SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

Project Title:	Black Sea Ensemble forecasting system
Computer Project Account:	spebulc
Start Year - End Year :	2021 – 2023
Principal Investigator(s)	Luc Vandenbulcke
Affiliation/Address:	MAST research group Université de Liège Belgium
Other Researchers (Name/ Affiliation):	Prof. Marilaure Grégoire (MAST / ULiege)

The following should cover the entire project duration.

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 thecoupled biogeochemical model as well.

Summary of problems encountered

(If you encountered any problems of a more technical nature, please describe them here.)

No particular technical problems were encountered during the project, although the project period covered the ECMWF computer change to Bologna.
When minor question appeared, ECMWF contact points were very helpful.

Experience with the Special Project framework

(Please let us know about your experience with administrative aspects like the application procedure, progress reporting etc.)

The procedures are clear, and not cumbersome.
One aspect which is more difficult to manage is the estimation of required computing time (at the project proposal stage). Although preliminary simulations were realized on the university's HPC, changes in the model itself, and differences between clusters, make the estimation complicated.

Another trap to avoid is to delay the simulations until the last minute (as there are continuous improvements to the model). Luckily the jobs submitted to the cluster launched almost immediately, without spending time in the queue.

Summary of results

(This section should comprise up to 10 pages, reflecting the complexity and duration of the project, and can be replaced by a short summary plus an existing scientific report on the project.)

This section is replaced by the scientific report submitted to the EU's Copernicus Marine, in the framework of the "ODESSA" service evolution project ; and joined at the bottom of this report.

In summary, during the project, an existing online-coupled model of the Black Sea physics and biogeochemistry based on Nemo 4.2, Bamhbi, and a radiative transfer model, was upgraded to an ensemble of models.
Using a degraded resolution (15km instead of 2.5 km), 40 small (10-member) ensembles with perturbations of individual model components were simulated for 1 year. Then, a 100-member ensemble with all perturbations together was also simulated.
The same ensemble was then simulated at nominal resolution (2.5 km).
Finally, in order to study predictability, the ensemble spread was reduced using a ensemble Kalman Filter both in the physical and biogeochemical parts, and a subsequent 1-month simulation was run again.

The cumulative effect of the different physical and biogeochemical perturbations was studied, and it was noted that perturbations of physics had a large effect on biogeochemistry (e.g. perturbing the incoming light can change the dominant plankton species). But the opposite is also true, e.g. perturbations on the plankton abundance leads to non-negligeable changes of sea surface temperature.

The creation of the ensemble is satisfactory, and will eventually lead to a stochastic system for the Black Sea biogeochemical forecasts in the Copernicus Marine service.

List of publications/reports from the project with complete references

* CMEMS service evolution 21036-COP-INNO SCI "ODESSA" final report (30/04/2024) (annexed below)
* L. Vandenbulcke, L. Macé, J.M Brankart, P. Brasseur, M. Grégoire. A stochastic forecasting system of the Black Sea (in preparation)

Future plans

(Please let us know of any imminent plans regarding a continuation of this research activity, in particular if they are linked to another/new Special Project.)

There are no imminent plans to submit a new Special Project to ECMWF next year.
However it is very likely that this will be realized in the next years, in order to allow the further development of the (ensemble) model.

ODESSA

Final report

Authors: L. Vandenbulcke, L. Macé, P. Brasseur, J-M. Brankart, M. Grégoire
Scientific advisor: A. Barth
Date: 30/04/2024
Final report

Foreword

This document is the final report for projects selected through the Service Evolution 21036-COP-INNO SCI. Final reports shall provide a comprehensive description of the study results and a detailed analysis of the potential impact of these results on the Copernicus Marine operational service.

As a guideline, the main body of the report should not exceed 20 pages. Appendices can be added for instance to detail methodologies, provide a peer-reviewed manuscript related to the project. The main body of the report should avoid technicalities and too many acronyms to ease its reading by TAC and MFC scientists, STAC members, etc. Figures should be integrated in the text (not at the end of the report). Guidelines are provided for mandatory sections of the report. Please delete them before sending your report.

A single report is asked per project, even though the project could be carried out by a consortium. It should be sent by email to sguinehut@mercator-ocean.fr.

Team presentation

The ODESSA project is carried out with contributions from the following persons:

- Luc Vandenbulcke, MAST, ULiege, Belgium
- Marilaure Grégoire, MAST, ULiege, Belgium
- Loic Macé, MAST, ULiege, Belgium and UGA, MEOM, Grenoble, France
- Pierre Brasseur, UGA, MEOM, Grenoble, France
- Jean-Michel Brankart, UGA, MEOM, Grenoble, France
- Alexander Barth, GHER, ULiege, Belgium

Contact: luc.vandenbulcke@uliege.be

Executive summary

The ODESSA project aimed at proposing a framework to transform a deterministic biogeochemical forecasting system in a stochastic one. Practically, the coupled model from Black-Sea BGC forecasting system was transformed into an ensemble of models.
During the project, new perturbation methods were implemented in NEMO's stochastic subroutines; multiple physical and biogeochemical model components were perturbed, and ensembles were simulated. The ENSDAM package to compute stochastic metrics was updated, and subsequently, stochastic metrics were computed in order to validate the ensemble, with respect to real observations (mainly BGC-ARGO). The predictability of the system was also studied.
A pilot uncertainty product is proposed for delivery in the CMEMS catalogue, but practical details still need to be decided (variables, resolution, statistical moments...)
Table 1 lists the specific objectives from the proposal, as well as the results at the end of the project and the potential impact on CMEMS.

Specific Objectives	Status	Scientific results	Potential impact on Copernicus Marine Service
1.1 Establish the list of model components affected by uncertainty	100 %	The list was established through a review of literature and the SESAME EU-project outcome	MFCs going to run an ensemble BGC model are able to compare the perturbation candidates with the ones listed
1.2 Create a database of relevant observations to validate the stochastic model	100 %	Observations from CMEMS in situ TAC, SST TAC and OC TAC are assembled and pre-processed	/
1.3 Develop the NEMO STOchastic module to generate an ensemble with all perturbations from objective 1.1 above	100 %	The processes identified above are perturbed using relevant methods.	MFCs using NEMO can directly use our perturbation schemes. The schemes are also shared with the NEMO STO working group for inclusion in future NEMO versions
2.1 Run small ensembles with 1 perturbation	100 %	Effect of individual perturbations is assessed	MFCs can choose whether to perturb given model parameters based on the impact on the ensemble spread for some model variables.
2.2 Stochastic assessment	100 %		
3.1 Run the complete ensemble (low resolution)	100 %	An ensemble of models affected by our best-estimate of input uncertainty, generates our best estimate of forecast uncertainty	MFCs can compare the magnitude of their forecasted uncertainty with the one obtained in ODESSA.
3.2 Run the complete ensemble (high resolution)	100 %		
3.3 Stochastic assessment	100 %		
3.4 Predictability analysis	80 %		
3.5 Production of a pilot uncertainty product	100 %	Uncertainty estimates are ready to be delivered to CMEMS	A pilot uncertainty product can be made available to users

Table 1. Specific objectives of the ODESSA project

Table of contents

1 Introduction

The overarching aim of ODESSA was to propose a framework for evolving the deterministic forecasting systems currently used, into a modeling system that explicitly simulates the main uncertainties in the boundary conditions and in the physical and BGC models, in a generic and multiform way (so that it can be applied to multiple CMEMS MFCs).

While doing so, the following scientific questions need to be answered: (1) what are the sources of uncertainties that matter for the forecasting of the green ocean, and how can they be represented? (2) What is the predictability of the obtained ensemble system? How does it evolve with the lead time? (3) How does a coarse resolution ensemble system compare with a high-resolution deterministic system, for different dynamical regimes? (4) Given the uncertainty on the optical forcing fields and parameters, what is the benefit of a full radiative transfer model?

This resulted in the specific objectives listed below.
1.1 Establish the list of model components affected by uncertainty
1.2 Create a database of relevant observations to validate the stochastic model
1.3 Develop the NEMO STOchastic module in order to generate different types of random perturbations, adapted to each model component discovered in objective 1.1
2.1 Run small ensembles with perturbation of 1 specific model component
2.2 Stochastic assessment of results of task 2.1
3.1 Run simulations using the fully-perturbed ensemble at low resolution (15km)
3.2 Run simulations using the fully-perturbed ensemble at nominal resolution (2.5km)
3.3 Stochastic assessment of results of tasks 3.1 and 3.2
3.4 Predictability analysis
3.5 Production of a pilot uncertainty product

All developments and implementation are realized in the context of the Black Sea BGC forecasting system, which is based on the NEMO model (first version 4.0.6, later version 4.2) online-coupled with the BAMHBI biogeochemical model. During ODESSA, a radiative transfer model (also called optical model) was also added to NEMO-BAMHBI.

The reason for running ensembles at a degraded horizontal resolution is related to computing resources, but also to study the impact of resolution on model perturbations.

The specific objectives listed above are described in detail in section 2. During ODESSA, all were addressed.

2 Scientific results

The scientific results are described in the order of the project tasks.

2.1 Task 1.1 – perturbation list

Task 1.1 consisted in generating a list of all components of the modelling system, affected by uncertainty. This includes boundary conditions, numerical schemes, and the physical and biogeochemical model parametrizations and/or parameter values. The identified perturbation candidates are described below.
The perturbed boundary conditions include the atmospheric forcing fields: wind, air temperature, cloud coverage (if one computes the incoming solar flux based on the position and date,time) or incoming solar radiation (if the incoming solar flux is read from files, e.g. from an atmospheric model), and precipitations.
Boundary condition perturbations also include the river freshwater flux, the river nutrient flux, and the atmospheric nutrient deposition flux. Other examples (not considered in ODESSA) are the grid itself, the lateral diffusion operator, the horizontal pressure gradient computation.
The physical model perturbation, and numerical scheme perturbation, concerned the bottom drag coefficient and the equation-of-state.
30 biogeochemical model parameters were identified to undergo perturbations. They were selected based on sensitivity studies in Gregoire et al (2008) describing the used BGC model called BAMHBI, in Wang et al (2020), in Garnier (2016), in Capet (priv. com. 2022) and in the SEAMLESS deliverable D3.2. Reassuringly, many selected parameters were present in more than one of the studies listed. A new parameter list was recently published in Kern et al (2024) but was not considered in ODESSA as it was not yet available when Task 1.1 was completed. It may be considered in future work.
During ODESSA, an optical model (based on the RADTRANS package of MiTgcm/DARWIN, see Dutkiewicz et al, 2015) was included in the NEMO-BAMHBI system. 2 more perturbations were added as well, affecting respectively the incoming light and the absorption of light by coloured dissolved organic matter (CDOM).

2.2 Task 1.2 – observation database

In order to assess the ensemble runs, a database of relevant observations was prepared. It includes ARGO temperature profiles, BGC-ARGO chlorophyll (CHL) and photosynthetically-active radiation (PAR) profiles, and satellite surface temperature and CHL maps.
The database was assembled entirely from the CMEMS catalogue for the years 2013 and 2016.

2.3 Task 1.3 – NEMO STOchastic module

The model components (listed in Task 1.1) need to be perturbed using relevant methods. NEMO includes a stochastic perturbation method based on autoregressive processes. Random 2D and 3D fields can be created with prescribed standard deviations, and spatial and temporal correlations. This feature was readily extended to 0D (scalar) random numbers (with a given standard deviation and temporal correlation).
For some types of perturbations, one may possess information about the spatial structure of the expected model uncertainty; it could be not completely random.

2.3.1 Perturbations as random linear combinations of prescribed fields

Therefore, a new routine was added, that allows to load prescribed fields from netcdf files at model startup. Then, during the model integration, random linear combinations of these prescribed fields are generated and added to the original (deterministic) variable.

These prescribed fields can be empirical orthogonal functions (EOFs) computed from a historical time serie, as is used in ODESSA for most atmospheric forcing fields. If the perturbation concerns a model variable (rather than a forcing variable), the prescribed fields could also be time lags (difference between model state at different instants) or assimilation increments from previous model runs.
The scalar random coefficients have a prescribed standard deviation and temporal correlation, e.g. 1 day for precipitation.
The prescribed fields are computed on their original grid, and are interpolated spatially on-the-fly by NEMO, exactly as the corresponding original field.

2.3.2 Perturbation of the wind field

In the case of wind perturbations, to account for spatial structures in the wind field, but also for the fact that in the atmosphere, different spatial and temporal scales co-exist, the Fourier transform of a time-serie of wind fields was computed point-wise and bi-variate (over U and V wind). During model integration, a random combination of Fourier modes is added to the original wind fields ; the random scalar coefficients multiplying each mode have a temporal correlation equal to the mode's period.

A practical problem appeared while trying to build these random combination of Fourier modes. Computing Fourier modes from hourly wind fields over 2015-2020 leads to over 100.000 modes; with over 100.000 corresponding periods, comprised between 1 hour and a few years. In order to build random combinations of only a reasonable amount of modes, the 100 most important ones (in a statistical sense) over the whole domain were selected.
Figure 1 shows the average amplitude of all modes with a period shorter than 4 weeks (or 672 hours). The most important mode has a period of exactly 24 hours (daily cycle); the second has a period of 12 hours; further important modes have no clear interpretation, and are distributed over the whole range of periods, except the shorter ones (2-96 hours).
A reasonable choice would be to select the most important modes (meaning, the ones with the largest average amplitude). The underlying hypothesis is that while all modes are affected by errors, the strong modes are the ones whose error will be more relevant in the final perturbation. This hypothesis is somewhat similar to fixed-base Kalman Filters (e.g. SEEK filter), where (EOFs of) time variability of model results is taken as a proxy for model errors.
However, to also represent the small-scale and fast-changing wind uncertainty, only the 75 most important modes were selected; and 25 more modes were randomly chosen within the under-represented range (2-96 hours). At model startup, NEMO loads the 100 selected Fourier modes (as for any other prescribed field) but also a netcdf variable containing the periods corresponding to the respective modes. The final wind perturbations are then built by randomly combining these 100 modes at each timestep.

2.3.3 Perturbation of scalar biogeochemical parameters

Concerning the biogeochemical model parameters undergoing perturbations, they are scalars in the original model. They are still scalars in the perturbed model (i.e. no spatial variability is added) and their perturbed value is chosen at model startup and kept constant (i.e. no temporal variability is added either). Perturbations of biogeochemical parameters are also supposed independent of one another. The perturbed values are drawn from given statistical distributions depending on constraints on the parameter (Prieur et al, 2019):

- parameters that need to be positive (or negative) are drawn from a Γ distribution.
- parameters that should be in the [0,1] interval are drawn from a β distribution
- for parameters that should be larger than 1, their logarithm is drawn from a Γ distribution

- parameters without constraints are drawn in a normal distribution

2.3.4 Using the new stochastic module

Technically, the developments of the NEMO stochastic module were started in NEMO v4.0.6, which was the version of the BLK-BGC forecasting system since the 202211 entry-in-service (EIS). Later, they were ported to NEMO v4.2.0 as well, as the BLK-BGC forecasting system also switched to this version for the 202311 EIS.
All perturbations listed above can be activated from the (existing) NEMO stochastic namelist or from the (new) NEMO-TOP stochastic namelist. To perturb the biogeochemical parameters, one just needs to specify the desired standard deviation in the namelist. The desired mean is taken as the deterministic parameter value specified in the BGC model namelist, and needs not to be repeated in the stochastic namelist. The STOpar module then computes the distribution parameters based on the mean and standard deviation.
The chosen method used for each perturbation, and some short explanations, are given in Table 2.

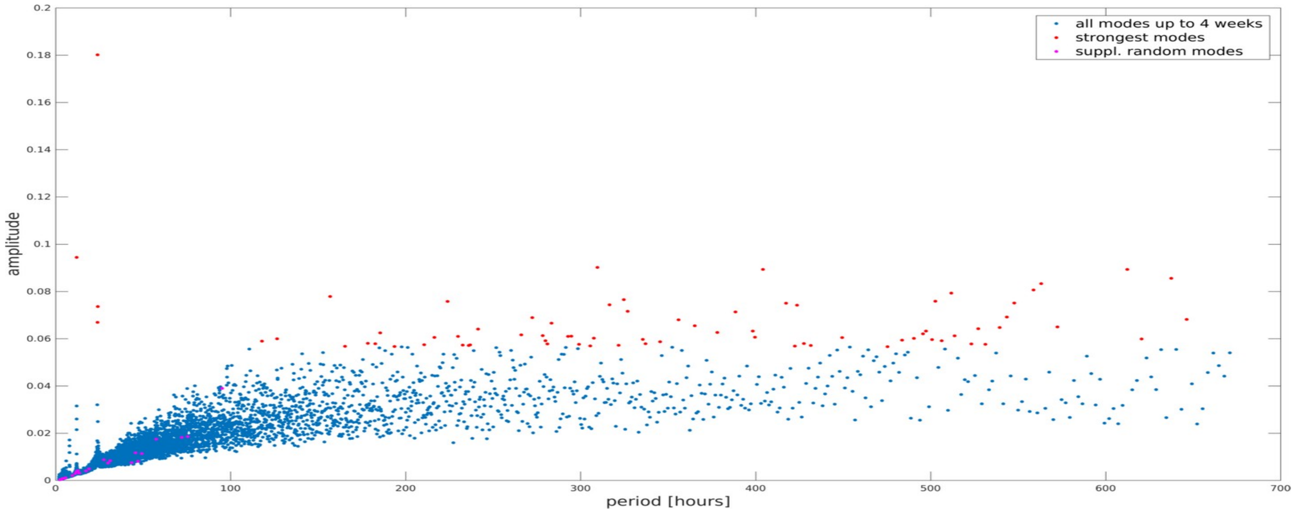

Figure 1. Basin-wide average of the amplitude of Fourier modes, as a function of their period. The modes with the highest amplitude are represented with a red dot ; other modes chosen randomly among the modes with a short period are represented

Perturbation	*Method*	*Note*
River water flux	AR(1)	Spatial correlation is 250km in order for Danube mouths to be perturbed consistently. Time correlation : 3 months
Bottom drag coefficient	2D AR(1)	Time correlation: 1 month
Equation-of-state	2D AR(1)	Time correlation: 1 week
Wind forcing	Fourier modes	Random comb. of 100 modes Time corr.: each mode's period
Air temperature forcing	EOF	Random comb. of 100 EOFs Time correlation: 5 days
Cloud coverage forcing	EOF	Random comb. of 100 EOFs Time correlation: 2 days
Precipitation forcing	EOF	Random comb. of 100 EOFs Time correlation: 1 day
River nutrient flux	2D AR(1)	Same as river water flux
Atmospheric nutried depo.	2D AR(1)	Time correlation: 1 week
Optical model: absorption by CDOM	2D AR(1)	Time correlation : 1 month
Optical model : incoming spectral radiation	log(EOFs)	Random comb. of 50 multi-var. EOFs Time correlation : 1 day
Biogeochemical initial conditions	Time-lags	Not applied in ODESSA
Meso-zooplankton grazing	Γ	stddev: 0.2
Diatoms α_{π}	Γ	stddev: 0.3
Zooplankton C assimilation efficiency	β	stddev: 0.2
Meso-zooplankton Net growth efficiency	β	stddev: 0.1
Meso-zooplankton maximum mortality rate	Γ	stddev: 0.5
Meso-zooplankton half-saturation rate	Γ	stddev: 0.67
Capture efficiency of mesozooplankton by gelatinous	β	stddev: 0.2
Gelatinous max. grazing rate	Γ	stddev: 0.2
Capture efficiency of diatoms by meso-zooplankton	β	stddev: 0.1
Zooplankton N assimilation efficiency	β	stddev: 0.1
μ_{max} diatoms	Γ	stddev: 0.3
Si:N ratio in diatoms	Γ	stddev: 0.05
Minimum N:C ratio in diatoms	Γ	stddev: 0.1
Diatoms mortality	Γ	stddev: 0.5
Messy feeding coefficient (micro-zooplankton)	β	stddev: 0.2
Messy feeding coefficient (meso-zooplankton)	β	stddev: 0.2
Minimum sinking rate of diatoms	Γ	stddev: 0.5
Maximum sinking rate of diatoms	Γ	stddev: 0.5
Mortality of Flagellates	Γ	same as for diatoms
Mortality of Emiliana	Γ	same as for diatoms
k_s for phytoplanton	Gaussian	stddev: 0.5
μ_{max} flagellates	Γ	stddev: 0.5
μ_{max} Emiliana Huxley	Γ	stddev: 0.3
Q10 phytoplankton	logΓ	stddev: 0.2
Q10 diatoms	logΓ	same as other phytoplankton groups
Q10 zooplankton	logΓ	same as for phytoplankton
ζ_{bio}	Γ	stddev: 0.05
η_{bio}	Γ	stddev: 0.05
Light Absorption coeff b (traditional light model)	β	stddev: 0.05
Q10 chemical reactions	logΓ	same as for phytoplankton

Table 2. Model components or parameters that undergo perturbations, and perturbation type

2.4 Task 2.1 – small ensembles with individual perturbations

For each type of perturbation, a small, 10-member ensemble was simulated for 1 year. It allows to study the propagation of parameter uncertainty into the model variables.

In the SEAMLESS project, the impact of perturbations on model variables was studied using 1D models. However, in the Black Sea, the biogeochemistry is strongly influenced

by lateral advection (e.g. nutrients are brought by the Danube, flow over the shelf, then into the deep sea), so that it seems reasonable to perform the experiments with a 3D model.
However, given the large number of simulations to perform, the simulations were carried out with a newly developed model configuration using 15-km horizontal resolution. Except for the horizontal resolution, the model configuration is the same as the nominal model.

Most individual perturbation were well-tuned using the chosen intensities. One notable exception was the perturbation of the atmospheric deposition of nutrients, which was very weak in the first iteration. Hence it was decided to triple the intensity of this particular perturbation.

An example of the chlorophyll spread in the ensemble where the light absorption by CDOM is perturbed, is shown in Fig. 2. Large spread can be seen in the surface, particularly close to river mouths (where the content of chlorophyll itself is high, so more prone to large uncertainty), but also at the depth of the deep chlorophyll maximum (DCM). Indeed, when light is absorbed more or less strongly, the depth of the DCM will vary.

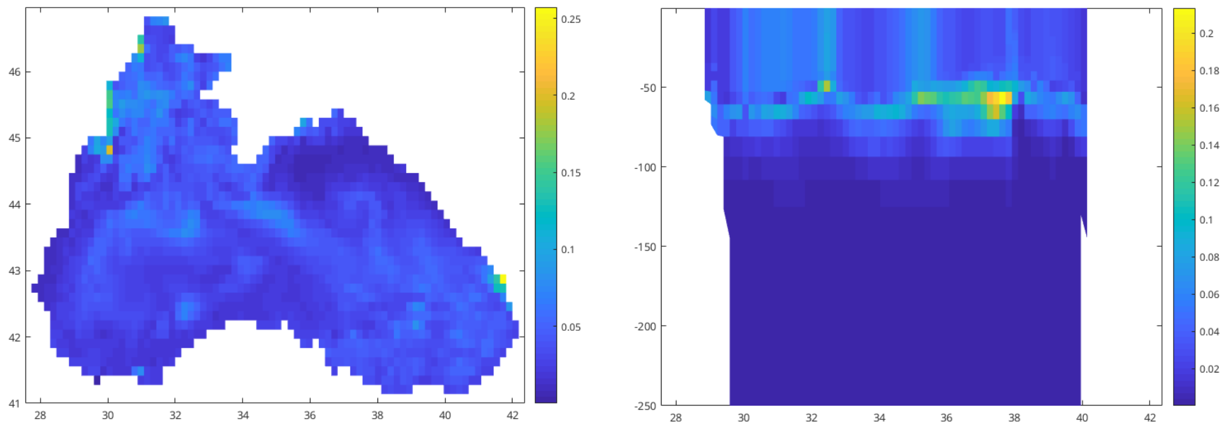

Figure 2: Example of ensemble spread of chlorophyll when perturbing aCDOM , (left) surface (right) vertical section at 43°N.

Another example concerns the perturbation of incoming spectral solar radiation mentioned in section 2.1 and Table 2. Multi-variate EOFs are computed from a time-serie of the logarithm of the incoming solar radiation (multi-variate in the sense that it is computed together over different wavelengths, and also over direct and diffuse solar radiation). A random linear combination of these EOFs is then computed during model integration (the random scalars have a prescribed temporal correlation, chosen here as 1 day), and finally the exponential of the random combination is computed, and multiplies the deterministic incoming radiation.
This perturbation of incoming radiation directly affects the model temperature field. Ensemble spread is large during daytime; whereas during night no supplementary temperature uncertainty is added (but spread remains from the previous days perturbations). The magnitude depends on the season, and other factors such as the mixed-layer depth and prevalent currents.
The perturbation also affects biogeochemistry. An example is shown in Fig. 3, again for chlorophyll, on 25/12/2013. Interestingly, the right panel in Fig. 3 seems to suggest that in this particular case, the perturbation of light generates a bi-modal distribution for surface chlorophyll.

It was verified that out of the 10 members in the ensemble, during the days preceding 25/12/2013 and in the area of the BGC-ARGO float, 3 members had more PAR in the surface layer (we call them high-PAR), and 7 members had less PAR (we call them low-PAR). The difference between the daytime PAR in the 2 groups is only a few $W.m^{-2}$ (corresponding to 1~2% of the actual irradiation). The low-PAR members all had high concentrations of Flagellates (around 7 mmol Carbon/m3) whereas the high-PAR members all had low concentrations (around 1 mmol Carbon/m3). Similarly, the low-PAR members also all had higher diatom concentrations (around 2 mmol/m3) than the high-PAR members (around 1 mmol/m3). However, the 7 low-PAR members had a very low coccolithophore concentration of 0.1 mmol/m3, whereas the 3 high-PAR members all had higher coccolithophore concentrations: around 1.5 mmol/m3.
Thus, the change in incoming light led to a change in dominant phytoplankton species.

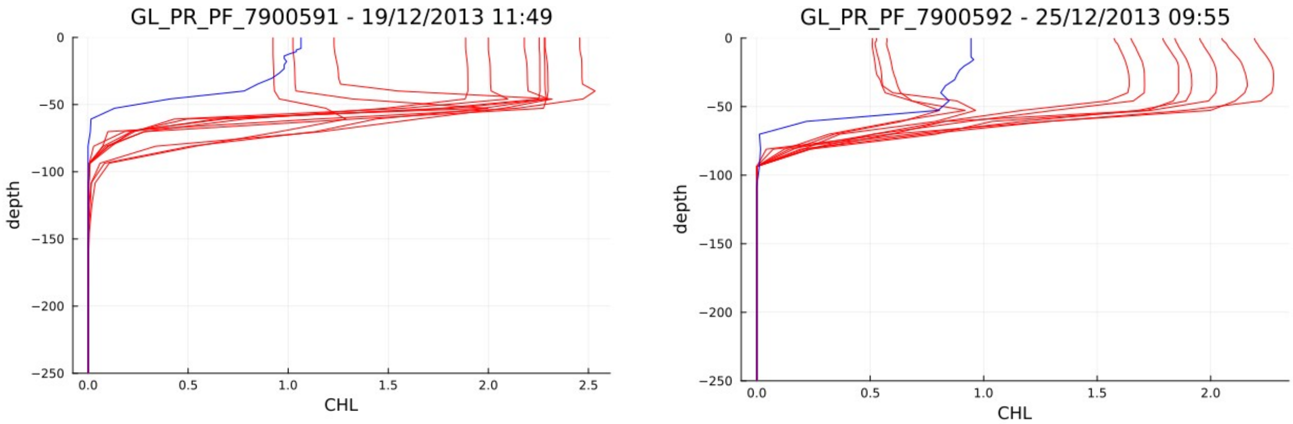

Figure 3: Chlorophyll concentration from ARGO profiles (blue) and corresponding model profiles extracted from the 10 members of a sub-ensemble with perturbed incoming solar radiation (red).

2.5 Task 2.2 – stochastic validation

In ODESSA, the ENSDAM package was used to assess the ensemble simulations, in particular the CRPS and RCRV scores, and rank histograms were computed.
ENSDAM is a Fortran library, usable in a variety of languages. In particular, during ODESSA, a Python interface was extended to cover all of ENSDAM's features. A Julia interface was also developed.
The simulations described in Tasks 2.1 were evaluated using this interface, as the deterministic validation of the BLK-BGC model is already scripted in Julia, incl. the construction of model-observation pairs.
More details about ENSDAM can be found online (https://github.com/brankart/ensdam).

As an example, the stochastic scores corresponding to the 2 cases of Fig. 3 are given in Table.

	CRPS	Resolution	Reliability	RCRV bias	RCRV spread
Left panel in Fig. 3	0.303	0.261	0.042	0.217	0.027
Right panel in Fig. 3	0.218	0.106	0.112	0.304	0.029

Table 3. Stochastic scores computed for the cases depicted in Fig. 3

2.6 Task 3.1 – fully perturbed ensemble (low resolution)

Compared to simulations with a single perturbation, in the fully-perturbed ensemble, the problem arises how the combined input uncertainty adds up and yields the forecast uncertainty, i.e. our best-estimate of the expected model forecast error.
Obviously, the spread in the fully perturbed ensemble is much lower than the sum of the spreads in every ensemble with 1 perturbed parameter. For example, looking at the

spread of surface chlorophyll concentration in the north-western shelf (NWS), the maximum spread in the fully-perturbed ensemble is around 3.4 mmol/m3, whereas the hypothetical sum of the spread of all ensembles would be over 8 mmol/m3.
Still, the surface chlorophyll spread obtained in the fully-perturbed ensemble may seem large, especially in summer. However, in the ODESSA set-up, we had no a priori aim regarding the desired ensemble spread ; we tried neither to have it small or large, or to keep its magnitude constant in time. Rather, the uncertainty is the result of our hypothesis on each of the perturbations; its relevance should be tested using stochastic metrics.
It should also be noted that, in the literature, ensemble sometimes suffer from spread “collapse” over time, and mitigation methods need to be used, such as uncertainty inflation. Thus, in some contexts, having a large ensemble spread could be an advantage, and furthermore it may be reduced by data assimilation (either state vector correction or parameter estimation).

Figure 4 illustrates the surface spread evolution during a 1-year simulation (year 2016), for each of the sub-ensembles (Task 2.1) and for the fully perturbed ensemble (Task 3.1).
The initial spread is zero. The spread then grows throughout the first months of the simulation, which is both due to perturbations starting to have an effect, and because the spread is larger in summer than in winter. After the summer, the spread reduces and reaches a lower value at the end of the year.

The perturbations most strongly contributing to surface chlorophyll spread are the wind forcing, (in the north-western shelf) μ_{max} of the Flagellates, and (in the deep sea) k_s(NHS); many other perturbations also have non-negligible effects.
As expected, the PAR spread in the surface layer is almost exclusively due to the (multiplicative) perturbation of the incoming solar radiation, and is important when the input itself is important (i.e. in summer). Finally, spread in surface temperature is influenced most strongly by the air temperature, by incoming solar radiation (mostly in summer), by the wind, by k_s(NHS) and by light absorption by CDOM. In summer, the contribution of the perturbation of the latter 2 biogeochemical processes is not much smaller than the contribution of the perturbation of the wind. This conclusion is somewhat similar to findings in Kara et al (2005), showing the importance of BGC on SST in the Black Sea.

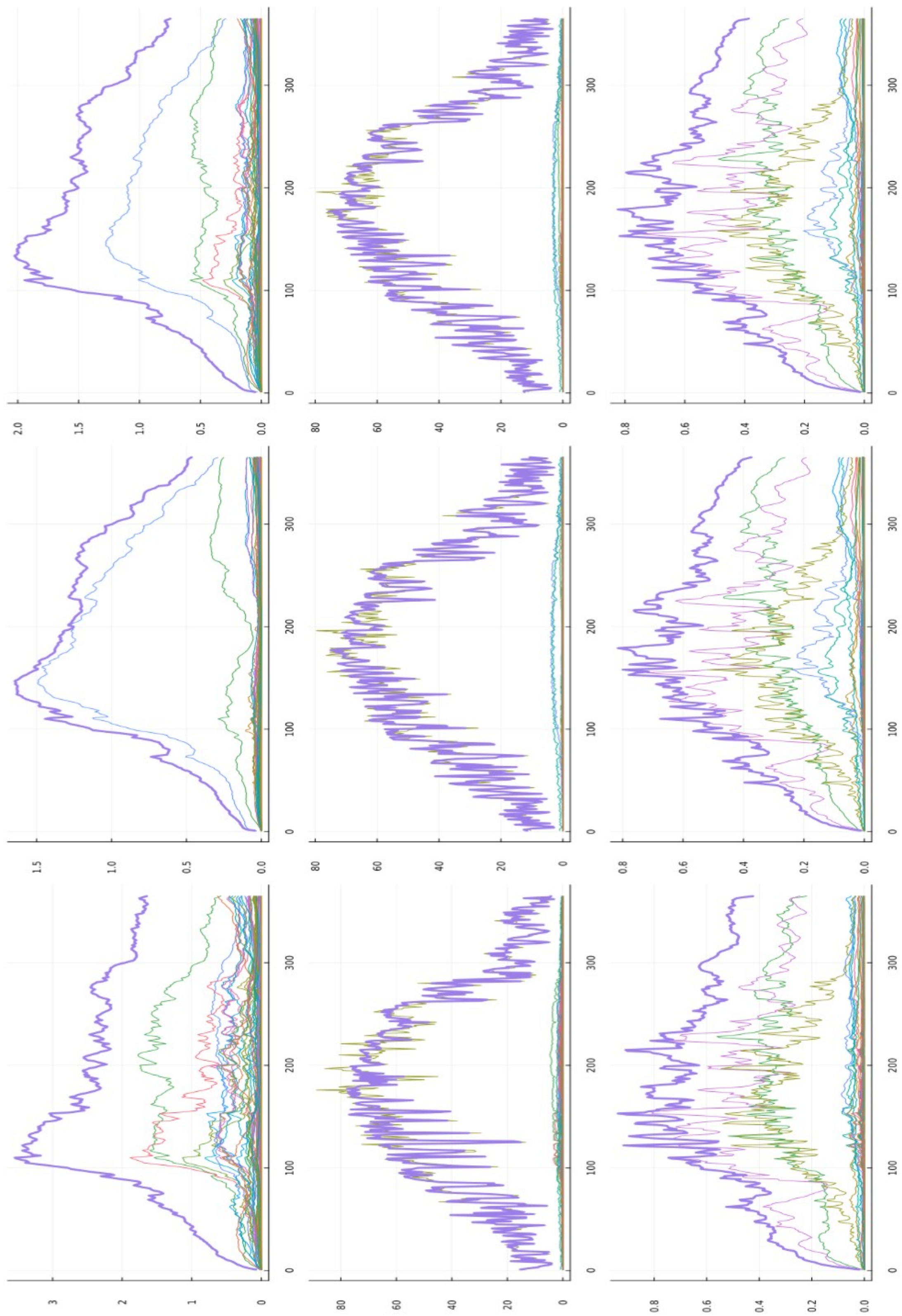

Figure 4. Ensemble surface spread for (upper row) chlorophyll, (middle row) PAR, (lower row) temperature, in (left column) the NWS, (middle column) the deep sea, (right column) all points. The bold purple line corresponds to the fully-perturbed ensemble

Spread below the surface is generally lower. However, at the DCM depth (for chlorophyll) or the MLD (for temperature), a distinctive area exists where the spread is also high (Fig. 5). Even deeper, spread tends to vanish.

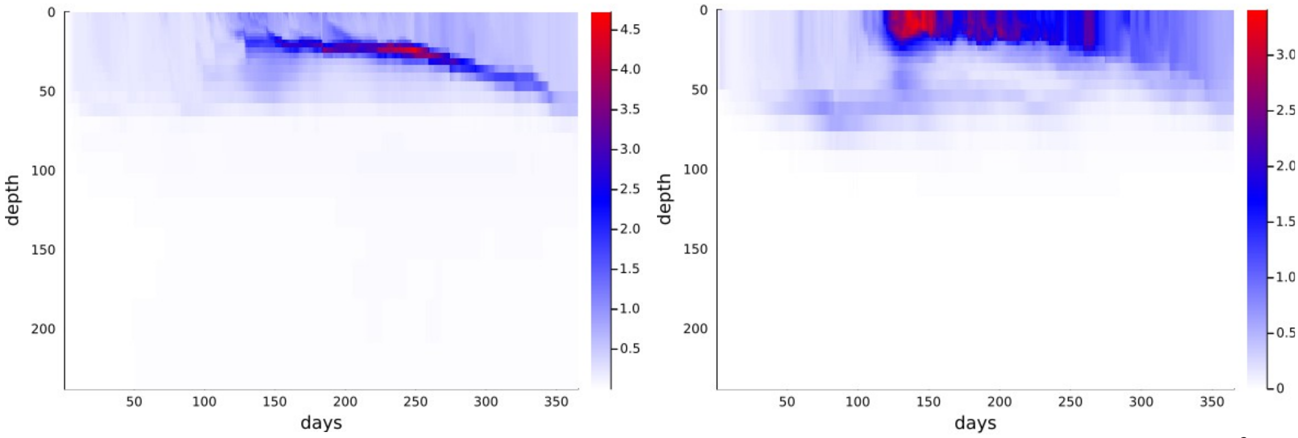

Figure 5: Ensemble spread Hovmoller diagram for (left) temperature [°C], (right) chlorophyll [mmol/m³]

2.7 Task 3.2 – fully perturbed ensemble (high resolution)

Fig. 6 allows to compare the ensemble spread between the low- and high-resolution fully-perturbed ensembles. The purple curve is copied from Fig. 4. Spread in the high-resolution is computed only on the 15th of every month and is shown as dots.

The perturbations in both ensembles are built using the same statistical parameters, i.e. the spatial correlations (which are given in grid points in the namelists) are identical when converted in kilometres.

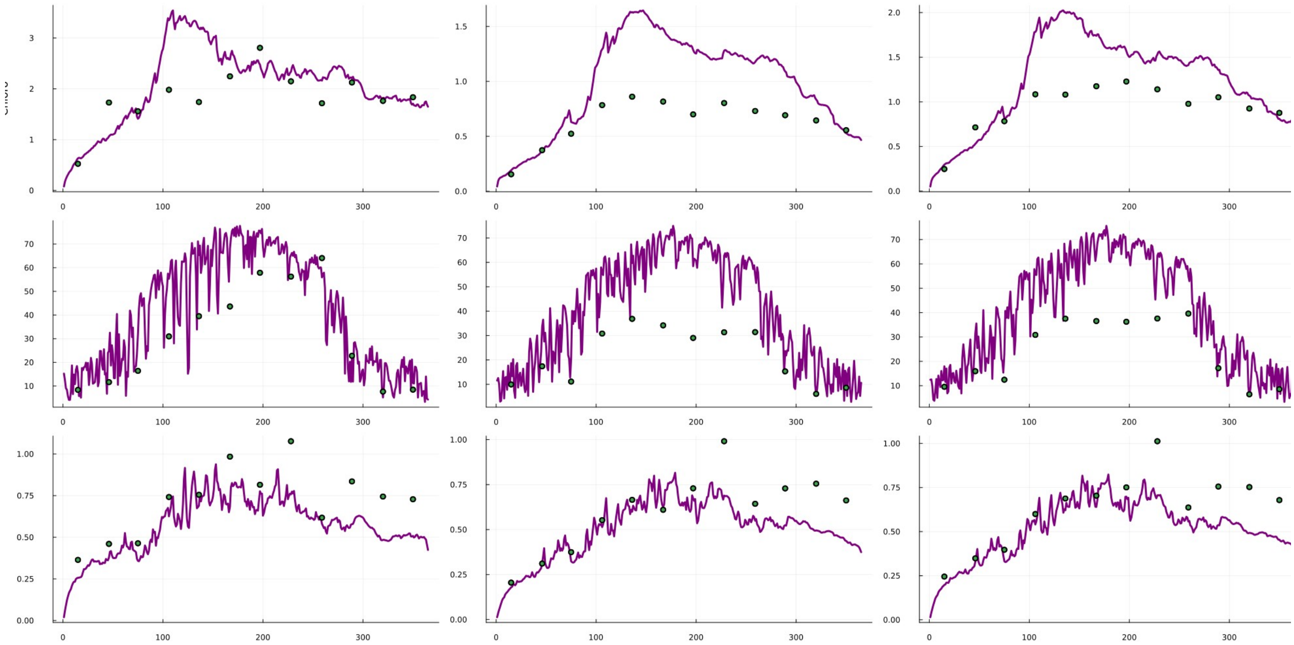

Figure 6: Time serie of surface spread in the low-resolution (purple line) and high-resolution (dots) ensembles. From left to right: NWS, deep sea, all points. From top to bottom: chlorophyll, PAR, temperature.

The resulting ensemble spreads are similar, but not identical. In particular, the uncertainty on light seems smaller in the high-resolution model, especially in the deep part of the Black Sea. Similarly, uncertainty on chlorophyll seems smaller in the high-resolution ensemble.
On the contrary, uncertainty on temperature seems larger in the high-resolution ensemble in the second half of the year.
This indicates that it may not be straightforward to estimate the uncertainty of the high-resolution model, using a low-resolution ensemble.

2.8 Task 3.3 – stochastic validation

The reduced centered random variable (RCRV) score is defined in (Candille et al., 2007). In a *reliable* system, the RCRV has a zero mean and unit standard deviation.
Fig. 7 shows the mean and standard deviation of the RCRV score for chlorophyll, PAR and temperature, with respect to BGC ARGO measurements over all available buoys and profiles in 2016.
Thus, in the plot, each point corresponds to 1 profile and is not necessarily originating from the same ARGO buoy as the next point nor is it spatially nearby.

The CRPS scores are shown in Fig. 8 for chlorophyll, PAR and temperature.
Both in the RCRV and CRPS scores, one can observe the initialization period during the first months, and the seasonality of the ensemble skill.

2.9 Task 3.4 – predictability

One objective of ODESSA was to assess the predictability of the forecasting system for short (10-day) and up to medium (30 days) lead-times.
Using the ensemble, the uncertainty affecting the model predictions can be estimated from the ensemble spread. Furthermore, predictability can be estimated by studying the temporal evolution of the CRPS score, and by looking at the final CRPS as a function of the initial CRPS, i.e. CRPS($t+\Delta t$) as a function of CRPS(t) (see H2020 IMMERSE project deliverable D7.2 ; Leroux et al, 2022).

The uncertainty affecting some of the model components is not expected to change over time, and hence the corresponding model perturbations are built with a constant standard deviation. Examples include perturbations of the equation-of-state and biogeochemical model parameters.
However, the uncertainty affecting other model components clearly increases during a medium-term simulation; the most obvious example is the atmospheric forcing, and in particular the wind.

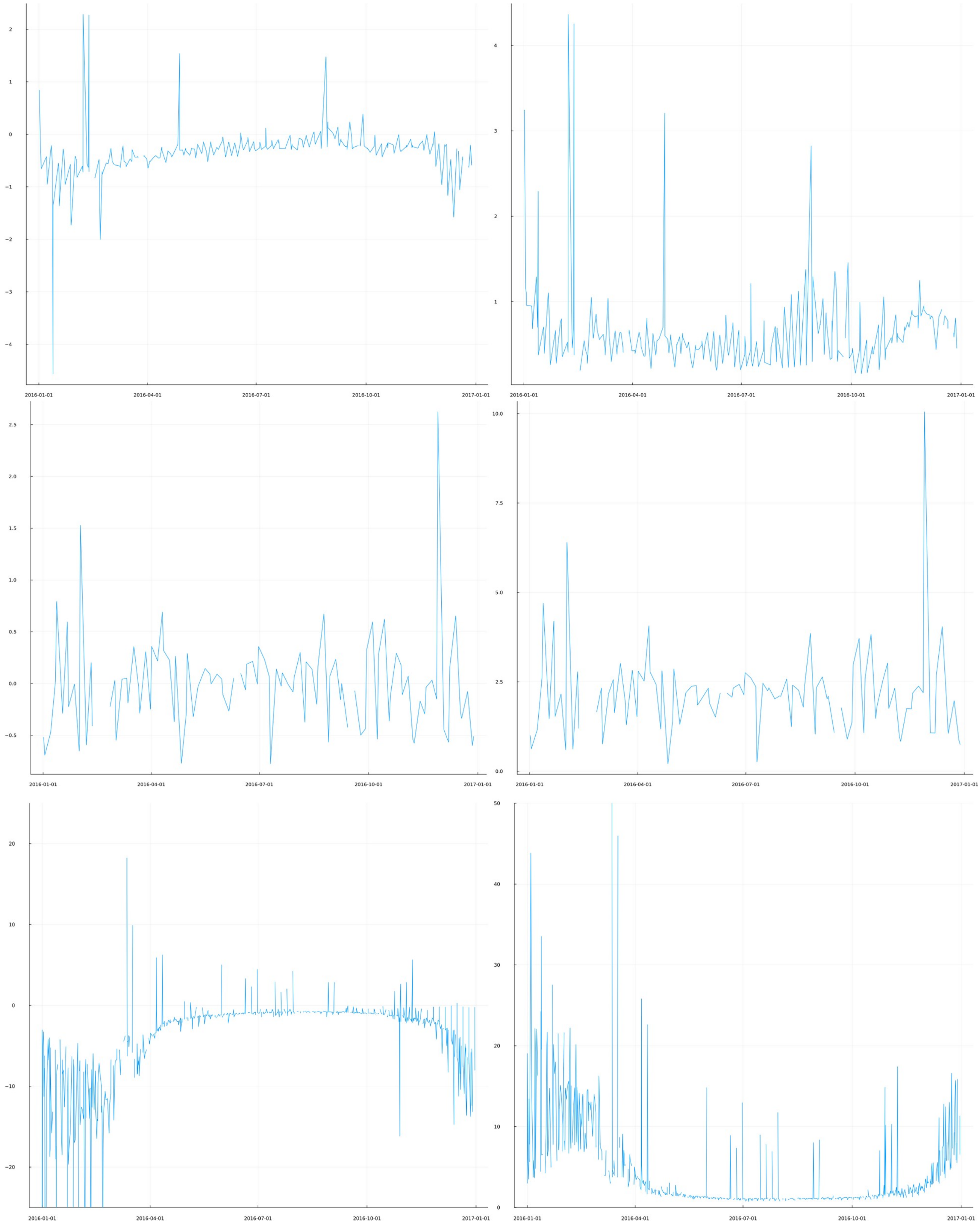

Figure 7: Time-serie of RCRV score (left column) mean and (right column) standard deviation. From top to bottom: chlorophyll, PAR, temperature. Scores are computed with respect to ARGO observations.

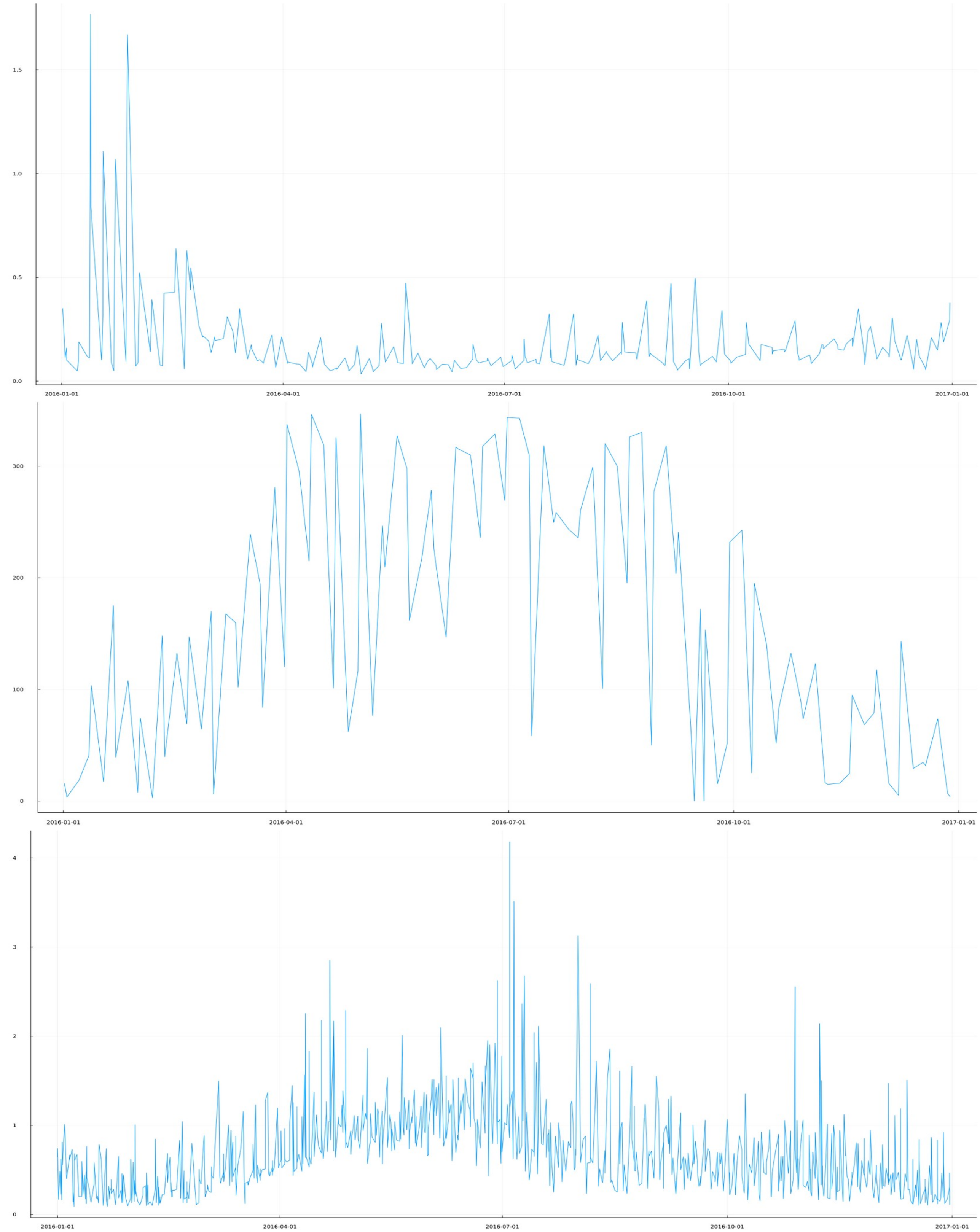

Figure 8: CRPS scores with respect to all available ARGO profiles in 2016, for (upper panel) chlorophyll, (middle panel) PAR, (lower panel) temperature

2.9.1 Methodolody

Some indications about the uncertainty increase in atmospheric forecasts are given in Dee et al (2013). At the time of that paper, the ECMWF operational model as well as the ERA-Interim simulation presented anomaly correlations sharply decreasing when lead-time increases: over 98 % (lead-time of 3 days), 90 % (5 days), 75 % (7 days), less than 50 % (10 days).

Internal ECMWF report 754 (Buizza and Leutbecher, 2015) shows how modern-day (atmospheric) predictability can be extended over the 10-day horizon by switching from deterministic to stochastic predictions. The ensemble can still carry useful information even when close to the limit of predictability (of the deterministic model). In fact, improvements between the start of operational ensemble predictions (2004 in the case of ECMWF) and the publication time (2015) have led to a gain of predictability of ~1 day per year.

Figures in the report show that the ECMWF ensemble CRPS score is very good for short lead time, but equal to the climatological ensemble after 15 ~ 20 days.

The report seems to concern mostly large-scale atmospheric processes, and little information is available in the literature on regional scales, and also on the surface wind in particular.

Furthermore, CMEMS MFCs currently use the (deterministic) ECMWF IFS atmospheric product for which the predictability limit is probably not higher than 10~15 days.

In ODESSA, the following pragmatic methodology was applied to run a medium-term simulation. In order to study the growth of the ensemble spread, the initial ensemble spread should be small; an Ensemble Kalman Filter (EnKF) assimilation step is applied to the initial condition taken from then ensemble members of the experiment described in Task 3.1.

Assimilation is applied separately for physics and BGC. Sea surface temperature and sea surface chlorophyll are observed; in the model, they correct respectively the 3D temperature and salinity, and the 3D phytoplankton groups (model chlorophyll is a diagnostic variable in this version of the forecasting system). The observation error (incl. the representativity error) is chosen in order to significantly reduce the ensemble spread (see Table 4). The assimilation is performed for 15 April 2016, given that there are few clouds on that particular day, allowing for a broad spatial coverage of both observed variables.

It is not the scope of this report to study in detail the data assimilation procedure; however it should still be mentioned that the ensemble is naturally representing the “uncertainty of the day”. Therefore, the EnKF does not require to take the same precautions as the SEEK filter used in the BLK-BGC system. For example, the BLK-BGC SEEK filter uses an errorspace computed once-and-for-all from a few years of model outputs, and, by construction, it cannot be expected to « know » anything about the current mixed layer depth. Hence, a vertical localization procedure is required; in the last versions of the BLK-BGC system, this procedure depends on the model MLD. However, the EnKF naturally propagates surface information downward only up the mixed layer depth, and does not require a vertical « localization » procedure.

The reduction in the ensemble spread of the initial condition is illustrated in Table 4. It should be noted that salinity spread is reduced by ~40%, although this variable is not observed.

	NWS	*Deep Sea*	*All*	*Units*
CHL forecast	3.46	1.32	1.83	mmol.m^{-3}
CHL analysis	0.28	0.17	0.19	mmol.m^{-3}
Temperature forecast	0.66	0.52	0.55	°C
Temperature analysis	0.13	0.18	0.17	°C
Salinity forecast	0.33	0.11	0.16	/
Salinity analysis	0.19	0.09	0.11	/

Table 4. Ensemble spread before and after assimilation on 15/04/2016

The 1-month simulation then starts from this new initial condition, and all perturbations are applied as usual, except the wind perturbation which increases in magnitude, consistent with the findings described above.
The short-term wind perturbation intensity is computed to have similar statistical features as the probability density function of the difference between ECMWF winds and observed winds (satellite scatterometer). In particular, the standard deviation is ~0.5 m/s.
However, in the medium term, i.e. after the 15-day lead-time, the uncertainty should be closer to the difference between the ECMWF wind and the wind climatology.
The hourly wind (intensity) climatology was computed over 2010-2020, and standard deviation between the ECMWF operational product and its climatology, computed during 15-04-2016 and 15-05-2016 was found to be slightly over 2 m/s.
Finally, the NEMO stochastic module was modified to allow for a time-varying expected standard deviation ; the standard deviation is the sum of a constant value and a sigmoid centered on day 15 after the simulation start. This feature was applied to the random scalar parameters used to multiply the Fourier modes for the wind perturbation. The resulting wind perturbation standard deviation is shown in Fig. 9.

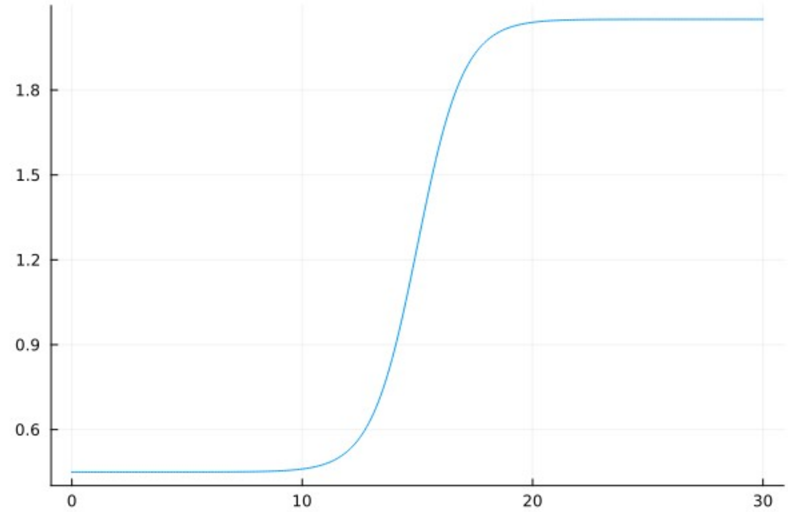

Figure 9. Wind perturbation intensity [m/s] as a function of lead-time [days], for the medium-term ensemble forecast starting on 15-04-2016

2.9.2. Results

Predictability is studied by the leave-one-out procedure, i.e. considering successively each member as the « truth » and using only the other 99 members as the ensemble. The CRPS is then computed over all spatial points, every day. The procedure is applied to the ensemble starting from the analyzed state on 15/04/2016, and simulating 1 month with « normal » perturbations, and then repeated for the ensemble with wind perturbations growing stronger.

The CRPS evolution in time is shown in Fig. 10. During the first 10 days of the simulation, by construction, both ensembles are statistically identical. During the last 10 days, the increased uncertainty on the wind forcing leads to higher uncertainty on PAR and temperature, whereas chlorophyll seems unaffected (at this particular time of the year).

This figure can be reshuffled to show CRPS(t+Δt) as a function of CRPS(t), as explained above. The result is shown in Fig. 11.
The temperature panels present expected conclusions. With short lags, the CRPS often remains similar from initial to final time, albeit with some exceptions that underline the importance of running a stochastic model. When considering longer time lags between the initial and final CRPS, the relationship becomes unpredictable, CRPS sometimes strongly increasing or decreasing. With very long lags (25 days), the initial time is always close to the initial condition (hence the CRPS is small), whereas the final CRPS is spread out from small to large values.
Regarding chlorophyll, the first panel shows almost no change in CRPS at short time lags; when the lags increases, the initial CRPS can take all possible values, whereas the final one is small up to very large.
Finally, regarding PAR, no conclusion can be given from the figure, except that a stochastic model is required to represent the uncertainty, which can be either small or large, and can either increase or decrease over time (both with short and long lags).

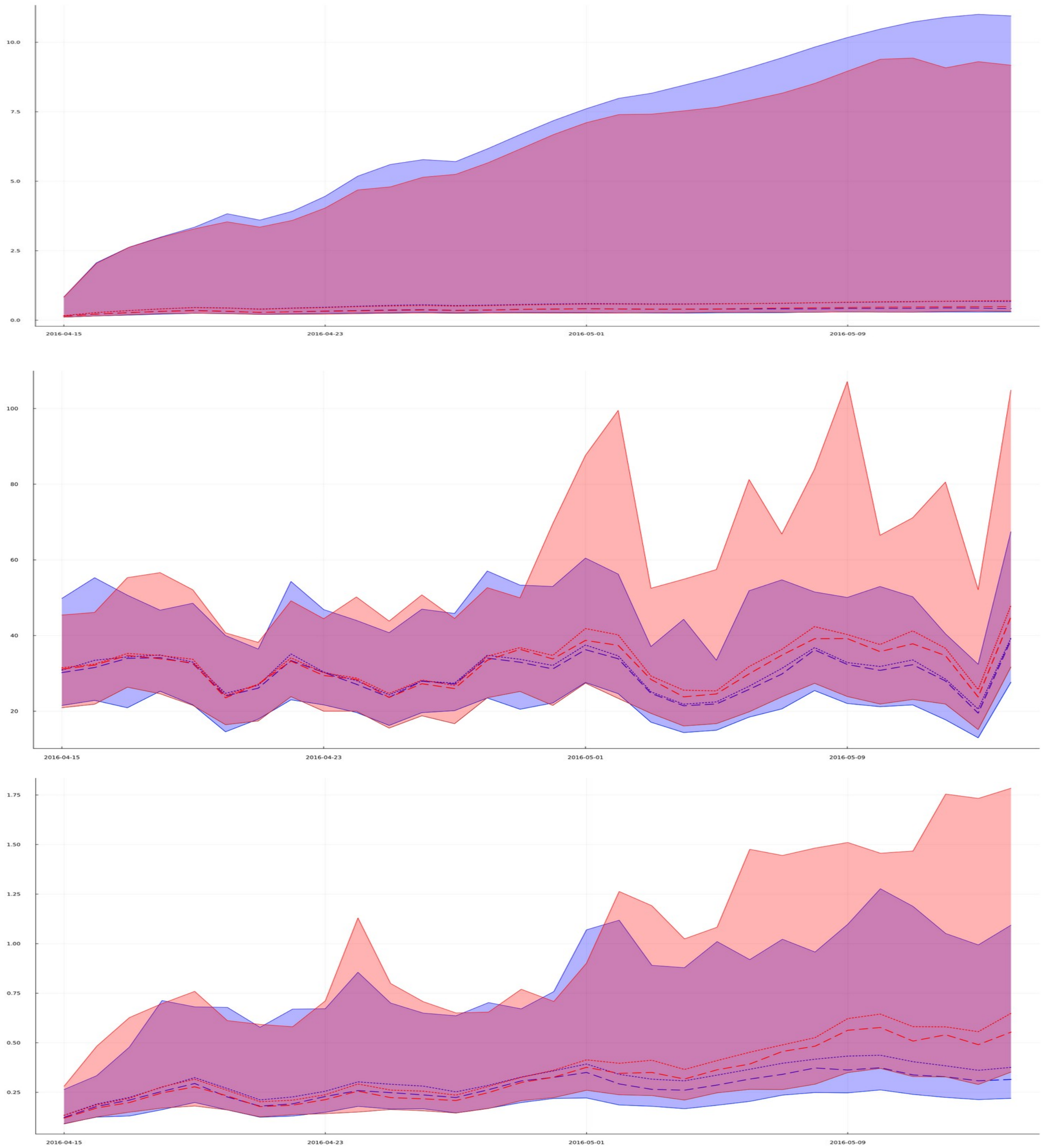

Figure 10: CRPS evolution in time for (top panel) chlorophyll, (middle) PAR, (bottom) temperature, for the ensemble simulation with standard perturbations (blue) and increased wind perturbation (red). The shaded area represents the range (minimum and maximum CPRS on a given day); the dashed lines the median, and the dotted line the mean

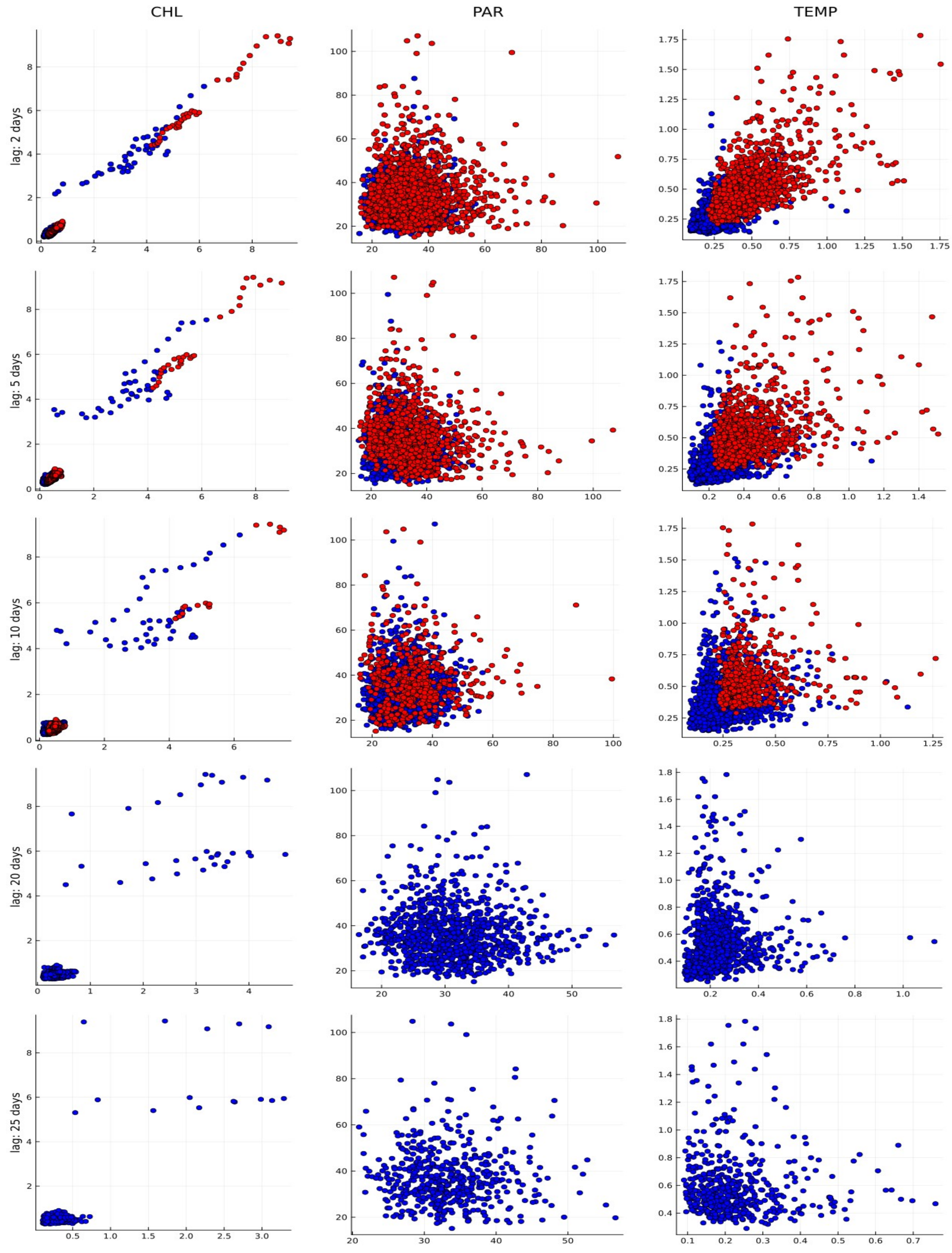

Figure 11. Final CRPS score as a function of initial CRPS score, for time lags (from top to bottom) of 2, 5, 10, 20, 25 days. Columns represent (from left to right) chlorophyll, PAR and temperature. Dots are blue when the initial CRPS score is obtained during the first 15 days of the simulation, else they are red.

2.10 Task 3.5 – sample product

The full results from the ensemble forecasting system have been saved, and samples can readily be delivered to CMEMS.

At the date of this report, there has been no final decision about which variables to deliver (e.g. chlorophyll), on the spatial extent (surface only, 3D fields…), on the temporal extent (more recent years can be simulated), on the spatial resolution, and on the choice of some statistics (e.g. ensemble mean and standard deviation, percentiles, all members).

Also, the practical way of sending the samples to CMEMS has not been established yet.

3 Identified issues

3.1 Closed issues

During the project, 2 issues were encountered, that were addressed timely.

- The radiative transfer module was not available yet during the first year of the project. Sub-ensembles with individual perturbations, and the fully-perturbed ensemble, were first obtained with the traditional light propagation algorithm. Later, all simulations were restarted using the fully coupled model (physics-biogeochemistry-spectral light).
- At the beginning of Tasks 2.1, 3.1 and 3.2, the year 2013 was simulated as it was the first year for which atmospheric radiative forcings became available. Afterwards, it was realized BGC-ARGO was available only starting from December 2013; the low-resolution simulations were then restarted in the year 2016, when 2 BGC-ARGO floats are available all year long.
  Most diagnostics and comparisons are performed during 2016. However, a few project results are described using model forecasts for 2013.

3.2 Open issues

- In the proposal, it was suggested that some conclusions would be presented regarding the transformation of *offline-coupled* biogeochemical forecasting system into a stochastic one. However, this issue was not addressed during ODESSA. It is clear that the necessary perturbation of physical forcing fields will be more complicated in offline systems.
- It was promised to include an optical model into the forecasting system, and to perturb it together with the other model components. This task was completed. However, the merit of including an optical model (and of its uncertainty) was not assessed. It was only shown that the uncertainty on two components of the optical model led to large forecast uncertainty, both on physics (temperature) and BGC (e.g. chlorophyll).

3.3 Supplementary questions

No other open issues remain at the end of the project. However, it would be very beneficial to further study the following open questions:

- the validation of the ensemble should be continued, using other observations, in particular satellite observations of temperature, chlorophyll, and eventually surface reflectance
- Talagrand histograms were computed using ENSDAM, but were not further analyzed, and were not presented in this final report. They would allow to further characterize the under- or over-dispersive nature of the ensemble, and potentially adjust the intensity of perturbations.
- the statistical distribution of the ensemble should be studied more in detail, in particular from the point of view of future data assimilation applications.

4 Uptake by Copernicus Marine Service

The main goal of the project is to help MFCs switch from the present deterministic BGC forecasting systems, to stochastic ones.

The research performed in ODESSA will be readily applied to BLK-BGC. However, it may also help other MFCs, particularly (but not only) if the forecasting system is using NEMO.

The list of model components whose perturbation has an impact of the ensemble spread, is relatively generic, and probably useful to all MFCs. In particular, it was shown that uncertainty on biogeochemical forecasts originates from uncertainty on physical *and* biogeochemical model components. It is not sufficient to perturb only BGC model parameters.

If a MFC is using a BGC model coupled to NEMO, the (new version of) the STOchastic module can be readily used to add perturbations during the model integration. These changes will also – hopefully – be included in the future versions of NEMO.

Potential paths of uptake by MFCs are shown in Table 5.

Development leading to potential impact on Copernicus Marine Service	*Status*	*Nature of the potential impact*	*Foreseen horizon of the impact*	*Targeted TACs and/or MFCs*
Perturbation list	Completed	Preparation of the next generation		BLK-MFC Other MFCs using similar BGC models
New version of NEMO STOPAR module	Completed in Nemo 4.0 and 4.2	Preparation of the next generation		BLK-MFC Other MFCs using BGC models online-coupled with NEMO (e.g. IBI)
Demonstration uncertainty product	Ready to be delivered	New product in the CMEMS catalogue	Months following the project end	BLK-MFC

Table 5. Uptake pathways

The switch from deterministic to probabilistic systems is encouraged by CMEMS. The Ensemble Working Group tries to harmonize the switch for all MFCs, and share experiences between them. ODESSA allowed quick progress in the framework of the BLK-MFC. Once technical and scientific issues (such as the ones addressed in ODESSA) are solved, a major issue will be the huge increase in required computing resources. Low-resolution ensembles may be one possible mitigation, but it was shown in ODESSA that the ensemble spread is not identical when horizontal resolution changes.

5 Communications

During ODESSA, the following communications were realized, or ar still foreseen:
* participation to CMEMS EAWG, incl. a presentation
* participation to CMEMS bioDAWG, incl. a presentation
* participation to the NEMO stochastic working group meetings in order to prepare the next NEMO version
* joint meetings with the MULTICAST service evolution project team.
* preparation of a peer-reviewed paper based on this report.

Project highlights:

The ODESSA project aimed at building a framework for transforming deterministic forecasting systems into stochastic ones; the methodology was applied to the Black Sea MFC biogeochemical forecasting system.

Various model components were perturbed: atmospheric forcing fields, bottom drag, the equation-of-state, river fluxes, biogeochemical model parameters, atmospheric deposition of nutrients, and components of the spectral radiative model. New methods were implemented in the NEMO model to generate some of these perturbations.

Running the ensemble is costly from a computational point of view. A model version with degraded horizontal resolution was developed and low-resolution ensembles were simulated.

Subsequently, the impact of the individual perturbations listed above was studied. It is very clear that uncertainty in biogeochemical forecasts originates from uncertainty both on the physical and biogeochemical model components and parameters.

Some surprising results were obtained, e.g. a reasonable perturbation of the incoming light may change the dominant phytoplankton species.

The spread in the fully-perturbed ensemble of models is an estimate of the uncertainty affecting forecasts. Further stochastic metrics, such as CRPS and RCRV, allow to quantify if the ensemble is reliable and its so-called resolution. It was shown that uncertainty is spatially varying (e.g. it is larger in the North-Western shelf area than in the open sea) and dependent on the season (e.g. it is larger in summer than in winter).

The spread in the low-resolution version of the ensemble, and the nominal version were compared; they are generally in good agreement, but some differences still appear. This may complicate the objective of using a low-resolution model to estimate the uncertainty affecting a high-resolution model.

The predictability can be studied by looking at the evolution of the CRPS score over time. During medium-term (30 days) simulations, the uncertainty affecting some model components increases over time (particularly so the atmospheric forcing fields). In the simulation realized in April-May 2016, this increase of the wind uncertainty did not seem to strongly affect the surface chlorophyll CRPS, whereas the light and temperature CRPS scores increased in the second half of the medium-term forecast.

The results obtained in ODESSA can be directly used by the BLK-BGC production unit, switching from a deterministic forecasting system to an ensemble-based system.

However, some results are generic enough and could help other forecasting centers. The list of perturbed model components list may serve as a first estimate for the equivalent list in other geographical areas or for other numerical models.

For NEMO-based systems, the implementation of the new stochastic module may be used directly. The latest version of the ENSDAM stochastic validation package can also be used readily by other centers.

Other results, e.g. related to horizontal resolution or to predictability, may also be of interest to the CMEMS community and forecasting centers in particular.

Finally, a demonstration uncertainty product can also be of interest to the CMEMS users.

6 Use of resources

The ODESSA project fully consumed the 15 person-month resources, as well as direct costs in terms of travel between the project partners ULiege and UGA. A budget has been foreseen for the article to be submitted.
The project benefited from extensive resources in terms of computing time, obtained both internally (at the CECI HPC consortium) and externally (at the ECMWF computing facility, and at the LUMI supercomputer).
One lesson learned during ODESSA is the need to estimate required storage resources and the associated costs. Ensemble forecasting systems consume vast storage space, of the order of tens of terabytes. In the case of ODESSA, during the proposal phase, this aspect had been neglected and only computing time requirements had been considered.

7 Miscellaneous

N/A

8 References

G. Candille, C. Côté, P. Houtekamer, G. Pellerin, Verification of an ensemble prediction system against observations. Mon. Wea. Rev, 135, 2688-2699, 2007. https://doi.org/10.1175/MWR3414.1
D. Dee, M. Balmaseda, G. Balsamo, R. Engelen, A. J. Simmons, J.-N. Thépaut, Toward a consistent reanalysis of the climate system. Bulletin of the American Meteorological Society, 2013. https://doi.org/10.1175/BAMS-D-13-00043.1
S. Dutkiewicz, A. Hickman, O. Jahn, W. Gregg, C. Mouw, M. Follows, Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model, Biogeosciences, 12, 4447-4481, 2015, https://doi.org/10.5194/bg-12-4447-2015
F. Garnier, Paramétrisations stochastiques de processus biogéochimiques non résolus dans un modèle couplé NEMO/PISCES de l'Atlantique Nord. PhD Thesis, Université Grenoble Alpes, 2006.
M. Grégoire, C. Raick, K. Soetaert, Numerical modeling of the central Black Sea ecosystem functioning during the eutrophication phase, Progress in Oceanography, 76, 286-333, 2008, https://doi.org/10.1016/j.pocean.2008.01.002
A. Kara, A. Wallcraft, H. Hurlburt, Sea Surface Temperature sensitivity to water turbidity from simulations of the turbid Black Sea using HYCOM. Journal of Physical Oceanography 35(1) 33-54, 2005, https://doi.org/10.1175/JPO-2656.1
S. Kern, M. McGuinn, K. Smith, N. Pinardi, K. Niemeyer, N. Lovenduski, P. Hamlington, Computationally efficient parameter estimation for high-dimensional ocean biogeochemical models, Geoscientific Model Development, 17, 621–649, 2024. https://doi.org/10.5194/gmd-2023-107
S. Leroux, J.-M. Brankart, A. Albert, L. Brodeau, J.-M. Molines, O. Jamet, J. Le Sommer, T. Penduff, P. Brasseur, Ensemble quantification of short-term predictability of the ocean dynamics at a kilometric-scale resolution: a Western Mediterranean test case, Ocean Science, 18, 1619-1644, 2022, https://doi.org/10.5194/os-18-1619-2022
R. Buizza, M. Leutbecher, ECMWF Technical Memorandum, report 754, June 2015.
C. Prieur, L. Viry, E. Blayo, J.-M Brankart, A global sensitivity analysis approach for marine biogeochemical modeling, Ocean Modelling, Volume 139, 2019, https://doi.org/10.1016/j.ocemod.2019.101402.
B. Wang, K.Fennel, L. Yu, C. Gordon, Assessing the value of biogeochemical Argo profiles versus ocean color observations for biogeochemical model optimization in the Gulf of Mexico, Biogeosciences, 17, 4059-4074, 2020, https://doi.org/10.5194/bg-17-4059-2020