecasts and coupled long-range climate prediction systems, where the boundary conditions perturbations may contribute to performance increases. Figure 1 shows the impact of the perturbation schemes on the SST anomaly and timeseriers, highlighting the significant RMSE percent decrease between 28°N and 50°N, with all schemes providing, on average, a positive impact.

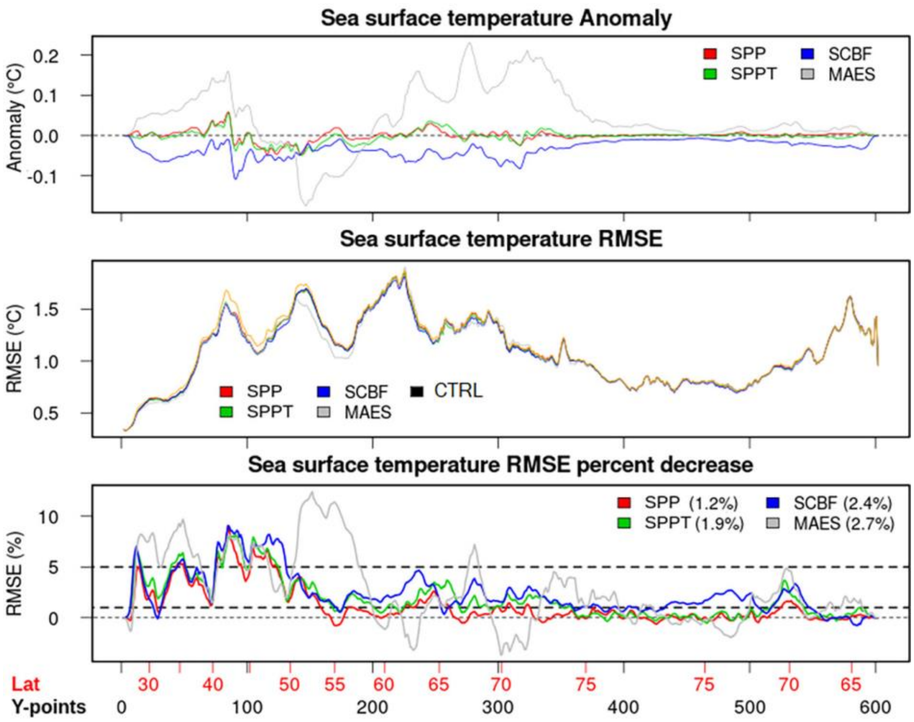

Figure 1. Sea surface temperature anomaly (difference compared to CTRL, top panel) and RMSE (middle panel) against the UKMO OSTIA daily SST analyses (Donlon et al., 2012), and RMSE percent decrease (bottom panel) compared to the CTRL experiment. The legend of the bottom panel reports the domain-averaged RMSE percent decrease. The x-axis reports the y-coordinate of the model as statistics are computed over the native model domain geometry, while red numbers show the corresponding zonally averaged latitudes (which are increasing in the Atlantic Ocean until reaching the North Pole and then decreasing until the Bering Strait).

Additionally, the SKEB scheme within STOPACK now includes a new energy term, which comes from mesoscale eddy dissipation and it is formulated as in Yang et al. (2019). Its advantage is that, unlike numerical and convection energy dissipation terms in STOPACK, its formulation does rely on horizontal diffusivity and not on resolution-dependent parameters, and it is therefore easy to port to different NEMO configurations. The full impact is evaluated in Storto and Yang (2023).

Coupled data assimilation experiments

At CNR ISMAR, we have recently developed a regional coupled modeling framework. The modeling system is called MESMAR (Mediterranean Earth System model at ISMAR), version 1, and is composed of the WRF atmospheric model, the NEMO oceanic model, and the HD hydrological discharge model, coupled via the OASIS coupler. The model is implemented at moderate resolution (about 1/12° for the ocean and river routing, while twice coarser for the atmosphere) for long-term investigations and coupled data assimilation experiments.

First, an implementation and tuning phase, achieved through the use of ECMWF HPC resources in spitstor, was performed. This implementation phase was mostly devoted to i) understanding the best vertical physics configuration for NEMO. For this point, verification results against both SST data from satellite and Argo profiles indicate that the GLS vertical mixing scheme with the Mellor and Yamada turbulence closure, and the stability function of Canuto, provides the best skill scores compared to other GLS configurations or the TKE scheme; ii) identifying the impact of the interactive river runoff through the HD model, which was found beneficial, mitigating the fresh bias that is found when using climatological river runoff; iii) choosing the best-performing physics-microphysics suite for WRF in the regional coupled system.

After the implementation phase, we implemented and tuned a weakly coupled data assimilation system, where the oceanic state is constrained by a three-dimensional variational (3DVAR) data assimilation system, and the atmospheric state by a spectral nudging scheme, which is already part of the WRF modeling system.

The 3DVAR scheme implements stationary background-error covariances estimated from the dataset of differences between two long-term simulations with different physics options in both the WRF and NEMO configurations. Several experiments were performed to identify i) the best-performing scheme for SST/SSS data ingestion; ii) the impact of spectral versus full-field nudging in the atmosphere; iii) the impact of oceanic data assimilation. Through a series of dedicated experiments, we have demonstrated the benefits of the spectral nudging on sea surface skill scores, oceanic eddy kinetic energy, and medicane event representation, while the ocean data assimilation is found crucial not only in the oceanic skill score metrics but also for medicane intensity predictions and, to some extent, to the low-troposphere skill scores. The final configuration including spectral nudging and ocean variational data assimilation will serve as the basis for regionally downscaling global atmospheric and oceanic reanalyses from ECMWF and as the basis for downscaling monthly to seasonal predictions. Table 1 reports a summary of skill scores (oceanic and atmospheric) for some experiments performed within the project. Full details in Storto et al. (2023).

Experiment Name	**Atmospheric Assimilation**	**Oceanic Assimilation**	**RMSE**					
			T850	**WS 1000-850**	**Z500**	**SST**	**T0-50**	**S0-50**
CTRL	No	No	2.07	3.58	29.5	0.63	1.13	0.32
AT0OC1	No	3DVAR+SRF	2.04	3.58	29.0	0.27	0.83	0.20
AT1OC0	Full-field nudging	No	0.82	2.08	10.7	0.71	1.10	0.32
AT1OC1	Full-field nudging	3DVAR+SRF	0.83	2.08	10.7	0.29	0.76	0.20
AT2OC0	Spectral nudging	No	1.0	2.77	10.7	0.63	1.11	0.29
AT2OC1	Spectral nudging	3DVAR+SRF	1.0	2.77	10.7	0.27	0.80	0.20

Table 1. List of experiments performed and shown in section 5 of the text, with different assimilation setups (AT0, AT1, AT2 refer to no atmospheric data assimilation, full-filed nudging, and spectral nudging, respectively; OC0 and OC1 to no oceanic data assimilation and variational ocean data assimilation, respectively). Right-side columns report total skill scores as RMSE for some selected parameters: air temperature at 850 hPa (K), wind speed in the layer 1000-850 hPa (m s-1), 500 hPa geopotential (m), SST (°C), seawater temperature (°C) and salinity (psu) in the top 50 m of depth.

Historical reconstructions

For early-period ocean reanalyses based on in-situ data, accounting for all sources of uncertainty is crucial to provide estimates with reliable uncertainty envelopes. In the present work, we aim to re-assess the OHC trends and their uncertainty, compare them with objective analyses, and draft a hierarchy of sources of uncertainty. We build upon a large-ensemble reanalysis system (32 members) that includes, in the ensemble generation, the major sources of uncertainty. We also take advantage of an objective analysis system (i.e. a statistical mapping of the observations using the same variational data assimilation system of the reanalysis), which is used to test if any of the reanalysis signals is spuriously given by the interaction between the ocean model and the data assimilation. The goal is thus to quantify and rank the sources of uncertainty in the OHC reconstruction.

On the Atos system, we run a moderate resolution (ORCA1) 32-member ocean reanalysis that covers the period 1961-2022 forced by ERA5, and we focus our analysis on the ocean heat content representation coming from this ensemble. The 1961-2022 warming is quantified in 0.43 ± 0.08 W m^{-2}. The acceleration is found significant and equal to 0.15 ± 0.04 W m^{-2} dec^{-1}. Regional patterns show a dominant trend and acceleration at high latitudes and near the Equator, with mid-latitudes exhibiting less pronounced accumulation, due to the large meridional heat transports therein. Patterns of OHC increase in 2022 are well spread over all the basins, and more than 11% of the global ocean show its highest OHC in 2022, almost doubling all previous years. Before the Argo era, the relative OHC uncertainty is the largest in the Tropical band and partly associated with El Nino events. On a global scale, the uncertainty reduces from 40% of its natural variability in the 1960s to about 15% during the last decade.

The system serves as a basis for further back-extension of the historical reanalysis system, from the end of the XIX century till the present, using the atmospheric forcing from 20CRv3. Preliminary comparison results (Figure 2) show that the two reconstruction systems (contemporary, ie forced by ERA5, versus historical, forced by 20CRv3) provide close ocean heat content reconstruction for the overlapping period, after independent spinup exercises; even a poor observational sampling before the 1950s can reproduce some notable warming events, like the one in the 1930s which is to some extent present also in other datasets (tide gauge reconstructions, not shown, etc.)

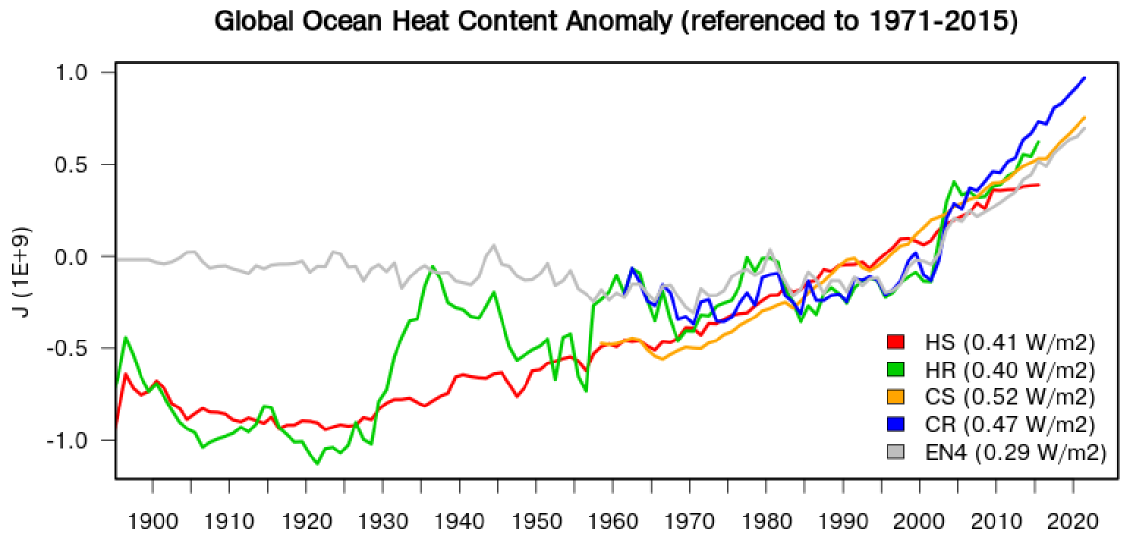

Figure 2. Global ocean heat content in a series of experiments. HS: historical simulation; HR: historical reanalysis; CS: contemporary simulation; CR: contemporary reanalysis; EN4: UKMO EN4 objective analyses. Contemporary runs start in 1958 and use ERA5 as atmospheric forcing, while historical runs start in 1860 and use 20CRv3 as a forcing. The legend reports the 1961-2015 ocean warming.

High-sampling rate altimetry

Altimetric radars measuring the sea level at millimetric precision have revolutionized our knowledge of the oceanic circulation, for more than 2 decades, at a large spectrum of scales ranging from mesoscale activity to the slowly varying basin-wide dynamics. Technological solutions are continuously needed and pursued to enhance the spatial resolution of the altimetric signal and enable the solution of the mesoscale dynamics, either in the design of the altimeter itself (e.g. wide-swath altimeters) or in the combined use of altimeter data from multiple bands. Newly reprocessed along-track measurements of Sentinel-3A, CryoSat-2, and SARAL/AltiKa altimetry missions (AVISO/TAPAS), optimized for the Arctic Ocean (retracking) and sampled at 5 Hz, have been recently produced in the framework of CNES AltiDoppler project. Our experiments are devoted to the exploitation of such satellite altimetry data in high-latitude regions. We investigate the benefits of the reprocessed altimetry dataset at 5 Hz with augmented signal resolution in the context of ocean and sea-ice coupled short-range forecasts. In particular, we compare the effectiveness of this dataset to improve the mesoscale details of the forecasts in comparison to the conventional altimetry sampling dataset and the altimetry-blind experiments, to assess the added value of the enhanced altimetry reprocessing in Nordic Seas. For instance, Figure 3 shows the increase of REKE in the Lofoten basin and other areas of the model domain. This has, in turn, a non-negligible impact on the high-latitude heat transports, whose mesoscale eddy component is enhanced (not shown). This comparison can motivate the assimilation of the high-resolution altimetry data in ocean re-analysis for the Arctic.

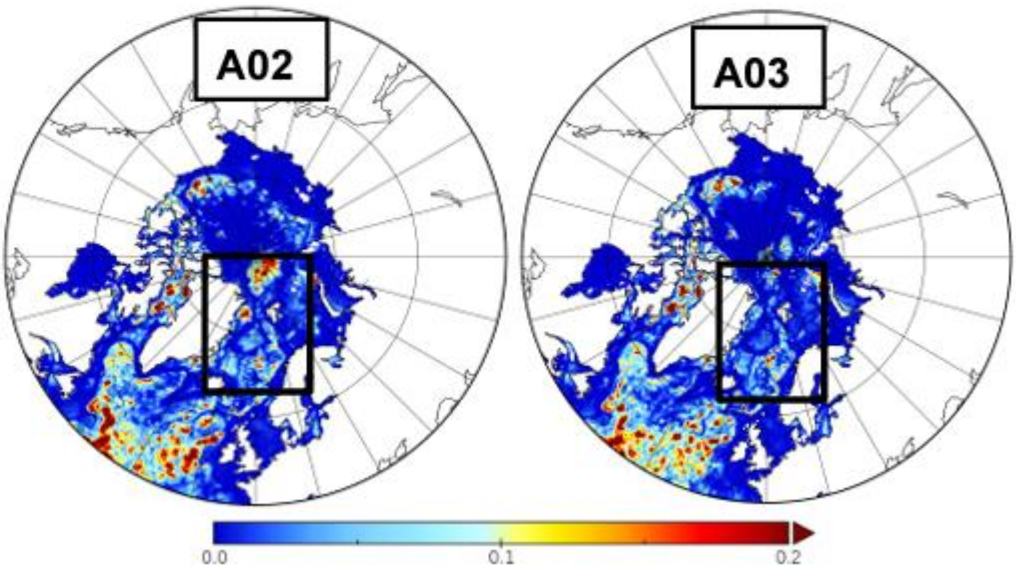

Figure 3. REKE: Relative eddy kinetic energy (fraction of EKE occurring in the presence of eddies, i.e. EKE due to eddies in compact form), for two experiments A02 and A03 for the period 2016-2019, where A02 includes the assimilation of high-sampling rate altimetry (5 Hz dataset) and A03 of conventional altimetry (1 Hz), on top of in-situ data assimilation.

References

Storto, A, Andriopoulos, P. A new stochastic ocean physics package and its application to hybrid-covariance data assimilation. Q J R Meteorol Soc. 2021; 1691– 1725. https://doi.org/10.1002/qj.3990

Storto A and Yang C (2023) Stochastic schemes for the perturbation of the atmospheric boundary conditions in ocean general circulation models. Front. Mar. Sci. 10:1155803. doi: 10.3389/fmars.2023.1155803

Storto, A., Hesham Essa, Y., de Toma, V., Anav, A., Sannino, G., Santoleri, R., and Yang, C.: MESMAR v1: A new regional coupled climate model for downscaling, predictability, and data assimilation studies in the Mediterranean region, Geosci. Model Dev. Discuss. [preprint], https://doi.org/10.5194/gmd-2023-77, in review, 2023.

Yang, Q., Nikurashin, M., Sasaki, H. et al. Dissipation of mesoscale eddies and its contribution to mixing in the northern South China Sea. Sci Rep 9, 556 (2019). https://doi.org/10.1038/s41598-018-36610-x