Project information	
Project full title	EuroSea: Improving and Integrating European Ocean Observing and Forecasting Systems for Sustainable use of the Oceans
Project acronym	EuroSea
Grant agreement number	862626
Project start date and duration	1 November 2019, 50 months
Project website	https://www.eurosea.eu

Deliverable information	
Deliverable number	D2.2
Deliverable title	Analysis of the physical and BGC design experiments
Description	Analysis of global numerical experiments with physical and BGC forecasting model to estimate the impact of new observing system design.
Work Package number	WP2
Work Package title	Observing System Design
Lead beneficiary	Mercator Ocean
Lead authors	Elisabeth Rémy
Contributors	Florent Gasparin, Alexandre Mignot
Due date	31 October 2021
Submission date	31 March 2022
Comments	Although the delivery of the synthetic input observations and the model fields for the experiments to the task partners has been conducted according to schedule, the departure of two persons co-responsible for this deliverable at the respective partner institutions has caused a delay in the completion of the report.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 862626.

https://doi.org/10.3289/eurosea_d2.2

Table of contents

Executive summary

The task objective is to assess the role of the in situ networks and their future extension at improving the accuracy of future global CMEMS physical and BGC analysis and forecasting system.

Data assimilation experiments for the physical in situ observations will focus on the impact of the increased density of floats in specific regions, as planned in the Argo-2030 project. The impact of such extensions was assessed at global scale on the analysis. The experiments were conducted with the future improved version of the CMEMS global operational system. Another aspect explored here is the complementarity between the satellite observing component and the in situ one in constraining the different scales of the ocean circulation.

Concerning the BGC in situ observations, comparison to the BGC Argo floats helps to identify regions with a lack of observations or large model error that should benefit from an increased number of BGC Argo floats. Thanks to the BGC model, key regions for the export of organic carbon were also identified.

1. Analysis of BGC experiments

The goal of this task is to identify the regions where BGC-Argo floats observations should be enhanced. First, a comparison between global CMEMS BGC forecast system and BGC-Argo floats was conducted to determine two types of regions where BGC-Argo floats observations should be enhanced: 1) regions where the actual density of BGC-Argo floats is not sufficient and 2) regions where the model error is the highest. Second, we identified key regions for the export of organic carbon. There are large uncertainties about the transport of organic carbon from the surface to deep ocean, total export being estimated between 5 and 15 PgC y^{-1} (Boyd et al., 2019; Boyd and Trull, 2007; Henson et al., 2011; Siegel et al., 2014). These uncertainties mainly result from a lack of observations and a lack of understanding of the processes that contributes to the export of organic carbon. Using the global CMEMS BGC forecast system, we identified regions that contribute the most to the export of carbon in order to deploy BGC Argo floats in those critical regions.

1.1. Data

BGC-Argo float observations

The float data were downloaded from the Argo Coriolis Global Data Assembly Centre in France (ftp://ftp.ifremer.fr/argo). The CTD and trajectory data were quality controlled using the standard Argo protocol (Wong et al., 2015). The raw BGC signals were transformed to biogeochemical variables (i.e., O_2, Chl*a*, NO_3 and pH) and quality-controlled according to international BGC-Argo protocols (Johnson et al., 2018b; Schmechtig et al., 2015; Thierry and Bittig, 2018; Johnson et al., 2018a).

In the Argo data-system, the data are available in three data modes: “Real-Time”, ”Adjusted” and ”Delayed” (Bittig et al., 2019). In the “Real-time” mode, the raw data are converted into state variable and an automatic quality-control is applied to “flag” gross outliers. In the “Adjusted” mode, the “Real-time” data receive a calibration adjustment in an automated manner. In the “Delayed” mode, the “Adjusted” data are adjusted and validated by a scientific expert. While the “Real-Time” and “Adjusted” data are considered acceptable for operational application (data assimilation), the “Delayed” mode” is designed for scientific exploitation and represent the highest quality of data with the ultimate goal, when time-series with sufficient duration will have been acquired, to possibly extract climate-related trends. However, for some variables, only a

limited fraction of data is accessible in "Delayed-Mode". Consequently, for each variable, we selected the highest data modes, where at least 80 % of the data are available (see Table 1). We removed data with missing location or time information and flagged as "Bad data" (flag =4). Depending on the parameter and the associated data mode, we also excluded data flagged as "potentially bad data" (flag=3) (see Table 1).

Table 1. Data mode and QC flags of the BGC-Argo observations used in this study. In the Argo data-system, the data are available in three data modes, "Real-Time", "Adjusted" and "Delayed". See section 2a for a brief description of each data mode. The flags "3" and "4" refers to "potentially bad data" and "bad data", respectively. See also Bittig et al. (2019), for a more detailed description of Argo data modes and flags.

Parameter	Data mode	QC flags
Chl*a*	Adjusted and Delayed	All flags except 3 and 4
O_2	Delayed	All flags except 3 and 4
NO_3	Adjusted and Delayed	All flags except 3 and 4
pH	Adjusted and Delayed	All flags except 3 and 4

CMEMS global BGC forecast system

The global model simulation used in this study originates from the Global Ocean hydrodynamic-biogeochemical model, implemented and operated by the Global Monitoring and Forecasting Center of the EU, the Copernicus Marine Environment Monitoring Service (CMEMS). It is based on the coupled NEMO–PISCES model and it is constrained by the assimilation of satellite Chl*a* concentrations. The BGC model is forced offline by daily fields of ocean physic, sea ice and atmosphere. The physical ocean and sea ice forcing come from Mercator Ocean global high-resolution 1/12° ocean model (Lellouche et al., 2018) that assimilates along-track altimeter data, satellite Sea Surface Temperature and Sea-Ice Concentration, and *in situ* temperature and salinity vertical profiles. The BGC model has a 1/4° horizontal resolution, 50 vertical levels (with 22 levels in the upper 100 m, the vertical resolution is 1 m near the surface and decreases to 450 m resolution near the bottom). It produces daily outputs of Chl*a*, NO_3, O_2, pH, and weekly outputs of particulate organic carbon (POC) (resampled offline from weekly to daily frequency through linear interpolation) from 2009 to 2017 from 2009 to 2020. Following the approach of Gali et al. (2021), the POC simulated by the model corresponds to the sum of the two sizes classes of phytoplankton, the small detrital particles and microzooplankton modelled by PISCES. This particular combination of phytoplanktonic and non-phytoplanktonic organisms has been shown match the small POC observed by the floats (Galí et al., 2021). The Black Sea was not considered in the present analysis because the model solutions are of very poor quality. Finally, the daily model outputs were collocated in time and the closest to the BGC-Argo floats positions, and they were interpolated to the sampling depth of the float observations.

1.2. Methods

Comparison between global CMEMS BGC forecast system and BGC-Argo floats

For the four variables of interest (Chl*a*, NO_3, O_2, pH), we identify two types of regions where BGC-Argo floats observations should be enhanced: 1) regions where the actual density of BGC-Argo floats is not sufficient and 2) regions where the model error is the highest.

For each variable, all vertical profiles available from 2009 to 2020 are aggregated in 4°x4° spatial bins. The 4° distance is an upper estimate of the autocorrelation length scales for O2, nutrients, and pCO2 (comprised between 300 and 400 km) between 20° and 40° of latitude in both hemispheres (Biogeochemical-Argo Planning Group, 2016). Bins with less than 4 vertical profiles correspond to the regions where actual density

of BGC-Argo is not sufficient. In the spatial bin with more than 4 profiles, we compute a vertical profile of RMS Difference between the model and the BGC-Argo data in 50 m-depth bins from 0 to 200 m. We then select the maximum value to obtain a spatial map of RMSD. Finally, spatial bins with RMSD larger than the global median value of RMSD are deemed as regions where the model uncertainty is the highest.

Key regions for the export of organic carbon

To estimate the key regions for the export of organic carbon, we follow the method proposed by Resplandy et al. (2019) to estimate the total export of organic carbon at z= 100 m. The total export of organic carbon is computed off-line as the sum of the gravitational, vertical advective and mixed layer pumps. The gravitational pump, $E_{gravitational}$ across the depth level z= 100 m, for a given day, is computed from the sinking speed of POC, W_p, (Aumont et al., 2015), and the daily concentration of POC,

$E_{gravitational} = - W_p * POC$ (1)

The vertical advective pump across the depth level z= 100 m, and for a given day is computed from the daily vertical velocity, W, and the daily concentration of POC,

$E_{advective} = W * POC$ (2)

Finally, the mixed layer pump across the depth level z= 100 m, and for a given day is computed using the daily vertical diffusion coefficient K_z and the daily vertical gradient in POC,

$E_{mixedlayer} = - K_z * dPOC/dz$ (3)

1.3. Results

Comparison between global CMEMS BGC forecast system and BGC-Argo floats

Figure 1 shows the regions where the BGC-Argo floats observations should be enhanced for Chl*a*, NO_3, O_2 and pH. The regions where the actual density of BGC-Argo floats is not sufficient (yellow areas) are the subtropical gyres, the equatorial band and the Indian Ocean. For pH, the lack of observations is even more critical, as pH data are only available, in the southern oceans and in a few places in the North Pacific. In the regions with sufficient data coverage, the model error is the highest (blue areas) in the western boundary currents, the Southern Ocean, the subpolar gyres and the Nordic Seas.

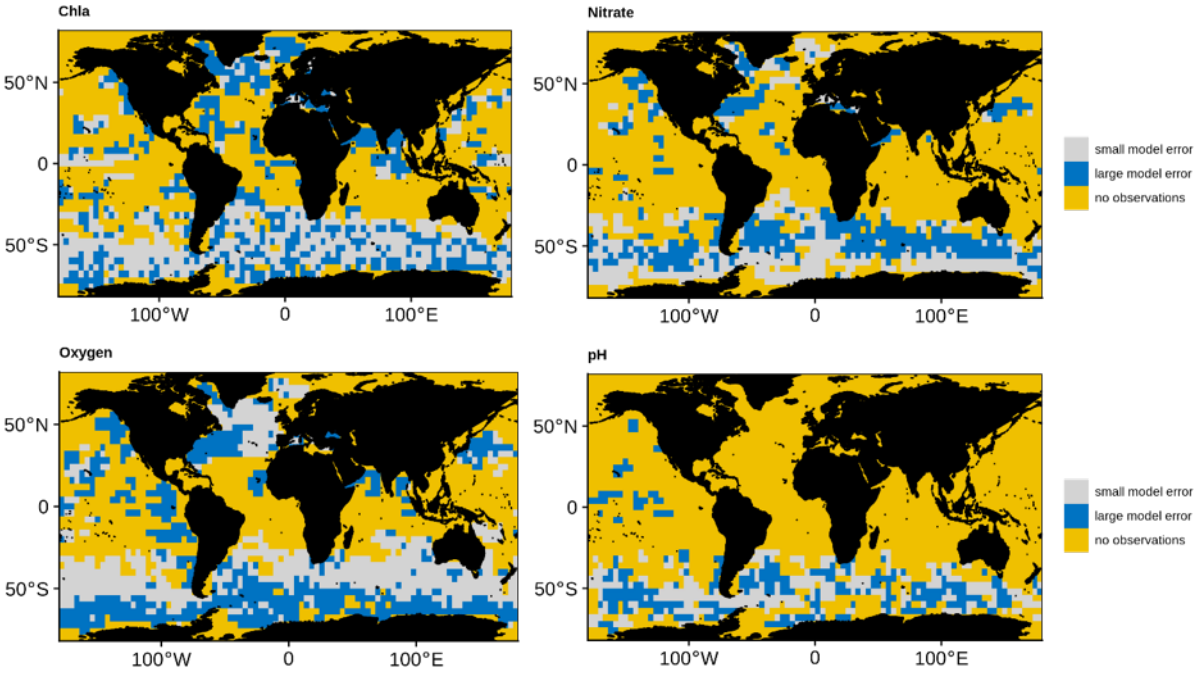

Figure 1. Regions where BGC-Argo floats observations should be enhanced. Yellow: Areas where there are less than 4 BGC-Argo vertical profiles per bins of 4° by 4°. Blue: regions where the RMS differences between the model and the observations are the highest. Grey: regions where the RMS differences between the model and the observations are the lowest.

Key regions for the export of organic carbon

Figure 2 shows the 2010 mean annual total export of POC. The hotspot regions for the export of organic carbon from the surface to the deep sea are the western boundary currents, the eastern boundary upwelling systems (Canary, Benguela, Humbolt, California), and the Falkland current. The export in these hotspot regions is dominated by the gravitational pump that accounts for 90 % of the total export, the mixed layer pump contributes for 10 % and the vertical advective pump is small and contributes for less than 1 % of the annual export at 100 m.

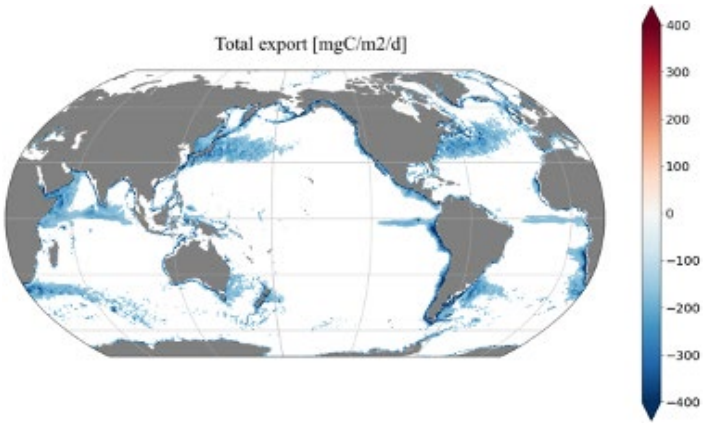

Figure 2: 2010 Mean annual export of organic carbon at 100 m computed from CMEMS global forecasting system (mgC/m^2/d). Only the regions where the export is the largest (< -150 mgC/m^2/d) are represented. Negative values indicate a downward export of POC whereas positive values indicate an upward export of POC.

2. Analysis of the physical experiments

As the ocean plays a fundamental role in regulating climate variability, it has been recognized in the 1980-90's that systematic ocean observations are essential to understand and monitor the changing climate of the Earth. Initially focused on capturing oceanic variability at large spatial and temporal scales mainly for academic research, the scope of sustained ocean observations is now expanded to multi-scales sampling, a multi-disciplinary approach and beyond only scientific research. These observations are now integrated in the global ocean observing system (GOOS), which has been essential for the development of ocean models to validate and optimize numerical representation.

Given these different sources of ocean information from satellites, in situ platforms and models, an initiative from the Global Ocean Data Assimilation Experiment (GODAE) has strongly supported the development of global ocean analysis and forecasting capabilities for operational oceanography. The objective was to build operational data assimilation systems able to provide to the scientific and broader communities the most accurate gridded fields of the main physical ocean variables. While benefits of such integrated approach are conceptually evident, it becomes however more and more difficult to determine how information from observations is used, as the diversity of in situ platforms assimilated in data assimilation system is continuously increasing. In addition, analyses and forecasts might be strongly sensitive to data assimilation choices and observation impact might strongly differ depending on systems (Fujii et al., 2019).

Ocean observations are intertwined in multiple ways through the entire effort to interpret the climate system. In addition to the direct constraint to the system through data assimilation, indirect contributions are also very important for the accuracy of the analysis, albeit less visible. They are fundamental for climatologies and long timeseries are necessary to develop, test and tune models. They are indispensable for the calibration and correction of other measurements (e.g., in situ data is essential for the ongoing accuracy of much satellite sampling). They are also at the heart of data assimilation techniques, to define prescribed error, establish a good consistency between local (in situ profiles) and integrated variables (altimetry), and to determine the most accurate mean dynamic topography for the assimilation of altimetry (Hamon et al., 2019; Gasparin et al., 2021).

The purpose of the task is thus to analyze the role of in situ observations for data assimilation systems, which are at the heart of operational systems and reanalyses. A first objective is to disentangle the added value of satellite and in situ observations and determine their ability to constrain specific temporal and spatial scales of the physical ocean state. This aims to demonstrate the complexity of merging ocean observations having distinctive characteristics of sampling, and to determine the ability of the data assimilation system to incorporate specific data sets. By identifying skills and weaknesses of the data assimilation system, this work should pave the way for future improvements to increase benefits of ocean observations for data assimilation systems. Another objective is to illustrate the need for continuing assessments and improvements of data assimilation techniques for in situ observations.

2.1. Data and methodology

The present study is based on a comparison of a series of numerical experiments, called observing system simulation experiments (OSSE), in which different designs of ocean observations have been assimilated in a data assimilation system. The present work follows the OSSE requirements proposed by Halliwell et al. (2014) as much as possible. The OSSE is thus composed of (i) an unconstrained simulation, named the Nature Run, assumed to provide a good representation of the "true" ocean variability over the space and time scales of

interest, (ii) a global data assimilation system ingesting the simulated observations, and (iii) a set of synthetic realistic observations simulating different observing system designs. Experiments are performed at global scale on a 3-yr period, 2015-2017.

Numerical experimental system

The Nature Run (NR) corresponds to the free-running version (i.e., without assimilation) of the GLORYS12 reanalysis (Lellouche et al., 2021). This unconstrained simulation has been developed at Mercator Ocean International, based on the NEMO ocean model (Nucleus for European Modelling of the Ocean, Madec and The NEMO Team, 2008), using a 1/12° ORCA grid (horizontal resolution of 9 km at the equator, 7 km at mid-latitudes and 2 km near the poles). The ocean model is forced at the surface with the atmospheric fields from the ERA-Interim reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al., 2011). The NR was initialized in October 1991, from the EN4 gridded fields of temperature and salinity (Good et al., 2013). Assuming that the velocity field is zero at the start, the model physics then spins up a velocity field in balance with the density field for 1 year. The NR was run up until the end of 2017, during which the period of 2015 to 2017 was used to generate simulated observations.

The data assimilation system (DAS), used to perform OSSEs, is based on the operational monitoring and forecasting system GLO4, which uses NEMO ocean model with a ¼° ORCA grid type (Lellouche et al., 2021). The DAS was initialized on January 07, 2015, using fields from a 4-yr spin-up run, and was run up until the end of 2017. The ocean model is forced at the surface by the operational atmospheric fields from the ECMWF-Integrated Forecast System (ECMWF-IFS). In addition to the ocean model, data assimilation procedures based on a reduced-order Kalman filter derived from a SEEK filter (SAM2, Brasseur and Verron, 2006) are used for the assimilation of satellite and in situ observations. A 3D-Var correction for the slowly evolving large-scale error of the model in temperature and salinity is applied. More details can be found in Lellouche et al. (2018). Note that, unlike Lellouche et al. (2018), no mean dynamic topography is used for referencing the altimetric sea level anomaly, since the total sea surface height is directly assimilated in the system.

Design Experiments and simulated data sets

A total of six global experiments has been performed to disentangle the role of ocean observations in the current ocean observing system design and demonstrate potential outcomes of Argo extensions (Table 2). By assimilating observing system components separately, we aim to identify the ability of observing arrays to resolve small and large spatial and temporal scales and highlight complementarity and redundancy of ocean observation information from an operational oceanography perspective. Each experiment is characterized by the assimilated data sets, which have been extracted from the NR. They are based on subsampling the daily fields of the free version of the 1/12° GLORYS12 reanalysis (Lellouche et al., 2021), called FREEGLORYS12 in the following, at the space and time location of each observation. As the study investigates the role of in situ observations as part of the integrated ocean observing system, simulated data sets include both satellite and in situ components, and numerical experiments use the same satellite configuration but differ in their in situ component.

Table 2. Experiments performed in this study.

Experiments	Assimilated observing arrays
FREE	No
ONLYSAT	Altimetry, SST
ONLYSITU	Argo_Nominal, Mooring_Nominal
NOMINAL	Altimetry, SST, Argo_Nominal, Mooring_Nominal
ENHANCED_AR	Altimetry, SST, Argo_Enhanced, Mooring_Nominal
ENHANCED_MO	Altimetry, SST, Argo_Nominal, Mooring_Enhanced
ENHANCED_ARMO	Altimetry, SST, Argo_Enhanced, Mooring_Enhanced

The common simulated satellite observations consist in sea surface height (SSH) and sea surface temperature (SST) variables (Figure 2), sea ice concentration is not assimilated. The SSH data set is built from a constellation of the three nadir satellites Jason-2, Jason-3 and Sentinel-3a. The positions (longitude, latitude, time) were extracted from Copernicus Marine Service Sea Level TAC (Thematic Assembly Center) multi-mission along-track L3 altimeter products. Each satellite provides around 50,000 measurements per date (10-day repeat cycle and 13 orbits per day for Jason-2 and Jason-3; 27-day repeat cycle and 14 orbits per day for Sentinel-3a). The SST data set consists in daily maps obtained from the Copernicus Marine Service ODYSSEA multi-sensor L3S product. This product consists in a fusion of SST observations from multiple satellite sensors, daily, over a 0.1° resolution global grid. For consistency with the resolution of the data assimilation system, SST observations are re-mapped on a 0.25° horizontal resolution grid.

The simulated in situ data sets consist of subsurface vertical profiles of temperature and salinity (T/S) from two historical global in situ networks; the Argo Lagrangian float array (www.argo.net) and the global tropical moored buoy array (www.pmel.noaa.gov/gtmba). Unlike satellite observations, simulated in situ observations are built on theoretical configurations to ease the interpretation of results, which are less regionally and temporally dependent on the availability of existing data (Gasparin et al., 2019). For each of those networks, two different designs are considered representing the current (NOMINAL) and enhancements (ENHANCED) as follows (Figure 3).

- Argo-NOMINAL mimics the standard configuration corresponding to one Argo float per 3°x3°x10-day square, sampling the 0-2000 m upper-ocean globally. Locations of T/S profiles are randomly distributed in a 3°x3° square, each square being sampled every 10 days. The day of the first 10-day cycle is randomly distributed in the first 10-day window and the space position is different for each 10-day cycle. This T/S configuration counts around 470 profiles per day from 3700 floats, with measurements located at the model vertical levels (including 22 levels within the upper 100 m, with 1-m resolution at the surface and 450-m resolution at the bottom).
- Argo-ENHANCED is based on the latter configuration, with one added float per 3°x3° square in high-activity regions, i.e, in western boundary currents and in equatorial/tropical regions (Figure 1) following international recommendations (Roemmich et al., 2019).
- Mooring-NOMINAL uses the position of tropical moorings during the 2020-2021 period (no TRITON moorings in the western tropical Pacific). Vertical levels of T/S profiles are based on "standard

instrumental depths" given by the GMTBA webpage[1]. Around 11 depth levels are located in between the surface and 500 m, with 20-m resolution in the upper 150 m. Note that both subsurface T/S profiles are found in the Indian and Atlantic Ocean, but only temperature in the Pacific.

- Mooring-ENHANCED is mostly characterized by an increase vertical resolution, which follows the recommendations of the tropical community (Foltz et al., 2019; Hermes et al., 2019; Smith et al., 2019): temperature sensors at 1 m, every 5 m from 5 to 30 m, every 10 m from 30 to 60 m, from 60 to 500 m, with vertical resolution similar to present (depending on the longitude in the basin) and salinity sensors at 1 m, every 5 m from 5 to 30 m, every 10 m from 30 to 80 m; 100 m. Strong modification of the spatial distribution of tropical moorings in the Pacific follows conclusions from the Tropical Pacific Observing System 2020 (Kessler et al., 2021).

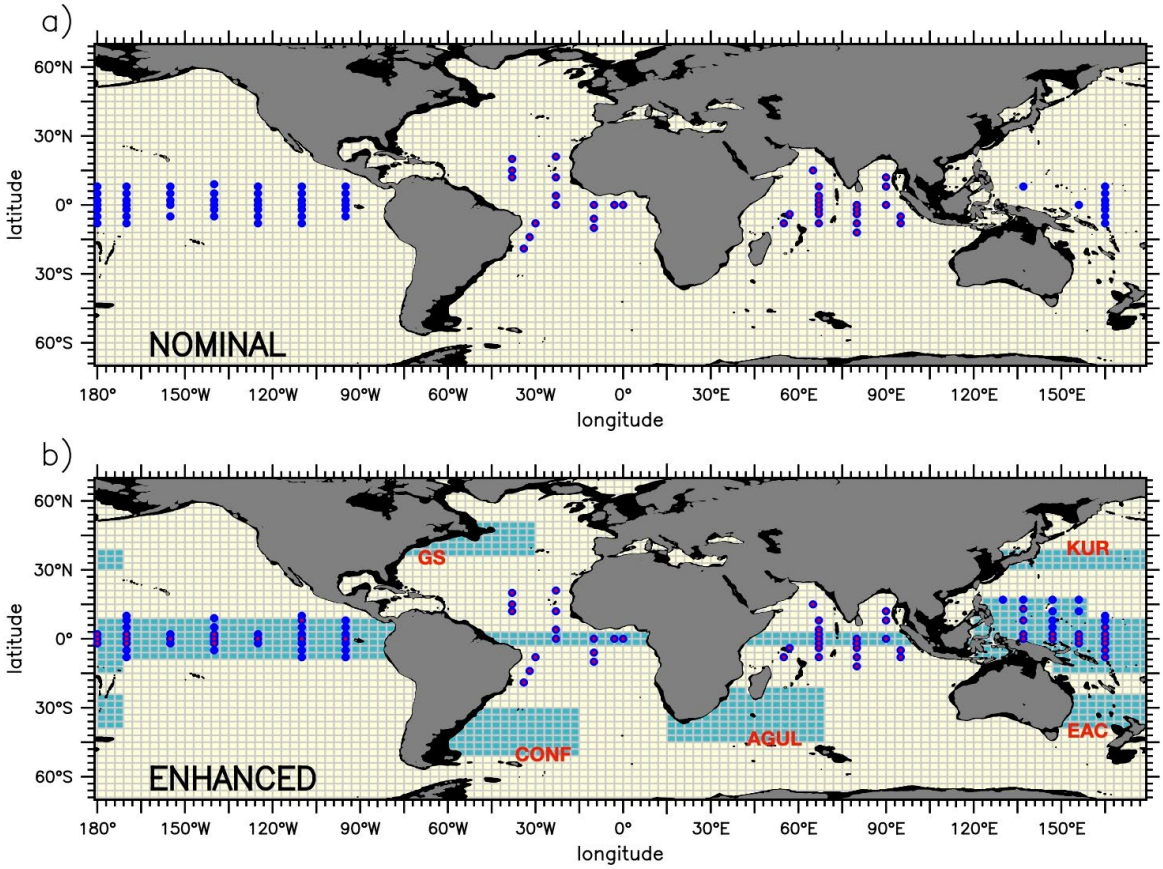

Figure 3: Nominal and enhanced configurations of simulated temperature and salinity (T/S) profiles for Argo floats (shading) and tropical moorings (dots). Argo sampling corresponds to 1 and 2 floats per 3°x3°x10-day square in yellow and green, respectively. Subsurface T/S observations from moorings are indicated by blue (T) and red (S) dots.

To mimic the noise included in existing observations, simulated data sets must deviate from the Nature Run realization. We follow the methodology of Gasparin et al. (2019), who included three different kinds of perturbation. First, simulated observations are generated from the Nature Run fields, which are at a higher horizontal resolution (at 1/12° resolution) than the data assimilation systems (at ¼° resolution). Small-scale variability embedded in the Nature Run (eddy-resolving system) is thus considered as noise in the experimental system (eddy-permitting system). Then, a representation error introduces vertically and

[1] https://www.pmel.noaa.gov/gtmba/moorings

horizontally correlated error resulting from unresolved or poorly resolved small-scale processes by the data assimilation system (e.g., inertial waves). A time-shifting technique, usually used by the atmosphere community (Huang and Wang 2018), generates weekly variability by randomly shifting the Nature Run fields by ± 3 days (following a uniform distribution, either 3 days before or 3 days after the given date). Finally, instrumental noise is added to each observation as an uncorrelated error following a Gaussian distribution with the standard deviation given by the instrumental uncertainty (0.35°for SST; 3 cm for SSH; 0.01°C/0.01 for Argo T/S profiles; and 0.02°C/0.02 for Mooring T/S profiles; Cabanes et al., 2013).

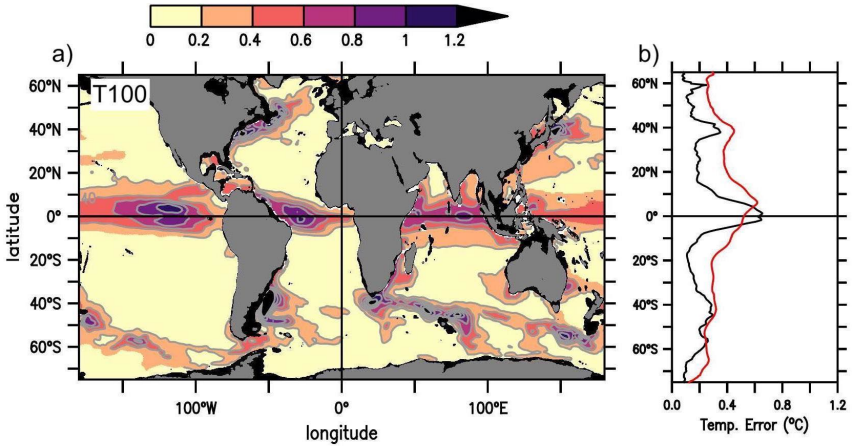

Figure 4: Amplitude of errors in simulated observations for 100-m temperature: (a) spatial maps, (b) zonally averaged. For comparison to the error estimate (black line), zonally averaged prescribed error in the operational system is shown (red line).

Calibration of experiments

As the reliability of observing system simulation experiments to correctly provide impact assessment partly lies in defining appropriate errors included in the synthetic observations, perturbation added to the Nature Run fields is evaluated for the 100-m temperature and 10-m salinity by computing Root-Mean-Square difference between the original Nature Run fields and the 3-day shifted Nature Run fields (Figure 4). Error due to time-shifting is higher in high variability regions such as western boundary regions, tropics, and in the Southern Ocean. Of the order of O (0.02°C) for 100-m temperature and O (0.05°C) for 10-m salinity (not shown), the representation error dominates the two others (Gasparin et al., 2019). The comparison with the amplitude of the prescribed error in the operational system demonstrates that perturbation introduced in simulated observations are appropriate (Figure 4). This is confirmed by comparing temperature perturbation at three mooring locations in the equatorial Pacific. The amplitude of the perturbation shows a maximum at the thermocline level, similarly to the prescribed error and the high-frequency variability (higher than 7 days) provided by mooring observations (not shown).

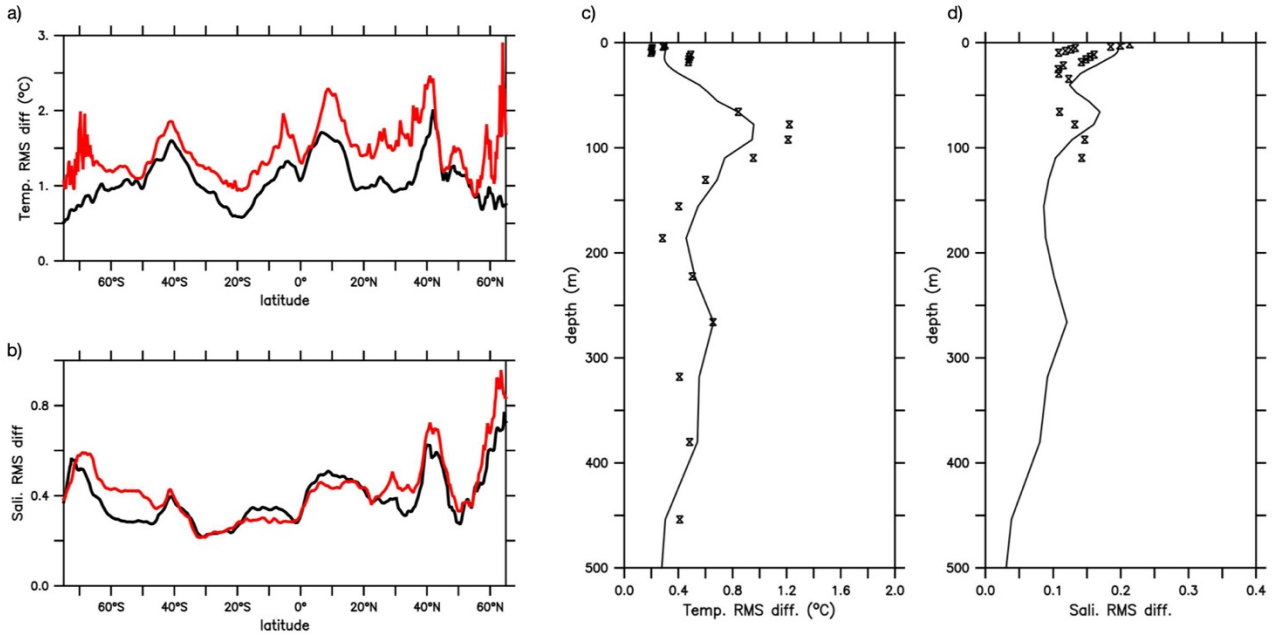

Figure 5: (a,b) 100-m temperature and 10-m salinity RMS difference, zonally averaged, between the free and assimilated simulations for the OSSE system (FREE and NOMINAL; black) and the reanalysis system (FREEGLORYS12 and GLORYS12; red). (c,d) Temperature and salinity RMS residuals at 23°W, 0° (Atlantic) from the OSSE system (distance between NOMINAL and simulated observations) and the GLORYS12 reanalysis.

Another validation of the good calibration of simulation experiments is to compare the distance between free and assimilated experiments with the distance between FREEGLORYS12 and GLORYS12. In Figure 5 ab, the 100-m temperature and 10-m salinity RMS difference between FREE and NOMINAL experiments, zonally averaged, ranges between 1.0 and 1.5°C for temperature and 0.25 to 0.75 for salinity. The consistency with the distance between FREEGLORYS12 and GLORYS12 demonstrates that a realistic behavior of experiments. In addition, the distance of NOMINAL simulation from the simulated observations at a mooring point in the equatorial Atlantic is compared with the distance of GLORYS12 from mooring observations and confirms the good calibration of our experimental framework (Figure 5 c and d).

2.2. Scale dependency of ocean observation contributions

As usual metrics in operational centers are mostly based on statistics in the observation space, data assimilation skill is often reduced to the predominant scales. To evaluate the ability of ocean observations to improve ocean state estimates from a ¼° data assimilation system, a detailed description of the signal produced by the eddy-permitting model will help to better understand the regional dependency of spatial and temporal scales. The first objective is thus to identify the relative contribution of scales to the total variability by separating the total signal into submeso-, meso- and large scales. This will allow to depict the role of observing system components on these specific space and time scales and help to better understand the regional dependency of data assimilation errors. Temporal and spatial scales shorter than 20 days and 100 km, respectively, are referred to as “submesoscale variability”. This variability, unresolved by the 7-day assimilation window and the ¼° horizontal grid of the eddy-permitting model, is isolated by applying a 1°x1°x20-day high pass filter on the gridded fields. Climatic signals are obtained by applying a 9°x9°x100-day running mean filter to represent “large scale variability”. Finally, the difference between the 9°x9°x100-day and 1°x1°x20-day smoothed series is referred to “mesoscale variability” to define intermediate signals such

as mesoscale eddy and intraseasonal signals, with temporal and spatial scales of 20-100 days and 100-1000 km, respectively.

Regional dependency of spatial and temporal scales

In Figure 6a, the variability of the surface steric height (SH) referred to the bottom, based on the unfiltered gridded fields of the FREE experiment, shows the regional dependency of ocean dynamics. High variability regions are clearly identified in western boundary regions and in the Southern Ocean, with amplitude reaching more than 12 cm. These regions are characterized by instabilities of the strong mean flow generating meanders and eddies (Ducet and Le Traon 2001). Moderate variability regions are seen in the tropical Indian Ocean and western Pacific, associated with instabilities of background tropical currents, also leading to the generation of eddies (Willett et al. 2006). Low variability regions are in the center of oceanic gyres. The standard deviation of dynamic height was then zonally averaged (Figure 6 b) to show the latitude dependence of the total signal. Spatial and temporal filters were applied to the unfiltered SH fields to compute the zonally averaged standard deviation of submesoscale, mesocale and large scales (Figure 6 c). At latitudes of the high and moderate variability regions, mesoscale variability represents almost 2/3 of the variability. At other latitudes, large-scale is equal or higher than mesoscale variability. Submesoscale variability which is usually dominated by coherent vortices, fronts and filaments, is unresolved here due to the ¼° horizontal resolution of the eddy-permitting model.

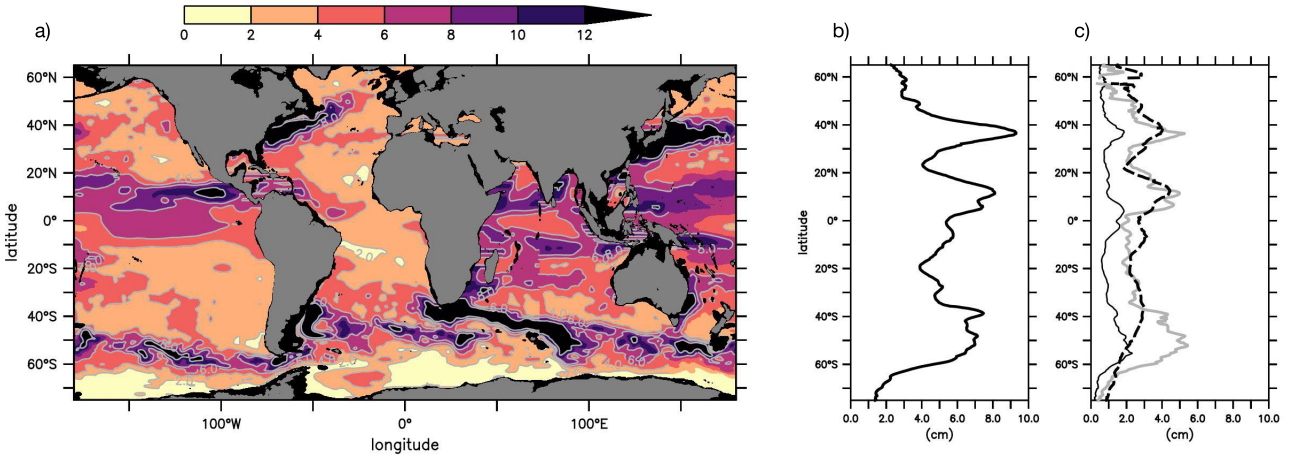

Figure 6: Standard deviation of the daily steric height (SH, cm) from the FREE experiment ((a) spatial map, (b,c) zonal-average, black line). For comparison, zonally-averaged standard deviation of the daily SH fields of the submesoscale (black line), mesoscale (gray line) and large-scale variability (dashed line) are also shown.

As numerical models forced by atmospheric fields can capture part of oceanic variability, error levels are likely not identical to energy levels, and observation impact must consider model skills. For that reason, the amplitude of the submesoscale, mesocale and large-scale signals is then compared to the errors of the FREE experiment, estimated from the RMS difference from the Nature Run (Figure 7). The amplitude of the total error is quite similar to that of the signal, with slightly higher error amplitude at latitudes of high variability regions and slightly lower in the tropical band, but the scale decomposition demonstrates that error is differently distributed over scales than the signal. For instance, while high variability regions of the Northern Hemisphere have similar mesoscale and climatic signals, errors are mostly dominated by mesoscale variability. Conversely, large-scale signal is significantly higher than submesoscale signal, but error is similar over these scales.

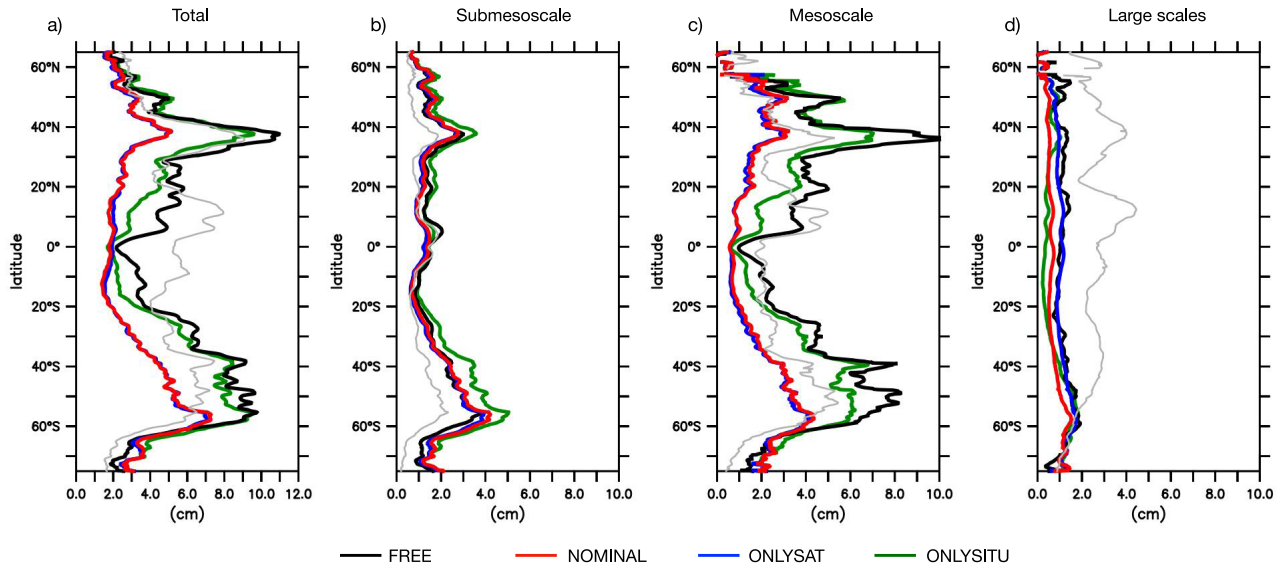

Figure 7: Zonally averaged steric height (SH, cm) RMS difference between the Nature Run and experiments (FREE, NOMINAL, ONLYSAT, ONLYSITU) for the submesoscale (smaller than 1°x1°x20-day), mesoscale (between 1°x1°x20-day and 9°x9°x90-day) and large-scale variability (larger than 9°x9°x90-day). For comparison, the standard deviation of the Nature Run SH, zonally-averaged, is also shown (gray).

Assessing impacts of an observing system component thus requires considering the scale dependency of errors. As analysis skills are mostly based on the distance from observations through the RMS difference between observations and analysis, statistics on the observation space mostly favors the dominant scale of the error, and thus, masks lower-amplitude scales. Unlike operational systems, numerical experiments allow to separate submeso-, meso-, and large-scale variability. In addition, this metric used alone does not inform on the error amplitude relative to the signal amplitude, and we use the term "percent of represented variance" as the comparison metric, calculated as 1 minus the proportion of the residual variance (e.g., the variance of the difference between independent data divided by the signal variance).

Dominant impact of observing system component

Based on the comparison of several experiments assimilating separately (ONLYSAT, ONLYSITU) and conjointly satellites and in situ data sets (NOMINAL), the aim here is to investigate the contribution of ocean observing system components to the data assimilation by disentangling impacts from satellites and in situ observations. It is likely that observation impacts based on experiments assimilating separately in situ and satellites components are not representative of impacts in experiment assimilated conjointly both data sets. But in any case, it can be seen as an upper bound for information brought by observing system components.

Added value of satellites for mesoscale activity

Figure 7c shows the RMS difference from the Nature Run of the ONLYSAT, ONLYSITU and NOMINAL for mesoscale variability. Even though the magnitude of error in ONLYSITU experiment is lower than the unconstrained experiment (FREE), it remains higher than the signal amplitude at all latitudes. In contrast, an important error reduction is seen in the ONLYSAT experiment, especially in western boundary current regions where error decreases to 3-4 cm, in comparison to 8-10 cm in the FREE experiment. With a similar SH error at mesoscale, the NOMINAL experiment recovers the entire gain from satellite observations. This is consistent with the scales of conventional one-dimensional nadir-looking altimeters having the ability to resolve wavelengths down to about 50–100 km depending on the specific satellite and geographic locations (Dufau

et al. 2016). However, it is noteworthy to compare with multi-mission product, with an effective resolution lower than 150 km (Ballarotta et al., 2019), plus around 30 days. Even if SST capture length scale, SH reflecting the integration of subsurface density fields, is mainly constrained by altimetry.

Sea surface height information provided by altimetry is a depth-integrated quantity, and is closely related to SH. A question arises on how this information is projected in depth. In Figure 8 a,b, the percentage of variance represented for each experiment is globally-averaged. The vertical shape of profiles indicates that surface layer is better represented compared to the deeper one, likely due to the decreasing signal with depth associated with a decreasing signal-to-noise ratio. However, impact of satellite observations is clearly seen in both temperature and salinity variables, with a gain of 15% of temperature variance in depth to 50% at surface (5% of salinity variance in depth to 30% at surface).

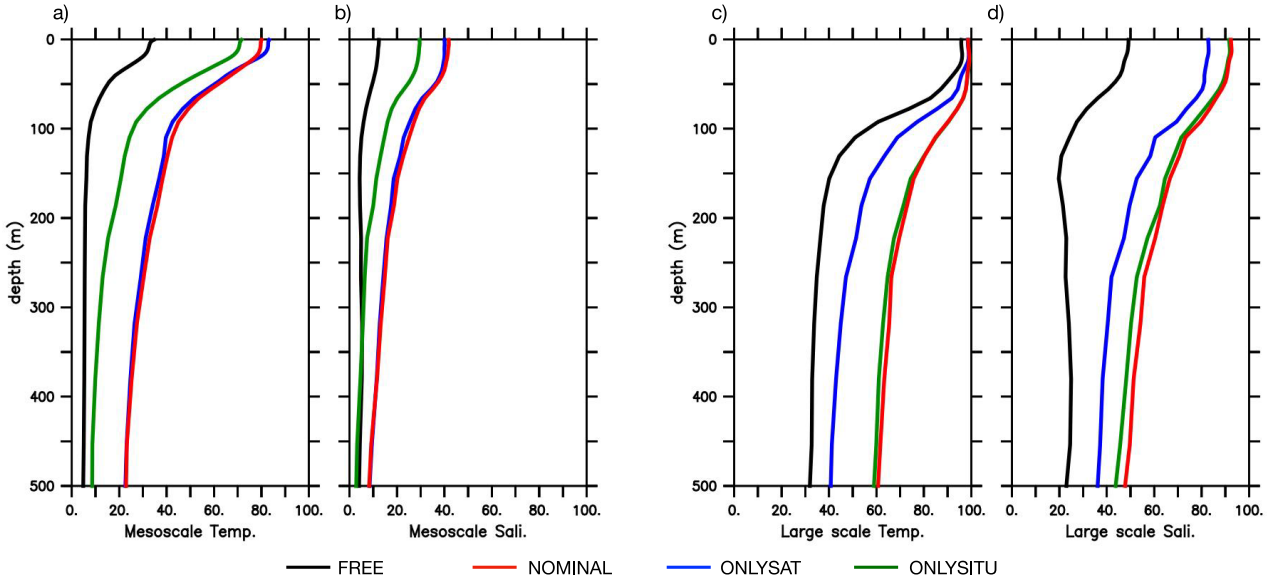

Figure 8: Globally averaged percentage of represented variance of the Nature Run for subsurface temperature and salinity at mesoscale and larger scales based on the FREE, NOMINAL, ONLYSAT, ONLYSITU experiments. Timeseries have been filtered with running mean filters (1°x1°x20-day, 9°x9°x90-day) as explained in the text.

Added value of in situ for climatic variability

In contrast to mesoscale variability, large scale SH error is significantly lower than the signal for all experiments (including the unconstrained FREE). While large scale error is slightly lower in ONLYSAT compared to FREE experiment, the error reduction in ONLYSITU experiment, likely explaining the error in NOMINAL. The contribution of in situ component for the large-scale signal is confirmed in Figure 8 cd. The variance explained by temperature is close to one at the surface decreasing to 30% in depth. ONLYSAT experiment increases the percentage of variance of around 10%, but the highest improvement is seen in the ONLYSITU experiment, which shows an increase of around 30%. Similar improvements are seen for salinity. This illustrates that in situ observations are essential to represent the thermohaline structure.

2.3. Potential outcomes of in situ observing system enhancements

The next objective of the present work is to evaluate potential outcomes of international recommendations regarding the evolution of the in situ observing system in determining the added value of doubling Argo in western boundary currents in equatorial regions and the gain of enhancing the vertical resolution of tropical moorings.

Western boundary currents are a fundamental element in the ocean circulation system and play a critical role in the poleward redistribution of accumulated heat in low-latitude regions. They are extensive regions of intense air-sea interaction, where the ocean loses heat and moisture to the atmosphere and absorbs carbon dioxide. Given their impact on weather and climate both locally and remotely, on time scales of days to decades, significant benefit can be achieved by improving their representation in numerical models of the atmosphere and ocean. With high levels of mesoscale variability, enhancing Argo sampling in western boundary current is expected to reduce noise in tracking the large-scale temperature and salinity fields.

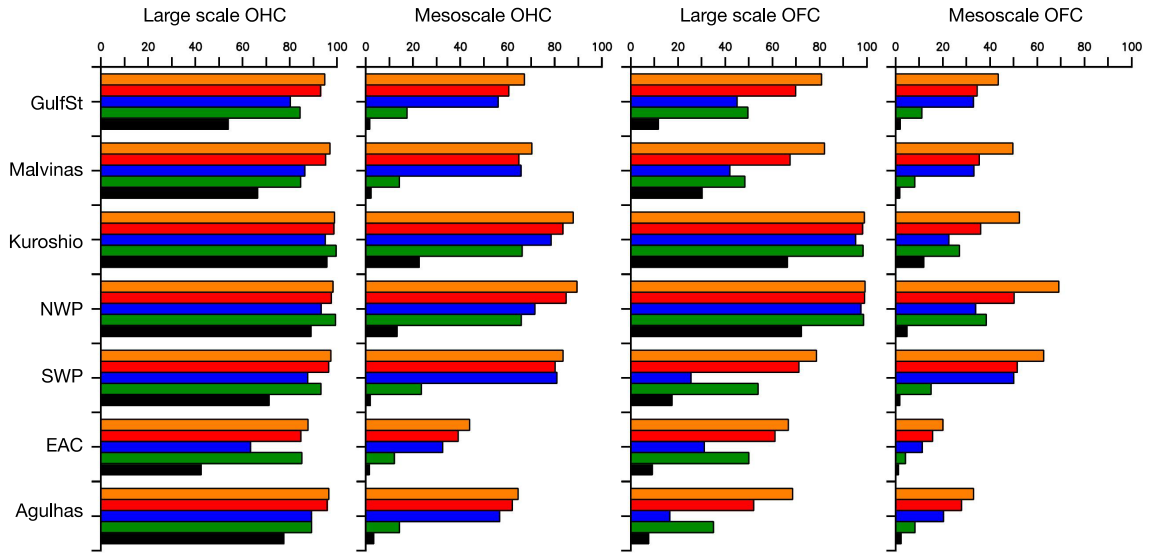

Figure 9: Percentage of represented variance of the Nature Run, area-averaged in western boundary current regions, for 0-700 m Ocean Heat (OHC) and Freshwater Contents (OFC) at mesoscale and larger scales based on the FREE (black), NOMINAL (red), ONLYSAT (blue), ONLYSITU (green) and ENHANCED_AR (orange) experiments.

A supplementary experiment, ENHANCED_AR, is thus considered here, in which the number of Argo profiles have been doubled in western boundary current regions. Based on temperature fields, the 0-700m ocean heat content is computed based on the depth-integration of temperature anomaly from the 2016-2017 temporal mean multiplied by the heat capacity (3900 J/kg/°C) and ocean density (1024 kg/m^3). Considering that about 3 cm of freshwater are needed to dilute 1 m of seawater by 1 psu, the freshwater anomaly, expressed in meters, is the depth-integral of salinity anomaly multiplied by -0.03. Similar spatial and temporal filters have been applied to the OHC and OFC fields to determine the area-averaged percentage of represented variance in different western boundary regions, depending on the experiments. Results are presented in Figure 9 for the Gulf Stream (GulfSt), Malvinas current, Kuroshio, North Western Pacific (NWP), South Western Pacific (SWP), East Australian Current (EAC) and Agulhas regions, those regions are shown in blue in Figure 3.

The doubling of Argo benefits to the representation of OHC and OFC at mesoscale and larger scales in the seven western boundary current regions. While a limited impact is only seen on large scale OHC estimates, OHC mesoscale is better represented in the ENHANCED_AR experiment. Important improvements of the order of 10-20% of the represented variance are preferentially seen in OFC estimates. The most important contribution of doubling Argo appears in mesoscale OFC.

Other diagnostics are currently performed in the tropical band. They are not finalized but will be available in the publication in preparation: *Gasparin, Lellouche, Cravatte, Rémy; Scale dependency of the upper-ocean observing system for a global data assimilation system, in prep.*

2.4. Discussion and conclusion

Based on a series of observing system simulation experiments (OSSE), a detailed description of the contribution of ocean observations from satellites and in situ networks for a global data assimilation system is presented here. The experimental system is composed of (i) an up-to-date ¼° data assimilation system, (ii) a Nature Run corresponding to the free version of the 1/12° GLORYS12 reanalysis, and (iii) the generation of simulated data sets by subsampling the Nature Run at the location and time of each observation of a given observing system. Additional errors have been introduced to the simulated observations to consider a realistic noise induced by unresolved or poorly resolved processes and instrumental noise. The good calibration of the system is given by the similar distance of the assimilation and free experiments with that of the GLORYS12 and its free version. The similar distance between observations and analysis fields in the experimental and reanalysis framework also provides a good confidence to the realistic calibration of the experimental system.

A decomposition of the signal, used to define assessment metrics focused on specific temporal and spatial scales (i.e., submesoscale, mesoscale and large scales), highlights the regional dependency of errors in a data assimilation system. In general, mesoscale activity dominates errors in simulations with no data assimilation, especially in western boundary currents, while they are equivalent to large scale errors in mid-latitude and low-latitude regions. These experiments demonstrate the data assimilation system's ability to represent 40-80% of the temperature variance at mesoscale (20-30% for salinity), and more than 80% of the variance for large-scale variability. An important complementarity of satellites and in situ observations is seen; satellites information, mostly through altimetric data, strongly constrain mesoscale variability, while in situ data, mostly through Argo floats, provide a large-scale information. Each observing system component provides a substantial gain for the data assimilation depending on the selected scales. In addition to integrated quantities, such as steric height, the added value of the observing system is clearly seen for subsurface temperature and salinity. While thermosteric dominates halosteric component for the steric height, the representation of salinity is significantly improved by satellite observations.

In addition to the evaluation of observing system configurations approaching the current observing system design, the numerical experimental system allows to assess the contribution of extra observations, as suggested by international recommendations. One of the main enhancements is the doubling of the number of Argo floats in western boundary currents regions. It is demonstrated that the representation of the ocean heat and freshwater content in these regions would be significantly improved, at both mesoscale and larger scales. Enhancements of moorings observations are mostly through additional salinity sensors in the tropical Pacific (and a modification of the mooring locations), and an increase of the vertical resolution in the surface layer in the Indian and Atlantic basins. Rather than substantial improvements at large scale, impacts of enhancements are mostly seen locally. As mentioned in several studies, characteristics of tropical moorings lead to local improvements. Further developments of data assimilation should make a better use of these observations.

Unlike many other data assimilation systems, benefits of in situ ocean observations are determined here as a complementary information of satellite data sets. The number of experiments, including simulations that assimilate observing system components separately, is a central point of this study and demonstrates the

complexity of impact studies in a multivariate system. As expected, benefits of in situ observations are lower in the context of the integrated observing system. The added value of ocean observing system is clearly demonstrated for specific temporal and spatial scales, but impact of ocean observations can be also redundant with regards to a data assimilation system. Redundant information is indispensable to intercalibrate measurements having their own uncertainty, but also, to help the development and qualification of data assimilation systems.

In conclusion, the study aims at highlighting benefits of observing system components for a data assimilation. Efforts have been made to extract important messages that can likely be applied for other data assimilation systems. However, it is noteworthy that observations impacts are strongly dependent on the data assimilation techniques, and more advanced techniques might increase the gain of a specific observing system. The question of ocean observation impact is addressed here from the assimilation angle, but it is important to point out that ocean observations strongly impact at the different steps of the development and qualification of a data assimilation system, with an accurate impact much more complex to determine.

Conclusion

CMEMS BGC and physical ocean monitoring highly rely on in situ and satellite observations to provide accurate analysis. Observations are important to constrain the analysis during the assimilation process but also to evaluate model and system development.

In this work, the BGC Argo float observations were used for model evaluation and to identify ocean regions with large forecast errors and regions that should be better sampled to allow better accuracy estimation of the model outputs. Using the BGC model simulation, regions that contribute the most to the export of carbon were also identify and be the one where BGC Argo deployment could be critical.

The data assimilation experiments for the in situ observation component clearly highlight the complementarity between altimetry, which constrain the meso-scale variability of temperature and salinity as the in situ observations constrain the large scale variability. They also show that the planned increased density of Argo floats in the Tropics and Western Boundary currents will have a positive impact on reducing the analysis error on temperature and even more salinity compared to the “today” Argo array sampling.

References

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M.: PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies, Geosci. Model Dev., 8, 2465–2513, https://doi.org/10.5194/gmd-8-2465-2015, 2015.

Ballarotta, M., Ubelmann, C., Pujol, M.-I., Taburet, G., Fournier, F., Legeais, J.-F., Faugère, Y., Delepoulle, A., Chelton, D., Dibarboure, G., and Picot, N.: On the resolutions of ocean altimetry maps, Ocean Sci., 15, 1091–1109, https://doi.org/10.5194/os-15-1091-2019, 2019.

Biogeochemical-Argo Planning Group: The scientific rationale, design and implementation plan for a Biogeochemical-Argo float array, https://doi.org/10.13155/46601, 2016.

Bittig, H. C., Maurer, T. L., Plant, J. N., Wong, A. P., Schmechtig, C., Claustre, H., Trull, T. W., Udaya Bhaskar, T. V. S., Boss, E., and Dall'Olmo, G.: A BGC-Argo guide: Planning, deployment, data handling and usage, Front. Mar. Sci., 6, 502, 2019.

Boyd, P. W. and Trull, T. W.: Understanding the export of biogenic particles in oceanic waters: Is there consensus?, Prog. Oceanogr., 72, 276–312, https://doi.org/10.1016/j.pocean.2006.10.007, 2007.

Boyd, P. W., Claustre, H., Levy, M., Siegel, D. A., and Weber, T.: Multi-faceted particle pumps drive carbon sequestration in the ocean, Nature, 568, 327–335, https://doi.org/10.1038/s41586-019-1098-2, 2019.

Brasseur, P., Verron, J.: The SEEK filter method for data assimilation in oceanography: a synthesis. Ocean Dyn. 56, 650–661, 2006.

Cabanes, C., Grouazel, A., von Schuckmann, K., Hamon, M., Turpin, V., Coatanoan, C., Paris, F., Guinehut, S., Boone, C., Ferry, N., de Boyer Montégut, C., Carval, T., Reverdin, G., Pouliquen, S., and Le Traon, P.-Y.: The CORA dataset: validation and diagnostics of in-situ ocean temperature and salinity measurements, Ocean Sci., 9, 1–18, https://doi.org/10.5194/os-9-1-2013, 2013.

Dee, D. P. and Co-authors: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137 (656), 553–597, https://doi.org/10.1002/qj.828, 2011.

Ducet, N., and P.-Y. Le Traon: A comparison of surface eddy kinetic energy and Reynolds stresses in the Gulf Stream and the Kuroshio Current systems from merged TOPEX/Poseidon and ERS-1/2 altimetric data, Journal of Geophysical Research: Oceans, 106 (C8), 16 603–16 622, https://doi.org/10.1029/2000JC000205, 2001.

Dufau, C., Orsztynowicz, M., Dibarboure, G., Morrow, R., and Le Traon, P.-Y.: Mesoscale resolution capability of altimetry: Present and future. Journal of Geophysical Research: Oceans, 121 (7), 4910–4927, https://doi.org/10.1002/2015JC010904, 2016.

Foltz, G. R. and Co-authors: The tropical Atlantic observing system, Frontiers in Marine Science, 6, 206, https://doi.org/10.3389/fmars.2019.00206, 2019.

Fujii, Y., and Co-authors: Observing system evaluation based on ocean data assimilation and prediction systems: On-going challenges and a future vision for designing and supporting ocean observational networks, Frontiers in Marine Science, 417, https://doi.org/10.3389/fmars.2019.00417, 2019.

Galí, M., Falls, M., Claustre, H., Aumont, O., and Bernardello, R.: Bridging the gaps between particulate backscattering measurements and modeled particulate organic carbon in the ocean, Biogeochemistry: Open Ocean, https://doi.org/10.5194/bg-2021-201, 2021.

Gasparin, F., and Coauthors: Requirements for an Integrated in situ Atlantic Ocean Observing System From Coordinated Observing System Simulation Experiments, Frontiers in Marine Science, 6, https://doi.org/10.3389/fmars.2019.00083, 2019.

Gasparin, F.,Cravatte, S., Greiner, E., Perruche, C., Hamon, M., Van Gennip, S., and Lellouche, J.-M.: Excessive productivity and heat content in tropical Pacific analyses: disentangling the effects of in situ and altimetry assimilation, Ocean Modelling, 160, 101 768, https://doi.org/10.1016/j.ocemod.2021.101768, 2021.

Good, S. A., Martin, M. J. and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, Journal of Geophysical Research: Oceans, 118 (12), 6704–6716, 2013.

Hamon, M., Greiner, E., Le Traon, P.-Y. and Remy, E.: Impact of multiple altimeter data and mean dynamic topography in a global analysis and forecasting system. Journal of Atmospheric and Oceanic Technology, 36 (7), 1255–1266, 2019.

Halliwell, G. R., Srinivasan, A., Kourafalou, V., Yang, H., Willey, D., Le Hénaff, M. and Atlas, R.: Rigorous evaluation of a fraternal twin ocean OSSE system for the open Gulf of Mex- ico. Journal of Atmospheric and Oceanic Technology, 31 (1), 105–130, 2014.

Hayes, S. P., Mangum, L. J., Picaut, J., Sumi, A., and Takeuchi, K.: TOGA-TAO: A moored array for real-time measurements in the tropical Pacific Ocean. Bulletin of the American Meteo- rological Society, 72 (3), 339–347, 1991.

Hermes, J. C. and Co-authors: A sustained ocean observing system in the Indian Ocean for climate related scientific knowledge and societal needs, Frontiers in Marine Science, 355, 2019.

Huang, B. and Wang, X.: On the use of cost-effective valid-time-shifting (VTS) method to increase ensemble size in the GFS hybrid 4DEnVar system. Monthly Weather Review, 146 (9), 2973–2998, 2018.

Henson, S. A., Sanders, R., Madsen, E., Morris, P. J., Le Moigne, F., and Quartly, G. D.: A reduced estimate of the strength of the ocean's biological carbon pump: BIOLOGICAL CARBON PUMP STRENGTH, Geophys. Res. Lett., 38, n/a-n/a, https://doi.org/10.1029/2011GL046735, 2011.

Johnson, Plant, J. N., and Maurer, T. L.: Processing BGC-Argo pH data at the DAC level, Ifremer, https://doi.org/10.13155/57195, 2018a.

Johnson, Pasqueron De Fommervault, O., Serra, R., D'Ortenzio, F., Schmechtig, C., Claustre, H., and Poteau, A.: Processing Bio-Argo nitrate concentration at the DAC Level, 2018b.

Kessler, W. S., Cravatte, S. and Co-authors: Final Report of TPOS 2020, GOOS-268, 83 pp. [Available online at https://tropicalpacific.org/tpos2020-project-archive/reports/], 2021.

Lellouche, Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Gasparin, F., Hernandez, O., Levier, B., Drillet, Y., Remy, E., and Le Traon, P.-Y.: Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1⁄12° high-resolution system, Ocean Sci., 14, 1093–1126, https://doi.org/10.5194/os-14-1093-2018, 2018.

Lellouche, J.-M., and Co-authors: The Copernicus Global 1/12° Oceanic and Sea Ice GLORYS12 Reanalysis, Frontiers in Earth Science, 9, https://doi.org/10.3389/feart.2021.698876, 2021.

Resplandy, L., Lévy, M., and McGillicuddy, D. J.: Effects of Eddy-Driven Subduction on Ocean Biological Carbon Pump, Glob. Biogeochem. Cycles, 33, 1071–1084, https://doi.org/10.1029/2018GB006125, 2019.

Roemmich, D., and Co-authors: On the future of Argo: A global, full-depth, multi- disciplinary array. Frontiers in Marine Science, 6, 439, 2019.

Schmechtig, C., Poteau, A., Claustre, H., D'Ortenzio, F., and Boss, E.: Processing bio-Argo chlorophyll-A concentration at the DAC level, Ifremer, https://doi.org/10.13155/39468, 2015.

Siegel, D. A., Buesseler, K. O., Doney, S. C., Sailley, S. F., Behrenfeld, M. J., and Boyd, P. W.: Global assessment of ocean carbon export by combining satellite observations and food-web models, Glob. Biogeochem. Cycles, 28, 181–196, https://doi.org/10.1002/2013GB004743, 2014.

Smith, N., and Co-authors, 2019: Tropical pacific observing system, Frontiers in Marine Science, 31, https://doi.org/10.3389/fmars.2019.00031, 2019.

Willett, C. S., Leben, R. R. and Lavín, M. F.: Eddies and tropical instability waves in the eastern tropical Pacific: A review, Progress in Oceanography, 69 (2-4), 218–238, 2006.

Thierry, V., Bittig, H. and the Argo BGC Team: Argo quality control manual for dissolved oxygen concentration, Version 2.0, http://dx.doi.org/10.25607/OBP-118, 2018.