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	D4.2
Deliverable title	Design of the glider assimilation experiments
Description	Investigations and preparation of glider observations to be assimilated in MED-MFC and WMOP systems.
Work Package number	4
Work Package title	Data integration, Assimilation, and Forecasting
Lead beneficiary	CMCC
Lead authors	Ali Aydogdu, Jaime Hernandez-Lasheras, Carolina Amadio, Baptiste Mourre, Gianpiero Cossarini, Jenny Pistoia
Contributors	
Due date	31 October 2021
Submission date	29 October 2021
Comments	

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_d4.2

Table of contents

Executive summary

This document presents preparations and experiment design strategy to assess the impact of glider observations on the MED-MFC and WMOP analysis and forecasting systems of the Mediterranean Sea. An extensive investigation has been carried out on different repositories providing glider observations. A special attention has been paid for the Copernicus Marine Service (CMEMS) repository managed by CMEMS in situ Thematic Assembly Center (TAC). Because the CMEMS Marine Forecasting Centers (MFCs) rely on the observational datasets provided by CMEMS TACs for their assimilation systems.

A couple of issues have been found related to the data availability and addressed in Task 4.2 via an internal milestone (IMS28) through a workshop organized by CMCC and SOCIB among the stakeholders in the observation and data assimilation communities. A report is made available summarizing the workshop outcomes on EuroSea cloud for the relevant communities.

Following the improvements made in the CMEMS repository after the workshop, in this report we outline the most recent status of the upstream data to be used and experiments that will be conducted in the following period of EuroSea in Task 4.2.

1. Introduction

One of the main goals of EuroSea is to support and improve the marine forecasting systems by increasing the quality and quantity of the observations that are ingested in them for better analysis and, eventually, more skillful forecasts. In WP4, the possible improvements are being investigated in Task 4.1 for the CMEMS Global-MFC and IBI-MFC and in Task 4.2 by CMEMS MED-MFC and WMOP analysis and forecasting systems.

In Task 4.2, the benefits of profiling floats from the Argo program and glider observations provided by the OceanGliders program are being studied. In the western Mediterranean Sea, the MED-MFC PHY (Clementi et al., 2021) and WMOP (Juza et al., 2016, Mourre et al., 2018) systems, developed and maintained by CMCC and SOCIB, respectively, are employed to investigate the impact of assimilating glider observations. These two systems currently ingest temperature and salinity (T-S) profiles, mainly from Argo floats. Therefore, the focus will be on assessing glider data impact. Besides, the MED-MFC BIO system (Feudale et al., 2021), developed and maintained by OGS, will be used to assess the impact that glider T-S data assimilation (DA) in the MED-MFC PHY system has on the biogeochemistry (BGC) of the Mediterranean Sea. The BGC-Argo floats profiles will be assimilated to assess their contribution in the analysis.

Gliders provide continuous, fine-resolution observations and have been used in a variety of applications, including observations of eastern boundary currents (Davis et al., 2008), coastal circulation (Todd et al., 2009), an eddy (Martin et al., 2009), and water formation (Houpert et al., 2016). The main limitation of gliders as platforms for ocean sampling is their relatively slow speed through the water (Rudnick and Cole, 2011) which, on the other hand, helps to provide high-resolution profiles.

The latter brings out challenges in the DA applications using glider observations due to the high correlation involved which has to be accounted for in the observation error covariance matrix, **R**, in the DA systems. Besides, the assimilation of glider observations in near-real-time operational forecasting systems have so far been a challenge due to other reasons; naming two are: Barriers in the data flow from the providers and implementation of suitable and efficient operational quality control algorithms. Currently, none of the

CMEMS marine forecasting systems assimilate glider observations operationally, to our knowledge, while some multi-year reanalysis systems use them. These, and further challenges require a coordinated action in Europe to achieve a best practice on the use of glider observations in the global, regional, and coastal systems.

Similar challenges are brought by BGC-Argo profiling floats. In particular, since profiling float data in near-real-time are not quality checked by the Argo data centres, the development of methodologies aimed at correcting and pre-processing data should be applied (Takeshita et al., 2013; Maurer et al., 2021). Furthermore, as for the glider's community, also the Argo community needs coordinated action to share best practices and knowledge and to fill the gap between the providers/data centres and modellers/data assimilators.

In this deliverable, we summarize the status of the Task 4.2 and activities performed so far including a joint workshop with various stakeholders. A non-exhaustive list includes researchers from modelling, data assimilation, in-situ and satellite observation communities in Europe representing CMEMS MFCs and TACs, OceanGliders, EuroArgo programs. We conclude by a discussion on suggestions and further steps.

In the following section, the status of the observations that will be used in this study are outlined. In Section 3, we present the systems used in this study and a demo of their capabilities. Then, in Section 4, we propose an experimental setup for the future investigations to be conducted in EuroSea. We discuss the interactions with other EuroSea work packages and tasks in Section 5. Finally, we discuss the outcomes and outline our future prospects.

2. Status of observations

The WP4 Task 4.2 targets to improve the Copernicus marine analysis/forecasting systems which rely on the in situ / float / satellite observations gathered in the CMEMS Thematic Assembly Centres. In Task 4.2, an initial assessment of the observation availability has been done for the gliders and profiling Argo floats which are discussed in Sections 2.1 and 2.2, respectively. The task team organized a workshop to address the issues encountered during these preparations and propose solutions. The summary and main outcomes of the workshop is given in Section 2.3. We close this section by outlining previous experiences on the assimilation of glider observations, especially on the pre-processing stage in Section 2.4.

2.1. Glider observation availability

Every year, more than 10,000 glider profiles are made available in CMEMS[1]. The data are all L2, i.e., have been quality controlled and flagged by Coriolis, who is the Production Unit integrating ocean glider data in CMEMS.

[1] INSITU_MED_NRT_OBSERVATIONS_013_035

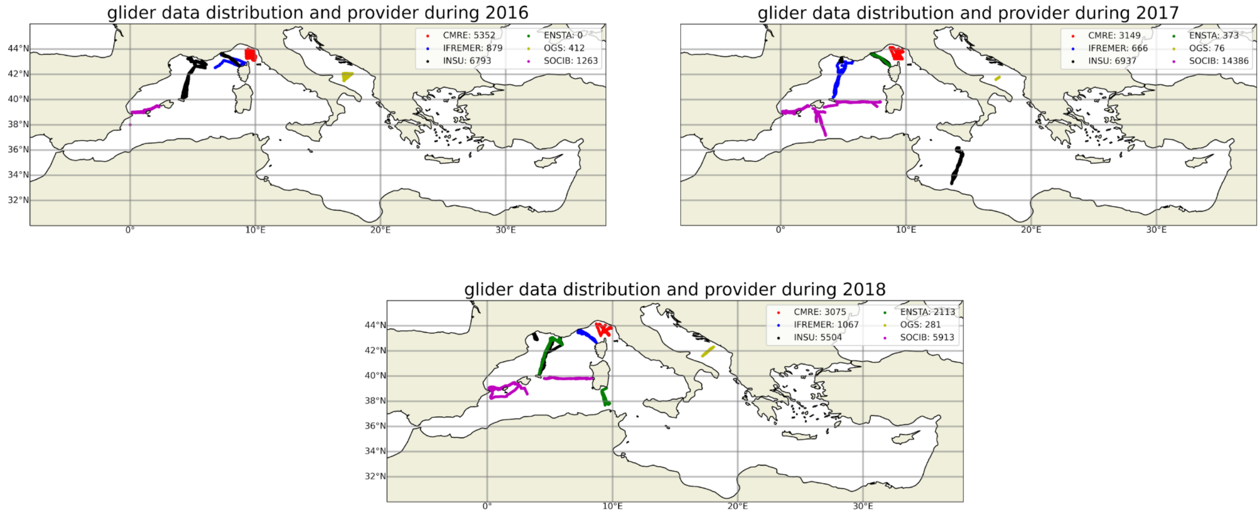

Figure 1. Distribution of glider observations in 2016 (top left), 2017 (top right) and 2018 (bottom) provided by six institutions: CMRE (red), IFREMER (blue), INSU (black), ENSTA (green), OGS (yellow) and SOCIB (purple). Data retrieved from CMEMS repository "monthly" catalogue on 10 October 2021.

Six institutions provided glider observations, as identified in the file metadata information in the central and western Mediterranean, between 2016 and 2018. Figure 1 shows the distribution of the data from each provider which are marked in different colours (see the legend). The glider observations in these three years are located in the western Mediterranean with some exceptions in the Adriatic Sea delivered by OGS in all three years and in the Sicily Strait by INSU only in 2017. There are more observations in 2017 and most of them are provided by SOCIB on the Ibiza channel and around the Balearic Islands. Some observations cover the area between the Gulf of Lions and the Ligurian Sea provided by CMRE, IFREMER, ENSTA and INSU. Table 1 summarizes the number of observations per provider (see also the legend in Figure 1) and their annual sum.

Table 1. Number of glider observations in three years provided by different institutions shown in Figure 1 and their annual sum in the central and western Mediterranean.

	SOCIB	**OGS**	**IFREMER**	**INSU**	**ENSTA**	**CMRE**	***Total***
2016	1263	412	879	6793	0	5352	14699
2017	14386	76	666	6937	373	3149	25587
2018	5913	281	1067	5504	2113	3075	17953

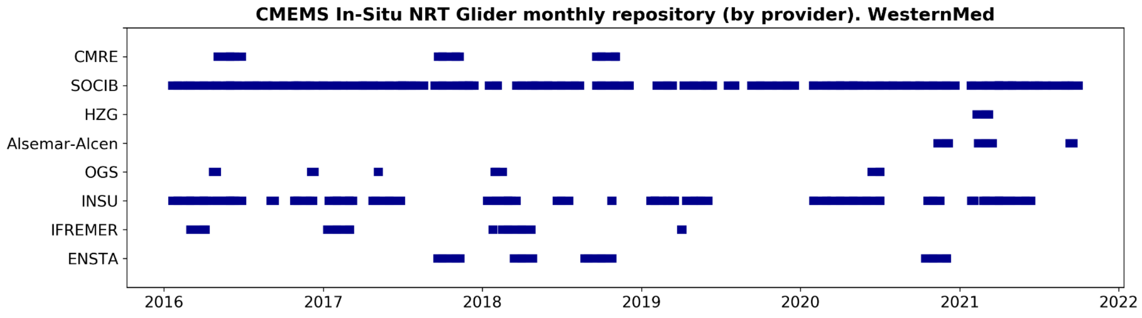

Figure 2. Timeline showing the Glider data availability in CMEMS "monthly" catalogue by providers. Only deployments in the Western Mediterranean Sea, as obtained on 20 October 2021.

Figure 2 shows the timeline of the glider observations available in CMEMS "*monthly*" catalogue[2] per data provider as is on 20 October 2021 in the western Mediterranean. This catalogue does not provide any data before 2016, as it is a dynamic catalogue in which only the last 5 years are distributed. The "history catalogue"[3], should, in theory, merge all historical observations of each platform into a single one and thus, data contained in the monthly catalogue should be also contained in the history one. However, this is not always the case and there are platforms for which, although having data in the monthly catalogue there is no aggregated file in the history one. The issue has been identified during the developments of CMEMS Mediterranean physical reanalysis in which the in situ (not only glider) observations from different resources has been merged to have a more complete dataset[4]. For instance, the SOCIB data which has recently been synchronized into CMEMS is, at the date of this deliverable, not yet incorporated into the "history catalogue" files. Thus, in this work we will be using data from the monthly catalogue as it should be more up to date for the period of our experiments.

SOCIB L1 RT glider data are synchronized with the CMEMS repository through Coriolis. First to the Global product and it is afterwards included to the Mediterranean one following internal protocol. Theoretically, CMEMS INSITU NRT product should have all the data from the different SOCIB missions, with all data for every single platform aggregated in the same file (for the case of files contained in history catalogue). Once the files are synchronized, CMEMS performs a cleaning of the data, followed by an interpolation and a quality control, according to its own standards. From 2016 CMEMS, through the Coriolis intermediate operator, established the necessity to use the EGO format[5] by the data producers before delivering it to them.

When this EuroSea task started we found that, due to operational issues, there was an almost 2-year gap, between 2016 and 2018, in which no data had been synchronized to the CMEMS database. From then onward real-time (RT) from SOCIB glider missions is available in CMEMS according to such standards. This RT product only contains profiles which are sent by the glider during the mission (about 2 profiles every 6-8 hours), which is not optimal for the experiments we intend to run in this task.

[2] ftp://nrt.cmems-du.eu/Core/INSITU_MED_NRT_OBSERVATIONS_013_035/med_multiparameter_nrt/index_monthy.txt

[3] ftp://nrt.cmems-du.eu/Core/INSITU_MED_NRT_OBSERVATIONS_013_035/med_multiparameter_nrt/index_history.txt

[4] https://resources.marine.copernicus.eu/product-detail/MEDSEA_MULTIYEAR_PHY_006_004/INFORMATION

[5] https://archimer.ifremer.fr/doc/00239/34980/71648.pdf

To address this problem, a special effort has been made from the SOCIB side to reformat all RT and DT data for the complete time period to EGO format and synchronize it to CMEMS. This iterative process with Coriolis has just been completed by the end of October 2021.

However, based on our experience, CMEMS in situ catalogue files should be carefully taken, as issues may arise: For instance, the coverage of both the history and monthly catalogues does not always match. We should be aware that these are dynamical catalogues and files are proven to be changed, reprocessed, etc. For instance, we have surprisingly found that some of the netCDF files that we downloaded four months ago (January 2020) from CMEMS and stored locally are not anymore available on CMEMS catalogue, while some additional historic files from other platforms have been uploaded meanwhile. This could be due to several reasons, for instance the In-situ TAC SRD (i.e., the global attributes considered mandatory[6]) might evolve. It is recommended to better check the GLO product, as it may contain some historical files which are not synchronized to the MED one.

Regarding the BGC data collected by the SOCIB gliders, this should actually be available in the CMEMS in situ GLO NRT database product. At least for the experiment period in the monthly catalogue. However, it must be noticed that the BGC observations do not have the same resolution as the physical ones and there are much less profiles available. We should also be aware that CMEMS also develops an in situ REP BGC product[7], which may contain this data, but which has not been analysed.

Based on the experience with SOCIB database glider observations, a similar evaluation should probably be applied to French and Italian gliders. CMEMS products are regularly updated and are prone to changes. In our experiments, it is crucial to make sure all participants store a local Backup of the data so that the assimilated datasets are the same in CMEMS and WMOP models.

A special attention must be given when using CMEMS observations. Although described as a gridded product with dimensions of TIME and DEPTH, the depths of the vertical levels are not the same for all the profiles and should be read in the pressure variable.

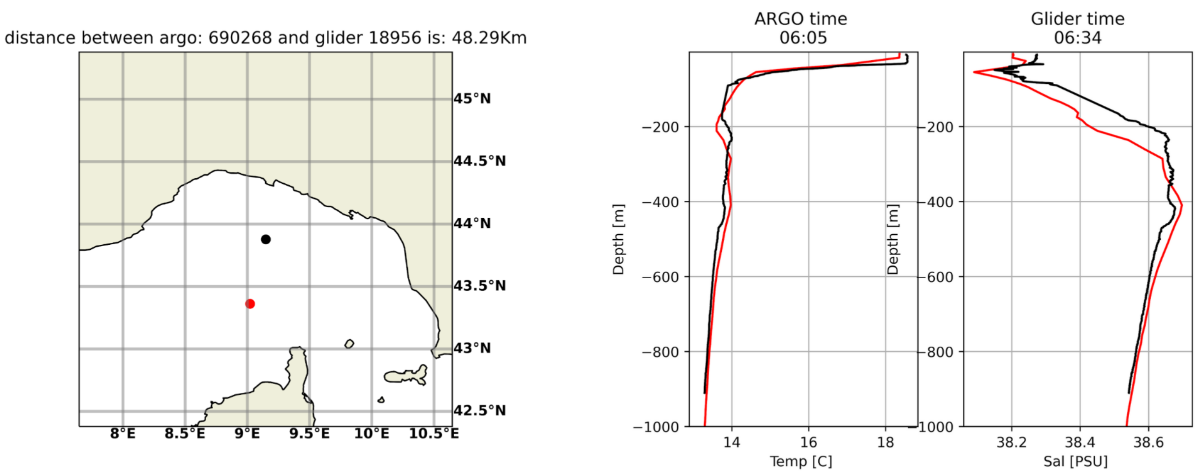

Figure 3. An example of the cross-control of profiles from Argo floats (red; no. 690268) and gliders (black; no. 18956) in the Ligurian Sea (map) for temperature (left) and salinity (right).

---

6 https://archimer.ifremer.fr/doc/00297/40846/

7 ftp://my.cmems-du.eu/Core/INSITU_GLO_BGC_REP_OBSERVATIONS_013_046/

In Figure 3, an example of comparison between the profiles obtained from Argo floats and gliders is given in the Ligurian Sea. In this example, the temperature measured by both platforms agree well while there are significant differences in salinity especially in the first 300 m. It can be delicate to find a consistency between the measurements when the water column is stratified.

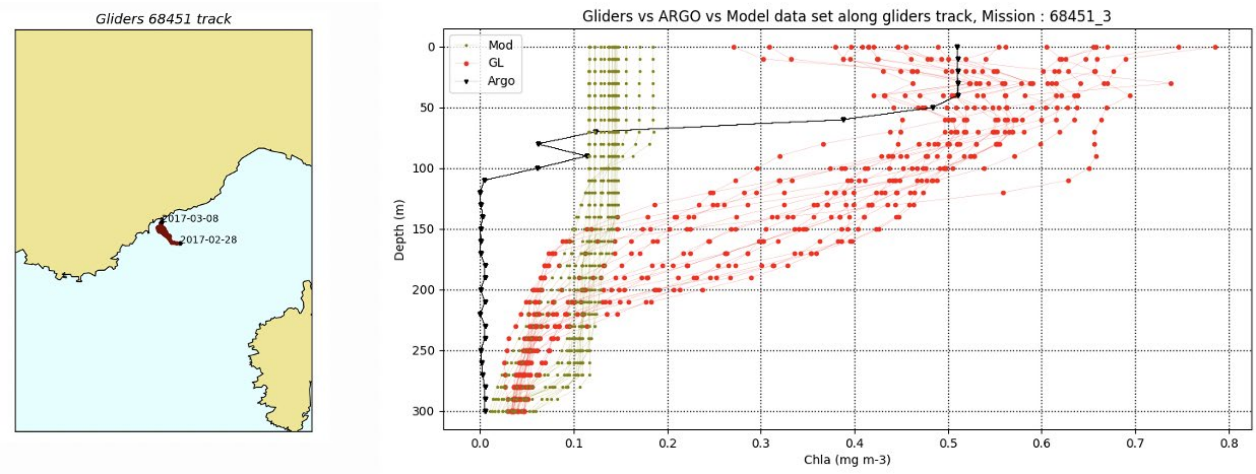

Figure 4. An example of the cross-control of profiles from BGC-Argo floats (black) and gliders (red; no. 68451) and the MED-BIO system (green) at the same locality in the Ligurian Sea (map) for chlorophyll (right).

More inconsistencies have been identified between the BGC-Argo and glider BGC observations. In Figure 4, the MED-BIO system is compared against the profiles from the Argo floats and the gliders in the Ligurian Sea for the chlorophyll. The spread of the measurements from gliders is large while they don't compare well with the BGC-Argo measurements especially between 50-200 m depth. While they compare well with the model solution below 150 m.

2.2. Joint workshop on sharing best practices: how to use novel sensors data for assimilation and validation in the CMEMS and SOCIB operational systems

The WP4 Task 4.2 team organized a workshop to address the issues related to the availability of the observations in the repositories (e.g., CMEMS). The workshop was held on 24 June 2021 with the participation of experts from OceanGliders, CMEMS INS-TAC, Euro-Argo and European glider data providers and EuroSea partners in WP3. The workshop aimed to open discussion on the best practices in the use of glider and floats in situ observations by operational forecasting systems, on the accessibility to the glider/Argo floats observations in NRT and DT mode and on the quality control (QC) in the assimilation systems. The minutes from the workshop can be found on EuroSea document cloud[8]. Major outcomes of the workshop can be summarized as follows:

- Specifically, for Task 4.2, it is made clear that the upstream data is ready to achieve the objectives of the task. There seems to be more time needed to assimilate the high-quality glider and BGC-Argo observations in the NRT systems; however, DM observations are already high-quality and synchronized to the required repositories.

[8] https://cloud.geomar.de/s/y769wEL8GPiM4Rs?dir=undefined&path=%2FEuroSea_info%2FInternal_Milestones&openfile=14450073

- There is a need to come up with a universal solution. CMEMS (European) and SOCIB (Balearic) systems involved in EuroSea can be taken as a base to detect the need for improvements and propose solutions for every step of the data flow and usage.
- There is a need for further communication between the communities, e.g., Argo vs. Glider communities to converge on coherent procedure and avoid inconsistencies, Argo + Glider vs. modelling + assimilation communities for the best practices on the use of observations in forecasting and reanalysis systems, e.g., on QC standards. There are channels, such as OceanGliders, CMEMS, OceanPredict to maintain the communication but a group of organized experts should take an initiative.

2.3. Pre-processing glider observations for data assimilation

Gliders are autonomous underwater vehicles of a small size that can go underwater along slightly inclined paths by changing their density (Davis et al., 2003) with a typical horizontal and vertical velocity of ~40 cm/s and ~15 cm/s, respectively. They provide very high-resolution observations in both space and time especially compared to the resolution of the ocean models used in operational centres. The assimilation systems used at CMCC (OceanVar) and OGS (3DVarBIO) assumes a horizontal correlation length scale about 10-20 km, while SOCIB data assimilation system uses an EnOI scheme in which correlations are calculated from an ensemble of model realizations, with an additional localization radius of 200km. Given the spatial scales of model correlations and the high-spatial resolution of gliders´ observations, these cannot be considered as independent from each other. However, as above mentioned, assimilation systems assume uncorrelated errors in observations, i.e., diagonal observational error covariances matrix, **R**. Therefore, glider measurements cannot be ingested into the assimilation systems without a pre-processing that takes these aspects into account.

Past experiences of the partners in Task 4.2 on the assimilation of glider observations suggest following considerations on data pre-processing before the assimilation step:

Choosing appropriate profiles:

- Using only up-casts (climb phase). The higher vertical speeds (up to 0.20 m s−1) during the start of the dive phase near the surface may cause some spurious salinity values as the glider passed through the thermocline (thermal lag issue).
- Discarding profiles with vertical gaps larger than 10m.
- Discarding profiles with low number of measurements.

Pre-processing to handle horizontal correlations in glider observations:

- Sub-sampling:
    - Removing profiles in the inference radius of the observation position (see Figure 5)
- Superobing:
    - May not be appropriate due to the diurnal cycle in surface/subsurface temperature and salinity

Pre-processing to handle vertical correlations in glider observations:

- Binning in vertical grid levels
- Discarding observations with large variance in vertical levels
- Estimating representativity error from observation variance in vertical levels

In Dobricic et al. (2010), the raw observations were averaged within a 12 h long time window, giving rise to observations spaced approximately 12km. They use a model configuration with a horizontal resolution of 1/16°, therefore, this spacing is represented by approximately two model grid points at model resolution. Since any corrections at spatial scales shorter than two grid points cannot be represented by the model finite difference scheme and would be removed as noise during the model integration, this justifies the averaging of observations. The observations are also averaged in the vertical direction by producing a single averaged observation at each model level which still does not produce completely independent observations, because the vertical dimension in the control space is reduced to 20 EOFs. These vertical dependencies, however, did not seem to have any impact on the rate of the convergence of the minimizer.

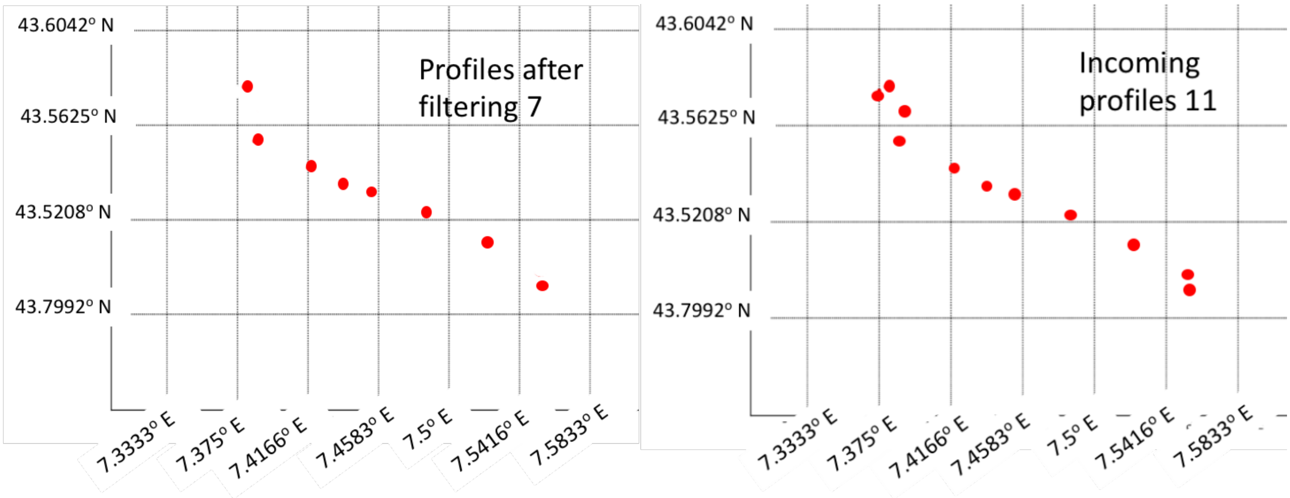

Figure 5. Example of horizontal sub-sampling according to the radius of inference

Mourre and Chiggiato (2014) consider the inclination of each profile from the vertical. These profiles are first interpolated on the vertical model grid before being incorporated. For each level, the observed variance in the vertical grid cell is used as an approximation of the vertical representation error. The horizontal representation error variance is assumed to be $(0.25)^2$ and $(0.05)^2$ for temperature (in ºC) and salinity (in psu) measurements, respectively. The observation error covariance matrix is specified as diagonal (i.e., **R** is exactly prescribed in the EnKF and not estimated from the ensemble of observations). To compensate for the lack of consideration of spatial error correlations, observation error variances are individually inflated by a factor equal to the number of neighbouring profiles within a model grid cell radius.

In Hernandez-Lasheras and Mourre (2018), the glider profiles are considered as vertical. The corresponding observations are binned vertically, and a single value is given for each model grid cell. The representation error is the addition of vertical and horizontal components. For each vertical level, the observed variance in the vertical grid cell is used as an approximation of the vertical representation error. In addition, the horizontal representation error variance is assumed to be 0.0625 and 0.0025 for temperature (in K) and salinity (in psu) measurements, respectively.

Hayes et al. (2019) consider that vertical speed during the glider was steadier on the upcast, since the glider adds buoyancy anytime the speed drops below the target speed. They also check the range of temperature and salinity measurements as for temperature to be between 10 °C-33 °C and salinity to be between 35-41 psu. Finally, temperature and salinity profiles were assumed vertical, and time stamped and assigned latitude and longitude based on the glider GPS fix, typically a few minutes after surfacing although the upcast was slanted with a horizontal displacement of approximately 3 km from bottom to top in the typical 1000 m dive.

3. System description and a demo of their capabilities

In Task 4.2, CMEMS MED-MFC and WMOP systems will be used. In Figure 6, the domain of the MED-PHY (left) and WMOP (right) systems are shown. The MED-BIO system forced by the MED-PHY shares the same domain and the resolution with the MED-PHY but in the Mediterranean Sea.

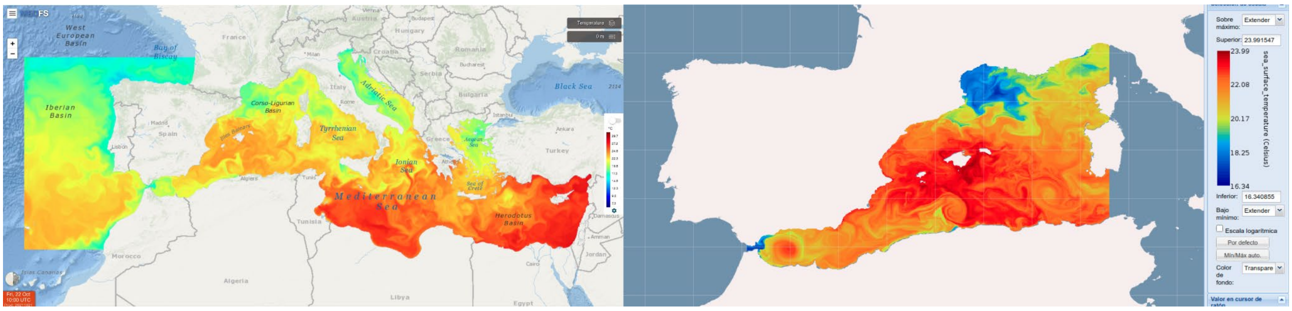

Figure 6. Left: MED-MFC systems domain including the Atlantic box. Right: WMOP system domain in the western Mediterranean. Both maps are colored by sea surface temperature. Last access 19/10/2021 from https://medfs.cmcc.it and https://socib.es/?seccion=modelling&facility=lw4nc , respectively.

The MED-PHY and MED-BIO systems share the same geometry in the Mediterranean Sea and the model resolution in the horizontal (1/24°) and vertical are common. The only difference is the Atlantic box in the MED-PHY system which requires the last 16 levels in the model grid to represent the depth. The WMOP system is implemented in the western Mediterranean Sea between the Gibraltar strait and Corsica / Sardinia Island with very high horizontal resolution (1/50°) and 32 vertical sigma-levels. Both systems differ also in the data assimilation (DA) approaches where MED-PHY uses a 3D variational (DA) algorithm (OceanVar) while WMOP uses an ensemble-based method (EnOI) to improve the model state using the observations. Other major characteristics of both MED-PHY and WMOP are significantly different and are listed in Table 2.

The MED-PHY and WMOP systems have been prepared to assimilate glider observations since the beginning of EuroSea. The observation operators implemented in the systems have been revisited. OceanVar at CMCC has the observation operator implemented by Dobricic et al. (2010) which has been updated in years for temperature and salinity observations. The WMOP data assimilation system has successfully been used to assimilate glider data in the past, as in Hernandez-Lasheras and Mourre (2018) where observations from up to 8 simultaneous gliders were assimilated.

A preliminary impact assessment of the assimilation of the glider temperature and salinity profiles has been done in the WMOP and CMEMS-PHY systems. In the WMOP system, glider superobing observations were generated and the same set-up as in previous works was employed. Figure 7 shows a sample glider trajectory and simultaneously gathered CTD profiles in a cruise to calibrate/evaluate the gliders (top panel). In the bottom panel, innovations and analysis residuals are compared in each observation location to assess the potential improvement brought by the assimilation of the glider observations for salinity (left) and temperature (right). The mean of the innovations in salinity is 0.014 with a standard deviation of 0.21 while of the mean of the residuals is 0.004 with a standard deviation of 0.049. This means that the model state is closer to the observations after the assimilation with a much lower variance. This is an indication that the assimilation pushes the system in the right direction, i.e., towards the observations. The same interpretation can be done for the temperature too.

Table 2. Main configuration

	CMEMS MED-PHY (CMCC)	WMOP (SOCIB)	CMEMS MED-BIO (OGS)
Domain	Mediterranean Sea (+ Atlantic box)	Western Med. Gibraltar to Corsica- Sardinia	Mediterranean Sea (+ Atlantic box)
Resolution	1/24^{o} degree (~4.5km) 141 z* vertical levels	~1/50^{o} degree (2km) 32 vertical sigma-levels	1/24^{o} degree (~4.5km) 125 vertical levels
Numerical model	NEMO v3.6 < - > WW3 v3.14	ROMS v3.4	MedBFM (OGSTM-BFM)
Time step	240 sec (Barotropic step 2.4sec)	120 sec (Barotropic step 6sec)	
Parameterizations	Tides, atmospheric pressure	No tides, No atm. pressure	plankton functional types: 4 phytoplankton groups, 4 zooplankton groups, 1 bacteria group Describes the biogeochemical cycle of N, P, C, Si and O. It includes the carbonate system dynamics
	climatological inputs from 39 rivers.	climatological inputs from 6 major rivers.	climatological inputs from 39 rivers.
	Richardson number-dependent vertical diffusion	Generic model of two-equations GLS turbulent closure.	
	Flather for barotropic currents and SSH. Orlanski for baroclinic currents	Flather for 2-D momentum. Chapman for surface elevation. Mixed radiation-nudging for 3-D equations.	
Atmospheric forcing	ECMWF HR 10km, 6hr resolution	AEMET (Spanish meteo agency) HARMONIE 2.5km 1hr	
Ocean forcing			MED-MFC PHY
Lateral open boundary condition	From CMEMS GLO-MFC NRT system	From CMEMS MED-MFC	
Data Assimilation	OceanVar: SLA along tracks, ARGO vertical T/S profiles	EnOI: SLA along-track, ARGO vertical T/S profiles, SST L4 satellite product, HF-Radar (Ibiza Channel)	3DVarBio: surface chlorophyll concentration from satellite observations

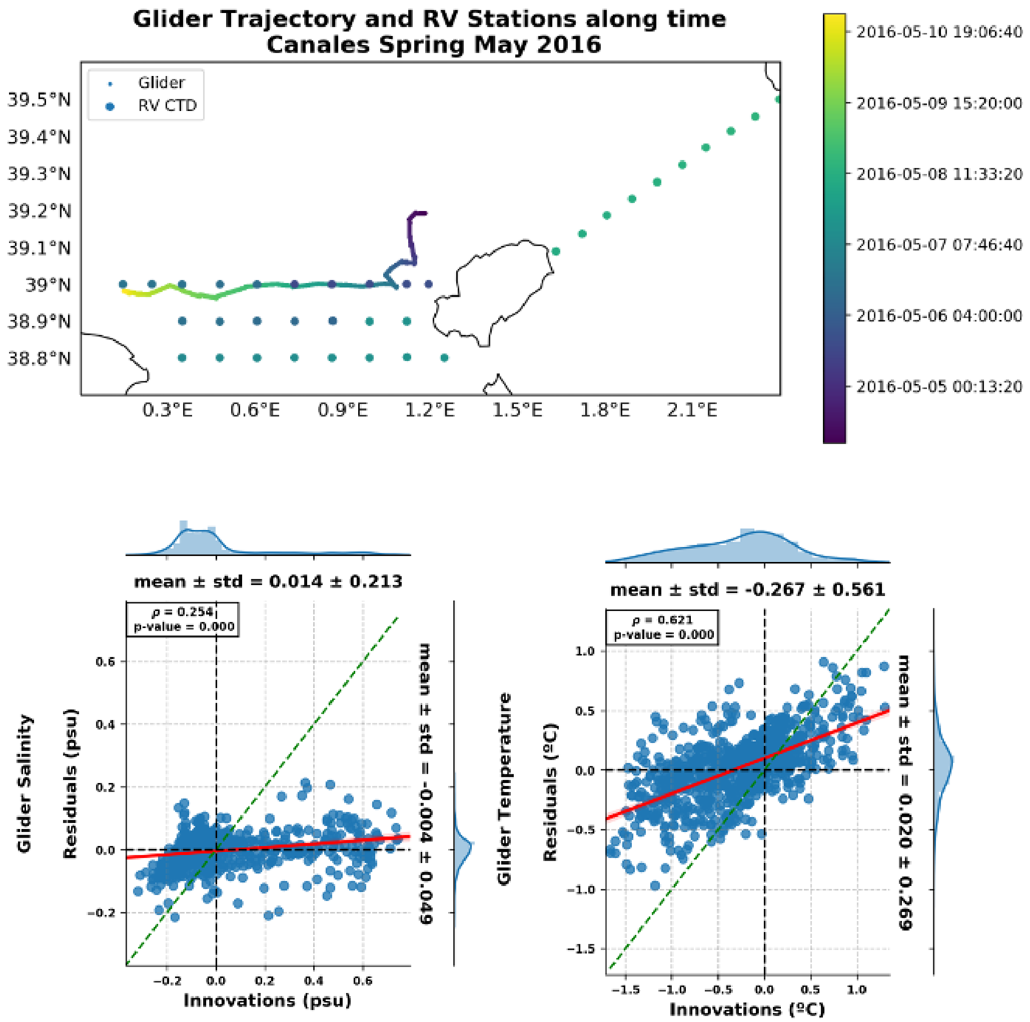

Figure 7. Top: A sample trajectory of a glider in the western Mediterranean Sea overlaid by the CTD sampling locations used for validation. Bottom: Comparison of the innovation and the residual for the salinity (left) and temperature (right) from the WMOP system.

In Figure 8, a sample experiment from the MED-PHY is shown in which an analysis is performed in 2017 with and without glider observations. The temperature and salinity RMS of misfits are presented between 15-45 m depth. The black (red) curve shows the experiment without (with) glider observations. The experiment is executed before the update of the data repositories, therefore, there are a very few numbers of glider observations assimilated (comparing the red and black shaded areas). Still, this demonstrates that the system ingests the glider observations successfully when they are available.

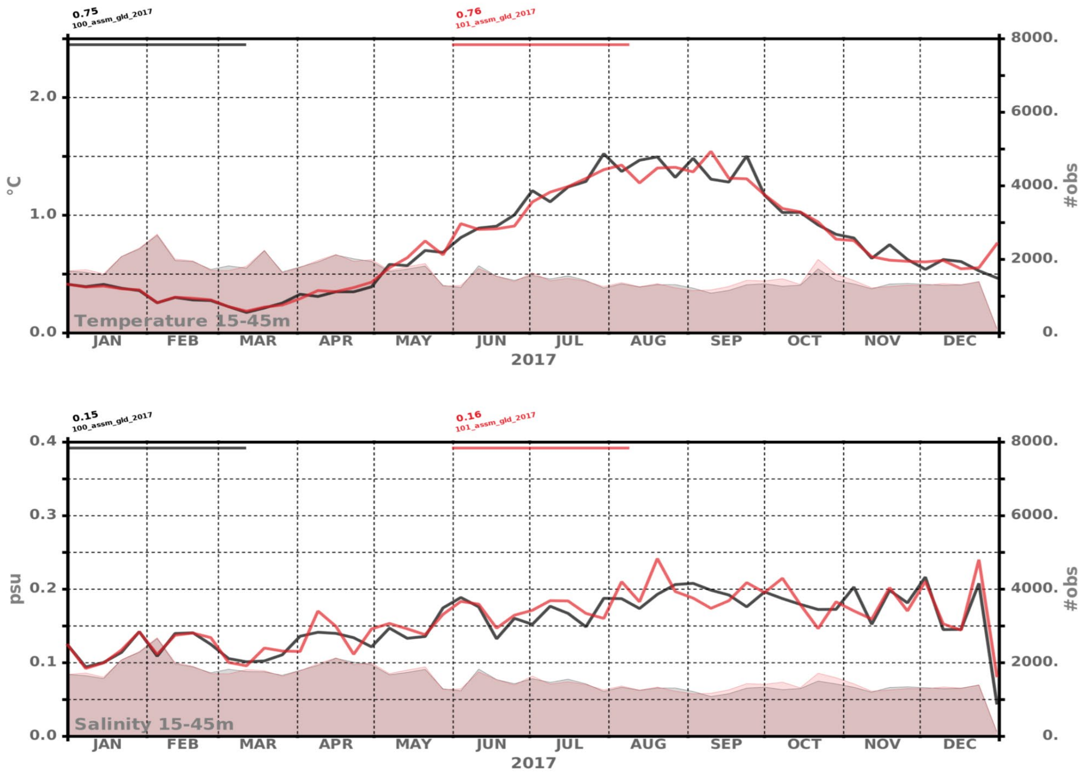

Figure 8. The MED-MFC PHY system preliminary results on the assimilation of the glider observations. The RMSE of temperature (top) and salinity (bottom) misfits between 15-45 m depth on the left y-axis. The number of in situ observations assimilated are shaded and shown on the right y-axis. The black and red are experiments without and with gliders, respectively.

4. Experiment setup

In Task 4.2, the impact of the glider and float profiles will be assessed using different methodologies. A direct evaluation will be performed for each individual system to examine the impact of new observations. Then, in the western Mediterranean Sea, where the majority of glider observations reside, an intercomparison will be performed between the MED-PHY and WMOP systems which will help to understand the changes under different configurations by parallel data assimilation experiments (D4.9). Moreover, the impact of physical assimilation -of glider profiles- on biogeochemistry will be studied by the CMEMS Med-BIO system (D4.10).

For the impact assessment of gliders, analogous experiments will be conducted with both MED-PHY and WMOP systems for the posterior intercomparison (Table 3). Two one-year-long simulations will be performed from 1 January to 31 December 2017. First, a reanalysis simulation in which each system will ingest the regular observing data sources, without any glider observations (i.e., satellite altimetry, sea surface temperature, Argo temperature and salinity profiles). Secondly, the same time period will be simulated, but including to the previously assimilated dataset all the available glider data from CMEMS database in the Western Mediterranean Sea. After the analysis of the CMEMS database and the latest update and synchronization of SOCIB glider data into it, we believe this should be the reference database used. Data from the same glider deployments will be assimilated in both systems to evaluate their impact in two different ocean modelling frameworks. A superobing/subsampling approach will be used, using a single averaged observation for each grid cell. The assimilated datasets will then be adapted to each system's characteristics, which will permit evaluating the influence of the data assimilation approach and model resolution.

Starting from a prototype of BGC-Argo float data assimilation, data assimilation experiments will be performed using the CMEMS Med-BIO system and BGC-Argo float data. The objectives are to evaluate the impacts of multivariate BGC-Argo float observations on the biogeochemical CMEMS products and to provide suggestions for the evolution of the BGC-Argo observing system. A multivariate covariance framework for increasing the spatial influence of the BGC-Argo network in the Mediterranean Sea at the basin scale will be used to support feasibility tests of biogeochemical assimilation of glider observations. Impacts of observations on the products from WPs 5,6,7 will be provided.

In task 4.2, CMCC, SOCIB and OGS aim to assess the impact of the available glider and Argo floats in the Mediterranean Sea on the MED-MFC PHY (CMEMS/CMCC), WMOP (SOCIB) and MED-MFC BIO (CMEMS/OGS) analysis/forecasting systems, respectively (Table 4). The temperature and salinity profiles obtained from Argo floats are already ingested in MED-MFC PHY and WMOP systems, therefore, both systems will be improved to assimilate high-resolution glider temperature and salinity profiles. The systems will be configured to assimilate the same set of observations to evaluate their impacts using different numerical models and data assimilation schemes. The development of the techniques to assimilate the BGC-Argo observations is an emerging research topic in the ocean biogeochemistry community. Taking advantage from the chlorophyll BGC-Argo assimilation developed in MASSIMILI CMEMS SE project, OGS will assess the impact of additional variables from BGC-Argo float (i.e., nitrate and oxygen) and will develop the assimilation of BGC-gliders observations in the MED-MFC-BIO system.

Table 3. Experiments to compare the MED-PHY and WMOP systems.

2017	**MED-PHY**	**WMOP**
WEST_MED_CTL_00	w/o glider	w/o glider
WEST_MED_GLD_01	with glider	with glider

Table 4. Experiments to assess the impact of the physical gliders and BGC-Argo profiling floats.

2017	**MED-PHY**		**MED-BIO**
ALL_MED_CTL_00	w/o glider	ALL_MED_BIO_00	w/o BGC-ARGO
		ALL_MED_BIO_01	with BGC-ARGO
ALL_MED_GLD_01	with glider	ALL_MED_BIO_02	w/o BGC-ARGO
		ALL_MED_BIO_03	with BGC-ARGO

5. Interfaces with other tasks and WPs

In Task 4.1, CMEMS IBI-MFC will evaluate the impact of glider observations too. The IBI-MFC system shares a common domain in the western Mediterranean with the MED-MFC and WMOP systems. CMCC and SOCIB will consider coordinating with MOI to have all three systems in the intercomparison experiments using the observation datasets prepared in Task 4.2.

6. Future prospects

In order to achieve the goals until the end of the project, we consider the following time schedule for the internal deliveries. The MED-PHY system will perform the experiment set in Table 2 until PM30 (May 2022) of EuroSea to have time for the preparation and execution of the experiments in the MED-BIO system. SOCIB and OGS will lead the upcoming deliverables on PM42 D4.9 on the intercomparison of physical systems and D4.10 on the impact assessment on the BGC, respectively.

References

Clementi, E., Aydogdu, A., Goglio, A. C., Pistoia, J., Escudier, R., Drudi, M., Grandi, A., Mariani, A., Lyubartsev, V., Lecci, R., Cretí, S., Coppini, G., Masina, S., and Pinardi, N., (2021). Mediterranean Sea Physical Analysis and Forecast (CMEMS MED-Currents, EAS6 system) (Version 1) set. Copernicus Monitoring Environment Marine Service (CMEMS).

Davis, R.E., Eriksen, C.E., Jones, C.P., (2003). Autonomous buoyancy-driven underwater gliders. In: Griffiths, G. (Ed.), Technology and Applications of Autonomous Underwater Vehicles. Taylor and Francis, pp. 37–58.

Davis, R.E., Ohman, M.D., Rudnick, D.L., Sherman, J.T. and Hodges, B., (2008), Glider surveillance of physics and biology in the Southern California current system, Limnol. Oceanogr., 53, 2151–2168, doi:10.4319/lo.2008.53.5_part_2.2151.

Dobricic, S., Pinardi, N., Testor, P. and Send, U., (2010). Impact of data assimilation of glider observations in the Ionian Sea (Eastern Mediterranean). *Dynamics of Atmospheres and Oceans*, *50*(1), 78-92. DOI: 10.1016/j.dynatmoce.2010.01.001.

Feudale, L., Bolzon, G., Lazzari, P., Salon, S., Teruzzi, A., Di Biagio, V., Coidessa, G. and Cossarini, G., (2021). Mediterranean Sea Biogeochemical Analysis and Forecast (CMEMS MED-Biogeochemistry, MedBFM3 system) (Version 1) set. Copernicus Monitoring Environment Marine Service (CMEMS). https://doi.org/10.25423/CMCC/MEDSEA_ANALYSISFORECAST_BGC_006_014_MEDBFM3

Hayes, D.R., Dobricic, S., Gildor, H. and Matsikaris, A., (2019). Operational assimilation of glider temperature and salinity for an improved description of the Cyprus eddy. *Deep Sea Research Part II: Topical Studies in Oceanography*, *164*, 41-53. DOI: 10.1016/j.dsr2.2019.05.015

Hernandez-Lasheras, J. and Mourre, B., (2018). Dense CTD survey versus glider fleet sampling: comparing data assimilation performance in a regional ocean model west of Sardinia, Ocean Sci., 14, 1069–1084, DOI: os-14-1069-2018.

Houpert, L.X., Durrieu de Madron Testor, P., Bosse, A., D'Ortenzio, F., Bouin, M.N., et al., (2016). Observations of open-ocean deep convection in the northwestern Mediterranean Sea: seasonal and interannual variability of mixing and deep water masses for the 2007-2013 period. *J. Geophys. Res. Oceans* 121, 8139–8171. doi:10.1002/2016JC011857

Juza M., Mourre, B., Renault, L., Gómara, S., Sebastián, K., Lora, S., Beltran, J.P., Frontera, B., Garau, B., Troupin, C., Torner, M., Heslop, E., Casas, B., Escudier, R., Vizoso, G. and Tintoré J., (2016). SOCIB operational ocean forecasting system and multi-platform validation in the western Mediterranean Sea, J. Oper. Oceanogr., 9 :sup1, s155-s166, doi:10.1080/1755876X.2015.1117764

Martin, J.P., Lee, C.M., Eriksen, C.C., Ladd, C., and Kachel N.B., (2009), Glider observations of kinematics in a Gulf of Alaska eddy, J. Geophys. Res., 114, C12021, doi:10.1029/2008JC005231.

Maurer, T.L., Plant, J.N. and Johnson, K.S., (2021). Delayed-Mode Quality Control of Oxygen, Nitrate, and pH Data on SOCCOM Biogeochemical Profiling Floats. Frontiers in Marine Science, 1118, DOI: 10.3389/fmars.2021.683207

Mourre, B. and Chiggiato, J., (2014). A comparison of the performance of the 3-D super-ensemble and an ensemble Kalman filter for short-range regional ocean prediction, Tellus A: Dynamic Meteorology and Oceanography, 66:1, 21640, DOI: 10.3402/tellusa.v66.21640

Mourre B., Aguiar, E., Juza, M., Hernandez-Lasheras, J., Reyes, E., Heslop, E., Escudier, R., Cutolo, E., Ruiz, S., Mason, E., Pascual, A., and Tintoré, J., (2018). Assessment of high-resolution regional ocean prediction systems using multi-platform observations: illustrations in the Western Mediterranean Sea. In “New Frontiers in Operational Oceanography”, E. Chassignet, A. Pascual, J. Tintoré and J. Verron, Eds, GODAE Ocean View, 663-694, doi: 10.17125/gov2018.ch24.

Rudnick, D.L. and Cole, S.T., (2011), On sampling the ocean using underwater gliders, *J. Geophys. Res.*, 116, C08010, doi:10.1029/2010JC006849.

Takeshita, Y., Martz, T.R., Johnson, K.S., Plant, J.N., Gilbert, D., Riser, S.C., Neill, C., Tilbrook, B., (2013). A climatology-based quality control procedure for profiling float oxygen data. Journal of Geophysical Research: Oceans, 118(10), 5640-5650. DOI: 10.1002/jgrc.20399

Todd, R.E., Rudnick, D.L. and Davis R.E. (2009), Monitoring the greater San Pedro Bay region using autonomous underwater gliders during fall of 2006, J. Geophys. Res., 114, C06001, doi:10.1029/2008JC005086.