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 …07 / 2021 – 06 / 2022………………………………

Project Title: ………………………………………………………….……
Black Sea Ensemble forecasting system
………………………………………………………….……

Computer Project Account: …spbebulc……………………………………………….……

Principal Investigator(s): …Luc Vandenbulcke…………………………………….……

………………………………………………………….……

Affiliation: …Université de Liège, Belgium……………………….……

Name of ECMWF scientist(s) collaborating to the project (if applicable) ……/…………………………………………………….……

………………………………………………………….……

Start date of the project: …01/01/2021……………………………………….……

Expected end date: …31/12/2023…………………………………………….……

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

(if applicable)
Please answer for all project resources

		Previous year (2021)		**Current year (2022)**	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	650.000	1.031.115	2.400.000	/
Data storage capacity	(Gbytes)	2.000	/	5000	/

Summary of project objectives (10 lines max)

The aim of the project is to convert a deterministic NEMO implementation used in the CMEMS Black Sea Forecasting Center, into an ensemble model. The ensemble members will have different initial conditions, scalar model parameters e.g. in the turbulence module, the light penetration scheme, the surface bulk formulae, etc The uncertainty obtained from the ensemble will be quantified and the ensemble reliability and consistency will be studied.
Next, SST and in situ temperature and salinity will be assimilated using an EnKF method.
The impact of the uncertainty coming from the physical model, on a coupled biogeochemical will also be investigated. In particular, it will be analyzed if spurious adjustments of the vertical velocity lead to artificial nutriment upwellings.
Finally, some biogeochmical variables will be assimilated in the coupled biogeochemical model as well.

Summary of problems encountered (10 lines max)

……………………………………………………………………………………………………
not applicable
………………………………………………………………………………………………………
………………………………………………………………………………………………………
……………………………………………………………………………………………………....

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

The project has not started on the ECMWF facility yet. The models preparation is ongoing (at Univ. Liege premises). The first runs are expected to take place on the ECMWF facilities during the second half of 2021.
The project is not late with respect to the schedule, as only a limited amount of SBU were scheduled on the ECMWF computers during the projects first year

List of publications/reports from the project with complete references

……not applicable
………………………………………………………………………………………………………
………………………………………………………………………………………………………
………………………………………………………………………………………………………

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.
……………………………………………………………………………………………………

During the first year (2021), the NEMO-BAMHBI model configuration used in the Copernicus Marine Environmnent Monitoring Service (CMEMS) Black Sea Monitoring and Forecasting Center (BS-MFC) was installed in the special project directory on XC40, and compiled with the intel compilers.

This deterministic model was then modified to become a probabilistic model, and as such the first major objective of the special project was fulfilled. The other objectives were not yet addressed: the implementation of data assimilation using the Ensemble Kalman Filter (incl. studying the generation of spurious vertical currents by data assimilation), the impact of uncertainty on the ecological and biogeochemical dynamics in the coupled biogeochemical model, and the data assimilation of biogeochemical variables.

The probabilistic model was implement using NEMO's stochastic (STO) module based on autoregressive processes.
The module is foreseen to use NEMO as a single executable running (at once) multiple members (double parallelisation on members and in the spatial domain). This could be implements on the ECMWF HPC, but not in an operational context at BS-MFC HPC. In order to be able to run different members in different (PBS or slurm) jobs, the STO random number generator's seed code was modified ; the seed can now be based on the current unix time.

The STO module and its upgrade implemented during the SCRUM project (using also computing time on the ECMWF HPC) allows to generate autoregressive processes as 2D or 3D with specified spatial correlation lenghts (expressed in spatial grid points) and correlation time expressed in timesteps). These fields of random numbers, generated in the STO module, are then used as additive or multiplicative perturbations that can be applied elsewhere in the code, e.g. to the equation of state (EOS), to the trend computation (TRD), to the drag coefficient, to the wind forcing field.

Additionally, we implemented a 0D autoregressive process in the STO source code, based on the code for 2D and 3D processes.

The following supplementary perturbation methods were then implemented to generate the ensemble members in the NEMO physical model.
* river discharge of freshwater (based on 2D autoregressive processses with a spatial correlation length of 250km, so that the different mouths of the Danube river are perturbed similarly)
* atmospheric fields of wind, T2M, TCC, precip, based on the EOF method implemented in the nemo surface fluxes bulk formula module. A previously-computed series of EOFs is loaded during NEMO initialization, and at each time, a random linear combination of these EOFs is added to the true ECMWF forcing field. Regarding wind, bi-variate EOFs are used (for U10 and V10).

The MY_TRC module allows to couple a biogeochemical model in NEMO. The biogeochemical model can now also access the random fields generated in STO and use them to perturb biogeochemical parameters and processes. Until now, perturbations were implemented for:
* river discharge of nutriments (2D stochastic field)
* atmospheric deposition of nutriments on the sea surface (2D stochastic field).

These perturbations were tried out individually, and the intensity and correlation lengths were tuned in order for the perturbations to have realistic magnitudes and spatial patterns.
For example, regarding the wind, the perturbation (i.e. the random combination of EOFs) should have a similar probability distribution as the real uncertainty on the ECMWF wind product (see Fig. 1). The latter is estimated by comparing ECMWF winds with real measurments.

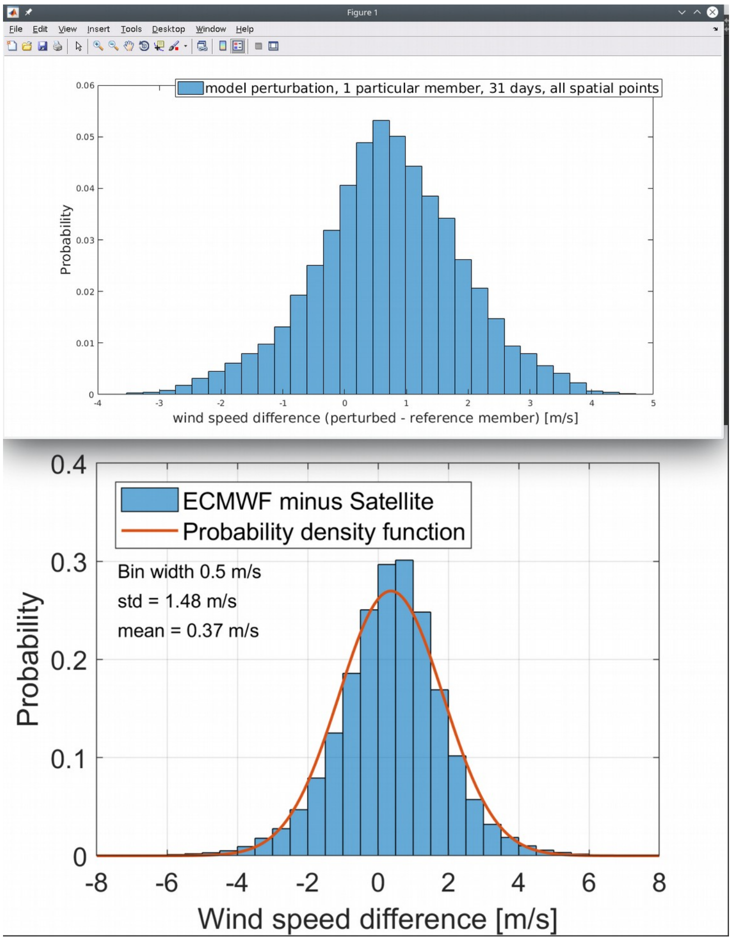

Figure 1. Upper panel: probability distribution of the generated perturbation for 1 particular member, for a 1-month simulation, and for all points in the spatial domain. Lower panel: probability distribution of the ECMWF ERA5 winds minus satellite-observations of surface winds, during the same month

For other variables, the probability distributions of the pertubations were checked qualitatively.

Then, using all these pertubations at once, an ensemble of 20 members was simulated during 1 month. The simulations were carried out on the XC40 machine during December 2021; however preleminary tests during November 2021 also consumed SBUs.

The probability distributions of the model variables in the ensemble were checked to be realistic, e.g. the sea surface salinity variability is large in the Danube plume due to uncertainty of river discharge.

In general, uncertainty in the physical model is realistic, even for variables only indirectly perturbed, such as the mixed-layer depth (MLD) or the Cold Intermediate Layer cold content (CiLCC). The delay (after model start) after which perturbations start to appear in different variables varies. For example, and as expected, the spread in CilCC only grows after convection events (more or less CIL formation) or summer heating (more or less CIL destruction), whereas salinity in the river plumes varies very rapidly.

During the first half of 2022, more perturbation methods were implemented in NEMO 4.0.6 but not yet tried out on the ECMWF HPC.

During the second half of 2022 we would like to move the modelling system from CCA/CCB to the new Atos system AB.

Presentations based on the BSENSE results:
Results from the preliminary ensemble run were presented during the CMEMS working group meetings (Ensemble runs and Bio Data Assimilation) on 15-16 June 2022.
The results will be presented again during the ODESSA kick-off meeting (July 2022).

New projects based on BSENSE:
ODESSA is a new CMEMS service evolution research project that was recently approved for funding. Its aim is to develop the ensemble forecasting capabilities of the Black Sea Monitoring Center (BS-MFC), and it will build on the methods developed, and simulations realized, during this Special Project.