REQUEST FOR A SPECIAL PROJECT 2022–2024

MEMBER STATE: ITALY

Principal Investigator[1]**:** ANDREA STORTO

Affiliation: CNR-ISMAR (Institute of Marine Sciences, National Research Council)

Address: Via del Fosso del Cavaliere 100, I-00044 Rome, Italy

Other researchers: Chunxue Yang
Eva Le Merle
Jacopo Busatto
…………………………………………………………………………

Project Title: Enhancing regional ocean data assimilation in high and mid latitude European seas

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP** ITSTOR	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2022	
Would you accept support for 1 year only, if necessary?	YES	NO X (not needed)

Computer resources required for 2022-2024: (To make changes to an existing project please submit an amended version of the original form.)		**2022**	**2023**	**2024**
High Performance Computing Facility	(SBU)	5 800 000	5 800 000	5 800 000
Accumulated data storage (total archive volume)[2]	(GB)	20 TB	30 TB	30 TB

Continue overleaf

[1] The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.

[2] These figures refer to data archived in ECFS and MARS. If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year etc.

Principal Investigator: Andrea Storto

Project Title: Enhancing regional ocean data assimilation in high and mid latitude European seas

Extended abstract

The completed form should be submitted/uploaded at https://www.ecmwf.int/en/research/special-projects/special-project-application/special-project-request-submission.

All Special Project requests should provide an abstract/project description including a scientific plan, a justification of the computer resources requested and the technical characteristics of the code to be used.

Following submission by the relevant Member State the Special Project requests will be published on the ECMWF website and evaluated by ECMWF as well as the Scientific Advisory Committee. The evaluation of the requests is based on the following criteria: Relevance to ECMWF's objectives, scientific and technical quality, and justification of the resources requested. Previous Special Project reports and the use of ECMWF software and data infrastructure will also be considered in the evaluation process.

Requests asking for 3,000,000 SBUs or more should be more detailed (3-5 pages). Large requests asking for 10,000,000 SBUs or more might receive a detailed review by members of the Scientific Advisory Committee.

1. Introduction and Motivation

Numerical models of the ocean general circulation are increasingly adopted tools for both climate monitoring and operational forecasting, in response to the societal demand, respectively, for a better understanding of the recent climatic changes and to serve operational activities such as oil spill monitoring, search and rescue, fishery, etc., including their use in long-range prediction systems. Data assimilation is the statistically optimal technique to combine numerical models and observations, in order to produce initial conditions for subsequent forecasts (operational oceanography) or data-aware reconstruction of the climate of the past (reanalyses). Oceanic data assimilation has dramatically advanced during the last decade, filling the methodological gap with respect to the atmospheric data assimilation community, although the typical sparseness of oceanic observations and the lack of sensors capable to sense the interior of the ocean pose several non-trivial problems to the oceanographic data assimilation problem. Hybrid ensemble-variational data assimilation methods are becoming the state-of.the-art data assimilation schemes also in the ocean data assimilation community (see for instance the outcomes of the recently held workshop on ocean data assimilation, available at https://www.ecmwf.int/en/about/media-centre/news/2021/scientists-review-advances-ocean-data-assimilation), which require advanced technologies for ensemble generation, affordable ingestion of flow-dependent variances, together with increased attention on optimal pre-processing of the observations, coupled algorithms (coupling with sea-ice, land and/or the atmosphere) and use of methods inherited from machine learning to improve selected operators in the data assimilation schemes.

CNR-ISMAR is consequently devoting several efforts in these fields, thanks to a newly formed group on oceanic modelling and assimilation, targeting in particular basin-scale ocean dynamics, stochastic ocean physics, and the impact of advanced data assimilation methods. The overarching goal of this project, which is in full continuity with the previous Special Project run by the P.I. (spitstor-2019) 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 project focusses in particular on several development tasks, which are detailed in Section 3. Section 2 provides an overview of the modelling tools. Section 4 summarizes the computational resources requested in this proposal. Section 5 puts our plans in the context of the ECMWF strategy, while Section 6 describes the background of the P.I., his colleagues collaborating to this project, and the mandate of CNR-ISMAR.

2. Modelling frameworks

The analysis system that will be used within this project includes two main elements. The ocean modelling framework is the NEMO ocean model (https://www.nemo-ocean.eu/), version 4, which also includes the sea-ice dynamical and thermo-dynamical model SI^3. NEMO is adopted by the majority of European research and operational institutions, including the ECMWF medium and long range prediction systems and its research activities, and it has been already ported to most computing architectures, including the cca/ccb clusters and the future Atos HPC architecture.

CNR-ISMAR is currently performing research activities using NEMO implemented over different configurations, and with a variational assimilation code developed over the years by the P.I. and colleagues. These tools are briefly reviewed below:

a) North Atlantic and Arctic oceans

North Atlantic-Arctic-Mediterranean configurations at either eddy-permitting (1/4° of horizontal resolution) or eddy-resolving (1/12° of horizontal resolution). These configurations are called CREG025 and CREG12 (CREG standing for Canadian REGional configurations), respectively, and are derived directly from the ORCA025 and ORCA12 grids through the removal of the north fold discontinuity of the global grids. Figure 1 shows a snapshot of surface current output in order to visualize the CREG025 domain. Specific grid data (geometry and bathymetry) were originally provided by Environment Canada, who initially developed these grids for operational purposes (see e.g. https://doi.org/10.5194/gmd-8-1577-2015). These grids present several technical and scientific advantages: they allow studying the North Atlantic climate, the impact of data assimilation techniques at mid and high latitudes, with a more computationally affordable tool than the global grid, while keeping exactly the same grid as that used operationally by ECMWF (CREG025 is as ORCA025 in the overlapping domain) and by Mercator Ocean (CREG12 is as ORCA12 in the overlapping domain) within the Copernicus programs. This implies that any development and finding can be readily ported to the global model system. The family of configuration also provides an easy way to test developments in eddy-permitting or resolving configurations, namely providing a benchmark configuration for resolutions typical of climate and operational models, respectively.

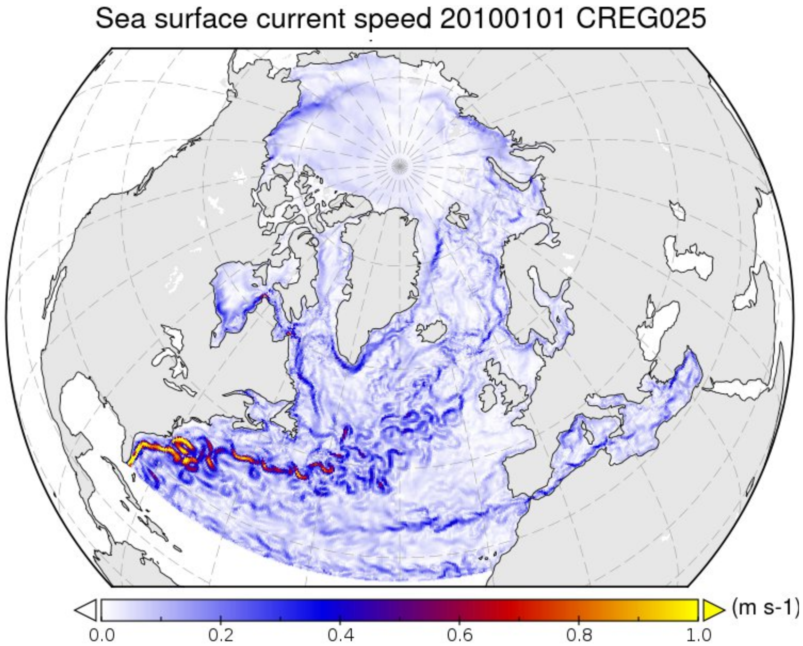

Figure 1*. A snapshot of surface current speed from the CREG025 configuration, showing its computational domain covering the North Atlantic (north of 26°N), the Mediterranean and the Arctic Seas (bounded by the Dardanelles and the Bering Strait, respectively).*

Computational costs for the CREG configurations (on cca/ccb) are as follows: 11650 SBU and 115 GB of output data per 1 year of simulation of the CREG025 configuration; 81600 SBU and 1 TB of output data per 1 year of simulation of the CREG12 configuration.

b) Global model configurations

CNR-ISMAR is also actively using global model configurations at both coarse (ORCA2) and eddy-permitting resolutions (ORCA025) for sensitivity analyses linked to stochastic physics, air-sea bulk formulations, coupled processes, wave-induced vertical ocean mixing and assimilation of historical datasets. However, we are not planning to specifically use these configurations during the project.

c) High-resolution regional configurations

CNR-ISMAR is also using high-resolution (about 1.5 km of horizontal resolution) regional configurations, mostly for benchmarking activities and sub-mesoscale data assimilation experiments. One NEMO implementation at such resolution covers the Ligurian Sea (western Mediterranean Sea) and has been used by the P.I. and colleagues to test new assimilation algorithms and the coupling with underwater acoustic models. This high-resolution regional system has also been coupled to the WRF atmospheric model for coupled data assimilation studies.

d) Variational data assimilation code

The analysis tool used in this project is the data assimilation system developed by the P.I. and colleagues during the last decade (also called OceanVar). It is a variational data assimilation system, formulated through the control variable transformation, where background-error covariances are modelled through vertical EOFs and horizontal Gaussian correlations are modelled through efficient recursive filter implementations. The system has been extended during the years to include: i) a hybrid ensemble-variational formulation, where stationary covariances are complemented by ensemble-derived ones. Such scheme uses the augmented control vector approach (a new control vector is associated to flow-dependent covariances), and the flow-dependent covariance matrix is parametric as in the stationary component (i.e. the ensemble is used to estimate flow-dependent horizontal correlation length-scales and vertical EOFs). The hybridization also include the possibility of using 3D EOFs derived from the eigen-decomposition of ensemble-derived sample covariances, which thus introduce some flavour of anisotropy in the background-error definition; ii) a tangent-linear and adjoint version of NEMO, which in turn allows running strong-constraint or weak-constraint 4DVAR experiments; iii) an experimental ETKF interface, built upon the same infrastructure; iv) several advanced observation processing techniques, such as variational quality control of in-situ observations and variational bias correction of the mean dynamic topography for altimetry assimilation; v) several balance options, including simplified barotropic balances, dynamic height formulas, purely statistical balances, and the possibility to gradually shift between different options depending on the specific area (e.g. switching to simplified barotropic balance or purely statistical one when approaching shallow areas); vi) the possibility to use observation operators derived externally (e.g. from linearised artificial neural networks).

The system has been widely used in operational global and regional analysis systems (e.g. the CMEMS MFCs in the Mediterranean and Black Seas) and historical or contemporary global and regional reanalyses (e.g. C-GLORS), including its use to initialize the sub-surface ocean within long-range coupled prediction systems (e.g. at CMCC and JAMSTEC).

Computational costs for 1 year of NEMO ocean model and assimilation experiments rise to about 28 000 and 196 000 SBU for CREG025 and CREG12 configurations, respectively, considering 5-day assimilation cycles.

3. Objectives of the project

The enhancement of regional analysis systems for climate and short-range applications is achieved through setting up four project tasks, whose scientific details are given below.

a. Stochastic physics experiments

In the numerical modelling community, stochastic physics schemes are emerging as an ensemble generation strategy for a large variety of applications, which include ensemble data assimilation, predictability studies and process-oriented energy transfer studies.

Recently, the P.I. has developed a new stochastic ocean physics package (STOPACK, Storto and Andriopoulos 2021, https://doi.org/10.1002/qj.3990), which enables the NEMO ocean model to embed three different stochastic physics schemes: stochastically perturbed parametrization tendencies (SPPT), stochastically perturbed parameters (SPP) and stochastic kinetic energy backscatter (SKEB) schemes. The schemes have been tested in high-resolution regional configuration (~2 km in the Ligurian Sea) and coarse resolution global configuration (~100 km), providing in both of them an increase of eddy kinetic energy (EKE) around the minimum scales resolved by the model at mid-latitudes, mostly due to SKEB. The schemes have been also implemented in the prepIFS system and are currently under testing at ECMWF, thanks to the support provided by the C3S_321b project. Figure 2 summarizes some preliminary results obtained with the ECMWF ocean model configurations at both 1° (ORCA1 grid) and 1/4° (ORCA025 grid) of spatial resolution. The plot indicates that stochastic physics schemes increase the eddy kinetic energy at mid latitudes, consistently for both configurations.

The increase of mid-latitude EKE is an important result, provided that most ocean model components used in climate projections (such as CMIP6) are still run at eddy-permitting resolution (~1/4° of horizontal resolution), which in turn hampers the eddy population and lifecycle to be fully resolved by the model at mid-latitudes.

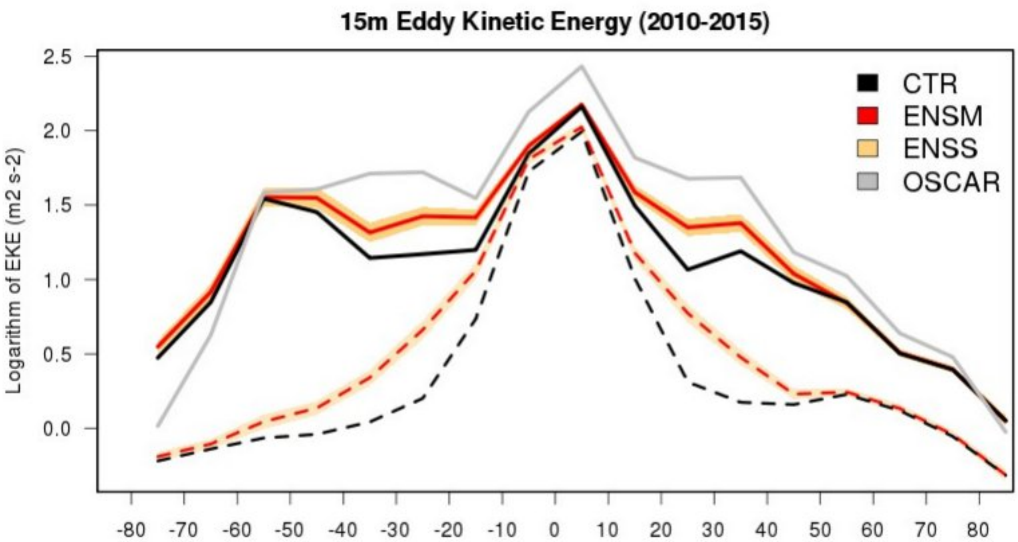

***Figure 2**. Eddy Kinetic Energy (EKE) at 15 m depth in logarithmic scale calculated over 2010-2015 from ORCA025 (solid lines) and ORCA1 (dashed lines) simulations, performed with the ECMWF ocean model system within the C3S_321b project. CTR: unperturbed deterministic simulation; ENSM: ensemble mean of EKE from stochastic physics experiments; ENSS: ensemble standard deviation of EKE from stochastic physics experiments; OSCAR: EKE from satellite-derived surface current data.*

 This form is available at:
http://www.ecmwf.int/en/computing/access-computing-facilities/forms

The main purpose of the task is to test the ability of the STOPACK scheme, and its SKEB component in particular, to provide an eddy activity as intense as eddy-resolving ocean model configuration. To this end, we plan to run several experiments using CREG025 with and without stochastic physics, and with other perturbation methods (inherent uncertainty through perturbation of initial conditions, and lateral boundary forcing perturbations). The CREG12 configuration will be used as high-resolution benchmark, to assess the increase of signal resolution due to the stochastic physics with respect to an eddy-resolving configuration. The experiments will be forced by ERA5 at the surface and by the Mercator Océan reanalyses at the lateral boundaries, while they will include a sea surface nudging to observed sea surface temperature and salinity data.

This section of the project will definitely shed lights on the possibility of using stochastic physics to enhance eddy activity in eddy-permitting ocean model configurations. As such resolution is at the moment still the highest one in use for climate applications (long-term prediction systems and CMIP projections), the benefits of our investigation go far beyond the analysis of the ensemble generation and include a large number of climate applications.

Resources needed for this task: a total of 7 000 000 SBU are required to run three 3-member ensemble experiment with CREG025 and one deterministic experiment with CREG12, both covering the latest 40 years (from 1981 onwards).

b. Altimetry data assimilation at high latitudes

The inter-annual changes of the Arctic Ocean features are well-known proxies of the global climate change, affecting the global climate through specific processes (e.g. dense water formation, meridional heat redistribution, Arctic amplification, etc.). The ocean circulation at high latitudes has significantly changed during recent decades, with an enormous impact on the socio-economic activities of the Nordic populations.

Monitoring the Arctic environment is however non-trivial: the Arctic observing network is notably lacking the capability to provide a full picture of the changing ocean due to technological and economical limitations to sample the seawater beneath the ice or in the marginal ice zones. This leads to the obvious need of optimizing the exploitation of data from space-borne sensors. Among these, altimetric radars measuring the sea level at millimetric precision have revolutionized our knowledge of the oceanic circulation, from more than 2 decades, at a large spectrum of scales ranging from the 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, like SWOT) 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. This task is devoted to the exploitation of such satellite altimetry data in high-latitude regions. We will 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 to the altimetry-blind experiments, in order to assess the added value of the enhanced altimetry reprocessing at high latitudes. This comparison can motivate the assimilation of the high-resolution altimetry data in ocean analyses and re-analysis for the Arctic Sea, and provide fundamental recommendations on the optimal way to use altimetry data for high-latitude process forecasting.

Resources needed for this task: a total of 3 200 000 SBU for four experiments (control, conventional altimetry, new altimetry and possibly wide-swath altimetry, each of them run over a 3-year period 2019-2021) run with both CREG025 and CREG12 configurations.

c. Coupled data assimilation algorithms

Coupled data assimilation is emerging as a target approach for Earth system prediction and reanalysis systems. Coupled data assimilation may be indeed able to minimize unbalanced initial conditions between different Earth system components and maximize the inter-medium propagation of observations. There is also the expectation that coupled data assimilation algorithms may significantly benefit the ingestion of satellite data in numerical forecasts through the formulation of coupled observation operators.

In this task, we plan to test strongly coupled assimilation algorithm for the optimal initialization of the ocean/sea-ice coupled system (in the CREG025 NEMO model configuration), and ocean/atmosphere coupled system (in the Ligurian Sea model system, where NEMO is coupled to WRF). We will test several strategies to formulate cross-medium covariances and balance operators, starting from an extension of the data assimilation systems to include multi-component variables through the use of multi-variate EOFs, which is already supported experimentally by our data assimilation system, to more sophisticated methods relying on cross-component balances from deep learning algorithms. The P.I. has indeed recently worked (https://doi.org/10.1175/MWR-D-20-0320.1) on the use of linearised artificial neural networks (ANN) to build complex observation operators, and he is planning to extend this approach to the coupled assimilation problem. This latest approach may be tested once the new Atos HPC system is available, as the Tensorflow library, used for the online ANN computation, is not available on the cca/ccb clusters.

In the CREG025 system, these ideas will be exploited to test the benefits of the ocean-ice coupled assimilation, which is a lower-dimensional assimilation problem compared to the ocean-atmosphere one, complicated however by the fact that the SI^3 sea-ice model embeds multi-category sea-ice variables which are in fact not observed. We plan to test the strongly coupled assimilation, either with EOFs or ANNs, in the problem of assimilating either sea-ice concentration only or sea-ice concentration and thickness observations, looking at the impact on the sub-surface ocean variables. This will be tested in multi-annual experiments, in order to assess the assimilation over different seasons and conditions.

In the Ligurian Sea system, the same coupled assimilation strategies will be adopted to test the impact of coupled assimilation of ocean observations from in-situ and remote sensing data into the coupled ocean-atmospheric forecasts, and validated against independent atmospheric observations. This can be tested for short periods (3-months, from August to mid-November), where observational data for assimilation and validation are available from campaigns.

Resources needed for this task: a total of 2 700 000 SBU for this task, divided as follows:

CREG025: 300 000 SBU per each 10-year assimilative experiment (for a total of 1 700 000 SBU considering preliminary tests)

Ligurian Sea: 200 000 SBU per each three-month coupled experiment, for a total of 4 experiments plus additional tests (total 1 000 000 SBU).

d. Weak-constraint data assimilation and algorithmic developments

Advancing oceanographic data assimilation can benefit several applications ranging from climate monitoring and predictions to operational environmental forecasting. Reaching the level of complexity of Numerical Weather Prediction (NWP) schemes is also required for future strongly coupled assimilation applications. Recently, we have formulated an oceanographic weak-constraint four-dimensional variational data assimilation scheme. Model-error covariances are i) flow-dependent, through the use of ensemble runs with stochastic physics and unperturbed initial conditions; ii) defined in EOF space, to allow anisotropic three-dimensional structures substantially different from the background-error covariances; iii) composed by the sum of non-linear and linear model errors, the latter aiming at capturing the uncertainty of the tangent-linear approximation.

Such formulation has been tested in low-dimensional models (Lorenz 96 model), providing promising results compared to strong-constraint 4DVAR or weak-constraint 4DVAR with stationary model-error covariances. Figure 2 summarizes the results of these Lorenz 96 assimilation experiments, performed through OSSE where only one variable every two model variables is observed. The tests suggest that the use of flow-dependent model errors significantly improves the forecast skill scores. The formulation is now embedded into the variational data assimilation system presented above, but needs to be tested with realistic settings such as the CREG025 configuration. Slightly different flavours of the scheme can be envisaged and will be tested depending on the results (e.g. inflation of flow-dependent model errors to account for the limited ensemble size, etc.).

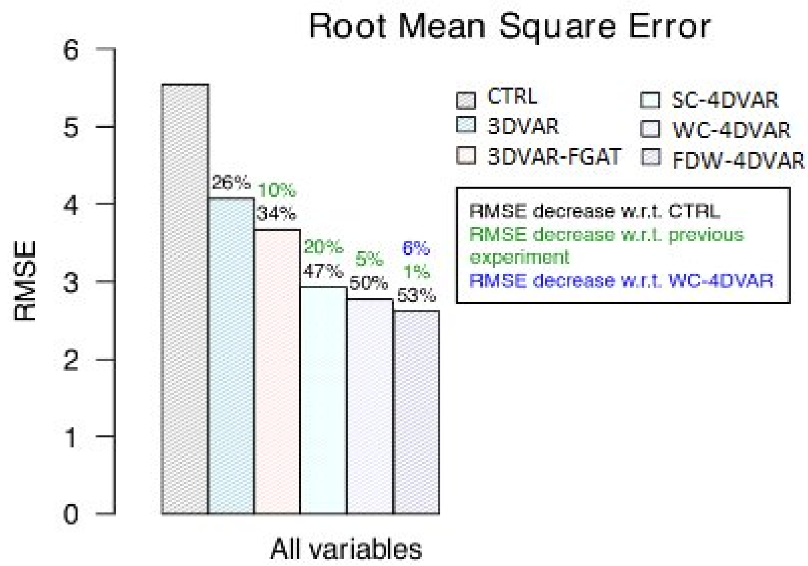

***Figure 3**. RMSE of OSSE performed with the Lorenz 96 model (against the truth over all variables) for different data assimilation schemes and the free-running model experiment: SC-4DVAR (strong-constraint 4DVAR), WC-4DVAR (weak-constraint 4DVAR, with stationary model errors), FDW-4DVAR (weak-constraint 4DVAR, with flow-dependent model errors). Percentage RMSE decrease on top of the bars are as indicated in the caption.*

Additionally, we have recently rewritten a simplified tangent-linear and adjoint (TA/AD) code for the NEMO ocean model, which allows for several degree of complexity, as for instance the possibility of having ocean currents as prognostic or diagnostic variables, the latter halving the computational costs associated with the TA/AD run and therefore with the 4DVAR scheme. The objective of this task is therefore a cost-benefit analysis of increasing complexity data assimilation schemes, which can contribute to the design of next-generation ocean analysis systems.

Resources needed for this task: about 4 500 000 SBU, which corresponds to five experiments each covering three years of simulations with strong constraint 4DVAR, stationary or differently-formulated flow-dependent model error covariances in the weak-constraint 4DVAR formulation (each 1-year 4DVAR experiment is estimated to cost 300 000 SBU), plus preliminary tests and optimization runs.

4. Overall computational resources requested to run the project tasks

The total estimated costs of the project is 17 400 000 SBU , which we equally divide into the three-year period of the project as it is difficult at the moment to foresee clear priorities. The actual priority of the tasks proposed here may also depend on the funding of specific research lines. The tentative plan is reported below in Figure 4, using a Gantt diagram. The first three tasks are spread around the 3-

year period of the project, while the weak-constraint data assimilation is planned to be developed throughout the project period.

We are planning to port both the NEMO model with our in-house developments (stochastic physics, and optimized relaxation routines) and our data assimilation code to the new Atos HPC system, possibly using the TEMS testing system beforehand, and assessing their scalability performances with the different compiling environments (amd versus intel compiling environments).

Note also that CNR-ISMAR has in-house high-performing data storage and archiving facilities, and it is connected through the Géant international network, therefore we will be able to move and archive data locally if our data occupancy exceeds the allocated one in the ECFS archive.

Task	1st Year		2nd Year		3rd Year		Description
	I	*II*	*I*	*II*	*I*	*II*	
a							Stochastic physics experiments
b							Altimetry data assimilation at high latitudes
c							Coupled data assimilation algortihms
d							Weak-contraint data assimilation

Figure 4. *Gantt diagram of the current project*

5. Contribution to the ECMWF strategic plan

Earth system data assimilation (Task c), ensemble forecasting and model errors (Task a), assimilation of new observations (Task b), use of AI methods (Task c), and data assimilation algorithmic developments (Task d) are all explicitly mentioned in the ECMWF Strategy 2021-2030 (http://dx.doi.org/10.21957/s21ec694kd). We are therefore confident that our research project will not only contribute to the CNR-ISMAR research priorities, but will benefit foreseen developments at ECMWF and will strengthen the international collaborations between CNR-ISMAR, ECMWF, Mercator Océan International (CNR being a member), the NEMO, NEMOVAR and EC-Earth (CNR being part of it) consortia.

6. The research team and CNR-ISMAR activities

The P.I. started his scientific career working on advanced methods for data assimilation in meteorological limited area models and algorithms for new satellite sensors, first at CNRM/Météo-France and then at the Norwegian Meteorological Institute (Met.no). During those years, he deepened aspects of optimal assimilation of satellite radiances, assimilation of cloudiness data, new metrics for assessing the impact of observations and ensemble data assimilation. After obtaining the doctoral degree from the University of Rome “La Sapienza” in close collaboration with Met.no, he moved to Bologna (Italy), shortly at the National Institute for Geophysics and Volcanology (INGV) as post-doc and later at the Euro-Mediterranean Center for Climate Change (CMCC), spending nine years, with qualifications from post-doc to junior scientist and finally scientist. The research has focused on developing and improving ocean data assimilation methods, especially, but not only, for reanalyses (global reconstructions of the ocean climate evolution). In particular, he devoted research efforts to altimetry assimilation, assessment and comparison of ocean reanalyses, observation impact studies, advanced assimilation methods (hybrid and four-dimensional schemes) and climate-oriented investigations of ocean heat content and sea level variability. Several studies of mines have attracted interest from the international climate community; he thus became responsible for a few reanalysis inter-comparison studies, and later he was selected to chair the CLIVAR panel on Global Synthesis and Observations (GSOP). He has also significantly contributed to the design of operational forecasting systems (in the global ocean, Black and Mediterranean Seas), seasonal and long-term prediction systems (especially for the CMCC and JAMSTEC/Japan seasonal prediction systems), and centennial reanalysis systems, with national and international collaborations (including ECMWF, Mercator Océan/France, JAMSTEC/Japan, NMEFC/China), testified by the numerous joint publications. For the ERA-CLIM2 project (FP7), he developed air-sea coupled data assimilation schemes for intermediate complexity Earth system models. For CMEMS (the Copernicus Marine Environment Monitoring Service), he led (P.I.) the SOSSTA research project on optimal SST data assimilation, and the GLO-RAN global ocean reanalysis production. He was also selected as Adjunct Professor through public call to teach Data Assimilation at the Climate Change Doctoral School of the University of Venice, for four academic years. Since 2018, he has been working as Research Scientist at the NATO/STO Centre for Maritime Research and Experimentation (CMRE), La Spezia (Italy), where his work focused on coupled oceanic-acoustic data assimilation, ocean stochastic physics, exploitation of the synergy of in-situ and remote sensing data, advanced ensemble-variational hybrid methods for data assimilation, and use of AI algorithms in data assimilation. On July 2020, he won a public call as Research Scientist at the National Research Council of Italy (CNR), and moved to the Institute of Marine Sciences (ISMAR), where he currently works. Recent activities focus on centennial reconstructions of the ocean climate, optimal assimilation of remotely sensed data in a coupled modelling framework, and weak-constraint assimilation methods. He coauthors more than 75 publications in peer-reviewed journals with impact factor, with a current h-index of 22 (Google scholar).

Chunxue Yang, the co-P.I., started her research career in the ocean and climate science field, doing her bachelor thesis on the teleconnections of the Atlantic Meridional Overturning Circulation. After two years of Ph.D study in China, she moved to Texas A&M University, USA, to pursue the Ph.D degree on the long term change of the El Niño Southern Oscillation (ENSO) by using an ensemble ocean reanalyses and climate coupled models. After her PhD study, she started the first post-doctoral research activities on producing and validating an ensemble of historical ocean reanalyses at the Euro-Mediterranean Center for Climate Change (CMCC), Bologna, Italy. Afterwards, she moved to CNR-ISAC, Bologna, to work on climate coupled models (EC-Earth) and develop climate simulations for the IPCC CMIP6 activities. She has also spent a few months in a consultant company (Fugro, UK) to produce ocean forecast simulations. After she moved to CNR-ISMAR, Rome, where she has pursued research activities in the North Atlantic ocean and tropical oceans. Currently in CNR-ISMAR, Rome, she is leading two projects funded by the European Copernicus program. Within these two projects, she is coordinating her own research group composed of three postdocs and one Ph.D student who are working on the ocean and climate dynamics. She has also been leading the intercomparison task in the international Group of High Resolution Sea Surface Temperature (GHRSST) and being a Science Member of the GHRSST science team. Her research experiences are closely related to the research topic proposed in this project.

Eva Le Merle is a postdoc at CNR-ISMAR. She graduated at CNRS-LATMOS (Paris) with a thesis on remote sensing of marine waves. Jacopo Busatto is a Ph. D. student at the University of Rome Tor Vergata, working in close collaboration with CNR-ISMAR on high-frequency coupled air-sea processes. Both are already using the NEMO ocean model in coarse resolution configurations at the local CNR-ISMAR clusters.

Brief overview of CNR-ISMAR

The National Research Council of Italy (CNR) is the largest research institution in Italy, committed to support scientific and technological research. The Institute of Marine Sciences (ISMAR) is devoted to inter-disciplinary research on physical, biological, chemical and geological oceanography, including paleoceanography studies and the role of the oceans in the contemporary changing climate.

CNR-ISMAR has a long-standing expertise in ocean observations, modelling of the sea state and design and exploitation of coastal and open ocean marine observatories. CNR-ISMAR has recently joint the Mercator Océan International (MOI) company, which is the entrusted entity by the European Commission to run the Copernicus Marine Environment Monitoring Service (CMEMS), thus interconnecting the national research activities in the context of European research and marine services. Among many activities, CNR-ISMAR coordinates the CMEMS assembly centres for the dissemination of sea surface temperature and ocean colour observations. Furthermore, CNR-ISMAR coordinates the Copernicus Climate Change Service (C3S) C3S_511 project about the "Independent assessment of essential climate variables". CNR-ISMAR has also developed an unstructured grid hydrodynamic model used for storm surge studies (SHYFEM) and has significantly contributed to the development of the WAM wave model.

Ten selected publications of the P.I. and the research team (about the topics proposed here)

1. Storto, A., De Magistris, G., Falchetti, S., & Oddo, P. (2021). A Neural Network–Based Observation Operator for Coupled Ocean–Acoustic Variational Data Assimilation, Monthly Weather Review, 149(6), 1967-1985. Retrieved Jun 14, 2021, from https://journals.ametsoc.org/view/journals/mwre/149/6/MWR-D-20-0320.1.xml
2. 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
3. Yang, C., Christensen, H.M., Corti, S. et al. The impact of stochastic physics on the El Niño Southern Oscillation in the EC-Earth coupled model. Clim Dyn 53, 2843–2859 (2019). https://doi.org/10.1007/s00382-019-04660-0
4. Storto, A., Oddo, P., Cozzani, E., & Coelho, E. F. (2019). Introducing Along-Track Error Correlations for Altimetry Data in a Regional Ocean Prediction System, Journal of Atmospheric and Oceanic Technology, 36(8), 1657-1674. Retrieved Jun 14, 2021, from https://journals.ametsoc.org/view/journals/atot/36/8/jtech-d-18-0213.1.xml
5. Storto, A., Martin, M. J., Deremble, B., & Masina, S. (2018). Strongly Coupled Data Assimilation Experiments with Linearized Ocean–Atmosphere Balance Relationships, Monthly Weather Review, 146(4), 1233-1257. Retrieved Jun 14, 2021, from https://journals.ametsoc.org/view/journals/mwre/146/4/mwr-d-17-0222.1.xml
6. Storto, A., Oddo, P., Cipollone, A., Mirouze, I. and Lemieux-Dudon, B. (2018) Extending an oceanographic variational scheme to allow for affordable hybrid and four-dimensional data assimilation. Ocean Modelling, 128, 67– 86. https://doi.org/10.1016/j.ocemod.2018.06.005.
7. Storto, A., 2016: Variational quality control of hydrographic profile data with non-Gaussian errors for global ocean variational data assimilation systems. Ocean Modell., 104, 226–241, https://doi.org/10.1016/j.ocemod.2016.06.011.
8. Storto, A., Masina, S., & Dobricic, S. (2014). Estimation and Impact of Nonuniform Horizontal Correlation Length Scales for Global Ocean Physical Analyses, Journal of Atmospheric and Oceanic Technology, 31(10), 2330-2349. Retrieved Jun 14, 2021, from https://journals.ametsoc.org/view/journals/atot/31/10/jtech-d-14-00042_1.xml
9. Storto, A., Dobricic, S., Masina, S., & Di Pietro, P. (2011). Assimilating Along-Track Altimetric Observations through Local Hydrostatic Adjustment in a Global Ocean Variational Assimilation System, Monthly Weather Review, 139(3), 738-754. Retrieved Jun 14, 2021, from https://journals.ametsoc.org/view/journals/mwre/139/3/2010mwr335
10. Storto, A., and Randriamampianina, R. (2010), Ensemble variational assimilation for the representation of background error covariances in a high-latitude regional model, J. Geophys. Res., 115, D17204, doi:10.1029/2009JD013111.