4DMED-sea

ESA Contract No. 4000141547/23/I-DT

D2- Algorithm Theoretical Baseline Document

D2 – ATBD

V2.0

Marie-Helene RIO

Digitally signed by Marie-Helene RIO
Date: 2025.02.25 14:48:54 +01'00'

Prepared by:

CNR ISMAR ISTITUTO DI SCIENZE MARINE

CLS

cnrs

CSIC

Aix*Marseille université

ANI

MERCATOR OCEAN INTERNATIONAL

INNUERE CONSULTING

Version History

Date	Version	Author	Notes
28/03/2024	1.0	Verbrugge Nathalie, Greiner Eric, Ballarotta Maxime, Meda Guillaume, Sammartino Michela, Hernandez-Carrasco Ismaël, Falcini Federico, Rossi Vincent, Guglielmo Lacorata	First release
30/09/2024	2.0	Verbrugge Nathalie, Greiner Eric, Jérémy Augot, Sammartino Michela, Colella Simone, Buongiorno Nardelli Bruno, Hernandez-Carrasco Ismaël, Falcini Federico, Rossi Vincent, Guglielmo Lacorata	Final version

Table of Contents

1 Scope of the document

This document is the Algorithm Theoretical Baseline Document (ATBD) for the algorithms description of the 2DMED altimetry and SSS, the 4DMED physical and combined physical-biological and the Lagrangian tools.

It describes the related data sources, processing steps and output data. It also includes a scientific analysis of the results driving to specific development choices and trade-offs (including technical considerations justifying the selected methodologies).

A two-fold approach has been investigated for the 2DMED altimetry (section 2.1): the L4 sea level products are based on the MIOST algorithm (Ublemann et al., 2021) for the first one and on the 4DvarNET machine Learning approach (Fablet et al., 2021) for the second one.

The 2DMED SSS (section 2.2) relies on the model proposed by Sammartino et al. (2022) to produce regional L4 2D SSS maps combining satellite and in situ observations.

The 4DMED physical algorithm (section 2.3) will allow the 4D reconstruction of temperature, salinity and geostrophic currents over the Mediterranean Sea. It is based on a Machine Learning approach, with a learning step using a Mediterranean Sea reanalysis and a production step using satellite observations. Then the data obtained are combined with in situ observations to correct the large-scale residual biases and geostrophic velocities are computed through the thermal wind equation.

The 4DMED combined physical-biological algorithm (section 2.4) provides joint reconstruction of key physical and biological variables (temperature T, salinity S, density D, zonal geostrophic current Ugos, meridional geostrophic current Vgos, chlorophyll-a, primary production). It is based on Deep Learning approach using satellites and in situ observations.

The Lagrangian analysis tools algorithms will be described in Section 2.5. In particular, it provides a detailed description of the algorithms and parameters used to compute the following Lagrangian diagnostics: i) finite-size Lyapunov exponents (FSLE), based on the Lagrangian model described in Hernandez-Carrasco et al., 2011, ii) the finite-time Lagrangian vorticity (FTLV) iii) Lagrangian flow networks (LFN) and iv) the II-Kind finite-scale Lyapunov exponents (II-FSLE).Finally, a Product Validation Plan, (section 3) describing the strategy of assessment/validation of the experimental 4DMed Sea products (WP2-WP3) is presented. This section includes summary tables of the observations and metrics used for this validation.

2 Algorithms description

2.1 WP2100 - Experimental 2DMED ALT product development

Mediterranean Sea Gridded L4 Sea Surface Heights and Derived Variables

The scientific objective of this work package is to improve the retrieval of finer-scale dynamics through two complementary approaches: 1) transitioning from 1Hz to 5Hz Level-3 (L3) along-track data (e.g., the exploitation of improved altimetry datasets (WP2110)) and 2) exploring MIOST and 4DvarNET Level-4 methods as alternatives to DUACS mapping (Optimal Interpolation) in the development of novel or improved mapping algorithms (WP2120). This improvement may have a significant impact on the 4DMED Project, fostering better consistency between 2D and discrete in situ vertical profiles and enhancing the resolution of small mesoscale ocean structures in the 3D products. Additionally, within WP2130, we will focus on developing a collaborative evaluation and intercomparison framework for Mediterranean Sea level grids.

2.1.1 Inputs, outputs & mapping methods description

2.1.1.1 *Data source & Processing step*

Sea level altimetry Level-3 product

The development of the experimental 2DMED ALT product relies on the "European Seas Along Track L3 Sea Surface Heights Reprocessed 1993-Ongoing Tailored For Data Assimilation Reanalysis" product, which is distributed through the Copernicus Marine Services[1]. These datasets serve as input for the mapping algorithms and are generated by the DUACS processing chain (Pujol et al, 2016, Taburet et al., 2019). This process involves several steps (as illustrated in Figure 1), including homogenization and cross-calibration (e.g., orbit and long-wavelength errors correction).

For our analysis, we used specific data from January 1, 2016, to August 22, 2022, comprising the altimetry constellation of missions Saral/Altika, Cryosat2, HaiYang-2A, HaiYang-2B, Jason2, Jason3, Sentinel3A, Sentinel3B, and Sentinel6A (Figure 2). This dataset is characterized by a 1Hz sampling rate.

It's worth noting that we also performed experiment with an alternative, high-resolution (5Hz) L3 dataset available from July 1, 2016, to December 22, 2018, for the Jason2, Jason3, Sentinel3A, Sentinel3B, and Sentinel6A altimeter missions (Figure 3). This latter dataset is distributed as experimental product on the AVISO+ portal[2] and is referenced as "SSALTO/DUACS Experimental products: along-track sea level anomalies 5Hz v0.2" (DOI: 10.24400/527896/a01-2021.003). However, this high-resolution dataset relies on different altimetry standards (specifically Mean Sea Surface, Dynamic Atmospheric Correction, Sea State Bias corrections) compared to the Copernicus Marine L3 1Hz products. As we will demonstrate later, these differences will impact the quality of the experimental gridded product.

[1] https://doi.org/10.48670/moi-00139

[2] https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/regional/along-track-sea-level-anomalies-5hz.html

For the Level 4 gridded product generation, we specifically used the unfiltered Sea Level Anomaly variable present in the L3 products.

Ocean numerical simulations

It is important to mention that the 4DvarNET algorithm is trained using a supervised learning strategy in an Observing System Simulation Experiment (OSSE) context, taking the SSH variable of an ocean model as ground truth. Once trained in OSSE, the 4DvarNET algorithm is ready to perform SSH reconstructions with real altimetric data (previously described) as input. Our studies involve training the 4DvarNET algorithm on realistic, high-resolution numerical simulations such as the eNATL60 simulation (eNATL60-BLB002 without explicit tidal motion) and a test was also undertaken with the assimilated CMCC MEDSEA model. The eNATL60 simulation is built upon the NEMO 3.6 ocean model, featuring a spatial grid spacing of 1/60°, equivalent to approximately 0.8km to 1.6km. This model is forced by surface atmospheric conditions from the ECMWF ERA-interim 3-hourly dataset. eNATL60 spans the North Atlantic from about 6°N up to the polar circle and fully includes the Gulf of Mexico, the Med Sea, and the Black Sea. More details on the eNATL60 simulation can be found at https://github.com/ocean-next/eNATL60. The CMCC MEDSEA system uses the NEMO model for the ocean modeling part and assimilates in-situ temperature and salinity profiles as well as satellite altimetry tracks with the OceanVar 3DVar scheme. The model horizontal grid resolution is 1/24˚ (ca. 4-5 km) and is forced by hourly ECMWF ERA5 atmospheric forcing fields. More details on the CMCC simulation can be found at https://doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1 .

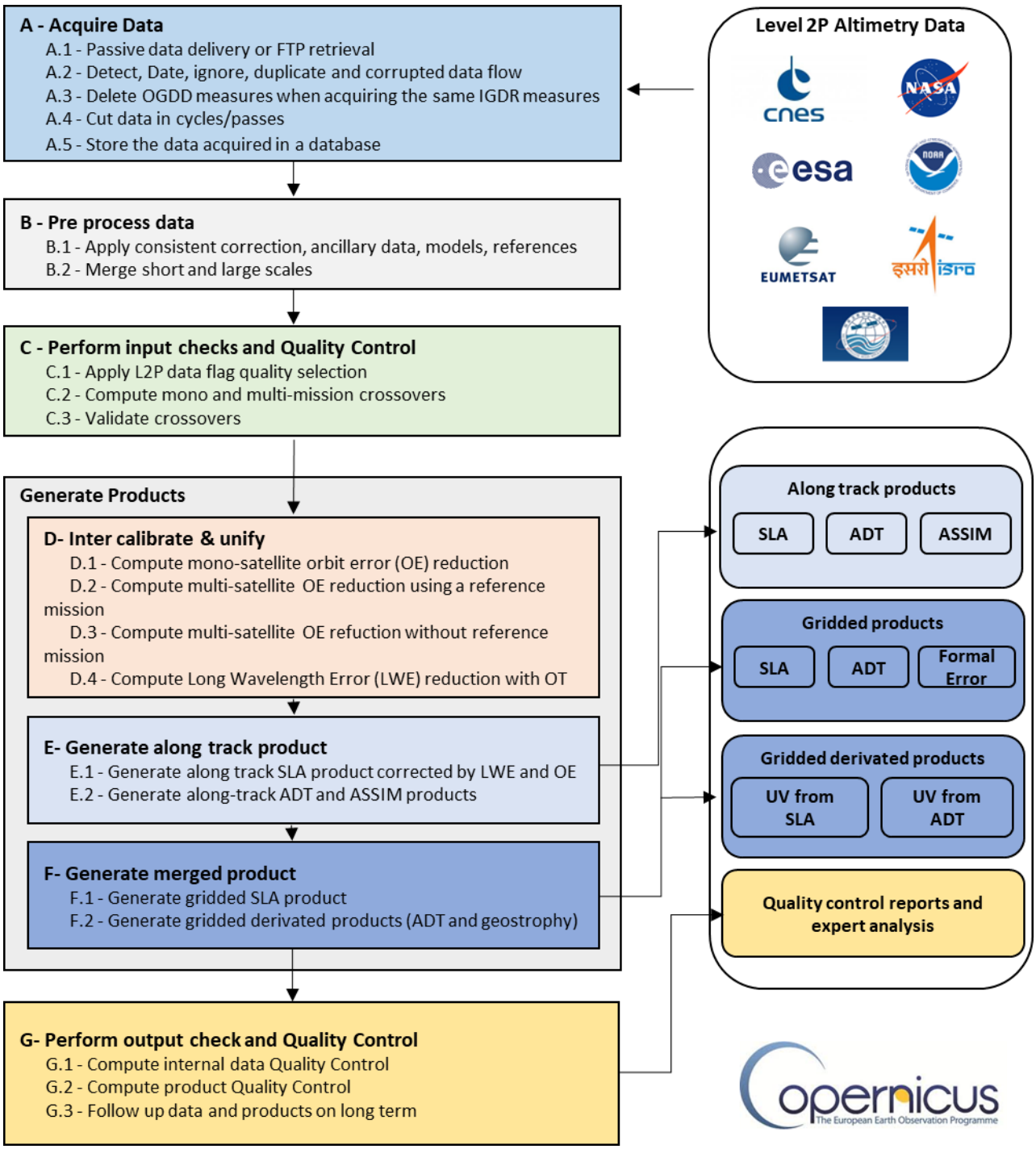


Figure 1: The DUACS processing chain

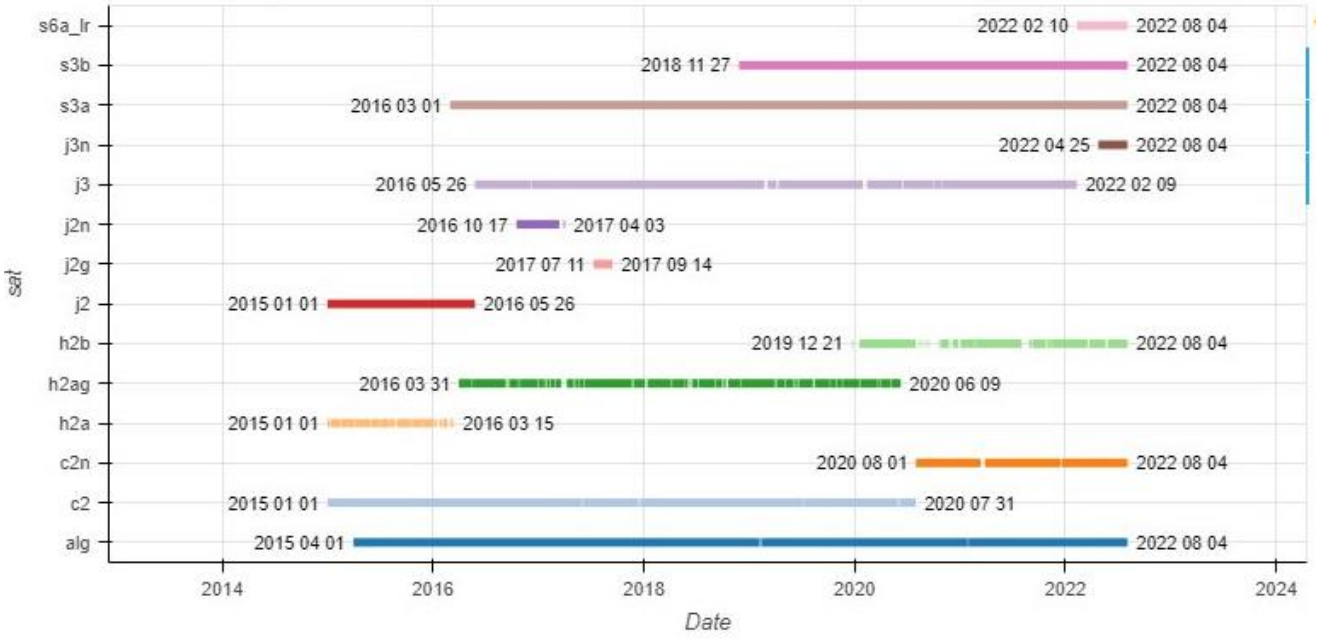


Figure 2: Altimeter constellation timeline over the period 2016 to 2023 (1Hz dataset)

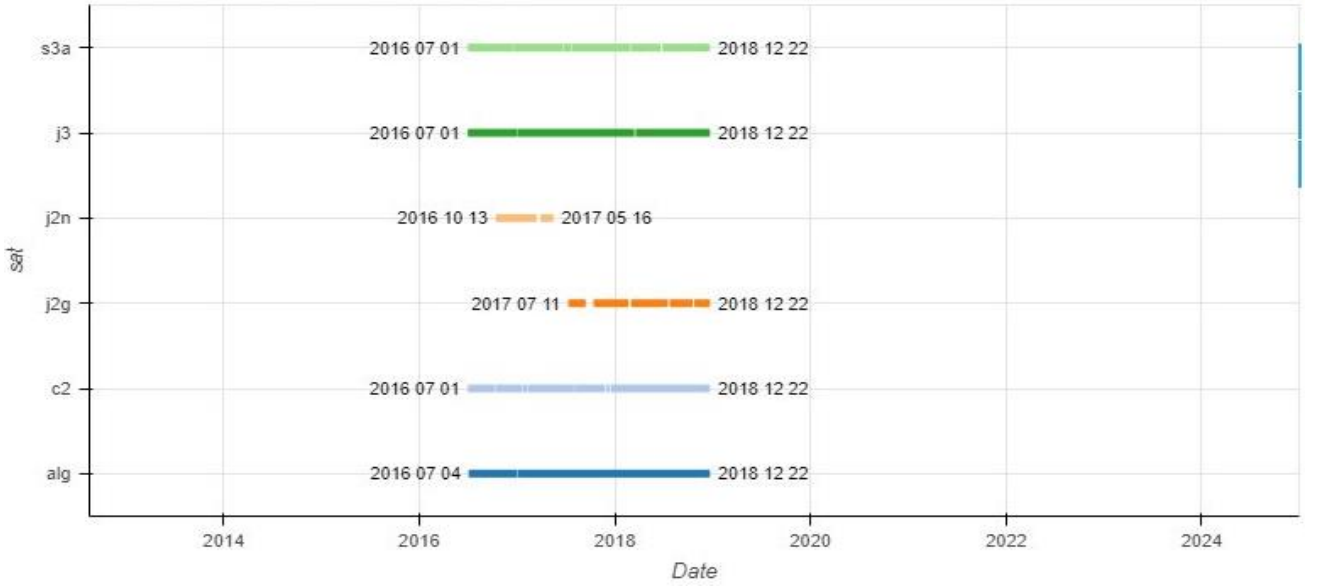


Figure 3: Altimeter constellation timeline over the period 2016 to 2023 (5Hz dataset)

2.1.1.2 *DUACS mapping description*

The DUACS system employs a linear optimal interpolation (OI) formulation (Bretherton et al., 1976; Le Traon et al., 1998), wherein a priori model specifies the covariance of sea surface height (SSH) in both spatial and temporal dimensions. Missing values are then estimated using the optimal linear least-squares estimator. The covariance is determined based on the Ahran and Colin de Verdière model (Ahran and Colin de

Verdière, 1985), which considers spatially varying zonal, meridional, and temporal decorrelation scales, specifically tuned to effectively map mesoscale ocean features.

2.1.1.3 *MIOST mapping description*

An alternative mapping algorithm, tested in the project, corresponds to the MultiScale Inversion of Ocean Surface Topography (MIOST) method (Ubelmann et al., 2021). The MIOST mapping method extends the linear mapping framework, using a wavelet decomposition to allow the construction of multiple independent components of the assumed covariance model. Here we considered a single component in the covariance model intended to represent the geostrophically balanced component of SSH evolution. This mapping solution has been tested globally and shows good mapping performance compared to the operational DUACS mapping approach (Ballarotta et al., 2023).

2.1.1.4 *4DvarNET mapping description*

Another alternative mapping solution is investigated using the 4DvarNET mapping algorithm (Fablet et al., 2021). 4DvarNET mapping is a data-driven approach combining a data assimilation scheme associated with a deep learning framework. This neural network framework involves the joint training of the representation of the ocean dynamic, as well as of the solver of the data assimilation problem. The 4DvarNET algorithm is trained using a supervised learning strategy in an OSSE context, taking the SSH variable of an ocean model (e.g. eNATL60) as ground truth. Once trained in OSSE, the 4DvarNET algorithm is ready to perform SSH reconstructions with real altimetric data as input.

2.1.1.5 *Output data characteristics*

Two Level-4 timeseries are distributed: one derived from the MIOST algorithm and the other from the 4DvarNET algorithm. For each method, the Mediterranean Sea Gridded L4 Sea Surface Heights and Derived Variables are generated on a regular grid with a spatial resolution of 1/24° in both longitude and latitude, and a temporal sampling interval of 1 day. The spatial coverage spans from 6°W to 36°E, while the temporal coverage extends from January 1, 2016, to August 1, 2022. This dataset is disseminated in the form of netCDF files, with one file per day, and includes the following variables:

Table 1 – Variables delivered in the sea level products

Variable name	Description
longitude	Longitude of the cell [degrees_east]
latitude	Latitude of the cell [degrees_north]
time	Time of the cell [days since 1950-01-01]
sla	Sea level anomaly [m]
ugosa	Geostrophic velocity anomalies: zonal component [m/s]
vgosa	Geostrophic velocity anomalies: meridional component [m/s]
adt	absolute dynamic topography [m]
ugos	Absolute geostrophic velocity: zonal component [m/s]
vgos	Absolute geostrophic velocity: meridional component [m/s]

Note that the anomaly of the geostrophic current disseminated to users is derived from gridded SLA field. It is computed using a 9-point stencil width methodology (Arbic et al., 2012). The absolute geostrophic current is obtained by adding to this anomaly the mean geostrophic current associated with the MDT field (MDT CNES CLS 18).

The MIOST gridded dataset (Ballarotta and Verbrugge, 2024) has been publicly released and is accessible via the following Zenodo address: https://zenodo.org/records/10648981

Table 2: Characteristics of the 2D MED MIOST gridded altimetry dataset

Product Name	ESA 4DMED-SEA - Mediterranean Sea Gridded L4 Sea Surface Heights And Derived Variables 4DMED_ADT_REP_2D
Geographical coverage	Mediterranean Sea [Lon -6° to 36°]
Horizontal resolution	1/24°
Variables	L4 Sea Surface Heights and Derived Variables (Table 1)
Temporal coverage	From January 2016 to August 2022
Temporal resolution	Daily field
Format	Netcdf 4.0 CF1.7

2.1.2 MIOST SL mapping in Med Sea

Several studies were undertaken to identify the performances of the mapping solutions. Our analysis is based on the metrics described in section 3 Product Validation plan.

2.1.2.1 Impact of using high resolution (5hz) data instead of conventional 1Hz data

In this study, we conducted two experiments using the MIOST method to identify the impact of switching from 1hz sampling input data to higher resolution 5Hz dataset:

Experiment #1 involved the use of only a 1Hz dataset without HaiYang-2A (H2A), while Experiment #2 incorporated a combination of 1Hz and 5Hz datasets without H2A. The main goal here was to assess how the inclusion of a high-resolution along-track dataset, alongside the low-resolution dataset, influences the mapping process compared to utilizing only the low-resolution dataset. HaiYang-2A served here as independent validation data, covering the period from January 1, 2016, to December 31, 2018. It's noteworthy that the 5Hz and 1Hz datasets exhibit differences in DAC, MSS, and SSB corrections.

The comparison of mapping errors between EXP2 and EXP1 is depicted in Figure 4 for all spatial scales and for scales < 200km. It is evident that mapping errors are larger when utilizing the HR dataset in mapping, both across all scales and, for scales < 200km

primarily near coastal areas. In terms of effective resolution (Figure 5), this corresponds to approximately a 20% degradation in resolution in the product using 5Hz along-track data as input. Note that the independent HY2A dataset used here originates from the 1Hz dataset. A similar experiment was conducted using an independent 5Hz dataset, which did not demonstrate significant enhancement with the high-resolution along-track data.

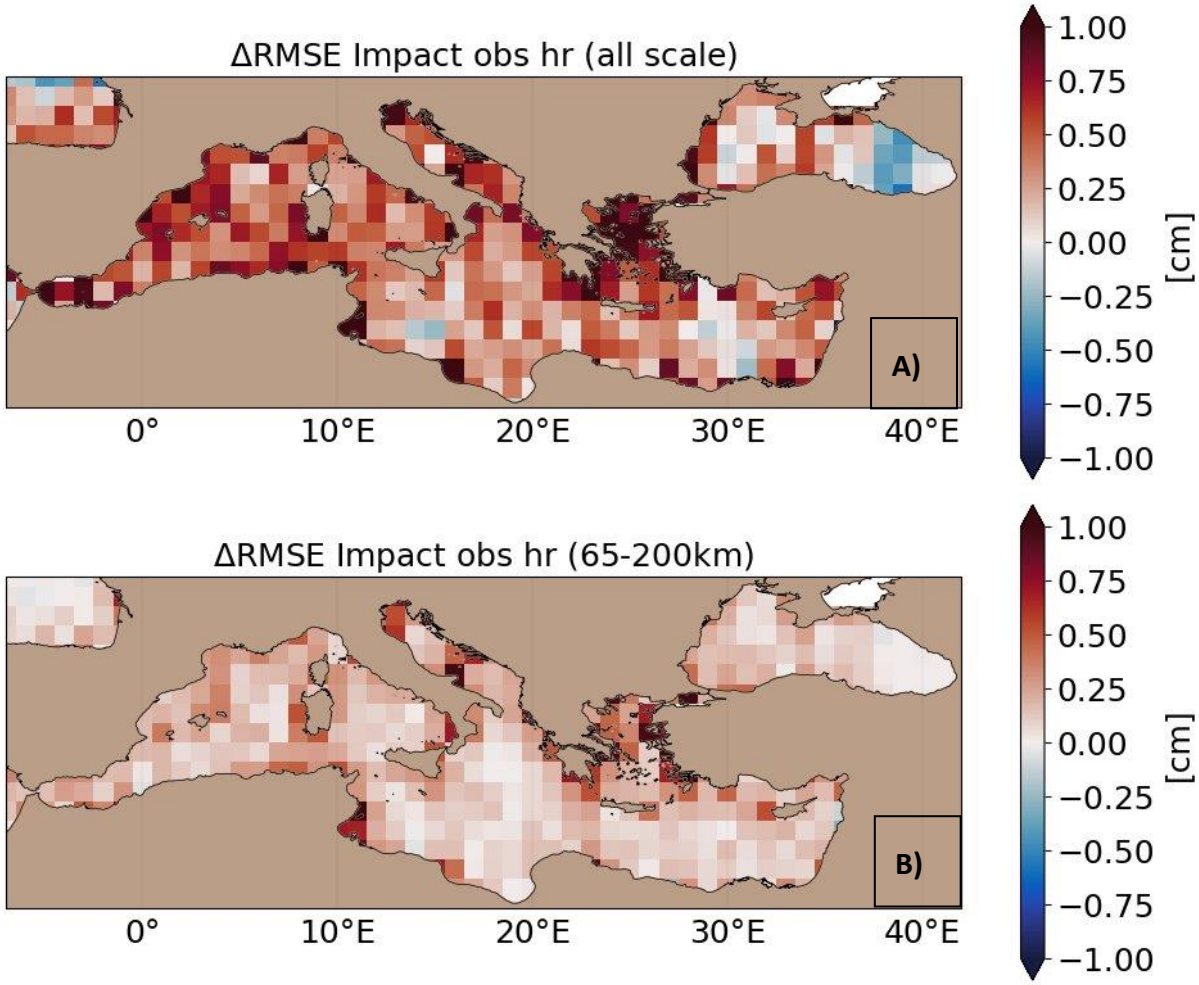


Figure 4: Intercomparing RMSE score. Gain / loss of performance (expressed in cm) when mapping is carried with 1Hz + 5Hz compared with 1Hz dataset only. Analysis is here for A) all spatial scale and B) spatial scale between 65km and 200km. Red color in RMSE means better scores when 1hz dataset only are considered in the mapping.

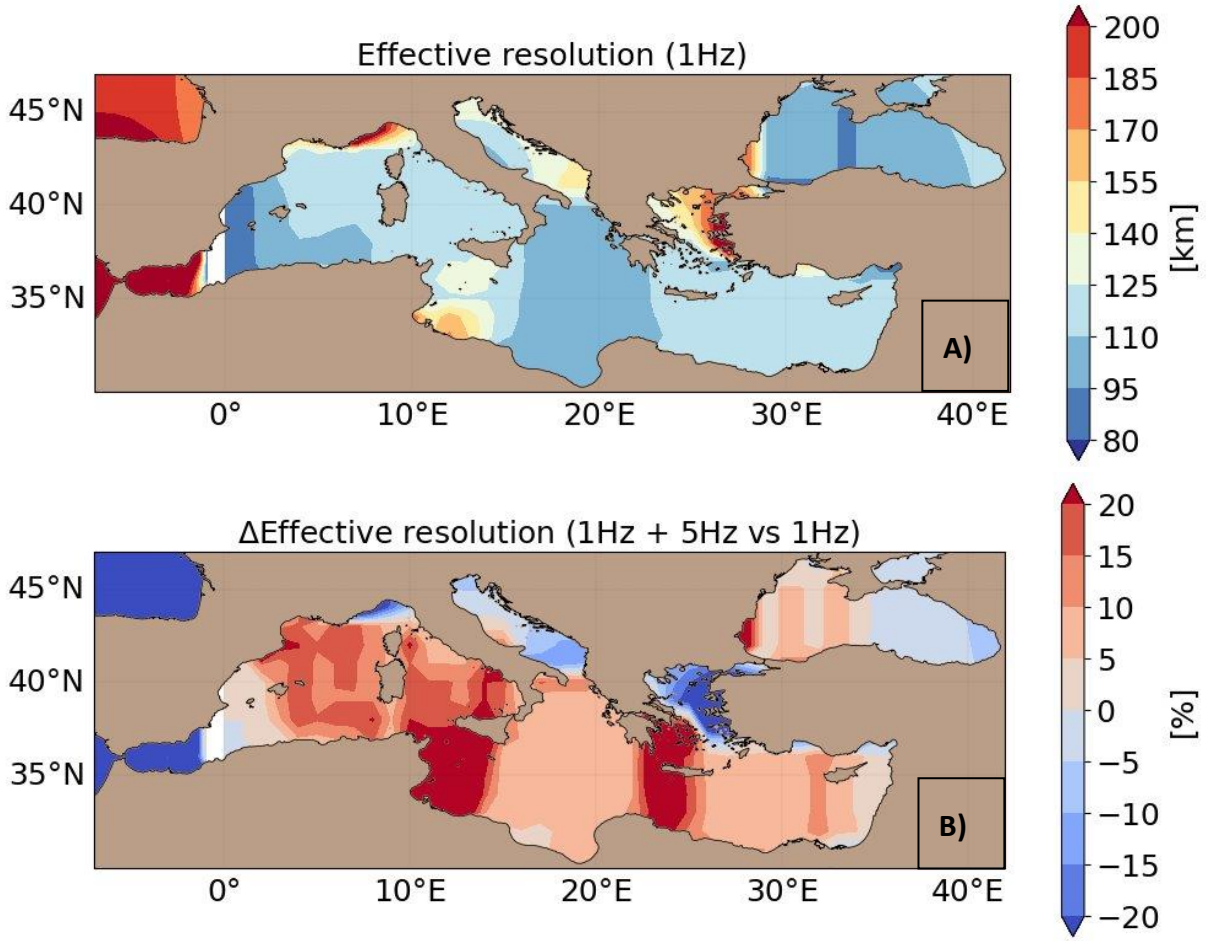


Figure 5: Intercomparing effective resolution score. A) Effective resolution of maps computed with 1Hz dataset only and B) Gain / loss of effective resolution (expressed in %) when mapping is carried with 1Hz + 5Hz compared with 1Hz dataset only. Red color means degradation of effective resolution when using 1Hz + 5hz dataset.

2.1.2.2 Impact of using MIOST mapping method instead of DUACS mapping

In this study, we carried out two experiments using the DUACS and MIOST systems to assess the impact of transitioning from DUACS mapping to MIOST mapping: Experiment #1 is done with the DUACS system and with all 1Hz along-track datasets except for HaiYang-2A (H2A), which was used for independent evaluation. Experiment #2 was performed with the MIOST system and with all 1Hz along-track datasets except for HaiYang-2A (H2A), also set aside for independent evaluation.

The mapping errors of the operational DUACS system (EXP1) are illustrated in Figure 6 for all spatial scales and scales < 200km, respectively. The most significant errors are observed in highly energetic regions and coastal areas. The mapping error in the Mediterranean Sea basin is approximately 20cm² for all spatial scales, while it is about 1.5cm² for scales less than 200km, particularly in the near of the coast and energetic regions.

The mapping errors are particularly reduced with MIOST system for all spatial scales (Figure 7a). The impact on scale < 200km is less than 0.6cm2 (Figure 7b) suggesting that

the DUACS and MIOST system map relatively the same sea surface topography at smaller scale. The impact on the effective resolution of the product is illustrated in Figure 8. The effective resolution computed from the DUACS maps ranges from 90km to almost 200km (Figure 8a). Contrasting this, MIOST exhibits a slightly finer resolution (~15%) than DUACS in the western basin, while it appears coarser (~15%) in the eastern basin (Figure 8b).

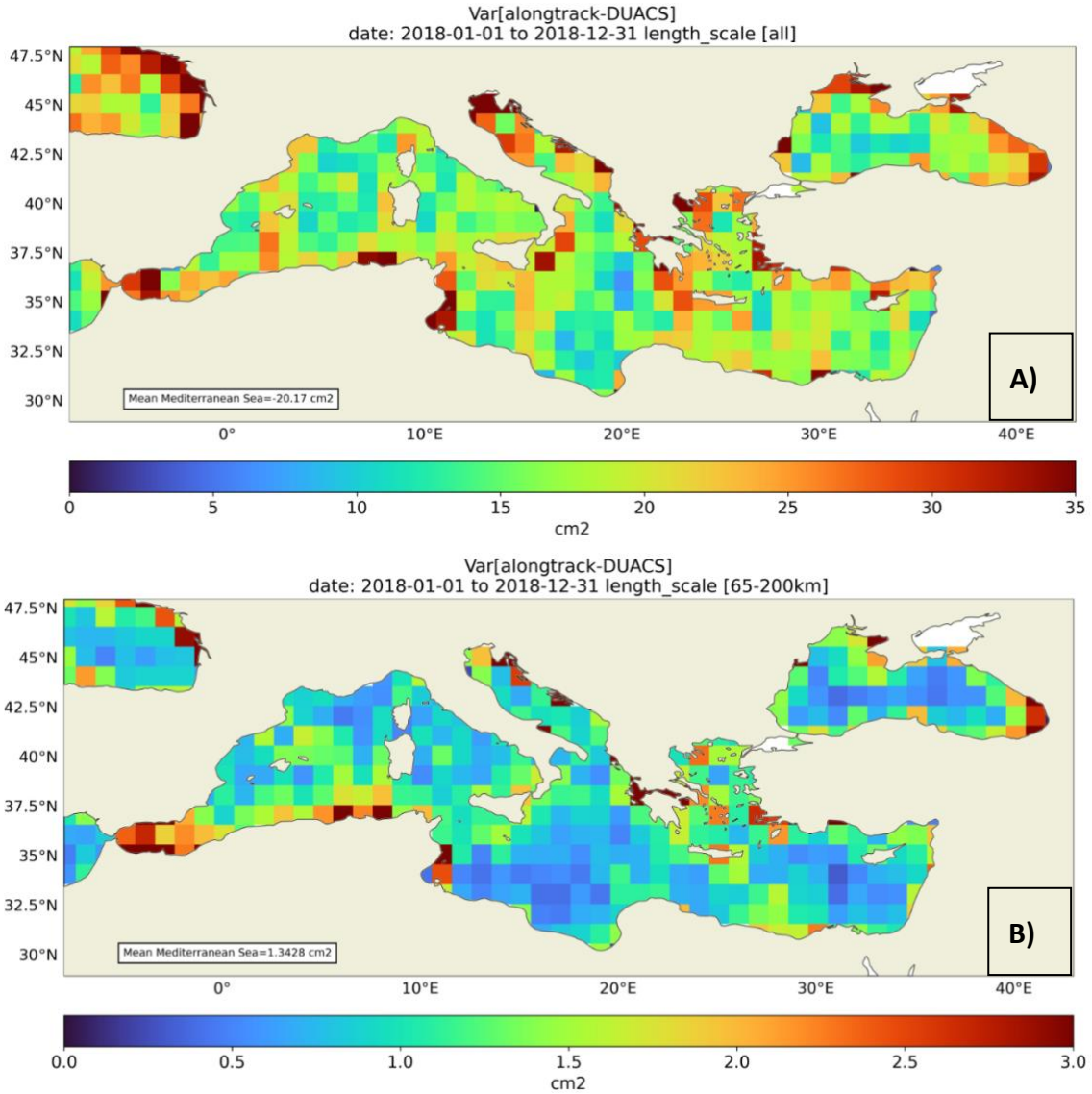


Figure 6: Variance of the mapping error with the DUACS system for A) all spatial scale and B) for spatial scale < 200km

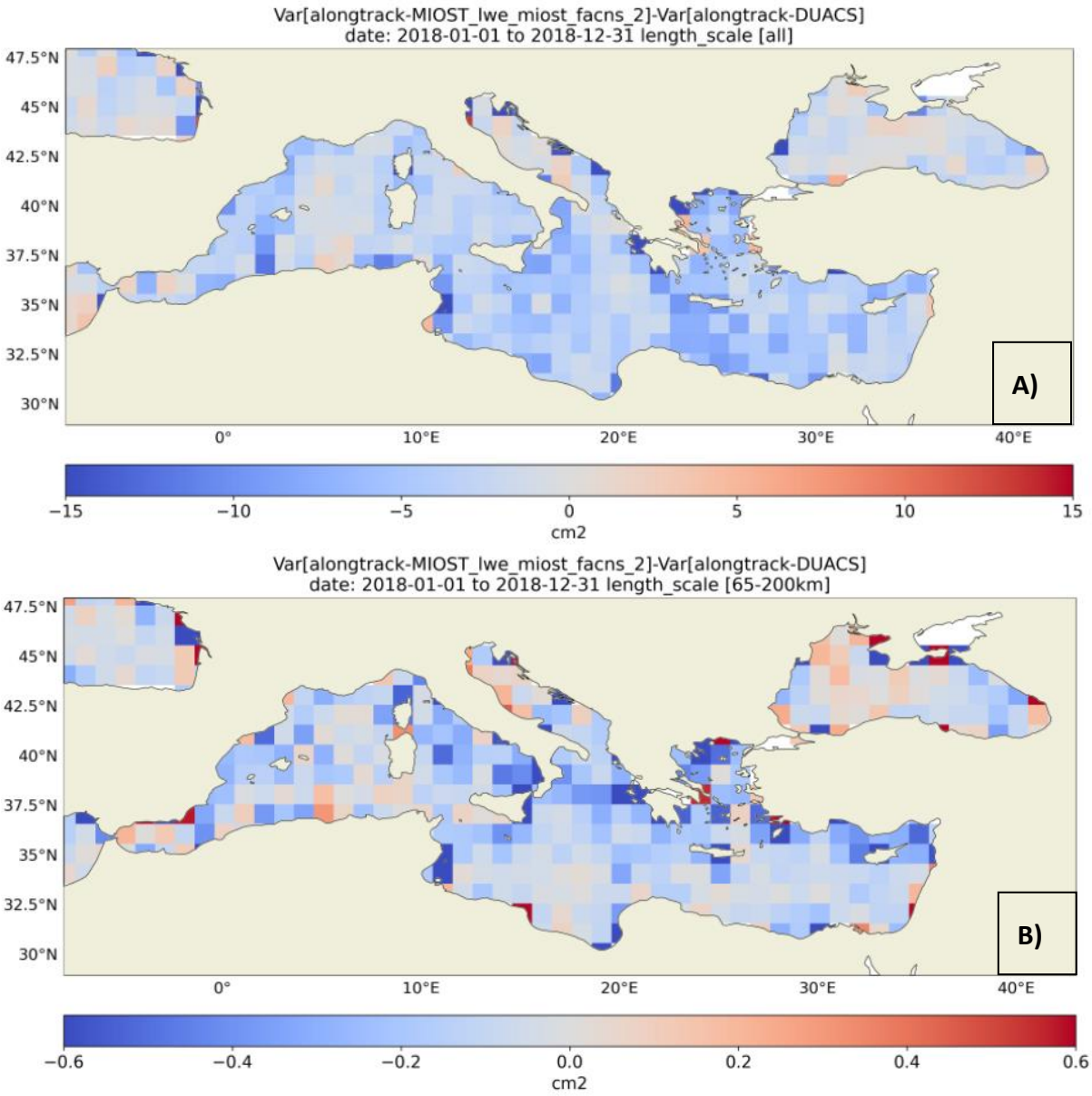


Figure 7: Difference in the variance of the mapping error between the MIOST and the DUACS systems for A) all spatial scale and B) for spatial scale < 200km. Blue means a reduction of the error with MIOST.

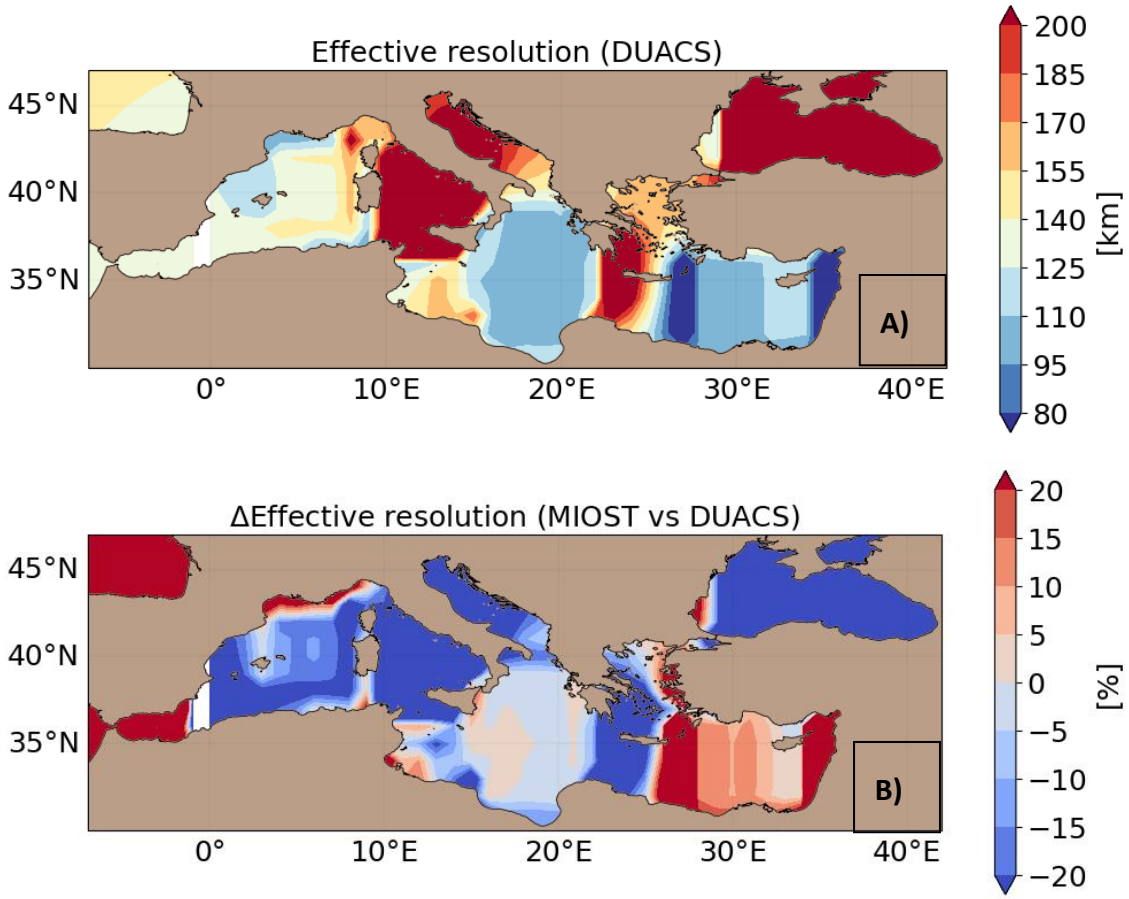


Figure 8: Effective resolution of DUACS (top) and % of gain/loss in MIOST relatively to DUACS (bottom)

2.1.3 4DVarNet SL mapping in Med Sea

Based on the encouraging results obtained in the North Atlantic basin (Meda et al., 2023), a configuration of 4DVarNet for the Mediterranean basin has been developed specifically for the project. A specific training of 4DvarNET was conducted in the Mediterranean basin, whereas previous training had been done in regions of the North Atlantic. The experiments conducted within an OSSE (Observing System Simulation Experiment) framework showed that this new 4DvarNET model tuned for the 4DMedSea project was capable of better mapping the dynamics of surface topography in the Mediterranean than MIOST and DUACS, considering both all spatial scales and fine oceanic scales. When considering simulated data, the gain compared to MIOST for fine scales (<200km) reaches 30% of error variance.

Transition to real data proved to be more complex, and several sub-studies were therefore conducted to identify the reasons for the modest performance of 4DVarNet in OSE (Observing System Simulation) and refine the processing step of the 4DVarNet chain. They are presented in the following sections.

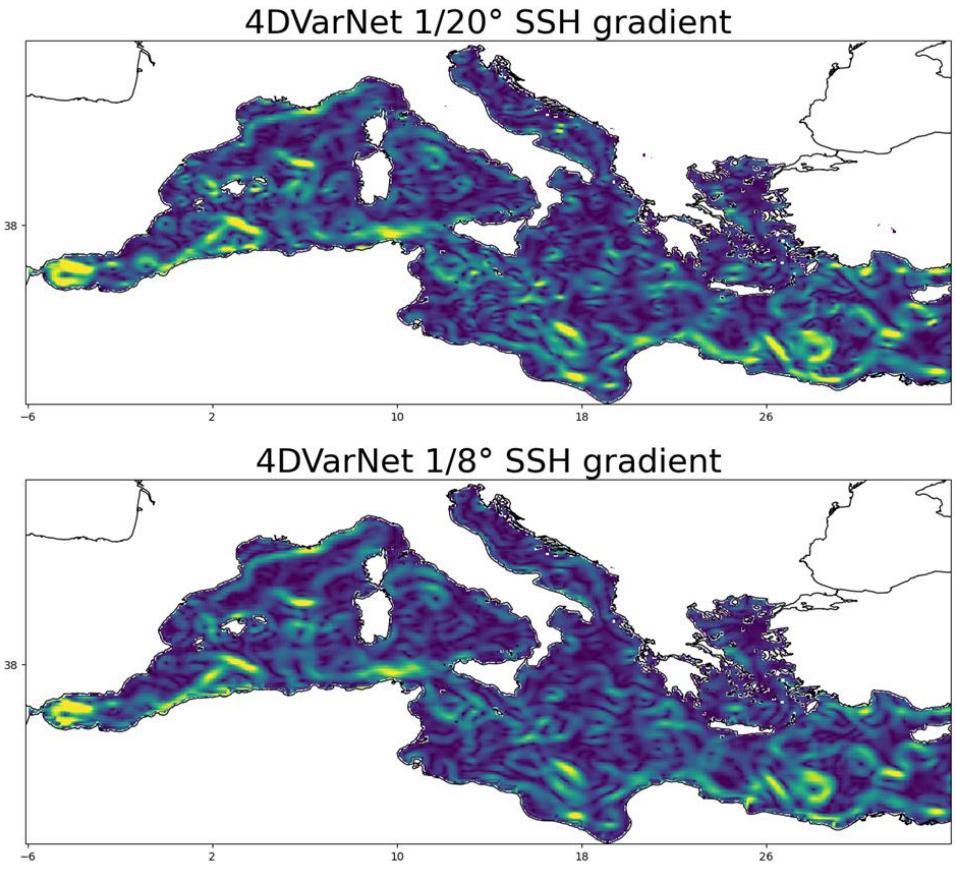
2.1.3.1 Learning model resolution: 1/8° vs 1/20°

Given that the 4DMedSea project focuses on reconstructing fine spatial scales, studies were conducted to quantify the impact of learning resolution in the context of a real

altimetry mapping in the Mediterranean Sea. Therefore, we compared the real SLA mapping performance obtained when training 4DVarNet at 1/8° and 1/20° resolutions.

On average, learning at 1/20° did not yield significant gains compared to 1/8°, as the mean scores across the basin show differences of less than 2% in terms of RMSE and a few kilometers in terms of effective spatial resolution. However, qualitatively, in Figure 9, daily snapshots of reconstructed SSH gradients show notable differences. The 1/20° version of 4DVarNet reconstructs fields significantly richer in small-scale eddy content compared to the 1/8° version of 4DVarNet or MIOST. While these small structures appear physically realistic in regions of energetic currents such as near Gibraltar or the Algerian current, this is probably not the case in the Adriatic Sea, for example. This is consistent with Figure 10, which shows that the 1/20° version, compared to a 1/8° version, improves fine structures (<200km) in these energetic regions (Algerian current) but degrades notably in the Adriatic Sea.

At this point, determining the best solution between 1/8° and 1/20° is quite challenging given our metrics, although there are significant differences in reconstructed physics at small scales. We propose providing both 4DVarNet solutions, allowing other teams to distinguish between them using complementary diagnostics such as Lagrangian diagnostics or comparison with drifters. It is worth noting that the 1/20° version will require further refinement and enhancement, especially to avoid the reconstruction of unrealistic fine-scale structures in certain areas of the basin. Some avenues to consider could be incorporating bathymetry as input data for 4DVarNet or conducting regional learning.



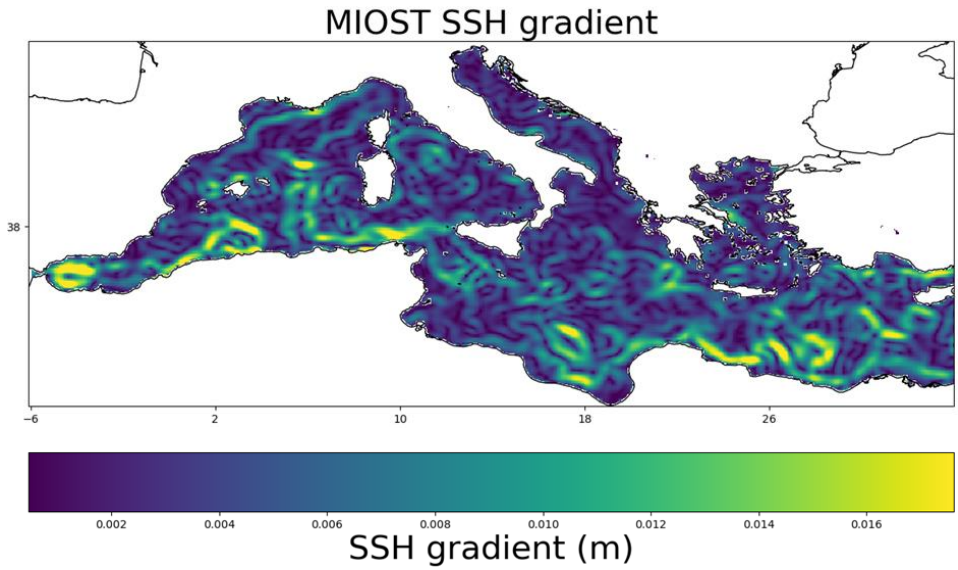


Figure 9: Daily snapshots of the SSH gradient reconstructions for day 2016-01-21. SSH fields are reconstructed from: 4DVarNet learned at 1/20° (top), 1/8° (middle), and MIOST (bottom). All SSH reconstructions are regridded to 1/24°.

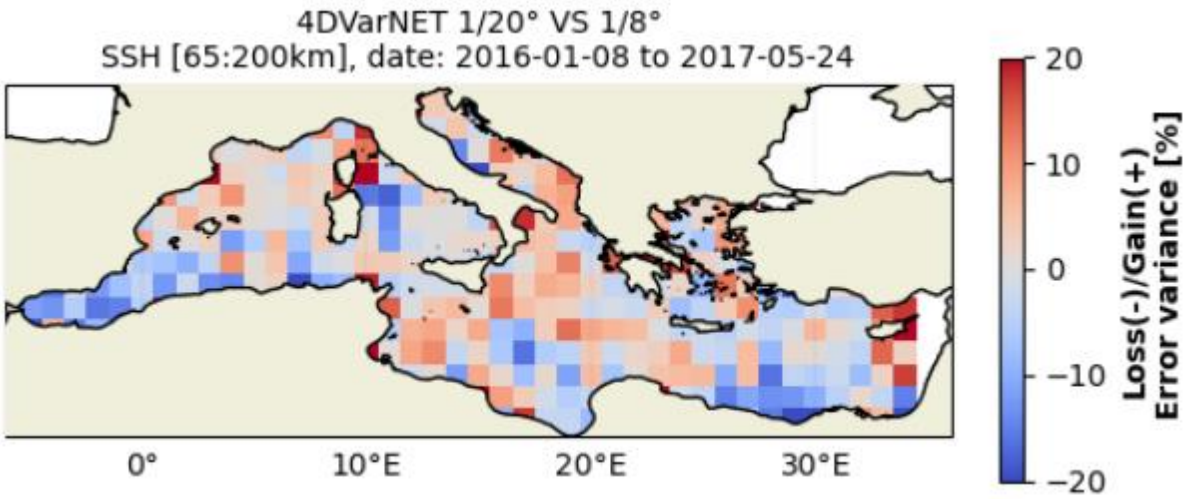


Figure 10: Estimated error variance, relative to an independent altimeter, between 4DVarNet learned at 1/20° and 1/8° resolution, for scales that range between [65:200km]. Blue means 4DVarNet 1/20° version outperforms the 1/8° version.

2.1.3.2 Learning ocean model: eNATL60 vs CMCC

Sensitivity analyses were also conducted on the impact of the ocean model used for training on the mapping performance of 4DVarNet in the Mediterranean. Indeed, the poor transfer of 4DVarNet good performance from OSSE to OSE could be attributed to an overfitting on the Enatl60 model. A poor representation of the ocean dynamics of the Mediterranean basin by the eNATL60 model might lead 4DvarNet to reconstruct unrealistic structures when mapping real surface topography in OSE. Therefore, another 4DvarNet model was learned from the assimilated CMCC MEDSEA model [https://doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1]. The results obtained showed that learning on MEDSEA CMCC led to a degradation in OSE mapping performance of around 15%, compared with eNATL60 learning. The eNATL60 was therefore kept as the reference simulation for the 4DvarNet training.

2.1.3.3 Removing the high frequency signal

The first 4DVarNet training carried out on the Mediterranean Sea with the eNATL60 simulation was challenging. The simulation includes basin scale inter-daily SLA beats over the Mediterranean basin. This large-scale, high-frequency (~ 1 day) signal, which was negligible in previous training areas (such as the Gulf Stream) dominated by mesoscale eddy variability, affects 4DVarNet learning in the Mediterranean. The 4DVarNet seems to be suited for reconstructing mesoscale or small-scale ocean structures, but not this type of high-frequency, large-scale signal. A pre-processing step was therefore carried out to remove this high-frequency signal from the pseudo-nadir observations. It consists in computing the daily mean SLA for the Mediterranean basin and removing it from the daily SLA data inputs to 4DVarNet. In both OSSE and OSE, this daily average is estimated under the Nadirs tracks (simulated or real). This pre-processing step is very effective in an OSSE context, delivering a performance gain of around 30% of error variance compared to a 4DVarNet model without this step. However, when considering real altimetric data, this preprocessing step only allows for a slight performance gain (a few percentage points of error variance). This might be attributed to this high-frequency signal being much weaker in real altimetry data than in simulated data, as it is removed by the various processing stages of the DUACS chain.

2.1.3.4 Size of the space-time 3D reconstruction 'patches'

Let us recall that in the 4DVarNet framework, the mapping over a given domain is carried out following a 'patches' strategy, which involves multiple space-time 3D sub-blocks. The overall solution (in our case, over the Mediterranean basin) is reconstructed by averaging multiple overlapping patches. The size of the patch chosen in time and space limits the maximum spatial and temporal scales for which the model will attempt to establish correlations between data from the altimeter constellation. Therefore, one critical aspect is the adjustment size of the 3D reconstruction 'patches' to the space-time scales of the typical structures to be reconstructed in a basin of interest. In the Mediterranean basin, it seems intuitive to reduce patch sizes in latitude and longitude, as the structures of interest are smaller and more localized than in other regions such as the North Atlantic basin. Also, due to the unique geometry of the Mediterranean basin, smaller patches would prevent patches with a land majority in areas like Gibraltar Strait or Adriatic Sea. On the other hand, the limited space-time coverage of the Nadirs constellation requires selecting large enough patches so that the 4DvarNet model has enough altimetric data inputs to perform the mapping successfully. Consequently, the optimal patch size must strike a balance between managing the coastline feature, the size of oceanic structures to observe, and ensuring enough altimetric data. The Mediterranean basin is thus particularly complex from this point of view, due to its geometry and the size of its oceanic structures. Empirically, a patch size of 16° in latitude and longitude, and 15 days in time has proven to be a good compromise.

2.1.2.5 Key scientific results and prospects

The 4DvarNet experimental learning design that enabled the best mapping performance in OSE is a learning on eNATL60, over the whole Mediterranean basin, and on one year of data from mid-2009 to mid-2010, with daily outputs. The simulated altimeter constellation is made up of 5 noise-free nadirs with 1hz sampling: H2B, J3, S3A, S3B, S6. The temporal and spatial lengths of the reconstruction patches are set to 16° in both latitude and longitude, and 15-day long in time (i.e. a data assimilation window of 15 days). The 4DVarNet framework takes as input the nadirs SLA normalized by the SLA standard deviation over the basin and the period studied. During preprocessing, the basin daily average of SLA is estimated under the tracks and removed to eliminate the large-scale and high-frequency signal.

2.1.4 Comparison 4DVarNet vs MIOST

To quantify the mapping error or intercompare with alternative mapping solution such as the 4DvarNET method, specific maps excluding one altimeter (H2A) have been carried out.

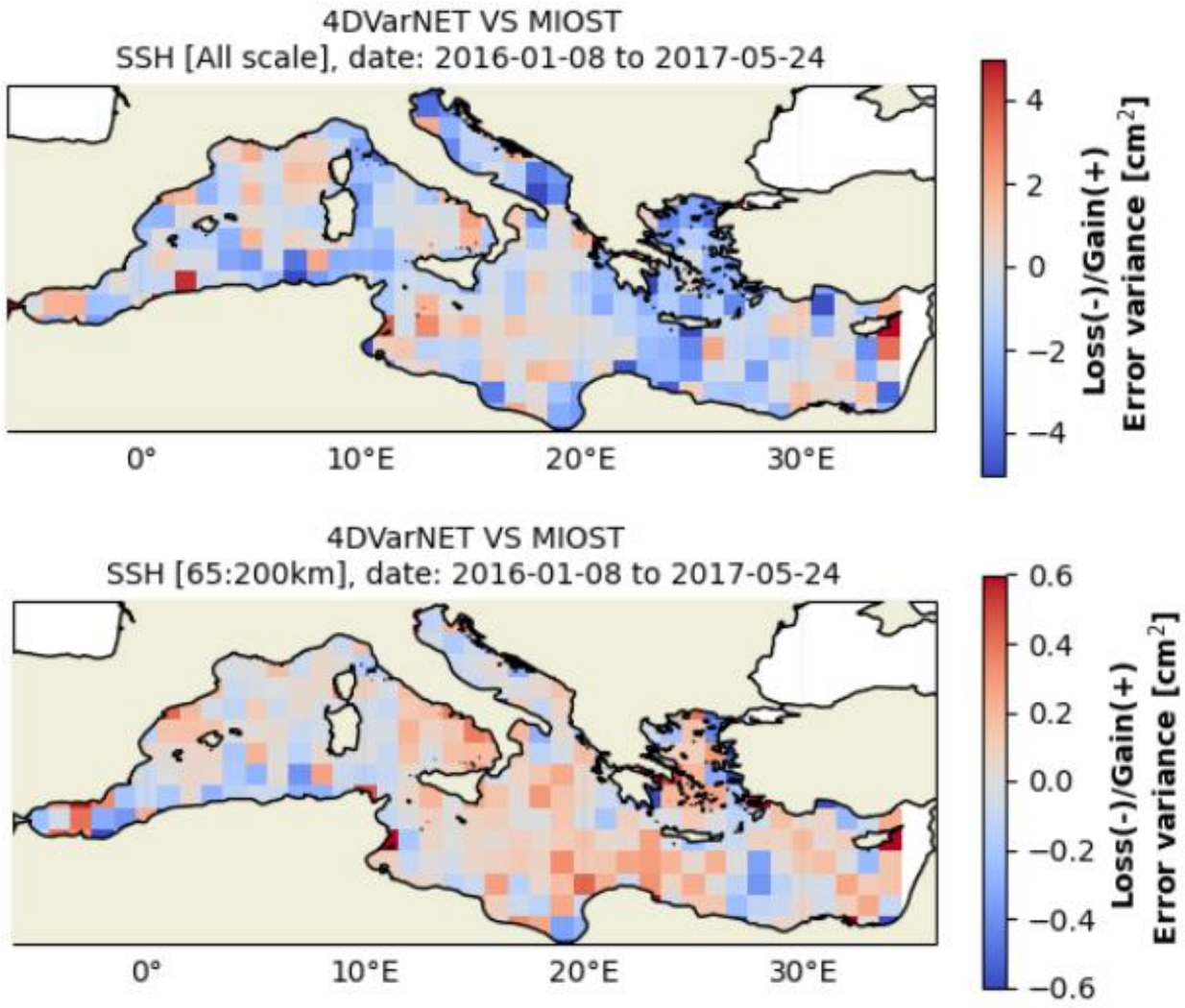


Figure 11: Estimated error variance, relative to an independent altimeter, between 4DVarNet 1/20° and MIOST for all scales (top) and scales that range between [65:200km]. Blue means that 4DVarNet outperforms MIOST.

Considering all scales, error variance comparisons in Figure 11 indicate that 4DVarNet mapping appears slightly superior to MIOST across most of the basin. When considering fine-scale structures (<200 km), the assessment is more nuanced, with 4DVarNet slightly outperforming MIOST in energetically active regions like the Algerian Current, and performing less well in the central basin, for example. It is worth noting that the locally observed differences in variance error between the two methods are very slight, being less than 0.4 cm². The average scores calculated over the entire basin, as depicted in Table 3, also show marginal differences.

Table 3: Whole Mediterranean basin average RMSE metrics, relative to an independent altimeter over the period 2016-01-08 to 2017-05-24, between 4DVarNet and MIOST for all scales and scales that range between [65:200km].

Scale	Method	$RMSE - cm$	$\sigma(RMSE) - cm$
All Scale	4DvarNET	4.402	1.254
	MIOST	4.464	1.275
65 - 200 km	4DvarNET	1.091	0.241
	MIOST	1.075	0.239

In conclusion, with the current metrics, it is challenging to determine definitively whether 4DVarNet or MIOST is the best solution for fine-scale mapping in the Mediterranean Sea. However, as mentioned in the section, there are differences in the small-scale physical content between the two solutions. Consequently, further analysis undertaken by the other working group (Lagrangian analysis tools, physical-biological analysis, etc.) may reveal which solution is the best in term of physical content.

Finally, a solution relying solely on conventional nadir altimetry seems limited for observing fine scales by 4DVarNet. Possible avenues for improvement in this regard could be considered:

- Inclusion of auxiliary data such as bathymetry, SST, and wind in the input data to enrich the model. As a data-driven mapping solution, 4DVarNet is easily configurable to integrate heterogeneous data from different sensors and extract useful information from them.
- Fine-tuning the model with real data (the target is now real independent nadir) to limit overfitting on an idealized physics represented in the model used for learning (here eNatl60).
- Inclusion of SWOT data, which observes very fine structures.

2.1.5 Data Challenge

The ESA-4DMEDSea Data Challenge is designed to assess experimental sea surface height gridded products produced within the ESA-4DMEDSea project. Through a

dedicated website[3] (Figure 12) and associated GitHub repository[4], participants can access essential resources, including input datasets, independent validation data, and comparison tools for evaluating their methodologies.

Currently, the evaluation process involves a comparison among the Copernicus Marine operational product, MIOST and 4DvarNET mapping techniques. Note that the MIOST and 4DvarNET product validation proposed in the data-challenge are based on specific mapping experiment where H2A altimeter is set aside for independent validation, while the Copernicus Marine operational product includes all satellites. This evaluation employs a range of metrics (RMSE (Root mean square error), effective resolution) to gauge the accuracy, reliability, and efficiency of the gridded sea surface height and currents. Leveraging independent datasets such as along-track altimeter and in-situ drifter data, the framework aims to provide insights into the efficacy of different mapping reconstructions.

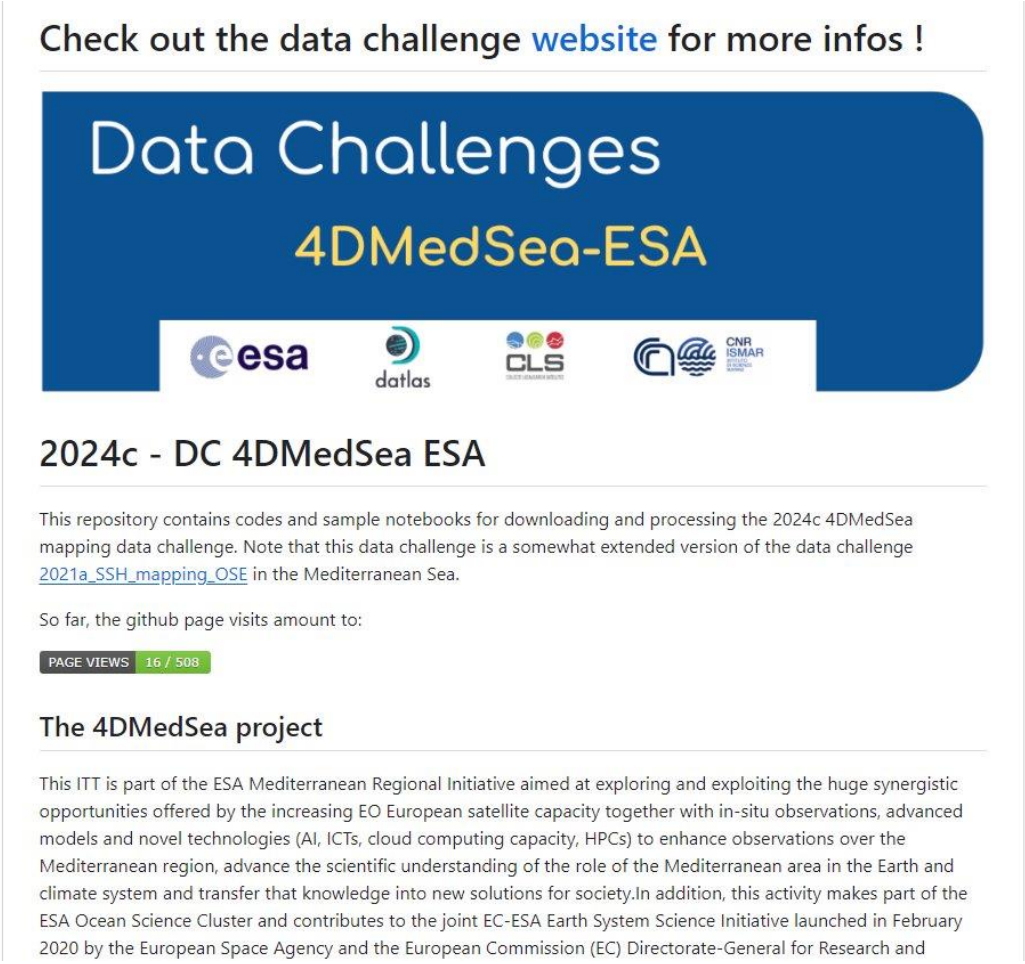


Figure 12: DataChallenge Website

[3] https://2024c-dc-4dmedsea-esa.readthedocs.io/en/latest/index.html
[4] https://github.com/ocean-data-challenges/2024c_DC_4DMedSea-ESA?tab=readme-ov-file

Main mapping scores are consolidated in a readily accessible scoreboard[5] facilitating informed decision-making for researchers and users concerning the selection and application of diverse mapping approaches.

SUMMARY

For this WP2100 work package, we firstly conducted experiments incorporating both High-Resolution (5Hz) and Low-Resolution (1Hz) along-track data into the MIOST mapping algorithm. However, merging the 1Hz and 5Hz datasets in the mapping process proved challenging, attributed to differences in corrections between the 1hz and 5Hz datasets. To ensure consistency in the timeseries, it is recommended to exclusively use the 1Hz dataset in the mapping procedure.

Mapping technique intercomparison indicates that the MIOST solution is slightly better than the operational DUACS method over the MED Sea, in particular when considering all spatial scale, but for scale <200km the mapping technique are relatively similar.

A 4DvarNET solution has been adapted and tested specifically for the Mediterranean Sea. A comparison between the MIOST and 4DvarNET mapping solutions revealed a general similarity in retrieving small-scale oceanic SSH structures over the Mediterranean Sea, as indicated by the diagnostics generated in our work package.

Overall, at our level, we found that the DUACS, MIOST and 4DvarNET are globally similar for retrieving fine scale structure in the Mediterranean Sea. However, it's worth noting that further analysis undertaken by the other working group, particularly WP3600, (in utilizing their own diagnostics such as Lagrangian analysis tools or physical-biological analysis), may present a different perspective on the finer-scale processes resolved by the two experimental mapping systems tested in our study.

The activities conducted within this work package have contributed to the refinement and enhancement of the regional MIOST system for Mediterranean/European Sea products that will be distributed in Copernicus Marine by the end of 2024. Furthermore, it facilitated the development of a 4DvarNET version tailored to the basin, although it requires further refinement to accurately capture the complex dynamics that take place here.

Specific timeseries maps excluding H2A data that have been produced for the experiment of this work package are accessible in the data challenge repository: https://github.com/ocean-data-challenges/2024c_DC_4DMedSea-ESA.

Additionally, MIOST and 4DvarNET maps covering all altimeter data from January 1, 2016, to August 1, 2022, have been created and are accessible on a 1/24° grid and are available online (MIOST: https://zenodo.org/records/10648981; 4DVARNET with learning at 1/20° : https://zenodo.org/records/10912777, 4DVARNET with learning at 1/8° : https://zenodo.org/records/10908416)

[5] https://2024c-dc-4dmedsea-esa.readthedocs.io/en/latest/2_eval_generic/overall_scores.html

2.2 WP2200 - Experimental 2DMED SSS product development

The objective of this task is to develop an improved 2D Sea Surface Salinity (SSS) product over the Mediterranean Sea, which is needed as input to the 4DMED experimental processing chains. To this aim, a specific effort has been dedicated to develop, optimize, test and assess the new SSS product (WP 2210, 2220).

The development of experimental 2DMED SSS product started from the model proposed by Sammartino et al. (2022). It provides daily L4 Mediterranean SSS fields at 1/16° grid resolution by applying multivariate optimal interpolation algorithm (Buongiorno Nardelli et al. 2012, 2016; Droghei et al. 2016, 2108) to integrated multi-sensor satellite (NASA's Soil Moisture Active Passive (SMAP) and ESA's Soil Moisture and Ocean Salinity (SMOS) platforms) and in situ SSS observations, combined with remotely-sensed Sea Surface Temperature (SST). This model has been optimized to produce the experimental 2DMED SSS product at 1/24° grid of resolution.

The optimization has been carried out refining the background field ingested as input in the OI (Optimal interpolation) processing chain limiting the use of climatological fields to main rivers. A new refined mask focused on specific most big river mouth areas and North Adriatic Sea has been created and then applied to the in situ monthly climatology and weekly global SSS means used to construct the background. The optimization included also the use of a higher resolution SST product (UHR) with respect to that (HR) used in Sammartino et al. (2022)

2.2.1 Input datasets

Five different datasets have been used for the development of the experimental 2DMED SSS product:

- Differing from the SST dataset used in Sammartino et al. (2022), within 4DMED-SEA project a product at higher resolution is used for the 2DMED SSS processing. The Mediterranean Ultra-High Sea Surface Temperature (SST) field is based on the SST_MED_SST_L4_NRT_OBSERVATIONS_010_004 dataset produced by the SST-CNR-ROMA-IT Production unit and distributed in near-real time by Copernicus Marine Service (https://doi.org/10.48670/moi-00172). The provided ultra-high SST is a regional daily gap-free (L4) product at 0.01° of spatial resolution available from 2008 to 2023 (for more details about the algorithm and processing see https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-SST-QUID-010-004-006-012-013.pdf).

- Satellite SSS observations include observations from both ESA's Soil Moisture and Ocean Salinity (SMOS) and NASA's Soil Moisture Active Passive (SMAP) satellite. The former consists of the L3OS MIR_CSF2Q debiased daily valid ocean salinity values at ~25 km from 2010 to present. It is disseminated by the Centre Aval de Traitement des Données SMOS (CATDS) (for more details, see http://dx.doi.org/10.12770/12dba510-cd71-4d4f-9fc1-9cc027d128b0). Ascending and descending orbits are merged and passed as an input to the Optimal Interpolation processing. The latter data are 0.25° longitude/latitude Level-3 gridded sea-surface salinity (SSS) daily mean, provided by the NOAA

CoastWatch/OceanWatch from 2015 to present (for more details, see https://coastwatch.noaa.gov/cwn/products/sea-surface-salinity-near-real-time-smap.html).

- The third dataset is used to build the background field needed as input to the optimal interpolation. It is based on the MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013 product (https://doi.org/10.48670/moi-00051), distributed by the Copernicus Marine Service platform. The weekly global gap-free Level-4 (L4) analyses at ¼° grid of resolution is obtained through a multivariate optimal interpolation algorithm that combines SMOS+SMAP satellite images and in situ salinity measurements with satellite SST information (Buongiorno Nardelli, 2012; Buongiorno Nardelli et al., 2016; Droghei et al., 2016; see reference to the CMEMS QUID at (https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-MOB-QUID-015-013.pdf).

- The monthly in situ Mediterranean climatology is extracted from a high-resolution atlas of salinity observations, covering the Mediterranean Sea (Iona et al. 2018). This is based on data collected within the pan-European marine data infrastructure SeaDataNet, containing the most complete quality controlled in situ data collection for the Mediterranean Sea. The dataset is based on in situ measurements, acquired between 1950 and 2015. The atlas consists of horizontal gridded fields produced by the Data-Interpolating Variational Analysis, in which unevenly spatial distributed measurements were interpolated onto a 1/8◦ × 1/8◦ regular grid on 31 depth levels. Along with the weekly gap-free Level-4 (L4) analyses, this monthly climatology has been used to create the background over river mouth areas and coastal ones.

- The in situ dataset comprises surface salinity values coming from different platforms and instruments as CTD, Argo profiler, thermosalinograph and drifter. In this work two datasets have been used: the first is an internal product of Copernicus Marine Service (INSITU_GLO_TS_ASSIM_REP_OBSERVATIONS_013_051 (REP)) and accounts for the years from 1992 to 2020; the second is used from 2021 onward and is based on the INSITU_GLO_PHY_TS_DISCRETE_MY_013_001 (EasyCORA) product. Both datasets provide quality controlled and homogenized profiles/time series of T/S, by concatenating the data per day and type to reduce the total number of data files. In this work, this dataset has been remapped on the same output grid and used as an input for the multidimensional OI, along with the background and SSS satellite data.

2.2.2 Multivariate Optimal Interpolation Algorithm and processing chain

The 2DMED SSS L4 daily maps are obtained by applying an Optimal Interpolation (OI) method to the synergic combination of in situ and multi-sensor satellite SSS data over the Mediterranean Sea. OI is one of the main and powerful techniques to interpolate geophysical fields, in case of discrete observations. The OI algorithm provides the desired values ($x_{analysis}$) at the interpolation grid point, as a weighted sum of the anomalies of N

observations ($y_{observed}$), with respect to the first guess represented by the background (x_{first_guess}):

$$x_{analysis} = x_{first\_guess} + C(R+C)^{-1}(y_{observed} - x_{first\_guess}) \quad (1)$$

where C represents the background error covariance, and R represents the observation error covariance (here, assumed diagonal, meaning that observation errors are constant values per each observation type or platform):

$$C = E\left\{\varepsilon_{fg}\varepsilon_{fg}^{T}\right\} = E\left\{\left(x_{first\_guess} - x_{true}\right)\left(x_{first\_guess} - x_{true}\right)^{T}\right\} \quad (2)$$

$$R = E\left\{\varepsilon_{obs}\varepsilon_{obs}^{T}\right\} = E\left\{\left(y_{observed} - x_{true}\right)\left(y_{observed} - x_{true}\right)^{T}\right\} \quad (3)$$

The OI method also allows us to compute the variance of the error of the optimal analysis field $x_{analysis}$:

$$OI_{error} = 1 - C(R+C)^{-1} \quad (4)$$

Based on the same OI scheme adopted by Droghei et al. (2018) and Buongiorno Nardelli et al. (2016), the observation error covariance R is expressed here as a noise-to-signal ratio (dividing it by signal variance). The background field is computed from the combination of the Mediterranean in situ monthly climatology (Iona et al. 2018) over coastal and specific river mouth areas and the previous day 2DMED SSS High Resolution (HR) analysis over open ocean (if previous analysis is not present the upsized Lower Resolution (LR, ¼°) Copernicus Marine SSS L4 weekly product V.8 (MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013) is used over open ocean). Both datasets (in situ climatology and global weekly product) have been, first, remapped on the output grid resolution (1/24◦ × 1/24◦), through a 2D cubic-spline interpolation, then, the weekly fields were linearly interpolated in time and successively combined with the in situ climatology. The blending of both fields has been carried out by using a specific new mask (Figure 13) at higher resolution (1/24°) with respect to that used in Sammartino et al. (2022) (1/16°). It was obtained following the same grid of a Mediterranean Copernicus Marine Product at 1/24° of resolution SEALEVEL_MED_PHY_MDT_L4_STATIC_008_066 and it has been also refined in some specific areas close to the main rivers and North Adriatic Sea. The comparison of the mask used in Sammartino et al. (2022) and the new mask used for the experimental 2DMED SSS product development is given in Figure 13.

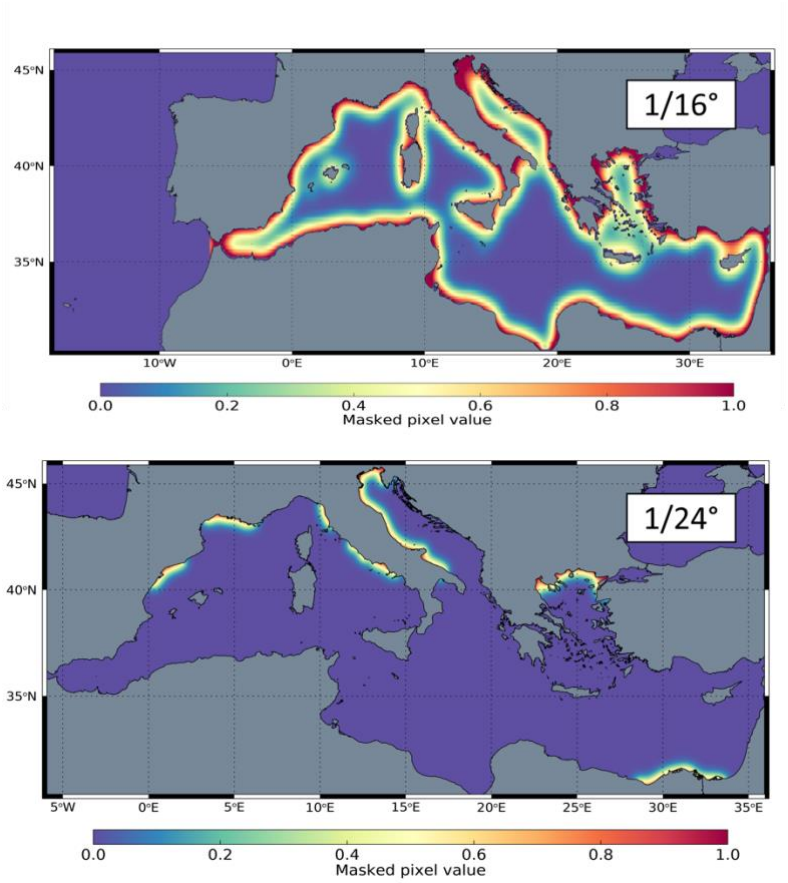


Figure 13: Comparison of the mask used in Sammartino et al. (2022) (left panel) and the new refined mask at higher resolution (right panel)

This refined mask includes values that progressively spans from 0 (offshore) to 1 (inshore). Following Equation (5), the previous analysis/weekly background data are gradually replaced by in situ monthly climatology in specific coastal areas (mainly corresponding to the main big rivers) moving from offshore toward inshore.

$$\text{background} = \left(\text{SSS}_{\text{climatology}} \times \text{mask}\right) + \text{SSS}_{\text{CMEMS}} \times (1 - \text{mask}) \qquad (5)$$

This background is also exploited to extract pseudo-observations (with 6-pixel subsampling step) to be included as input into the OI process. As detailed in previous works (Droghei et al. 2016) and in the quality information document of Copernicus Marine MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013 product (https://doi.org/10.48670/moi-00051), the use of pseudo-observations allows to "re-modulate" the background field, using the information extracted from the SST pattern. By this way, this approach can reproduce mesoscale patterns, also, when too sparse observations are available.

In this work, the in situ observation noise-to-signal ratio was set to a constant value of 0.05, while for pseudo-observations two different errors have been considered, 0.1 and 0.4 for the in situ monthly climatology and previous analysis/weekly means pixels, respectively.

In principle, the estimation of background error covariance should encompass all available observations. However, with satellite data, this can become too complex and can require big computational resources. Consequently, the background error is typically approximated using an analytical function based on the distance between samples.

Nevertheless, depending on the system under consideration and the available data, covariance models may also extend to multidimensional spaces, such as incorporating space/time distances into the covariance function definition.

In this study, we employ the same model as developed by Buongiorno Nardelli et al. (2012). This multidimensional covariance model includes high-pass-filtered thermal differences (ΔSST) alongside space-time differences (Δr, Δt), see Equation 6. The inclusion of SST forces the interpolated field to follow the surface isotherms giving more weight to observations located on the same isotherm as the interpolation point compared to those with similar spatial and temporal differences but differing in their sea surface temperature (SST) values. The integration of high-resolution satellite SST data proved to be an effective strategy for enhancing the resolution of Level 4 sea surface salinity (L4 SSS) fields, giving insight on mesoscale dynamics.

$$C(\Delta r, \Delta t, \Delta SST) = e^{-\left(\frac{\Delta r}{L}\right)^2} e^{-\left(\frac{\Delta t}{\tau}\right)^2} e^{-\left(\frac{\Delta SST_{filtered}}{T}\right)^2} \quad (6)$$

where Δr, Δt and ΔSST refer to the spatial, temporal, and thermal distances, respectively; L, t, and T represent the spatial, temporal, and thermal decorrelation terms, respectively.

In the present algorithm, decorrelation and filtering values have been set as the same as those already used in Droghei et al. 2018, namely L = 500 km, τ = 7 days, and T = 2.75 K. To minimize the computation effort, in situ and satellite input data are sampled within a searching radius of 1500 km and a time window of 15 days and 3 days for in situ and satellite observations respectively. Finally, among satellite data, a priority is given to the SMAP passages with respect to those of SMOS in the filling of the input matrix.

2.2.3 2DMED SSS product characteristics

The final experimental 2DMED SSS time series has been processed and make available on a regular grid with a spatial resolution of 1/24° in both longitude and latitude, and a temporal sampling interval of 1 day. The spatial coverage spans from 6°W to 36°E, while the temporal coverage extends from 1st January 2016 to 31st December 2022.

The 2DMED SSS dataset (Sammartino and Buongiorno Nardelli, 2024) has been publicly released and is accessible in a netcdf format via the following Zenodo address: https://doi.org/10.5281/zenodo.13753090 .

The structure of each file is that reported in the following example:

Name	Long Name	Type
dataset-sss-ssd-MED-daily_20220101T1200Z_P20231229T0000Z.nc	Mediterranean Analysed Sea Surface Salinity and Density	Local File
depth	depth	—
dos	sea surface density	Geo2D
dos_error	sea surface density error	Geo2D
lat	latitude	1D
lon	longitude	1D
sos	sea surface salinity	Geo2D
sos_error	sea surface salinity error	Geo2D
time	time	—

2.3 WP2300 - Experimental physical variables 4DMED product development

The objective of this task is to develop an experimental 4DMED algorithm providing a 4D-reconstruction of key physical variables (temperature T, salinity S, zonal geostrophic current Ug, meridional geostrophic current Vg), that will be later used to provide the 4DMED physical product within WP3300. To this aim a 2-step approach is chosen. To provide the combined physical 4DMED experimental product, a 2-step approach is chosen. It consists firstly of a Machine Learning method to retrieve the temperature and salinity 4D fields with a learning based on the CMCC's 1/24° MEDSEA_MULTIYEAR_PHY_006_004 reanalysis (WP2310); and, secondly, of a merging with in situ observations to correct the residual large-scales biases of the first step (WP2320 and WP2330).

2.3.1 Input model data description and preparation

- The CMCC's 1/24° MEDSEA_MULTIYEAR_PHY_006_004 reanalysis (MEDSEA hereafter) will be used as a learning base in this project to reconstruct temperature and salinity from 0m to 300m on a daily basis. MEDSEA system uses the NEMO model for the ocean modeling part and assimilates in-situ temperature and salinity profiles as well as satellite altimetry tracks with the OceanVar 3DVar scheme. The model horizontal grid resolution is 1/24˚ (ca. 4-5 km) and is forced by hourly ECMWF ERA5 atmospheric forcing fields. More details on the CMCC's simulation can be found at https://doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1.
- The DT2021 DUACS L4 ADT (Absolute Dynamic Topography) or SLA (Sea Level anomaly) used to compare with MEDSEA corresponds to the Copernicus Service ¼°gridded and daily product SEALEVEL_GLO_PHY_L4_MY_008_047 (https://doi.org/10.48670/moi-00148).

Note that the model SSH is the sea surface height above the geoid and it corresponds to the absolute dynamic topography in altimetry.

A preliminary study over the period 2005-2020 has therefore been carried out to ensure that the reanalysis is not affected by problems that could be detrimental to Machine Learning.

MEDSEA accurately represents all spatio-temporal processes in terms of sea surface height and dynamic height (HDYN). The differences between MEDSEA height and DUACS altimetry were examined, and the following key points were identified:

- Absolute sea level from MEDSEA and DUACS compare well.
- Large scale and structure positioning may regionally differ. Adriatic and Aegean Sea should be discarded from the study because the agreement between the model and altimetry is poor in these two regions.
- High frequency (< 20 days) barotropic signal is absent from DUACS and dominating in MEDSEA. Small scale 70km noise is present in DUACS and absent in MEDSEA.

- Interannual in HDYN (MEDSEA) and SLA (DUACS) are fairly similar.
- Long term trends are very coherent.

The model data contains some numerical noise at very small scales, and small temporal jumps in certain regions at certain dates. This is expected from an assimilation system using a high-resolution time-splitting version of NEMO (https://www.nemo-ocean.eu/). This noise is sufficiently small and rare that it should not affect the learning method.

2.3.1.1 Spatial resolution

The model grid spacing is dx=0.04° i.e. 1/24°. No obvious anomalies or visible small-scale noise were identified in the MEDSEA SSH. The energy spectra at 34°N of the SSH and the meridional geostrophic velocity (Figure 14) are very regular, with a linear decrease in energy.

The Figure 15 shows the power spectrum of MEDSEA SSH and DUACS ADT. The power spectrum at 34°N reveal a similar behaviour, except DUACS has a flat spectrum (noise) for periods smaller than 70km. This flat tail is mostly due to the noise from the MDT (Mean Dynamic Topography) CNES-CLS-2019 and not the altimetry.

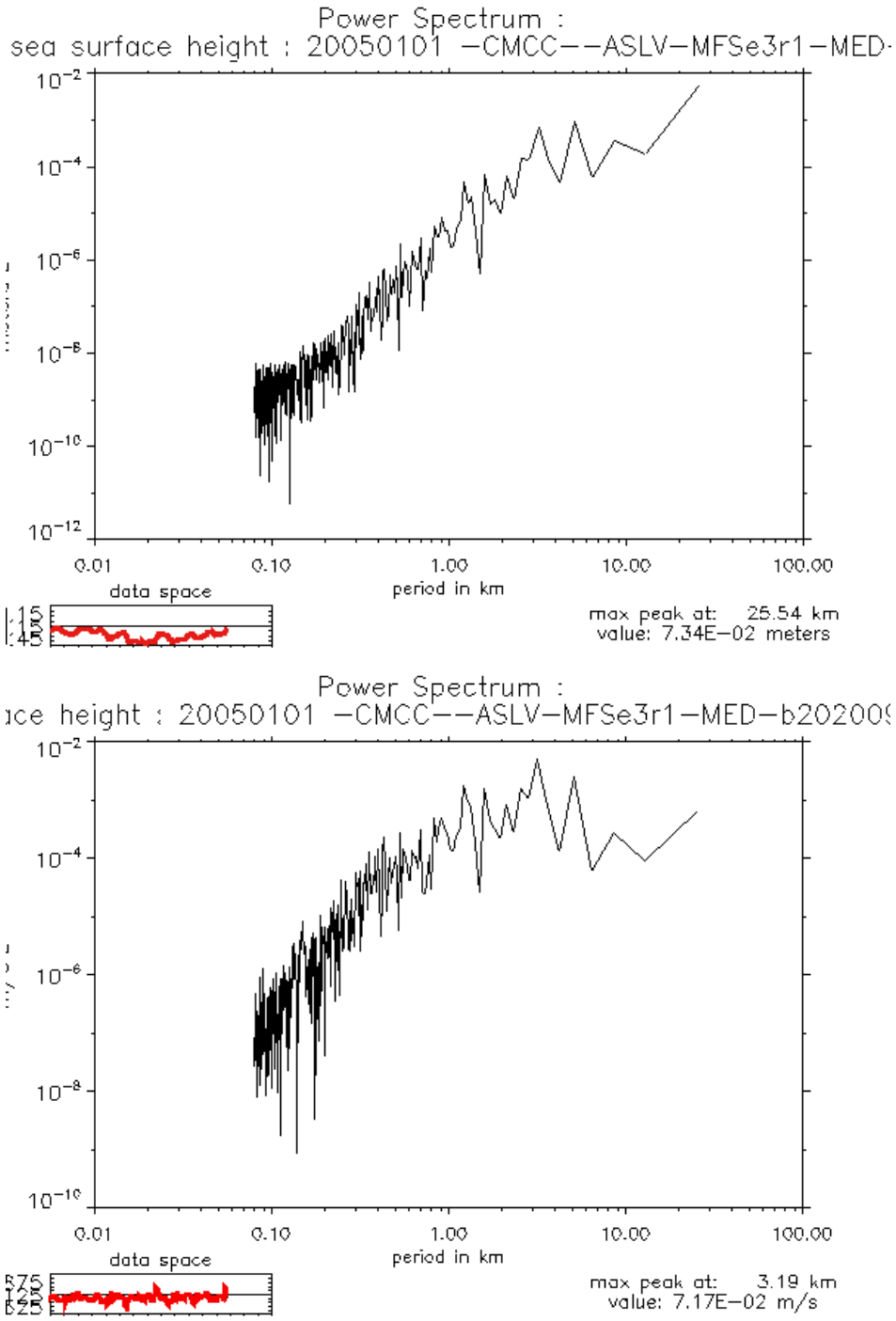


Figure 14: Power spectrum along 34°N for the 20050101 MEDSEA SSH (top) & meridional geostrophic velocity (bottom)

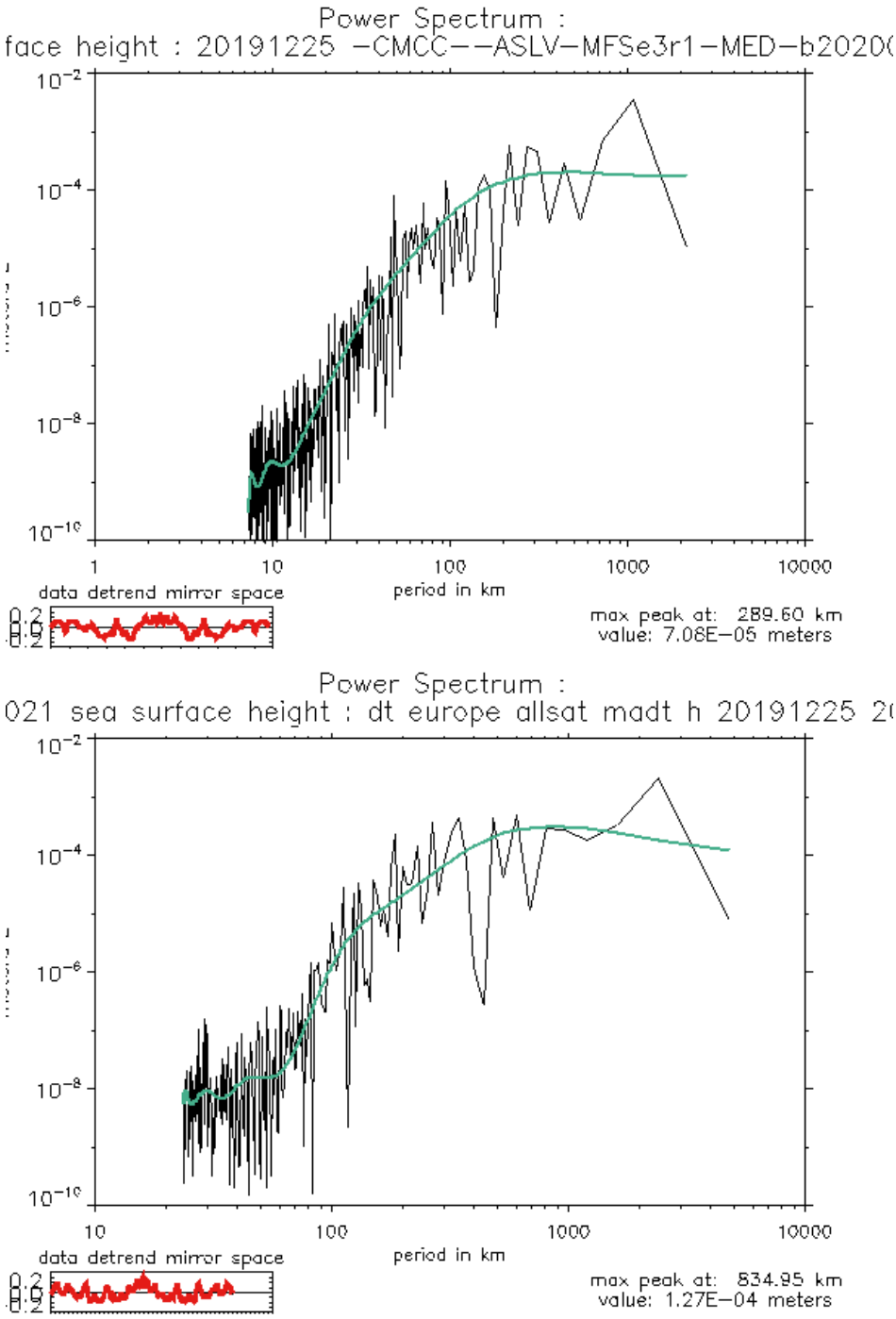


Figure 15: Power spectrum along 34°N of the MEDSEA SSH and DUACS ADT for the 25/12/2019

2.3.1.2 Spatial coherence

The skill pattern (computed as the correlation in 2° moving boxes for the 25/12/2019 - Figure 16) between the MEDSEA SSH and the DUACS ADT is a good indicator of the positioning of the mesoscale structures (it is not too sensitive to large scale MDT or barotropic biases). The result is satisfying in most of the domain. The skill is rather low in the Adriatic or in the Aegean Sea. Similar results are obtained with various dates and seasons (not shown).

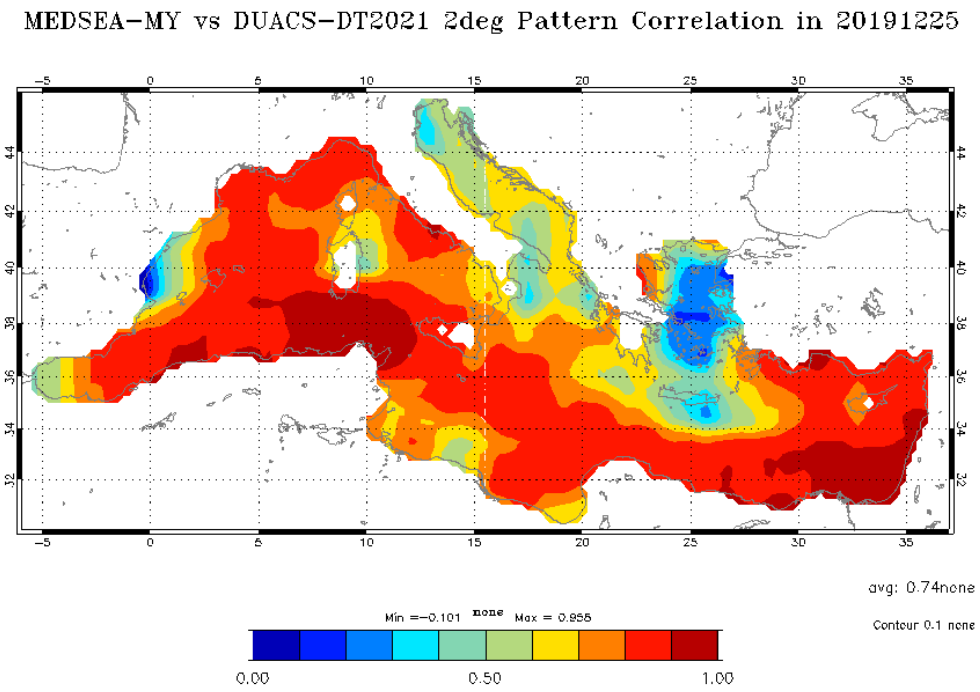


Figure 16: SSH MEDSEA vs ADT DUACS-DT2021 skill pattern for the 25/12/2019

2.3.1.3 Space-time signal

The analysis of Extended Orthogonal Functions computed in MEDSEA SSH reveals that the signal is dominated by the fairly high-frequency HBAR (barotropic signal), which is typical of a strong response to the wind in semi enclosed seas (see the time series icon at bottom right, Figure 17). The signal is separated into 3 sub-basins, as often noticed.

The main consequence of this barotropic signal in MEDSEA is that MEDSEA SSH and DUACS ADT are not consistent. Indeed, the barotropic signal from MEDSEA is annoying and must be eliminated. As we don't have the HBAR variable of the MEDSEA reanalysis, the dynamic height (HDYN) relative to the bottom will be used instead of the SSH in the machine learning reconstruction. The MEDSEA SSH variable "zos" should therefore not be used as it is polluted by the barotropic signal. The variable to be used as an absolute sea surface height value comparable with the DUACS ADT will be:

$$\text{proxy absolute height}(t) = t\_ssh + [\, hdyn(t) - t\_hdyn] \qquad \text{Equation 2.3.1}$$

where t_ssh is the mean of the MEDSEA SSH over 2005-2019, and t_hdyn is the 2005-2019 average of the MEDSEA's hdyn(t) [time average over the study period].

The seasonal signal (the annual cycle) is the dominant signal with an amplitude of 20 cm (not shown). It corresponds mainly to a steric signal with heating of the western MED and the Levantine basin. This signal is well reproduced by MEDSEA. Conversely, as shown in Figure 18, the cumulative trend is weak and unorganised, which is somewhat surprising. The Mediterranean is known to have risen by more than 3mm/year over the

last 20 years (Calafat & al., 2022), but, on average over the basin, MEDSEA shows a decrease of -0.7mm/year. This clear discrepancy is probably linked to the nature of the "SSH" field provided: the reduction in the flow of the Nile and other effects of evaporation probably lead to a loss of mass. The steric water height would therefore be missing (quite naturally for this Boussinesq model).

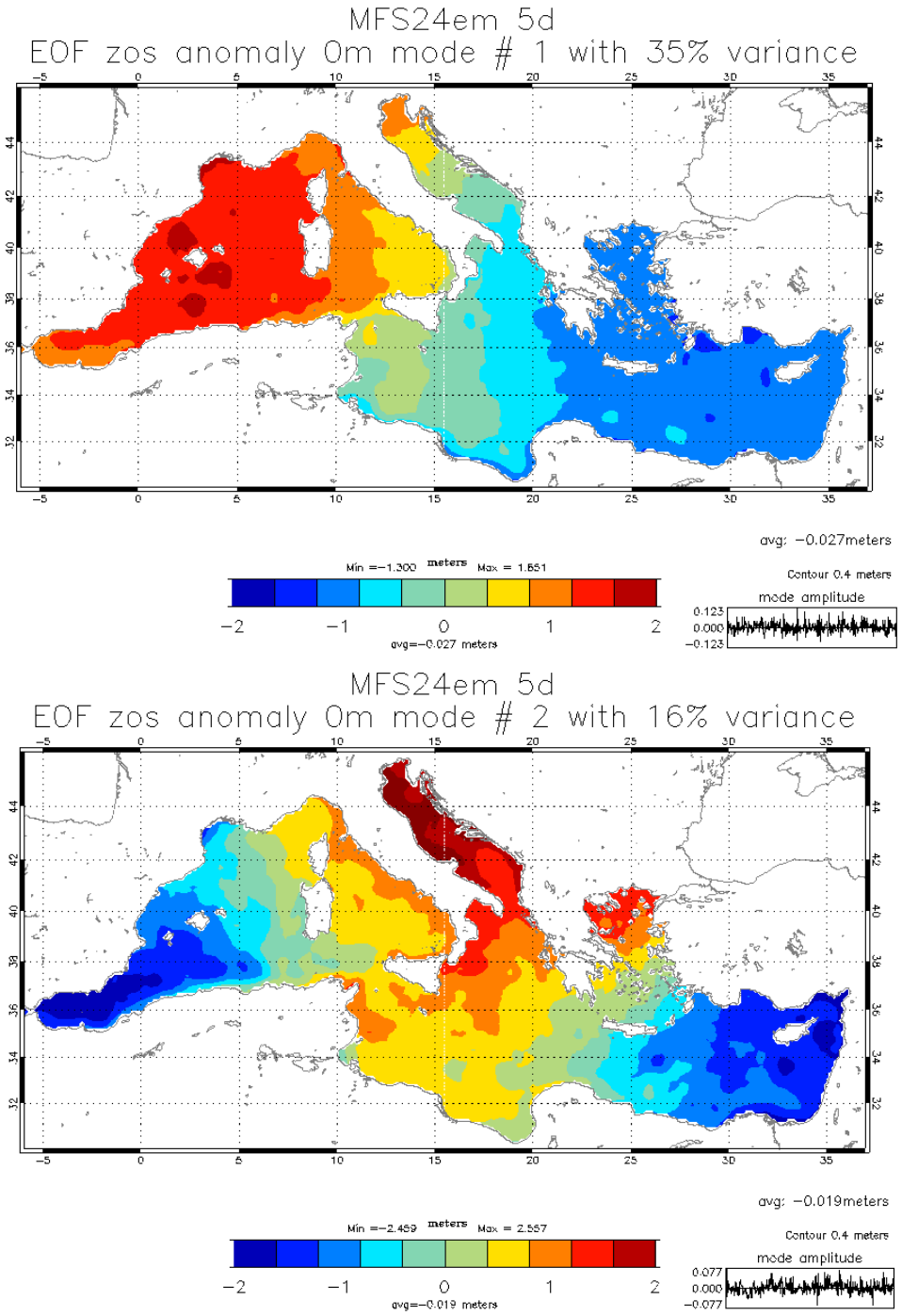


Figure 17: EOFs 1 (top) and 2 (bottom) in MEDSEA SSH over the period 2005-2020

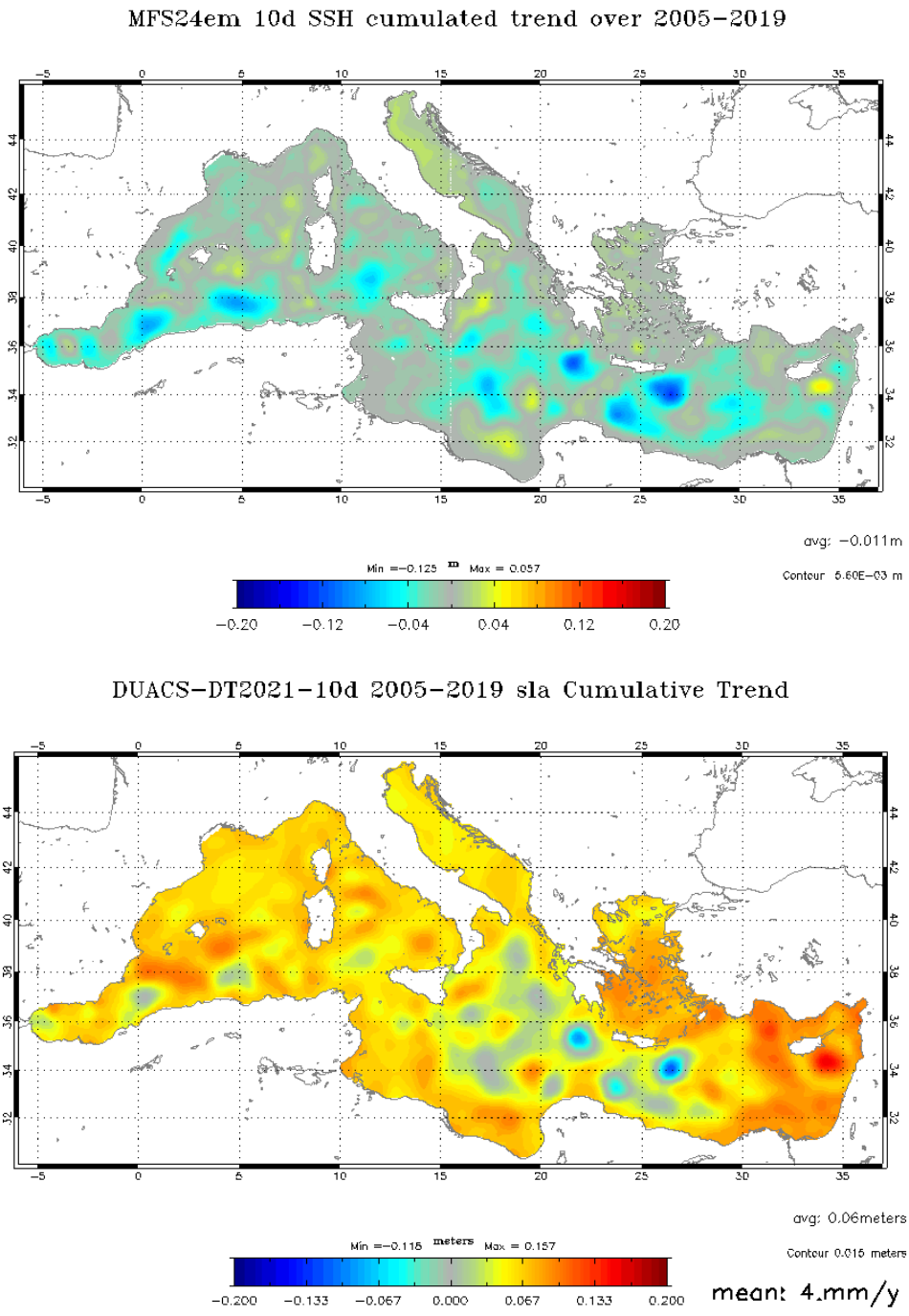


Figure 18: Sea level cumulative trend over 2005-2019 in MEDSEA (top) and DUACS (bottom)

The DUACS sea level trend is 0.5 mm/year higher than the MEDSEA trend. If we correct DUACS for this effect (steric basin effect), the agreement is really good in terms of trends with the exception of the Adriatic and Aegean Seas (Figure 19).

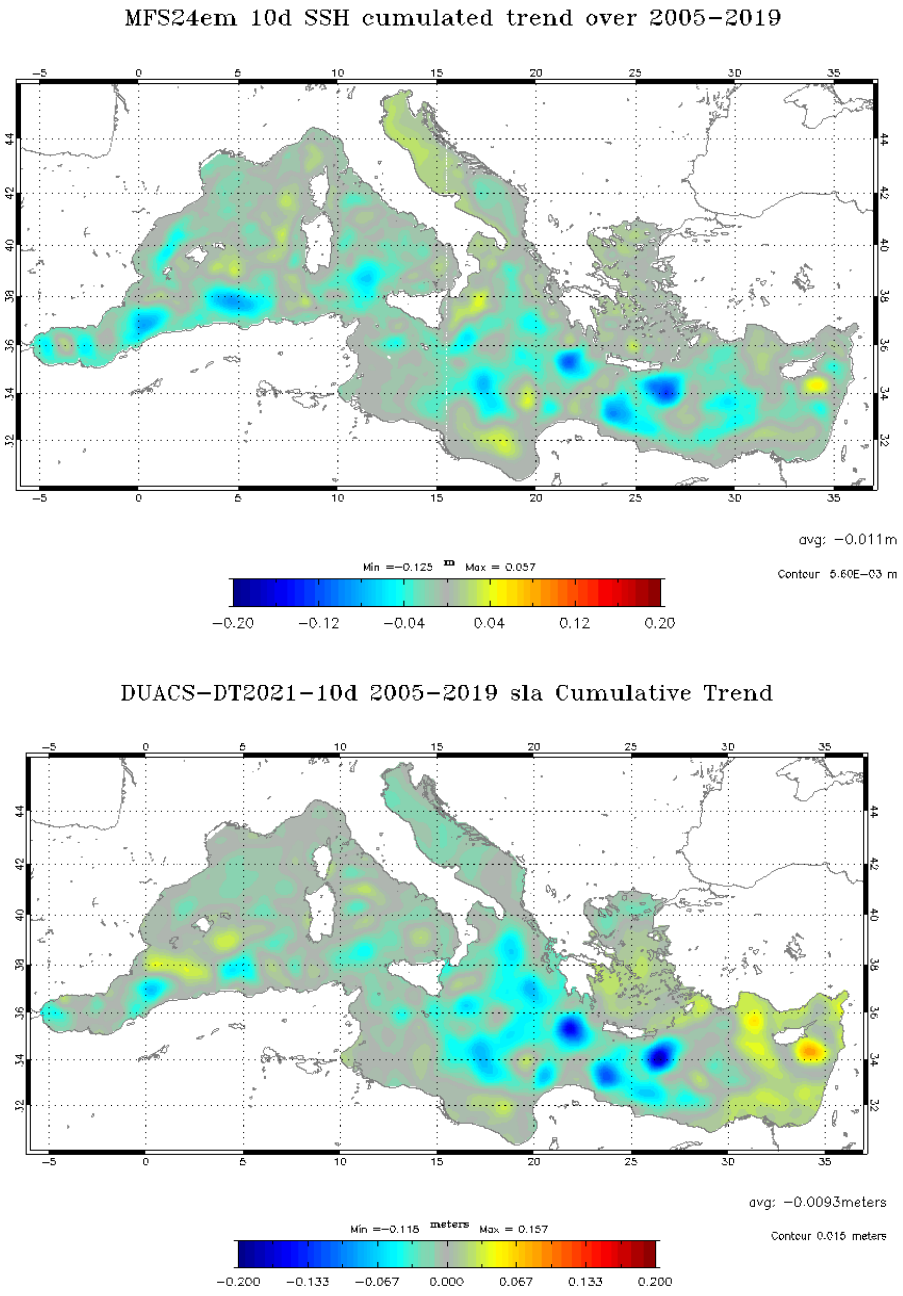


Figure 19: Sea level cumulative trend over 2005-2019 in MEDSEA (top) and DUACS corrected of the steric basin effect (bottom)

2.3.1.4 Interannual signal

Examination of the inter-annual and low-frequency signals in MEDSEA SSH did not reveal any anomalies. The low-frequency variability is familiar with well -known patterns such as the ones of the Algerian current, Ierapetra gyre, … The time series of the low-frequency EOFs indicates that the climatic signals are numerous and more chaotic than in the equatorial band or at high latitudes.

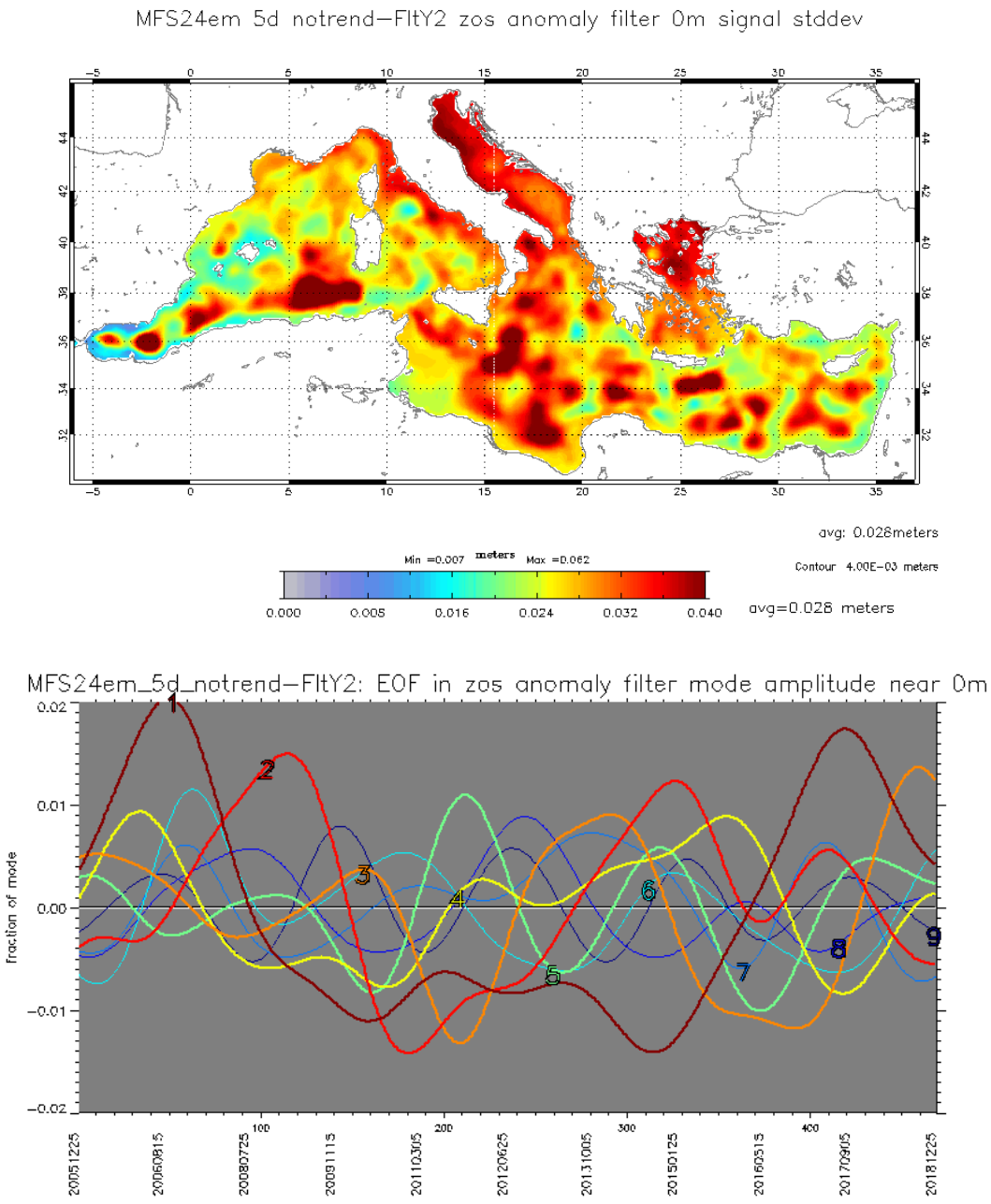


Figure 20: Standard deviation of the low-frequency & detrended MEDSEA SSH (top) and times series of the low frequency EOFs of the low-frequency & detrended MEDSEA SSH (bottom) over the period 2005-2018

2.3.1.5 Trends in steric height

The NEMO model is not sensitive to the basin average of the steric signal (Boussinesq assumption). Hence, the trend in SSH only reflects the mass and the barotropic trend. The steric height must be computed from the 3D temperature and salinity fields (here we used a linear formula verifying HDYN=HDYNT+HDYNS). The Figure 21 shows the cumulative trend in steric height.

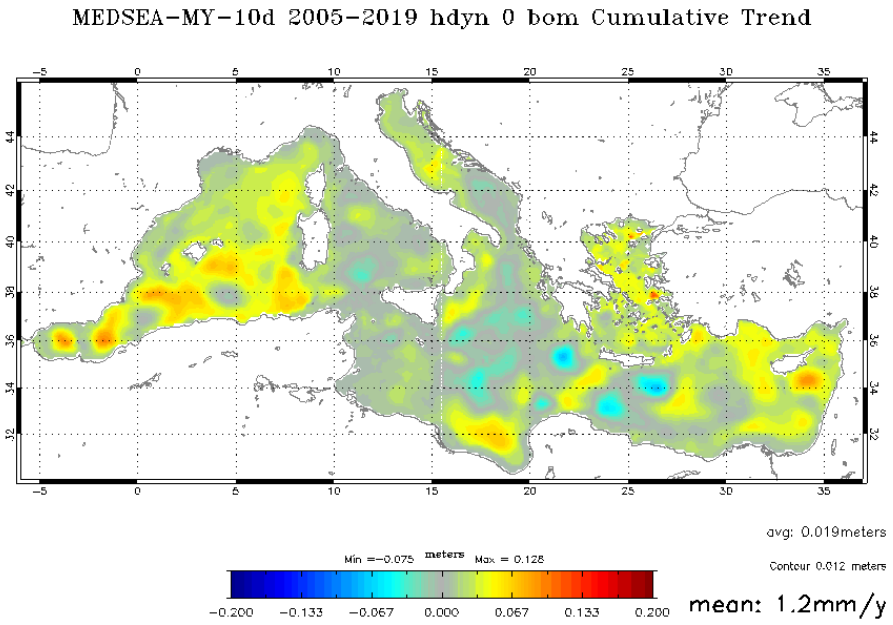


Figure 21: Cumulative trend of the computed MEDSEA dynamic height from 0m to bottom and over 2005-2019

The pattern is similar to the MEDSEA SSH trend. The Figure 22 reveals the decomposition into thermo and halo steric parts.

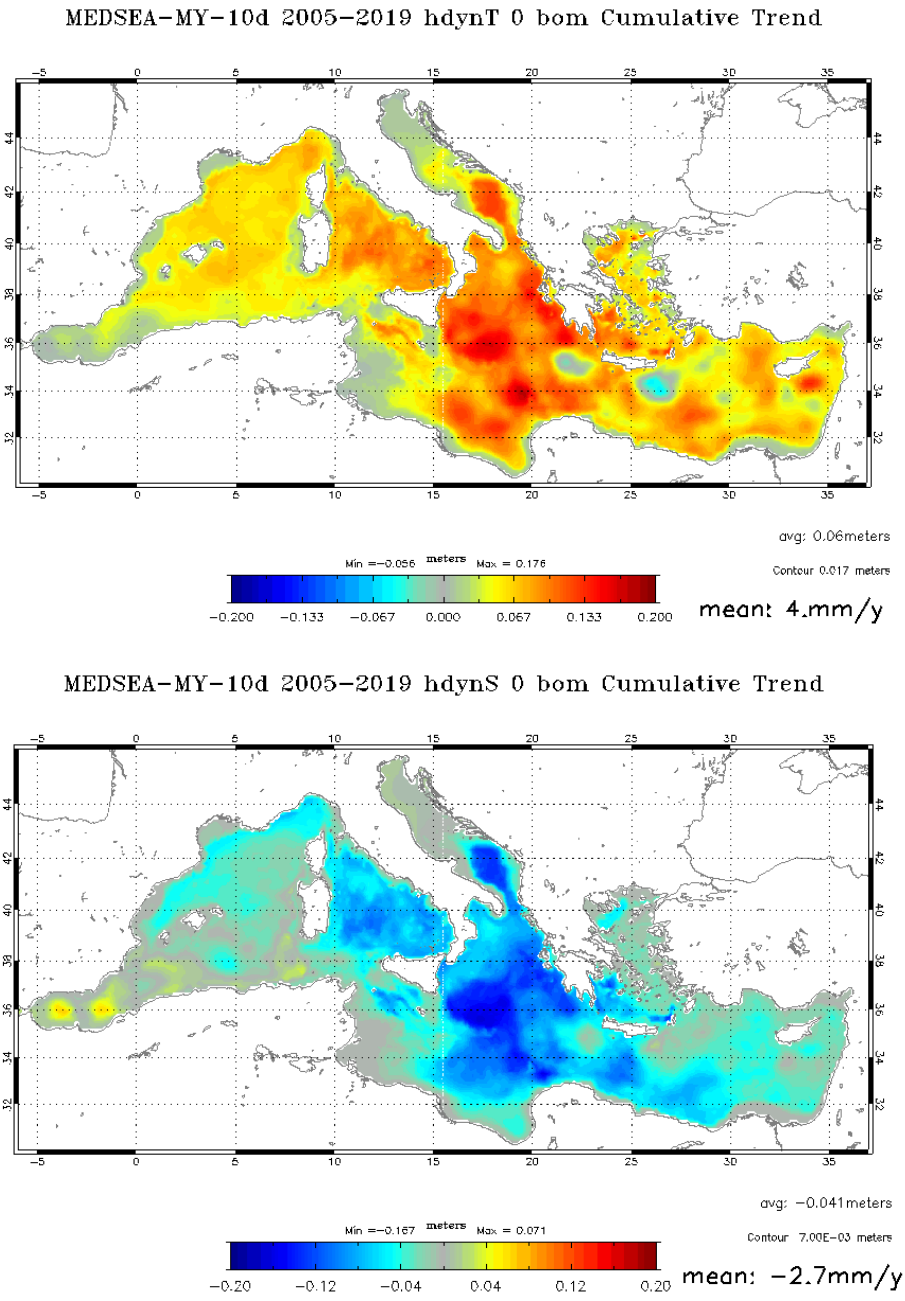


Figure 22: Thermosteric and halosteric contributions to the total steric signal for MEDSEA over 2005-2019

MEDSEA indicates a large warming and salinification, both effects leading to a moderate steric increase (1.2mm/year over 2005-2019, Figure 21). This has been noticed by Aydogdu et al. [2023], but it's not a new phenomenon (Millot et al., 2006). The deep waters of the western Mediterranean Sea have become saltier and warmer for at least the past 40 years at rates of about 0.015 psu and 0.04 °C per decade. It has been reinforced after 2013, the last year of deep convection (Margirier et al., 2020). Besides the many "short" events (Pingyang & Toste, 2020) under forcings from the North Atlantic Oscillation (NAO) and Atlantic Multidecadal Oscillation (AMO), reduced solar activity due to volcanic eruptions, the recovery from the Eastern Mediterranean Transient (Incarbona et al., 2016) is the more visible in MEDSEA trend. In short, the trends from MEDSEA seems very reasonable.

2.3.1.6 Empirical Orthogonal Functions

Various decompositions into Empirical Orthogonal Functions (EOF) show a rather similar behaviour between MEDSEA HDYN and DUACS SLA. This is illustrated below by the comparison of the 10-day time series (subsampled time series with one date every 10 days), filtered from the seasonal signal.

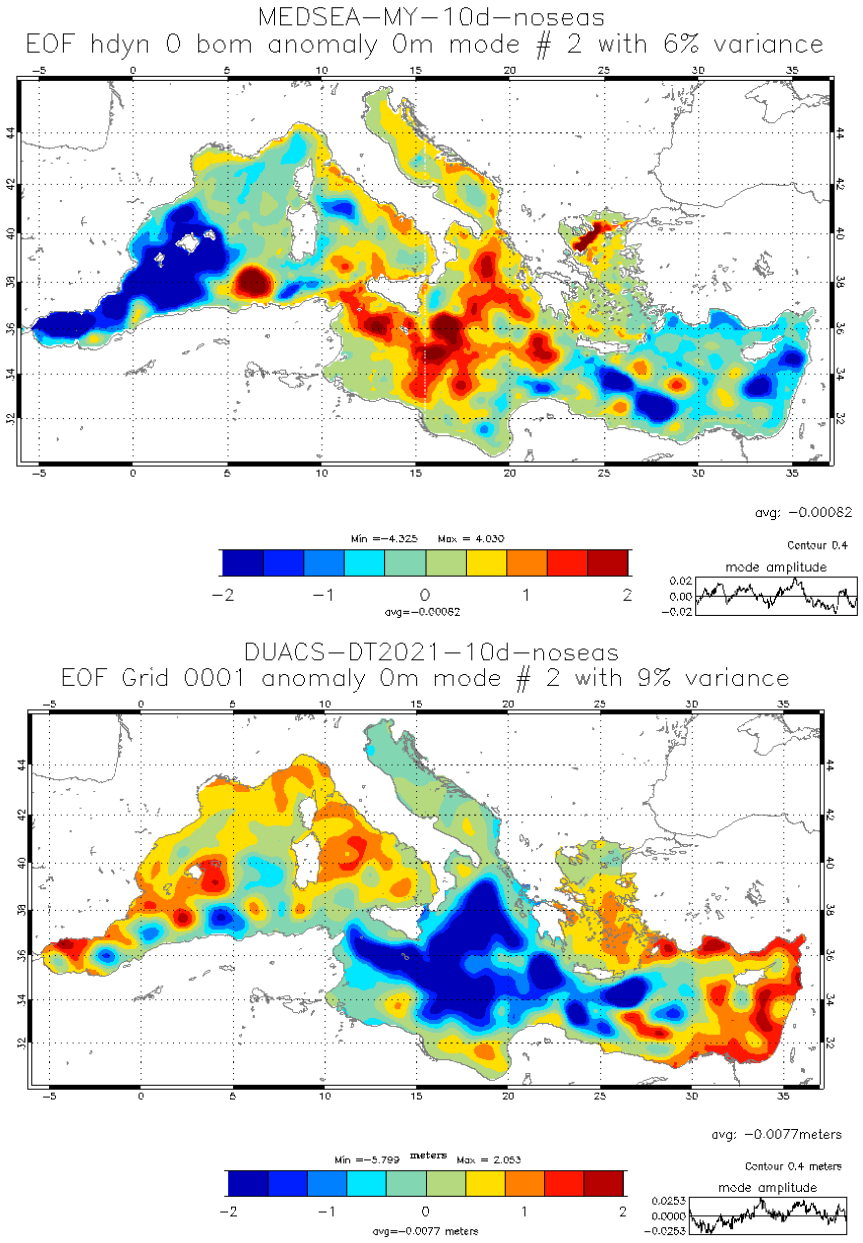


Figure 23- EOF#2 of the MEDSEA hdyn (top) and DUACS SLA (bottom), both averaged over periods of 10 days and without the seasonal cycle.

The first 5 modes are fairly similar in terms of time amplitude (icon at the bottom right of each figure) and patterns. More high frequency noise is observed in DUACS, probably due to the DAC+IB filtering near 20 days. The Figure 23 presents the second EOFs, which are interesting because it clarifies the relation between the 3 Mediterranean sub-basin (center, the west and the east). The trend is most marked in the central zone. Besides

the arbitrary sign of the EOFs, there is a good correspondence in term of size and positioning of the inter annual signal. This last result confirms MEDSEA's ability to accurately represent all spatio-temporal processes in terms of sea surface height and dynamic height.

2.3.1.7 Temperature

Examination of the temperature at various levels does not reveal any anomalies. Only two examples are shown below. At the surface (Figure 24, top), the seasonal cycle is clear and fairly regular (little affected by global warming). At 200m (Figure 24, bottom), the phase (maximum of the seasonal cycle) clearly separates the coastal current from areas of deep or modal water formation (Gulf of Lion, Tyrrhenian Sea). This accurately reflects the various physical processes at work in the basin.

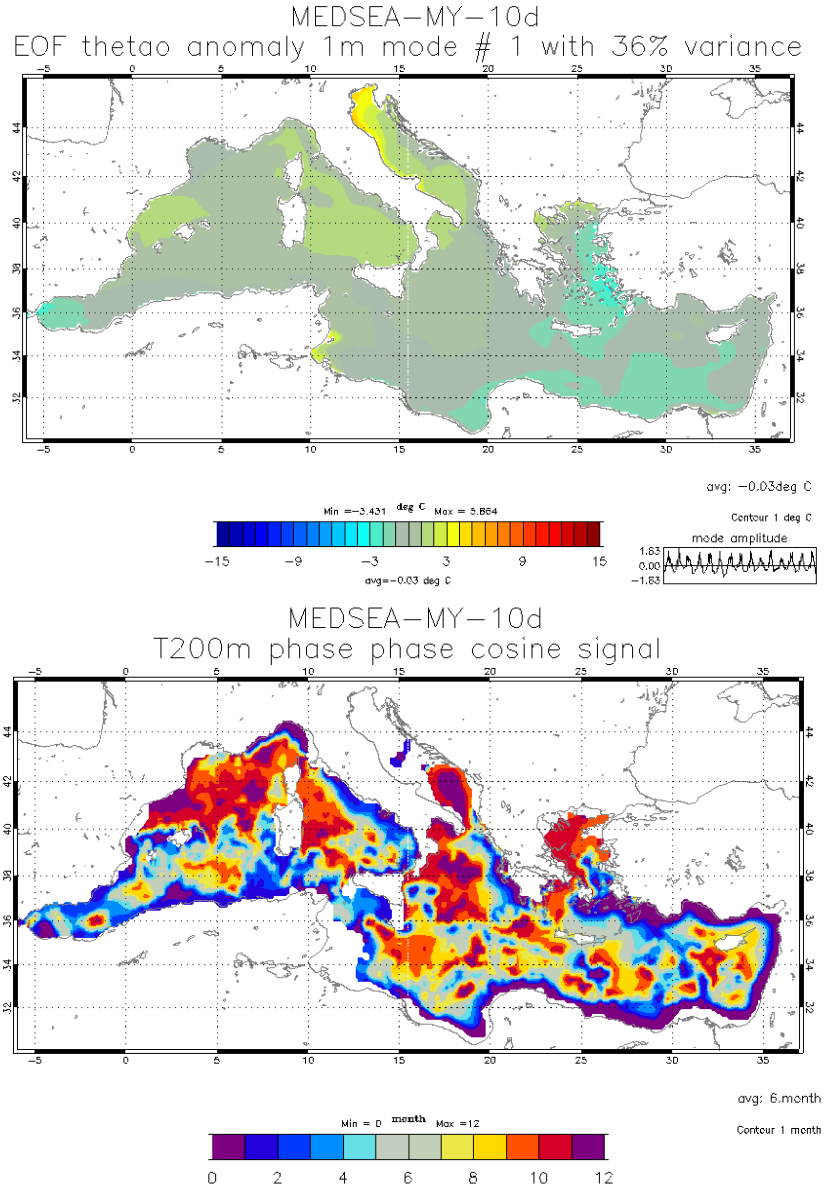


Figure 24: EOF#1 of the MEDSEA temperature at 1m (top) and phase of the maximum of the seasonal cycle for the MEDSEA temperature at 200m-deph

2.3.2 Machine learning model

To estimate in-depth temperatures and salinities in the Mediterranean Sea, two different approaches were explored. The first one is based on the CAREHeat project and consists of training multiple small models, i.e. one model every 0.25 degree. It is a very local approach, allowing the use of very small models like Random Forest, LightGBM or 1-layer Perceptron. This is a simple and effective way of dealing with local climatic conditions and biases. Based on CAREHeat results, the latter was chosen for this project.

The second approach is more global and consists of training only one model for the whole Mediterranean Sea. The motivation behind it is the possibility to use Deep Learning models that have been proven very successful in computer vision such as Convolutional Neural Networks (CNN). They have the ability to extract spatial structures thanks to their convolutional and pooling layers.

For both approaches, the models were trained on MEDSEA data, using the same temporal splits to be able to compare them. Then, one was chosen to infer in-depth temperatures and salinities on (1) unseen input MEDSEA data (test set) and (2) with observed SSS, SST and SLA fields measured by satellites. The transfer from MEDSEA to satellite data will be detailed.

2.3.2.1 Input data

For the training, the following variables are used:

- Input variables: MEDSEA Dynamic height (HDYN, see Equation 2.3.1), MEDSEA Sea Surface Temperature (SST), MEDSEA Sea Surface Salinity (SSS), cos(day of the year), sin(day of the year), latitude, longitude. The HDYN values are computed from the temperature and salinity 4D fields of the MEDSEA reanalysis from surface to bottom.
- Target: MEDSEA T(z), S(z) over 0-150m

For the production (inference): the Absolute Dynamic Topography (ADT) from the experimental 2DMED ALT product, the SSS from the 2DMED SSS product and the Mediterranean Ultra-High Sea Surface Temperature (Copernicus Marine Service product: SST_MED_SST_L4_NRT_OBSERVATIONS_010_004, https://doi.org/10.48670/moi-00172) fields are used to reconstruct the temperature and salinity profiles up to 150-m depth.

The SST_MED_SST_L4_NRT_OBSERVATIONS_010_004 is a regional daily gap-free (L4) product at 0.01° of spatial resolution available from 2008 to 2023.

2.3.2.2 Machine Learning model

Several models will be tested. The first one is a legacy of the ESA-CareHEAT project and it is based on a Multi-Layer Perceptron (MLP) which is a simple fast-forward neural network with only one hidden layer of 100 neurons. This approach consists of having one model/set of parameters per point.

The second approach is to use a more global model, i.e. one set of parameters for the whole Mediterranean Sea. Here, models such as MLP and Convolutional Neural Network

can be considered. The latter is widely used in image processing because of its ability to find patterns through spatial correlations.

2.3.2.3 Training/inference

First, the model is trained on MEDSEA to best fit the training data. Once this is done, the trained model can be directly used for inference, i.e. predict new temperature and salinity profiles on (1) unseen input MEDSEA data (test set) and (2) with observed SSS, SST and ADT fields to produce the final T & S profiles.

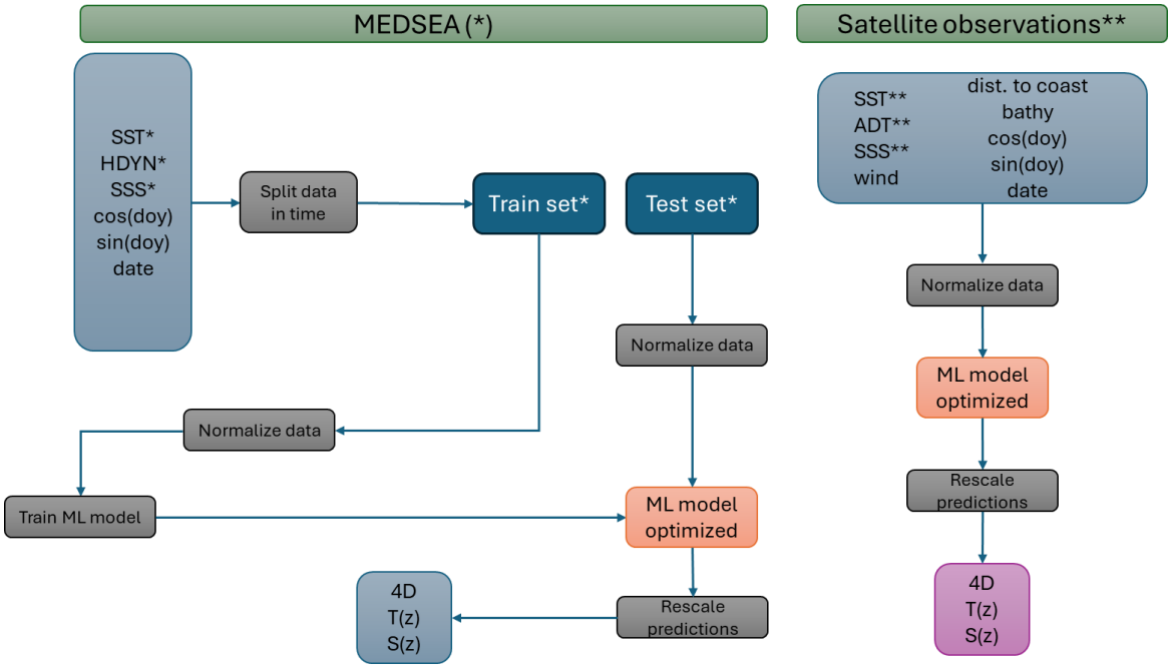


*Figure 25: Schematic view of the model training from MEDSEA and final production of the experimental 4D dataset from satellite observations. The * indicates the variables from MEDSEA; the ** indicates the variables from satellite observations.*

2.3.2.4 Training procedure

The same temporal split was used for both approaches. Since we are to produce temperature and salinity fields from 2016 to 2022, all these years were left out and kept as a test set. Among the remaining years (2005 - 2015), 2 years were kept as a validation set (2008 and 2014). This validation set is very important since it is the one that will be used to monitor and prevent overfitting.

Figure 26: Temporal split used for training. Blue: Training years, Orange: Validation years, Green: Test years

For the local approach, models were trained on CPU. To speed up training, several models were trained in parallel.

For the global approach, the use of a GPU was necessary, and MEDSEA data was downgraded to 1/8° to accelerate the training. In total, it took approximately 12 hours.

2.3.2.5 Models

Local approach

Based on CAREHeat project, the model used for each point is a one-layer neural network, with 100 neurons.

Global approach

Several architectures were tested and the one that gave the best results is a MobileNet (Howard et al, 2017). MobileNets are light weight deep convolutional neural networks thanks to their depthwise convolutions, allowing to reduce the number of parameters and therefore speed up calculations. The following table shows the architecture of the first MobileNet published by Howard et al (2017).

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5\times$ Conv dw / s1	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
$5\times$ Conv / s1	$1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

Figure 27: MobileNet architecture (Howard et al, 2017)

Overall, the model used has 3.1M parameters.

2.3.2.6 Comparison of the 2 ML model approaches

The figure below shows the temperature and salinity RMSE profiles for both approaches, calculated on the test set.

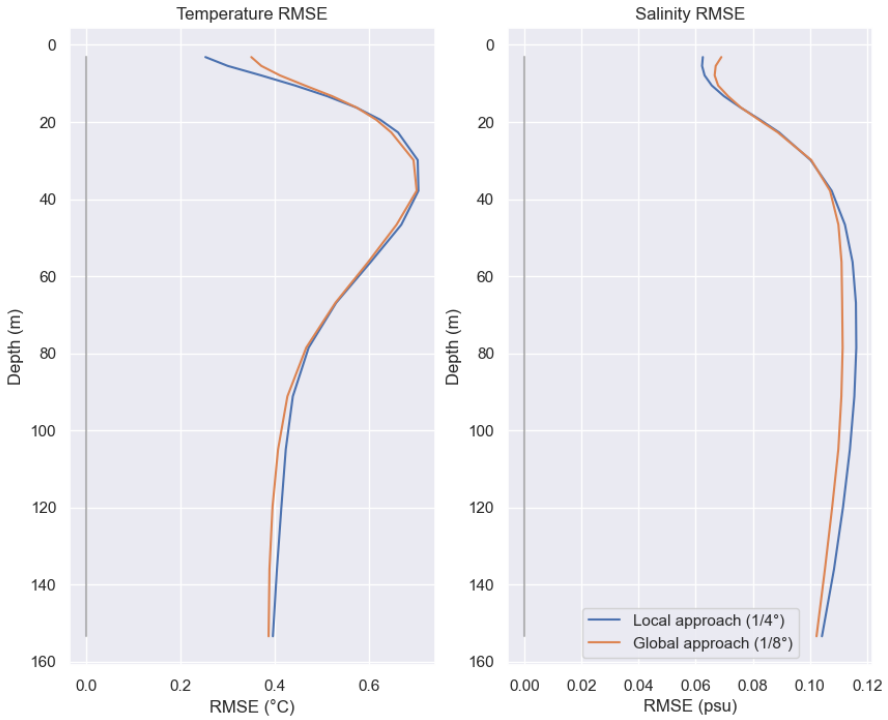


Figure 28: Temperature (left) and Salinity (right) RMSE profiles for the local (blue) and global (orange) approaches, calculated on the test set.

Both approaches give overall very similar performances. Note that the CNN was trained on 1/8° MEDSEA data, whereas the local approach had one model every 1/4°. But still, the CNN performs as well.

Although the CNN training time is higher than the local approach, having only one model to cover the whole Mediterranean Sea is very convenient since inference is much faster. For this reason, the global approach was preferred for the rest of the study.

2.3.2.7 Inference at 1/24° on MEDSEA test set

The CNN model (global approach) was trained on 1/8° MEDSEA data to speed up the learning process. However, the resolution of the final product is 1/24°. We first tried to directly use the model trained on 1/8° MEDSEA to infer on 1/24° MEDSEA test set, but it led to a drop in performances. Therefore, we finetuned on 1/24° MEDSEA data the CNN which was first trained on 1/8°. This finetuning process took another 19 hours. The results are shown below. Without finetuning, the model does not perform as well on 1/24° data than 1/8°, but the performances become similar or even slightly better at 1/24° after finetuning. This two steps approachs is particularly efficient in salinity. This type of multi-grid method is often used in data assimilation.

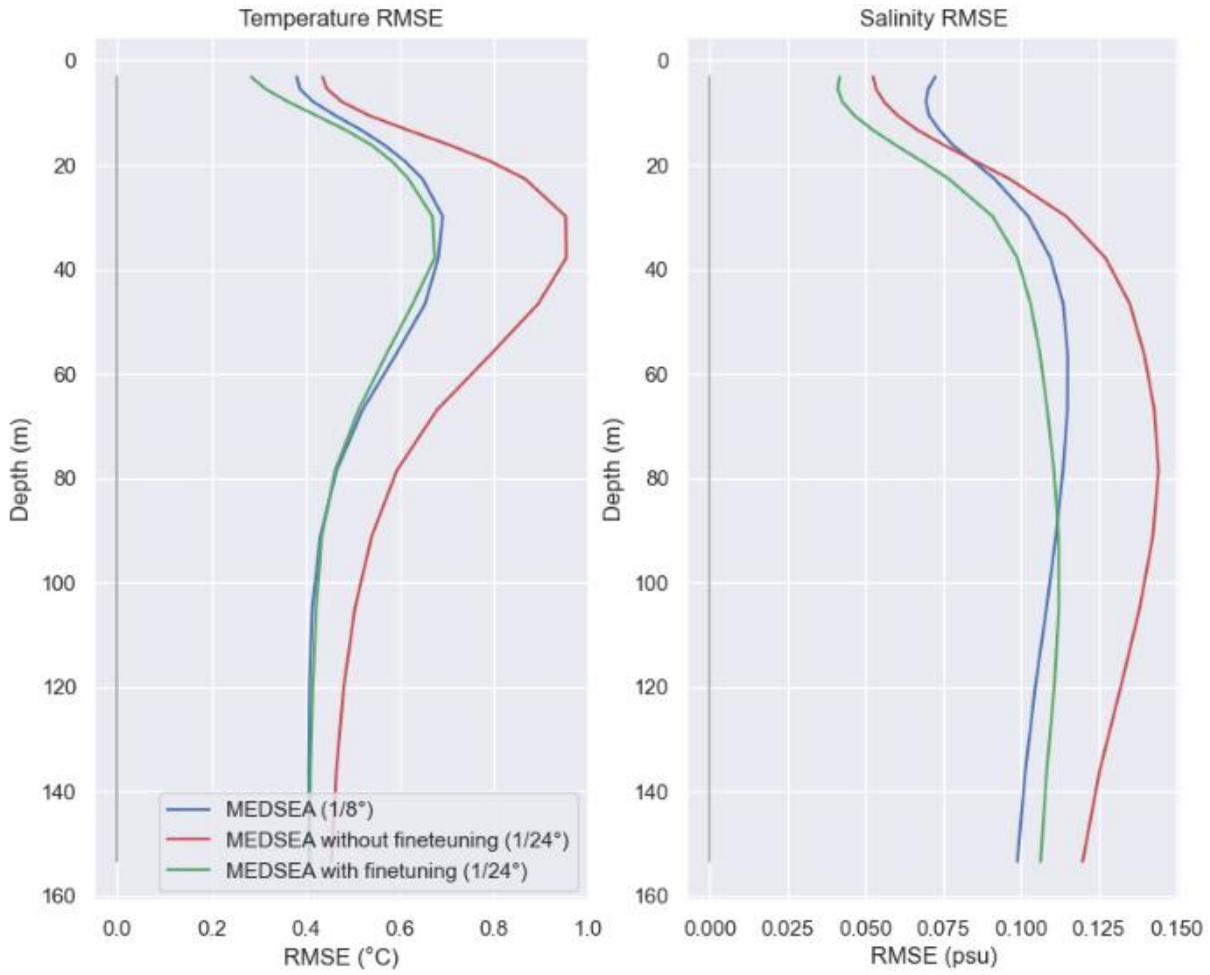


Figure 29: Temperature and salinity profiles in 3 scenarios. Blue: Model trained on 1/8° MEDSEA, and inferred on 1/8° MEDSEA test set. Red: model trained on 1/8° MEDSEA, and inferred directly on 1/24° MEDSEA test set. Green: model trained on 1/8° MEDSEA then finetuned on 1/24° MEDSEA, and inferred on 1/24 MEDSEA test set.

Below are represented the 2D RMSE maps at several depths of the finetuned model. Maximum errors are found as expected in regions of strong variability (Alboran, Algerian current, North Balearic Font...).

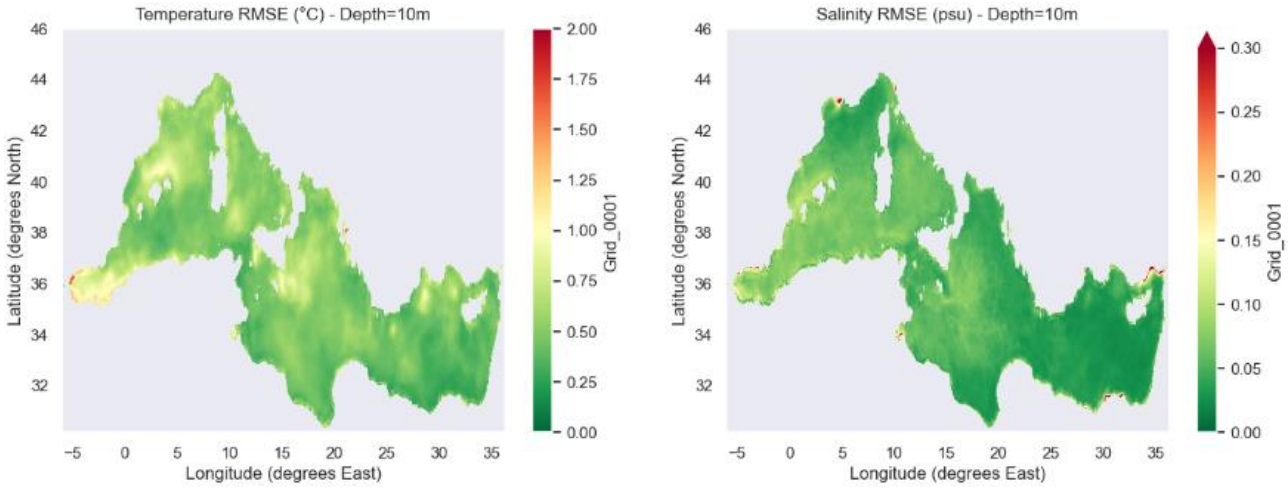


Figure 30: RMSE maps at 10m depth calculated on the test set (2016-2021)

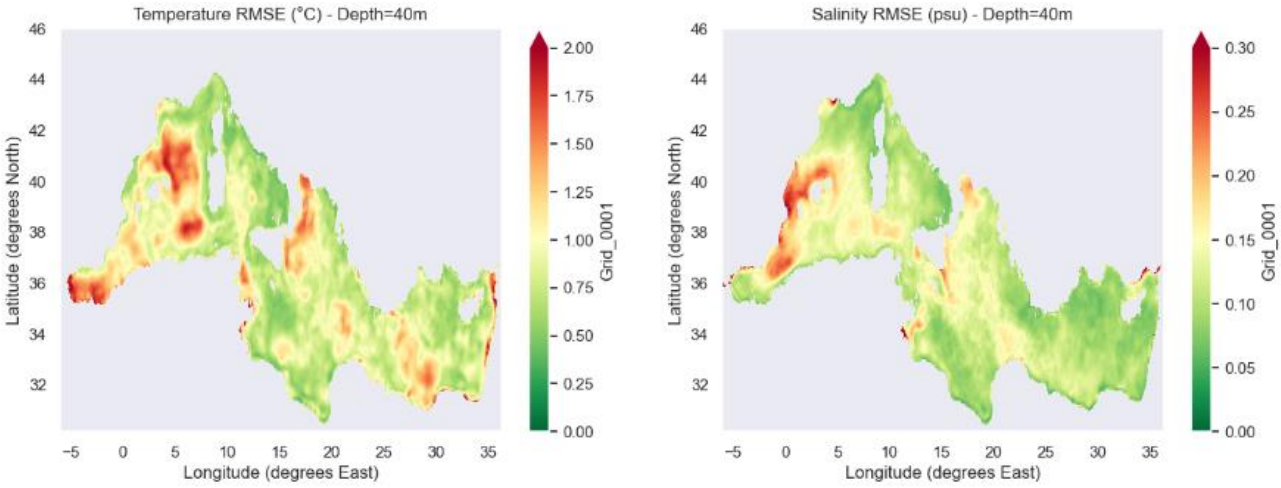


Figure 31 : RMSE maps at 40m depth calculated on the test set (2016-2021)

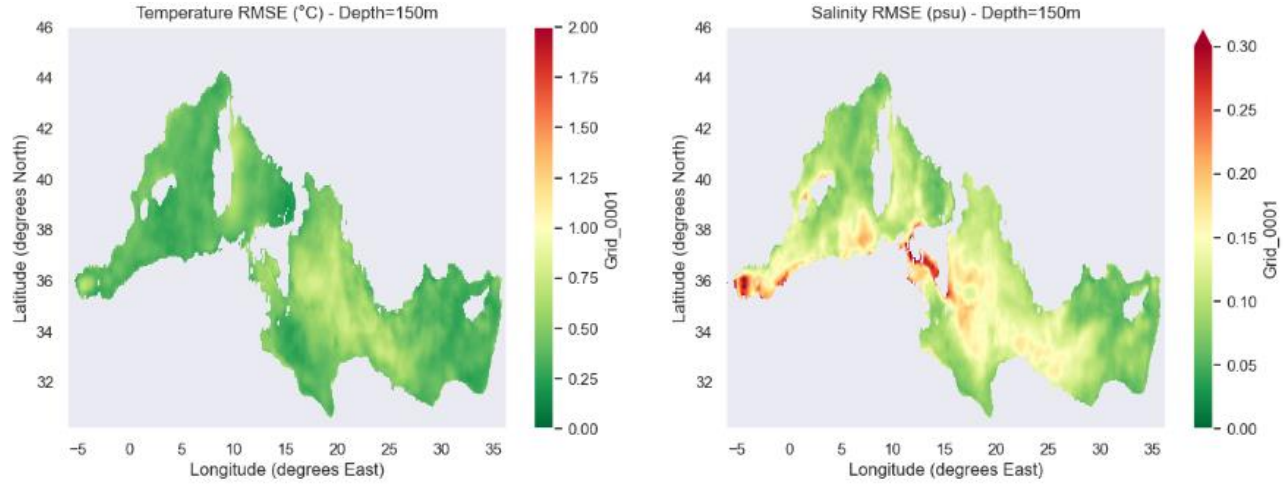


Figure 32: RMSE maps at 150m depth calculated on the test set (2016-2021)

Some higher error points can be seen along the coast, this is due to edge effects in convolution operations. This problem was tackled later and is presented in the next section.

Here are a few illustrations of temperature and salinity fields produced at a given date (in the test set). Inference is performed with the surface level from MEDSEA, mimicking satellite observations.

Temperature

At 3m, the CNN DL can keep most of the mesoscale information coming from the surface which is important to be consistent with satellite observations.

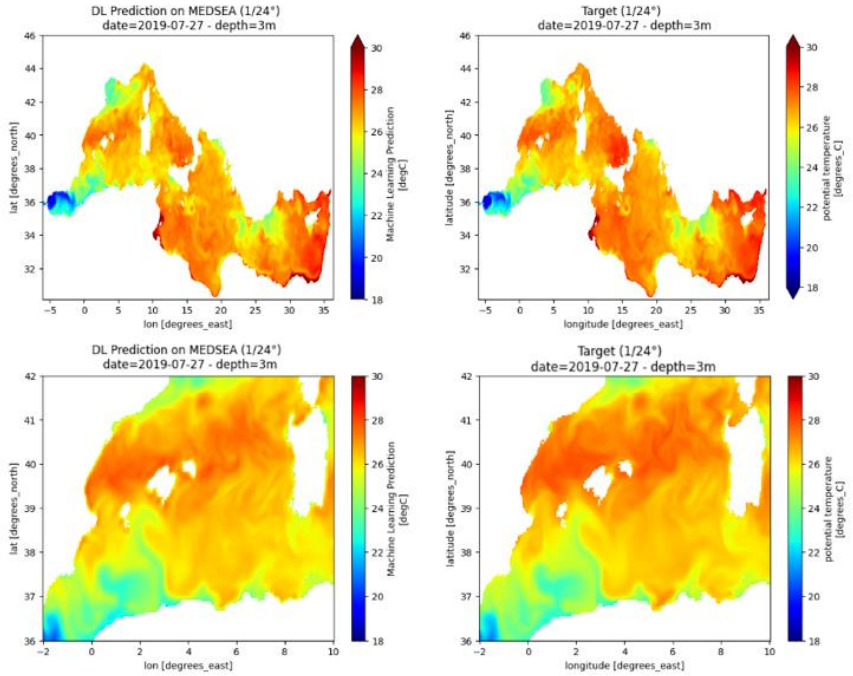


Figure 33: Visualization of the temperature CNN DL prediction on 2019-07-27 and the corresponding target, at 3m depth. The bottom row is a zoom on a specific region

At 40m, which is often the depth at the base of the mixed layer, the CNN is able to maintain good spatial coherence with good resolution, while not exhibiting the noise that could come from the satellite SST.

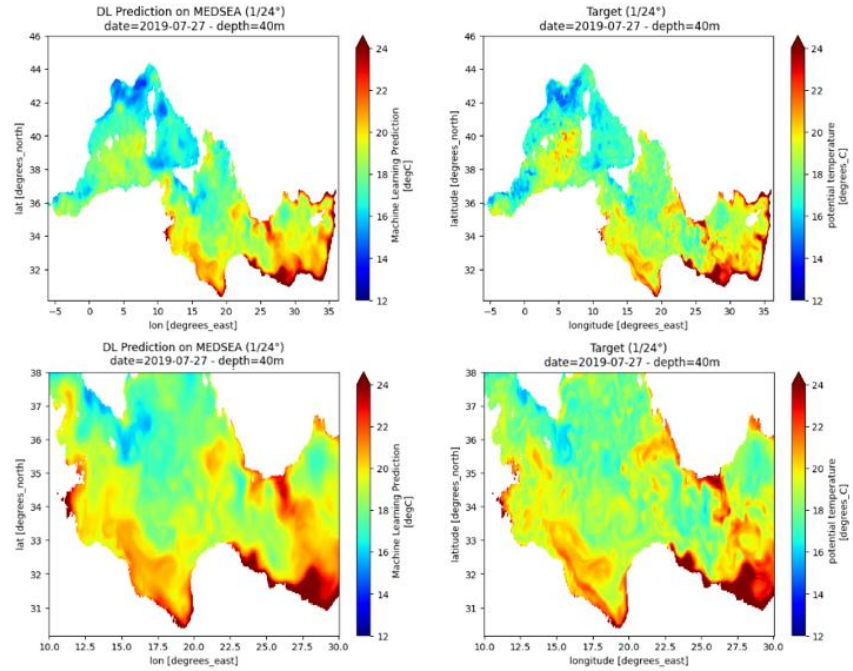


Figure 34: Visualization of the temperature CNN DL prediction on 2019-07-27 and the corresponding target, at 40m depth.

At 150m, the impact of sea level data becomes dominant and spatial resolution decreases compared with the subsurface.

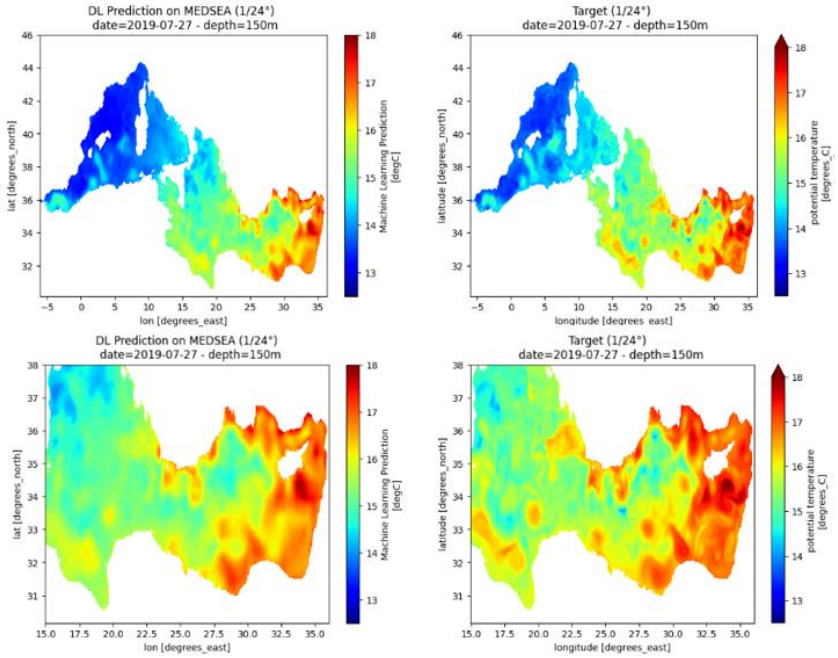


Figure 35: Visualization of the temperature CNN DL prediction on 2019-07-27 and the corresponding target, at 150m depth.

Salinity

Salinity behaviour is very similar to temperature behaviour. Close to the surface, the salinity of the CNN is very similar in structure to the SSS data, while below 50m, the salinity of the CNN is closer to sea level data. This is an expected result given that surface observations are mainly linked to the second baroclinic mode trapped at the surface, whereas sea level data allows other modes to be observed, in particular the first barocline mode.

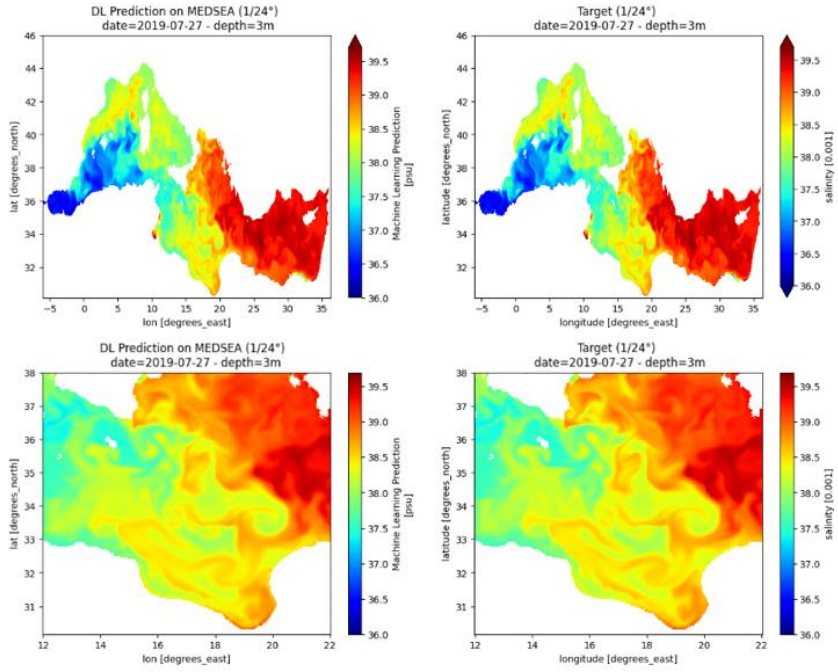


Figure 36: Visualization of the salinity CNN DL prediction on 2019-07-27 and the corresponding target, at 3m depth.

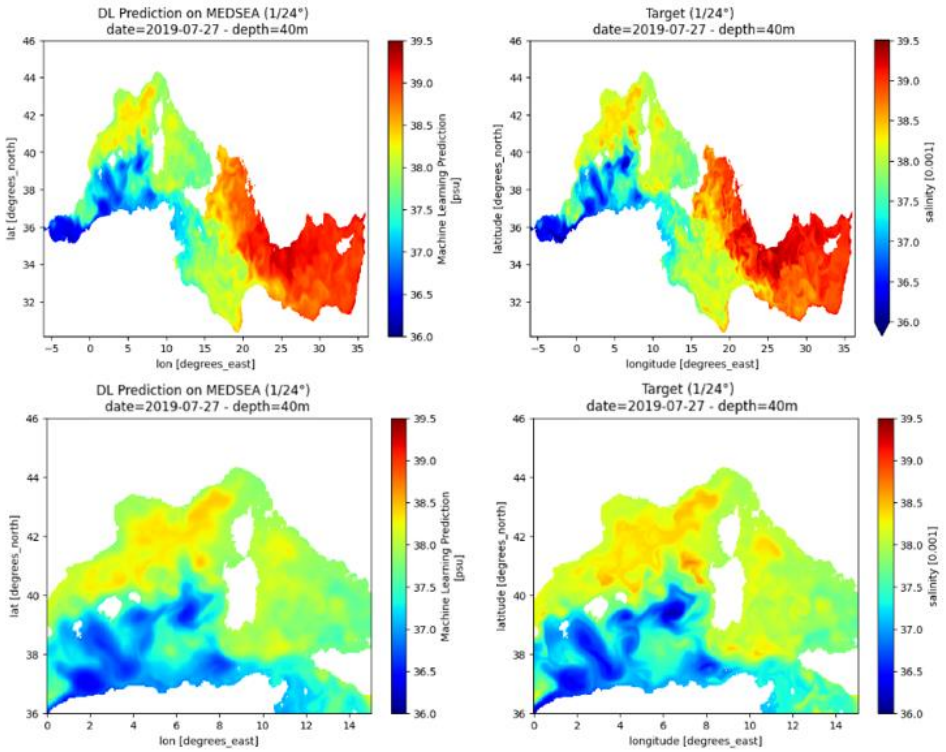


Figure 37: Visualization of the salinity CNN DL prediction on 2019-07-27 and the corresponding target, at 40m depth.

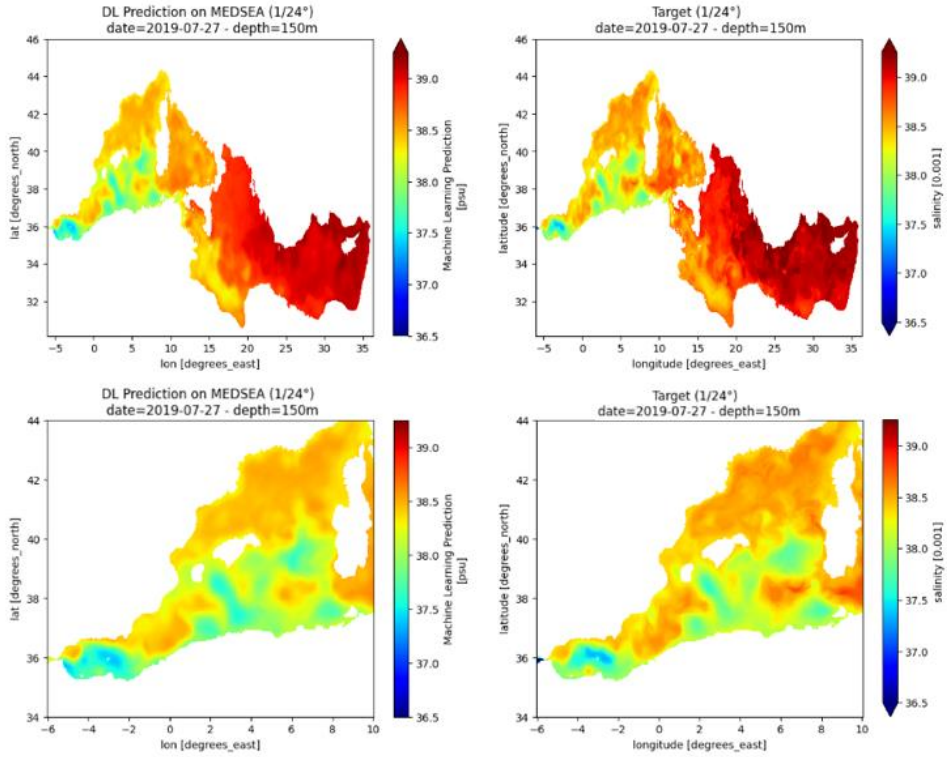


Figure 38: Visualization of the salinity CNN DL prediction on 2019-07-27 and the corresponding target, at 150m depth.

2.3.2.8 Inference at 1/24° on satellite test set

Input data

After training the CNN and assessing its performances on MEDSEA data, the final step was to evaluate its transferability to satellite data. Although SST and SSS observations images are quite similar to MEDSEA, there is a strong distribution shift for SSH, as illustrated in the figure below.

Figure 39: Input distributions for MEDSEA (red) and satellite data (blue) on the test set. The vertical lines show the distribution means.

It is very important that all satellite data have the same distribution than MEDSEA, otherwise the output predictions will be shifted. To recalibrate the satellite SSH distribution, the following transformation was applied.

$$NEW\ SSH_{sat} = \mu_{MEDSEA} + \sigma_{MEDSEA}\frac{SSH_{sat} - \mu_{sat}}{\sigma_{sat}}$$

With μ and σ being the mean and standard deviations. Figure 40 shows the SSH distribution after calibration.

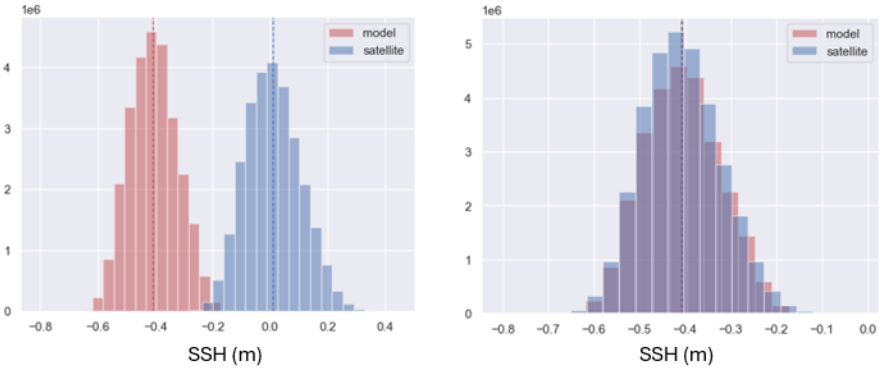


Figure 40: MEDSEA (blue) and satellite (red) SSH distribution before (left) and after calibration (right)

Edge effects (boundary conditions)

As mentioned previously, the coastlines can affect the quality of the predictions since land pixels are set to 0 before going into the CNN. This problem is linked to the fact that not all entries have exactly the same land mask (or coastline). To mitigate this effect, each satellite image was extrapolated on a few kilometres, then sent into the CNN for prediction, where land pixels were masked again afterwards.

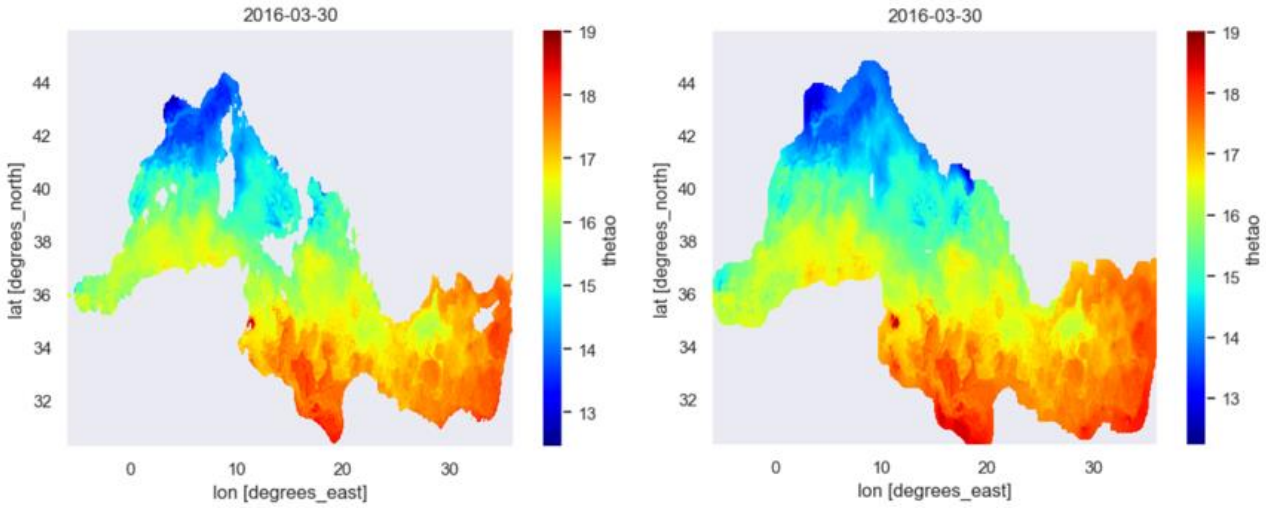


Figure 41: Illustration of the satellite SST coast extrapolation on 2016/03/30. Left: original satellite SST data, right: SST with extrapolated coasts.

The difference in the resulting predictions is shown below.

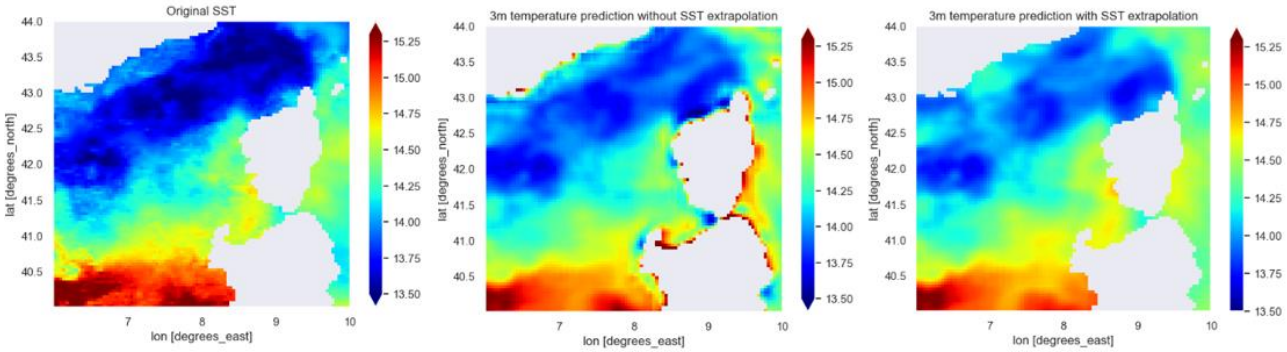


Figure 42: Impact of the coastline extrapolation on 3m prediction, on 2016/03/30.

On the left is the original satellite SST, in the middle the CNN output at 3m depth when using the original SST as input, and on the left the 3m prediction when using extrapolated SST. In the middle, some edge effects are visible around Corse and Sicily that are not in the input. They are the result of convolution kernels passing over land data. These effects were removed with SST extrapolation, as shown on the right.

Prediction errors with regard to MEDSEA targets

The first step to assess the CNN transfer from MEDSEA to satellite data is to calculate the prediction errors with regard to MEDSEA targets. This gives an idea of how close MEDSEA and satellite inputs are, by looking at the prediction errors. If both data sources are close, the CNN predictions will be similar, and so will be the errors.

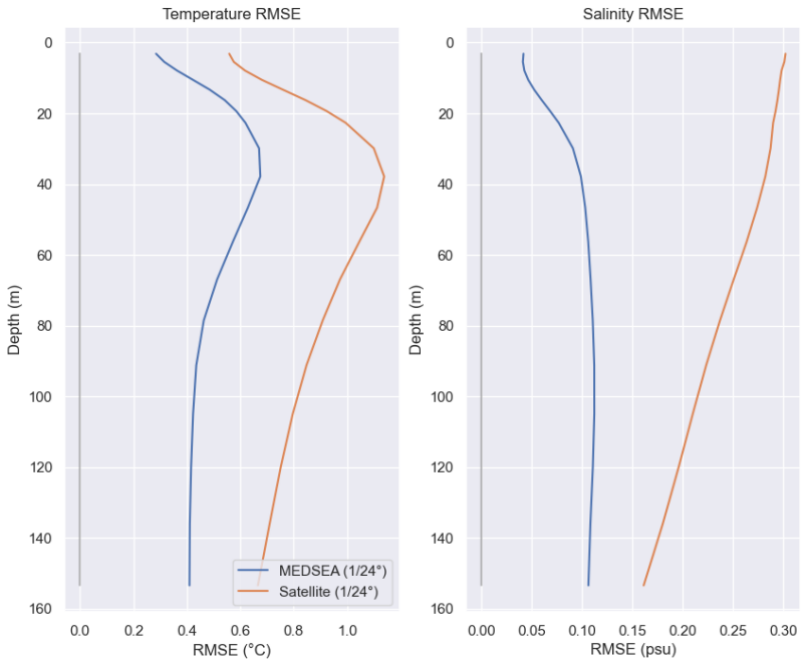


Figure 43: Temperature and Salinity RMSE profiles for CNN trained on MEDSEA and inferred on MEDSEA (blue) and CNN trained on MEDSEA and inferred on satellite (orange). The reference used to computed the RMSE is the MEDSEA target.

Here, since both RMSE profiles are calculated with regard to MEDSEA targets, the error differences are due to the differences in inputs, such as:

- MEDSEA model that can sometimes differ from real satellite observations.
- Effective resolution that is lower for satellite data than MEDSEA, especially for the SSS (see figure below).
- Barotropic or mass differences between MEDSEA SSH used for training and satellite SSH.

The latter seem to be the strongest difference between training and inference data, and has a big impact on the CNN outputs, which can explain the performance drop when inferring on satellite data. The figure below illustrates these differences.

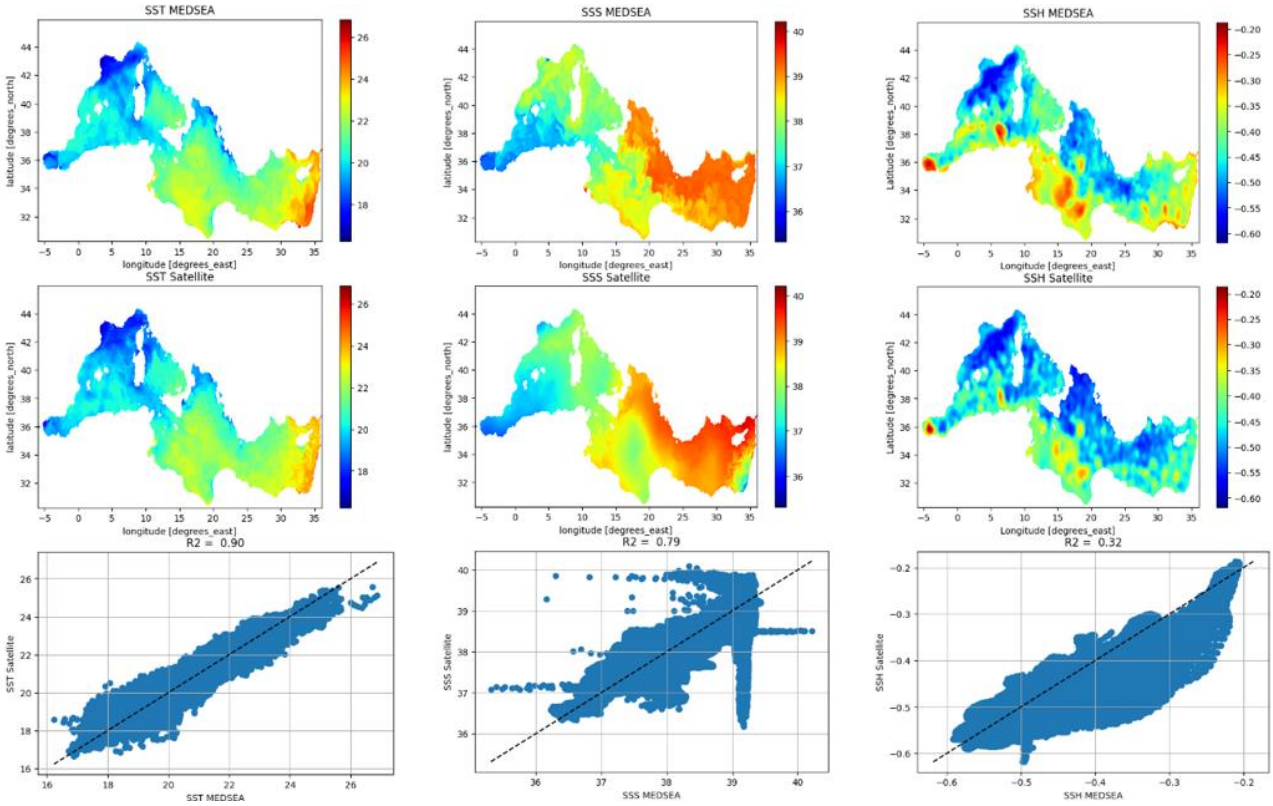


Figure 44: MEDSEA and satellite input comparison, on 2021/06/01. First row: MEDSEA inputs. Second row: Satellite inputs. Third row: Scatter plots between MEDSEA and satellite data.

2.3.3 Merging with in-situ observations and geostrophic velocities estimates

The SLA is only capable of representing the first and second vertical baroclinic modes of the ocean (Menemenlis et al., 2000). The SST and SSS are correlated with the upper layers of the ocean at a depth that depends on the local stratification. Indeed, inference from the surface alone will certainly miss some of the water mass variability. If necessary, that's means after validation of the results of the ML, the temperature and salinity reconstruction with the machine learning (ML) approach will be combined with in situ profiles of temperature (T) and salinity (S).

This second step consists in combining the ML T & S profiles with in-situ T & S profiles (Argo floats, moorings, XBT, CTD, Sea mammals, drifting buoys, …) using an optimal interpolation (OA) method (Bretherton et al., 1976).

The in-situ profiles are subsampled to avoid redundant information for the OA. This selection consists in keeping only one profile (T and/or S) for a same platform (platform_number identification) per 24 hours and in a 0.1° longitude per latitude spatial volume. If more than one profile is found in this volume, a selection to determine the best representative profile is done following the criteria:

- Keep only profiles which have more than 20% of valid data on the vertical (If no profile fulfils this requirement, then this test is not applied);

- The profile must have valid measurements up to 500-m depth (If no profile fulfils this requirement, then this test is not applied);
- Keep the profiles that have the largest number of valid values on the vertical;
- Keep the profile with the best vertical distribution of its measurements (mean distance between two consecutive measurements must be minimal);
- If after these tests more than one profile is listed, then the first T/S couple meeting these requirements is selected if exists, else, the best salinity profile and its collocated temperature is selected.

The main objective of the combination is to correct the large-scale part of the ML fields using surrounding in-situ profiles. The correlation scales of the first guess error have been computed at all depths over the Mediterranean basin by using the innovations with a method based on the algorithm of Hollingsworth and Lönnberg (1986). At 10-m depth these new correlations radii range between 80 and 200 km. The time radius is set at 7 days. To gain maximum benefit from the qualities of both data sets, namely the accurate information given by in-situ T profiles and the mesoscale variability given by the temperature ML profiles (from SLA and SST input fields), a statistical description of the errors of these observations has been introduced in the optimal interpolation method. For the in-situ profiles, the errors have been adjusted to avoid an overfitting of the product to the observations that leads to some discontinuities in the product. We apply a white noise equivalent to 20% of the variance. For the ML profiles, inferences from remote-sensing (altimeter, SST) observations, since these observations are not direct measurements but are derived fields, correlated errors are applied to correct long-wavelength errors or biases introduced by the ML method. This second step is implemented as anomalies from ML fields.

Finally, the geostrophic velocities will be deducted from the T & S gridded fields and the Absolute Dynamic Topography from the experimental 2DMED ALT product by using the thermal wind equation referenced at the surface. The method is described in the Quality Information Document (QUID) of the ARMOR3D product of the Copernicus Marine Service (See https://doi.org/10.48670/moi-00052)

2.4 WP2400 - Combined physical-biological 4DMED experimental product development

The objective of this task is to develop an experimental 4DMED algorithm providing a joint 4D-reconstruction of key physical and biological variables (temperature T, salinity S, density D, zonal geostrophic current Ugos, meridional geostrophic current Vgos, chlorophyll-a, primary production), that will be later used to provide the 4DMED physical/biological product within WP3400. To this aim, a specific effort has been dedicated to developing, optimize and test the algorithm used to provide the combined physical-biological 4DMED experimental product.

The development of the 4DMED product has been based on previous data-driven models (Sammartino et al., 2020 and Buongiorno Nardelli 2020) aimed to project the surface

SST, SSS, ADT, Chl-a at depth, by combining satellite and in situ observations. The 4DMED final product will be successively coupled with surface geostrophic currents to retrieve the geostrophic components of the 4D currents (WP2410), and with additional ancillary information to estimate 4D primary production (PP) (WP2420).

2.4.1 Input dataset description

To create the reference database for the training and test of the deep learning algorithm, surface and subsurface T, S and Chl measurements have been extracted by satellite regional products and in situ database. For the latter, two different data sources have been considered: a set of data coming from oceanographic cruises and a second set of bio-hydrographic profiles coming from Bio-Argo floats (Figure 45).

2.4.1.1 In situ observations

2.4.1.1.1 Oceanographic cruise database

The in situ oceanographic cruise database includes all concurrent profiles of temperature, chlorophyll and salinity collected during 26 oceanographic cruises carried out in the Mediterranean Sea from 1997 to 2017 (red crosses in Figure 45). Most of these data were acquired and processed by the GOS group (Global Ocean Satellite monitoring and marine ecosystem studies) of the Institute of Marine Sciences (ISMAR) of Rome of the Italian National Research Council. The estimate of chlorophyll has been done calibrating the fluorometer signal with bottle samples that were concurrently sampled in each cruise.

2.4.1.1.2 Bio-Argo database

The second in situ database includes hydrographic (T, S) and bio-geochemical (Chl-a) profiles acquired by BGC-Argo floats, over the Mediterranean Sea from 2012 to 2022 (blue crosses in Figure 45). These data were collected and made freely available by the International Argo Program and the national programs that contribute to it (http://www.argo.ucsd.edu, http://argo.jcommops.org). The Argo Program is part of the Global Ocean Observing System. For temperature and salinity profiles, both real-time and delayed mode have been considered, while for chlorophyll only the delayed mode data (CHLA_ADJUSTED) have been selected. All profiles have been downloaded from ftp server Coriolis ftp.ifremer.fr.

The real time profile files are available within 12 – 24 hours of the float completing its profile. In the real time profile files, the quality control flags are set by several automatic tests which aim to detect gross errors. If scientific calibration is applied to the measured parameters in real time, the real time adjusted values will be recorded in <PARAM>_ADJUSTED. In general, these data should be consistent with ocean climatologies. Monthly climatology tests are performed, along with visual comparisons for profiles that fail the tests. For hydrographic profiles, Argo delayed mode files are available 1 – 2 years after a profile is taken; sometimes earlier. These have been subjected to detailed scrutiny by oceanographic experts and the adjusted salinity has been estimated by comparison with high quality ship-based CTD data and Argo

climatologies using the process described by Wong et al., 2003; Böhme and Send, 2005; Owens and Wong, 2009; Cabanes et al, 2016.

For BGC parameters, delayed mode files can be available within 5 – 6 cycles after deployment. This is because the BGC sensors often return data that are out of calibration, but early adjustment methodologies exist that can significantly improve their accuracy. Additional delayed mode quality control occurs when a longer record of float data is available.

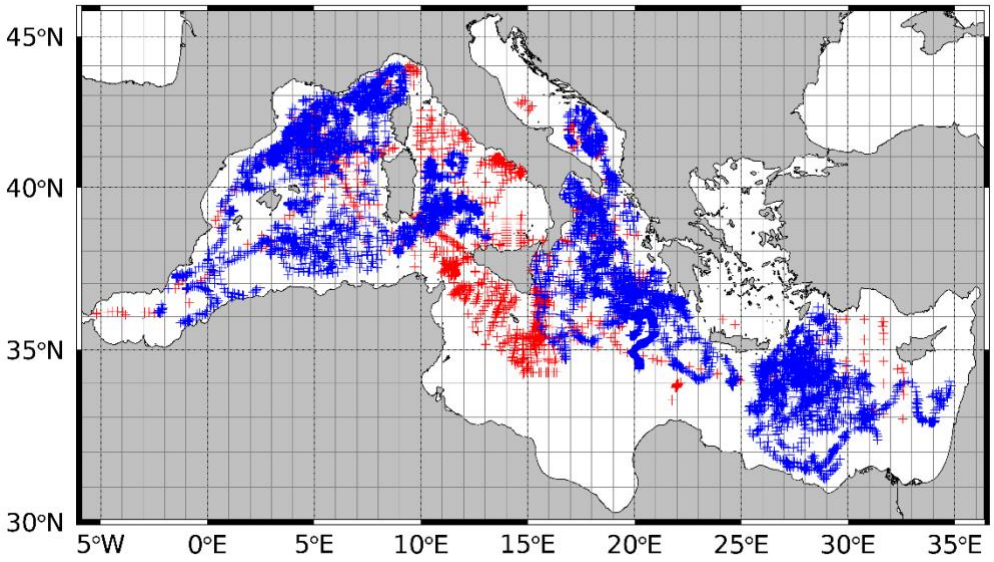


Figure 45: Spatial distribution of the in situ database comprising oceanographic in situ (red crosses) and Bio-Argo (blue crosses) stations.

2.4.1.2 Satellite observations

2.4.1.2.1 Satellite Chlorophyll L4 dataset

The satellite Chl-a L4 dataset used as input in the network for the reconstruction of the 4D Bio-physical product are based on the Mediterranean multi-sensor product (OCEANCOLOUR_MED_BGC_L4_MY_009_144, https://doi.org/10.48670/moi-00300) provided by the CNR Production Unit (OC-CNR-ROMA-IT) and available at Copernicus Marine Service portal (https://marine.copernicus.eu/). This product includes daily gap-free Chl-a at 1km of resolution with a temporal extent from Sept 1998 to Aug 2023. The dataset is obtained from the merging of multi sensor data (SeaWiFS, MODIS, MERIS, VIIRS-SNPP & JPSS1, OLCI-S3A) (for more details, see https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-OC-QUID-009-141to144-151to154.pdf). Finally, the interpolation procedure is based on the DINEOF algorithm described in Volpe et al. 2018.

2.4.1.2.2 Satellite Absolute Dynamic Topography dataset

The Absolute Dynamic Topography (ADT) and absolute geostrophic currents (zonal and meridian components, U & V respectively) are additional variables included in the

altimeter satellite gridded data of Sea Level Anomalies (SLA) available in the Copernicus Marine Service web-portal, produced by the production unit of SL-CLS of Toulouse in France. The ADT is the instantaneous height above the Geoid and it is obtained from the sum of the SLA_N and MDT_N (Mean Dynamic Topography, (Rio et al., 2014), where N, here, is the mean reference period (1993–2012). The ADT is a delayed time optimally interpolated product at a resolution of 0.125◦ × 0.125◦ processed by the DUACS multi-mission altimeter data processing system and coming from the merging of different altimeter missions (for more details about the processing see the Quality Information Document at https://doi.org/10.48670/moi-00141 or https://duacs.cls.fr). The absolute geostrophic currents are obtained by the SLA and ADT product, above mentioned. In this project, this ADT product has been used for the training and test of the neural network employed to produce the combined bio-physical 4DMED product. The future operational production of the combined physical-biological 4DMED experimental product, instead, will be based on the WP2100 - Experimental 2DMED ALT product (MIOST product).

2.4.1.2.3 Input dataset preparation and pre-processing

All in situ oceanographic cruise data passed the quality controls based on SeaDataNet standards (see quality control manual at https://www.seadatanet.org/Standards/Data-Quality-Control) and further requirements stated for the specific goal of this work. The same procedure adopted in Sammartino et al. (2020) was applied on either in situ chlorophyll and temperature/salinity profiles and included in the following points:

- Visual check (to guarantee the consistency of the database);
- Check of the first acquisition depth (accepted if ranging between 3 and 4 m to avoid any noise on the first acquired measure);
- Check of missing points along the profiles: a specific-case evaluation was done observing each sample. The profile showing too many missing points were discarded, otherwise they were linearly interpolated.
- Check maximum acquisition depth: only those profiles with a depth greater or equal to 150 m were included.

About the Bio-Argo dataset, besides the quality control and post-processing applied by Argo providers (http://www.argodatamgt.org/Documentation), these profiles have been undergoing additional controls. Only those profiles greater than 150 meters of depth have been considered and linearly interpolated on a regularly spaced vertical grid (with 1 m intervals).

From both oceanographic and bio-argo datasets, the optical weighted pigment (OWP) concentration has been computed from the chlorophyll profile, following a modified version of the Morel 1989 model. This allowed us to obtain from in situ data a chlorophyll concentration compliant to that observed by satellites.

Successively, the final in situ database has been matched up with satellite ADT and CHL products. The value for each satellite CHL matchup was obtained by computing the median value of the cloud free pixel within a 3 × 3 box, with a minimum of at least 45% of good pixels.

In this project the training and test of the network has been carried out by using the surface value of the T and S profiles as proxy for satellite SST and SSS and the OWP as proxy for satellite CHL.

The density profiles have been computed from in situ T and S profiles and their surface values used as proxy for SSD.

The profiles containing CHL constant values within the first 15 meters of depth have been filtered out.

At the end of the quality controls, the final 8624 profiles have been randomly divided into 80% used for training and the remaining 20% for the test of the network on an independent dataset.

All profiles were scaled within the 0–1 interval before feeding the network by using the following formula:

$$x_{norm} = \frac{(x - x_{min})}{(x_{max} - x_{min})}$$

2.4.2 Machine Learning model for 4D chlorophyll, temperature, salinity and density reconstruction

The algorithm used to infer the 4D temperature, salinity, chlorophyll and density from surface data has been developed starting from previous works of Sammartino et al. (2020) and Buongiorno Nardelli (2020). The input parameters of the selected model are: Day of the Year (cos(doy)/sin(doy)), Latitude, Longitude, SST, SSS, SSD, CHL-a, ADT, Ugeos and Vgeos.

To select the best deep learning model, different networks have been tested such as: Feed-Forward Neural Network (FFNN), Bi-directional Long Short-Term Memory Network (Bi-LSTM), LSTM, LSTM plus 1-Dimensional Convolutional Layer (LSTM+Conv1D).

Among them, the best performing resulted a modified version of the architecture proposed in Buongiorno Nardelli (2020). The final model is based on a two layer stacked Long Short-Term Memory Network (LSTM), including two 1-Dimensional Convolutional Layers (Conv1D) as depicted in Figure 46.

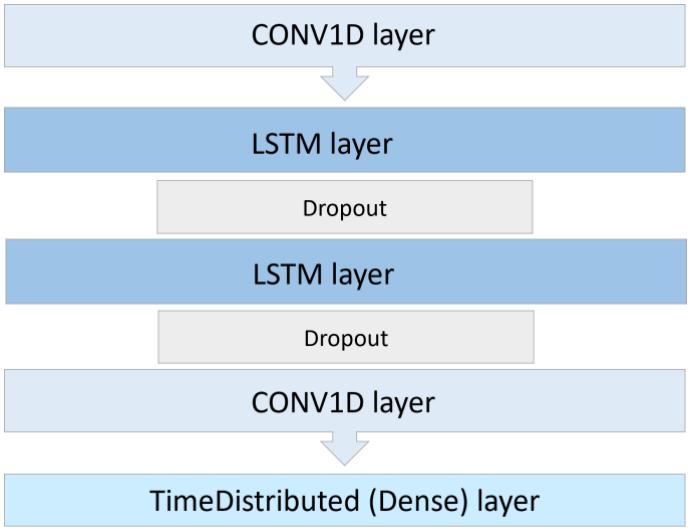


Figure 46: 4DMED Reconstruction Deep Neural Network architecture.

The LSTM belongs to the so-called Recurrent Neural Network (RNN), particularly suitable to model ordered sequences of data, as time series or, in this case, vertical profiles. They also are designed to avoid vanishing gradients. The principal unit of an LSTM is the "cell", that allow the flow of the information from one cell to the next one acting as a memory.

Within LSTM cells, the external input vector is concatenated to the previous cell hidden state and then passed through different "gates", each one aimed at carrying out a specific task to update both the present hidden state itself and the cell state, that is directly transmitted to the next cell (Buongiorno Nardelli, 2020).

On vertically stacking the LSTM cell sequences, the information flows across activation function connections, until the dense layer that assimilates all of that to make the final prediction (Kar et al., 2023).

The convolutional layer, instead, is used to extract latent features and creating informative representations of the vertical profile. It consists of a linear operation that takes a sliding portion of the input, multiplies it element-wise with a matrix named Convolution filter (or simply kernel), and sums the elements to provide the output (Xiao et al. 2022). This operation gives unique values, creating a feature array that summarises the presence of detected features in the input from previous layers.

Once identified the Conv1D-LSTM-Conv1D as the best performing architecture, several tests have been performed to select the optimal configuration of the hyperparameters. Specifically, a different number of hidden cells in each LSTM layers has been tested, the addition/removal of Dropout layer, number of filters and kernel size, batch size and type of optimizer.

The best set up of hyperparameters was the following: a first Conv1D layer (filters=30, Kernel=1), followed by a stacked LSTM layer containing 35 hidden units, then a Dropout layer (dropout fraction of 0.01), a second LSTM layer of the same size of the previous one, then an identical second Dropout layer, followed by a Conv1D layer with a number of filters=30 and a kernel=3. The last layer of the network is a Dense output layer (TimeDistributed) with four nodes.

In Table 4, a summary of the hyperparameters of the final neural network model is given.

Activation function	**tanh**
Optimizer	Adam
patience	40
N_epochs	2000
Validation_split	30
Bacth_size	68
Loss fnc	Customized Loss_tot

Table 4: List of the hyperparameters used to train the 4DMED Reconstruction Deep Neural Network

In addition to the tests on the hyperparameters, different customized loss functions have been tested with different configurations.

The first test started with a loss function (loss_tot1), where the minimization of the error accounted for two different weights: a first one multiplying the errors in the first 10 meters of depth (weight_surf=2), and a second lower weight (weight_profile=1.) used in the computation of the errors along the entire profile.

Successively, a second test has been done using the same loss function with the addition of a third weight (weight_deep=0.5) accounting for the errors computed at deeper layers (last 10 levels at the bottom). The final formula was:

$$Loss_{depth} = w_{surf}\frac{1}{N}\sum_{n=1}^{N}(y_{true\_profile}[z_1{:}z_{10}] - y_{predicted\_profile}[z_1{:}z_{10}])^2 + w_{profile}\frac{1}{N}\sum_{n=1}^{N}(y_{true\_profile} - y_{predicted\_profile})^2 + w_{deep}\frac{1}{N}\sum_{n=1}^{N}(y_{true\_profile}[z_{141}{:}z_{150}] - y_{predicted\_profile}[z_{141}{:}z_{150}])^2$$

A further step in the network tuning process has been represented by the introduction of two different constraints in the density reconstruction. This approach transformed the initial network into a so-called physically informed network:

1. a first one accounting for the error computed between the observed density profile and that obtained from the temperature and salinity profiles inferred by the network (loss_*density*). The formula implemented was:

$$Loss_{dens} = \frac{1}{N}\sum_{n=1}^{N}(\sigma_{true} - \sigma_{predicted})^2$$

2. a second constraint was implemented to avoid unrealistic vertical density inversions (*loss_phy*). Indeed, based on the work presented in Pauthenet et al. 2022, we added a constraint of monotony on the density profile to penalize the predictions that contain density inversions. The formula implemented was:

$$Loss_{phy} = \frac{1}{N}\sum_{n=1}^{N} RELU\left(\frac{\Delta\sigma}{\Delta z}\right)$$

By this way, the final customized loss function resulted in the addition of all the terms above mentioned as in the following formula:

$$Loss_{TOT} = Loss_{depth} + Loss_{dens} + Loss_{phy}$$

In this work the constraint on density errors and inversions was implemented to make the prediction product as much as possible physically consistent.

Several tests have been done to make the different customized losses of the same order of magnitude. The result of this analysis demonstrated that the losses were comparable each other without multiplying to any constant coefficient.

2.4.3 Computation of 4DMED experimental product and 4D geostrophic currents

The final combined physical-biological 4DMED experimental product are obtained by applying the 4DMED algorithm on the surface satellite observations.

The 4D geostrophic currents are computed on the steric height obtained from 4DMED density following the Arbic et al. (2012) method and adjusting them to MIOST data provided by WP2100. More in detail, from the 4D reconstructed density field, the 4D steric height is computed by integrating the specific volume anomaly with respect to a standard density 4D field. Then, the geostrophic velocities are obtained by using the Arbic et al. (2012) method applied on Steric Height, with x-point stencil centred differences, where x is the stencil width that progressively decreases from nine, seven, five and three.

Finally, the 4D geostrophic velocity fields are obtained by adding at each depth level the difference between the MIOST and surface geostrophic reconstructed velocity field.

2.4.4 Computation of Primary Production

The estimations of the marine primary production are based on an upgraded version of the Morel (1991) model.

This model combines an atmospheric (Tanré et al., 1979) and a bio-optical (Morel, 1991) models, allowing to estimate the photosynthetic radiation on the surface of the sea and its attenuation through the water column, and to compute the productivity through a parameterization of the main physiological processes, starting from the distribution of the algal biomass concentration and temperature. The original model version is updated in the atmospheric compart by a revised version of the multispectral Ocean Atmosphere Spectral Irradiance Model (OASIM; Gregg and Casey, 2009) that simulates the downward

spectral irradiance (direct and diffuse) over the ocean surface for clear-sky (Gregg and Carder, 1990) and cloudy (Slingo, 1989) conditions. This updated OASIM model provides the direct and diffuse irradiance over the ocean surface with 5 nm of spectral resolution ranging from 400 to 700 nm, 4 km of spatial resolution and daily estimations.

Table 5 summarized the ancillary data (and their source) needed by OASIM to achieve the downward spectral irradiance.

Table 5: List of the ancillary data (and their source) needed by OASIM to achieve the downward spectral irradiance

Variable	**Source**
Cloud Cover	ERA5
Cloud liquid water path	ERA5
Cloud droplet effective radius	MODIS data
Aerosol optical thickness	MODIS data
Aerosol asymmetry parameter	MODIS data
Aerosol single-scattering albedo	MODIS data
Surface pressure	ERA5
Wind speed	ERA5
Relative humidity	ERA5
Ozone	ERA5
Precipitable water	ERA5

Once the photosynthetic radiation on the surface is obtained, PP is estimated with the bio-optical model compart.

The assessment of the daily primary production of the water column within the euphotic layer is thus performed through the following expression:

$$P = 12\ Chl(z,t)a^*(z,t,\lambda)PAR\ (z,t,\lambda)\varphi_\mu(z,t,\lambda)$$

where 12 is the carbon molar weight to transform P in mass, a^* is the chlorophyll specific absorption spectrum, PAR is the photosynthetic available radiation and φ_μ is the quantum yield for net growth.

Following Antoine and Morel (1996) the above expression can be simplified as following:

$$P = 12\ a^*_{max}\varphi_{\mu,max}\int_0^L\int_0^D\int_{400}^{700} Chl(z)\ PUR\ (z,t,\lambda)\ f(x)\ d\lambda\ dz\ dt$$

where a^*_{max} is the maximal value of the chlorophyll specific absorption spectrum, $\varphi_{\mu,max}$ is the maximal value of the quantum yield for net growth, L is the day length, D is the euphotic depth, $Chl(z)$ is the vertical chlorophyll profile, $PUR\ (z,t,\lambda)$ is the

fraction of available energy (PAR) which can potentially be absorbed by algae and $f(x)$ is a function that reproduces the light photosynthesis curve when the irradiance is expressed in terms of PUR and x is the ratio PUR/KPUR.

KPUR depends on temperature following a Q10=1.88, as proposed by Eppley (1972):

$$KPUR(T) = KPUR(20°C)\ 1.065(T - 20°)$$

The original bio-optical model was here revisited to increase efficiency and to include new parameterizations for the Inherent Optical Properties such as absorption and scattering.

As shown in the previous formulas, PP is strictly dependent on chlorophyll and temperature vertical 4D profiles.

2.4.5 4DMED-SEA BIOPHYS product characteristics

The final combined physical-biological 4DMED experimental product are obtained by applying the 4DMED algorithm on the surface satellite and re-gridded fields of UHR-SST (the same used as input in the 2DMED SSS product), SSS and SSD from WP2200, ADT, Ugeos and Vgeos from WP2100 product and CHL as described in section 2.4.1.2.1. The 4D reconstructed fields are available on a regular grid with a surface spatial resolution of 1/24° in both longitude and latitude, 148 vertical levels (from 3m to 150 m of depth) with a step interval of 1m and a temporal sampling interval of 1 day. The spatial coverage spans from 6°W to 36°E, while the temporal coverage extends from 1st January 2016 to 1st August 2022.

The final physical-biological 4DMED experimental products are available in NetCDF format at two different thredds links: 1) including T, S, Chl, D, Ugos and Vgos; 2) including the Primary Production.

https://4dmed.artov.ismar.cnr.it/thredds/catalog/4dmed/catalog_NN.html
https://4dmed.artov.ismar.cnr.it/thredds/catalog/4dmed/catalog_PP.html

The link to the whole dataset is the following:

https://4dmed.artov.ismar.cnr.it/thredds/catalog/catalog.html

Table 6: characteristics of the experimental combined physical-biogeochemical 4DMED product

Product Name	4DMED_BIOPHYS_REP_3D
Geographical coverage	Mediterranean Sea [Lon -6.06° to 36.10°, Lat 30.27° to 45.99°]
Horizontal resolution	1/24°
Variables	Temperature, Salinity, Chlorophyll-*a*, *Density* and Geostrophic Velocities
Temporal coverage	From 2016 to 2022
Temporal resolution	Daily field
Number of vertical levels	148 levels (from 3 to 150 m, every 1 m depth)

Format	Netcdf 4.0 CF1.7

Table 7: characteristics of the experimental combined physical-biogeochemical 4DMED product

Product Name	4DMED_PP_REP_3D
Geographical coverage	Mediterranean Sea [Lon -6.06° to 36.10°, Lat 30.27° to 45.99°]
Horizontal resolution	1/24°
Variables	Primary Production, Euphotic Depth
Temporal coverage	From 2016 to 2022
Temporal resolution	Daily field
Number of vertical levels	148 levels (from 3 to 150 m, every 1 m depth)
Format	Netcdf 4.0 CF1.7

2.5 WP2500 - Development of Lagrangian analysis and validation tools

Background

Transport and mixing of water masses are basic processes that control the global thermohaline circulation (and thus the climate of our planet) and shape marine ecosystems. By constantly stirring and redistributing both horizontally and vertically key oceanic tracers of climatic-relevance (such as temperature, salinity and nutrients), multiscale ocean currents are fundamental to better understand the earth system and the global changes it currently undergoes (Ser-Giacomi et al. 2020). Owing to its inherent turbulent nature, ocean circulation is highly complex. We can model the main transport patterns of the general circulation at large scales, but the extreme richness of mesoscale O(10-100 km, 10-100 days) and submesoscale O(0.1-10 km, hours-few days) circulation patterns, makes the assessment of water pathways and the study of oceanic transport phenomena an extraordinarily complicated task. This problem can be approached from a Lagrangian point of view which studies the transport and mixing following the trajectories of fluid particles. Adopting this approach, we developed tools that exploit the improved representation of physical dynamics and variability (provided by the new 4DMED products) to help understanding the control it exerts on marine biodiversity, biogeography and population dynamics within WP4000.

In this section the algorithms of the Lagrangian diagnosis used to exploit and improve the representation of physical dynamics and variability (provided by the new 4DMED products) are described. These have been grouped into the following four key Lagrangian analysis tools according to their applications:

- Finite Size Lyapunov Exponents (FSLE) and Lagrangian averaged vorticity for the detection of Lagrangian Coherent Structures (LCS).
- Lagrangian Flow Network model for connectivity analysis.

- Bio-Lagrangian backtracking and Kinematic Lagrangian Model for regional studies
- II-Kind FSLE analysis for velocity fields intercomparison.

2.5.1 Finite-Size Lyapunov Exponents and Lagrangian averaged vorticity for the detection of Lagrangian Coherent Structures.

2.5.1.1 Background

The last two decades have seen important advances in the Lagrangian description of transport and mixing in fluid flows driven by concepts from dynamical systems theory and complex systems. One of the most widely used techniques contributing to this understanding in marine flows are the ones known under the name of Lagrangian Coherent Structures (LCS) computed from the finite-size Lyapunov exponents (Boffetta et al., 2001; d'Ovidio et al., 2004; Hernandez-Carrasco et al., 2011). They are Lagrangian analyses of the velocity fields that identify lines acting as transport barriers, e.g. dynamical fronts or eddy boundaries. These circulation coherent structures greatly organize the motion of particles around them, and, identifying their location and motion, provide a kind of 'skeleton' or 'template' of the main routes of transport in a region.

2.5.1.2 Lagrangian model for the computation of Finite-Size Lyapunov Exponents

The value of the FSLE, denoted here by λ, for a particle located at **x** and at time t_0 is given by the following formula:

To compute the horizontal FSLE (Eq. 2.5.1) we need to obtain the time, τ, required for two neighbouring fluid particles (one of them placed at **r**) initially separated by a distance δ_0 to a fixed final separation distance δ_f (see Figure 47).

$$\lambda(r,\ t_0,\ \delta_0,\ \delta_f) = \frac{1}{|\tau|}\ln\frac{\delta_f}{\delta_0} \qquad \text{Eq. 2.5.1}$$

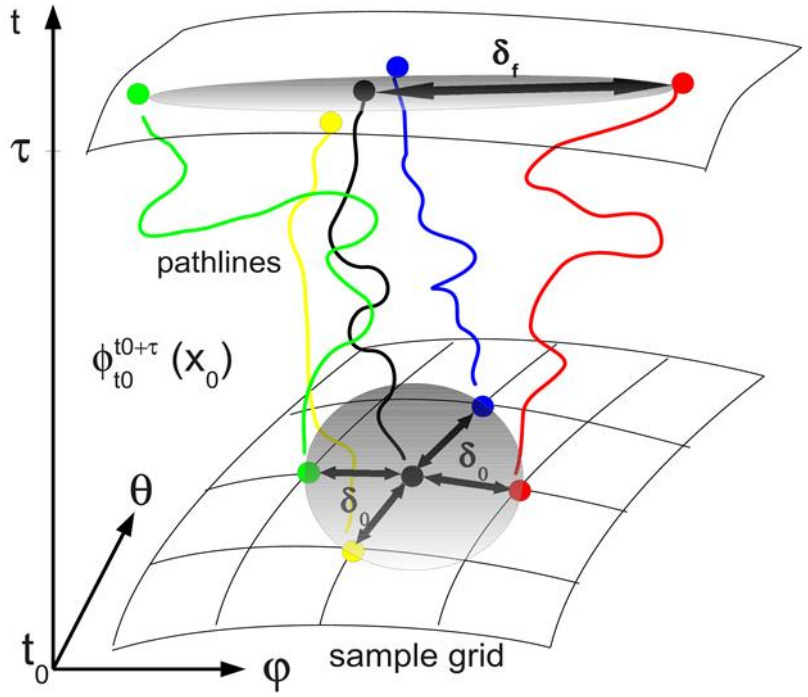


Figure 47: Sketch showing how the FSLE is obtained. Particles are virtually deployed in a regular grid of spatial separation δ_0 and advected in the flow during the time t0 until they reach a fixed separation distance δ_f with respect to one of its four neighbours. $\Phi^{t0+\tau}_{t0}$ is the flow map, represented here as the temporal evolution and displacement of the particle advected in the velocity field from t_0 to $t_0 + \tau$.

Following Hernandez-Carrasco et al., 2011, the algorithm of the Lagrangian model used to compute the FSLE for a discrete space-time grid of the 4DMEDSea velocity field is divided into the following steps:

1. Integration of the equation of motion. To compute the τ time, we need to know the trajectories of the particles. The FSLE are computed for the points **r** of a lattice with spacing coincident with the initial separation of fluid particles δ_0. The equations of motion that describe the horizontal evolution can be expressed as:

$$\frac{d\varphi}{dt} = \frac{u(\varphi,\theta,t)}{R\cos\theta}, \qquad \text{Eq. 2.5.2}$$

$$\frac{d\theta}{dt} = \frac{v(\varphi,\theta,t)}{R}, \qquad \text{Eq. 2.5.3}$$

where *u* and *v* represent the eastward and northward components of the velocity field coming from the 4DMEDSea velocity field described in sections 2.3.3 and 2.4; R, is the radius of the Earth (6400 km in our computations), φ is longitude and θ latitude. Numerically we proceed integrating the ordinary differential equations (ODE) given by the Eqs. (2.5.2) and (2.5.3), using a standard, fourth-order Runge-Kutta scheme, with an integration time step dt, which value will be selected depending on the relation between the time and spatial resolution of the data.

2. Interpolation. Since information is provided just in a discrete space-time grid, spatiotemporal interpolation of the velocity data is achieved by bilinear interpolation. However, we notice that bilinear interpolation requires an equally spaced grid. If our data input is expressed in spherical coordinates, and the linear separation is constant in the latitudinal and longitudinal directions, then the grid is not uniformly spaced in the angular latitude coordinate. In the case of the 4DMEDSea velocity field this does not occur as the grid is equally spaced, and the grid does not need to be transformed.

Once we integrate the equations of motion, Eqs. 2.5.2 and 2.5.3, we obtain the particle trajectories in the form:

$$r(t + \Delta t) = r(t) + \int_t^{t+\Delta t} U(r(t),\, t)dt, \qquad \text{Eq. 2.5.4}$$

where U=(u,v), and we can compute the FSLEs with Eq. 2.5.1 for the points **r** of a lattice with spacing δ_0. Initial conditions for which the prescribed final separation δ_f has not been reached after using all the available times in the data sets are assigned a control value $\lambda = 0$.

3. Parameter selection. To compute particle trajectories, the equation of motion must be solved for the characteristics of the 4DMEDSea velocity field.

The first preliminary step is to adapt the Lagrangian model used to calculate the trajectories in 4DMEDSea velocity fields. The integration of the equation of motion is based on these two methodological components:

- **ODE Integration scheme**: We consider the following 4th order Runge-Kutta (RK) parameters: i) integration time-step (dt), whose value depends on the time and spatial resolution of the 4DMEDsea velocity data. RK time-step depends on time/spatial resolution of the 4DMEDsea velocity data. ii) total integration period (T), duration of the evolution of the virtual particle trajectories. Larger values of T allow more pairs of virtual particles to separate the fixed final distance of separation, obtaining more FSLE values closer to zero
- **Interpolation method**: We have selected trilinear interpolation because it offers the best compromise between accuracy and computational cost, while not introducing spurious values near the coast. Because the grid of the 4DMEDsea velocity fields derived from altimetry is equally spaced, it doesn't need to be transformed during the interpolation.

2.5.1.3 Lagrangian coherent structures from FSLE.

The Finite Size Lyapunov Exponent is a metric that has an explicit spatial and time dependence (i.e. the initial time and position of the Lagrangian particle trajectory) in contrast to the original definition (Aurell et al, 1997), as can be observed from Eq 2.5.1. This definition allows to obtain maps of FSLE field allowing the analysis of the spatial organization of ridges of the local Lyapunov exponent. The largest Lyapunov values

concentrate along characteristic lines which are the LCSs (d'Ovidio et al., 2004, 2009; Hernández-Carrasco et al., 2011).

We compute the FSLE by backwards time integrations. In this way we quantify the fluid deformation by past stirring, identifying attracting LCS, which have a more direct physical interpretation.

Parameters for the Lagrangian Coherent Structures computations.

- **δ_0: initial separation**. Values of δ_0 provide the spatial resolution of the FSLE field. Smaller values of δ_0 than the velocity resolution allows capturing subgrid dynamics. Smaller values of δ_0 implies larger computational cost and bigger output files.
- **α : amplification factor.** The amplification factor determines the value of δ_f ($\delta_f = \alpha\delta_0$). Values of α larger than 10 are used to identify regions of high stretching, and thus identify relevant dynamical structures. Larger values of α result in longer computations.

Maps of FSLE over the Ionian Sea for different values of these parameters are shown in Figure 48 to Figure 51.

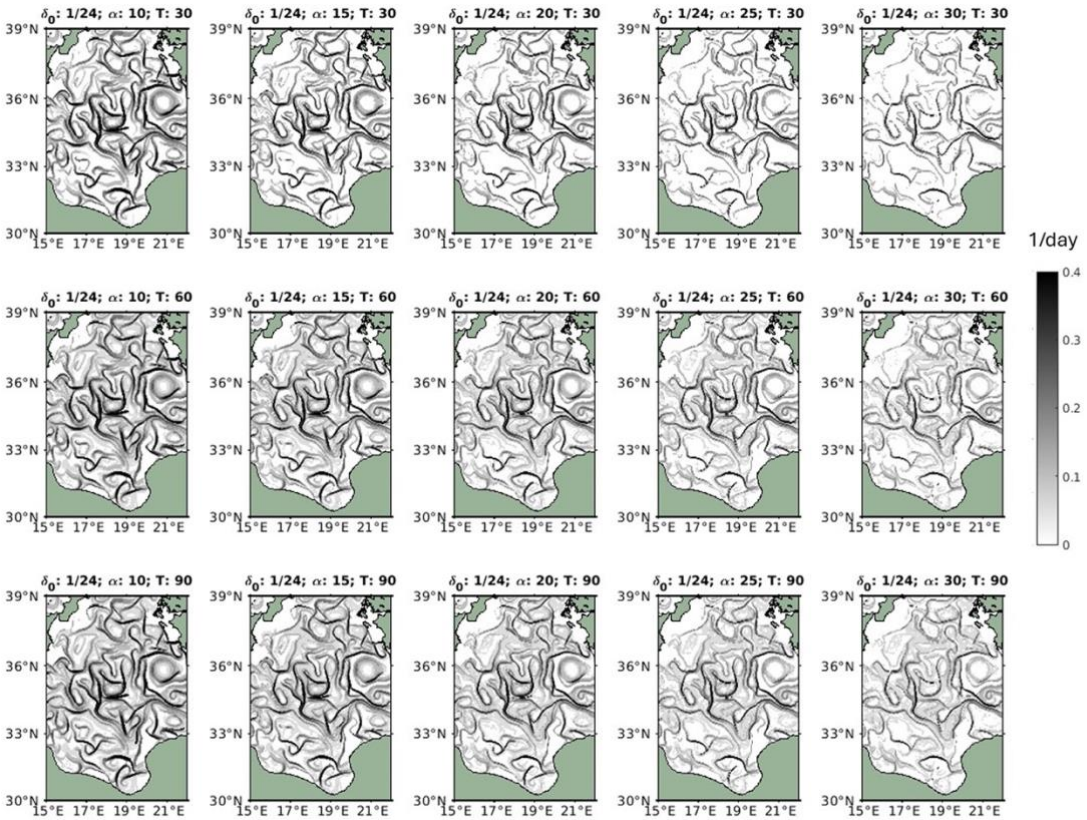


Figure 48: FSLE maps over the Ionian Sea at 01/04/2016 for different parameter configurations computed using currents from the L4 MIOST dataset for α = 10,15,20,25,30 (various columns) and T = 30, 60 and 90 days (various rows) and an initial distance separation d0 = 1/24 deg.

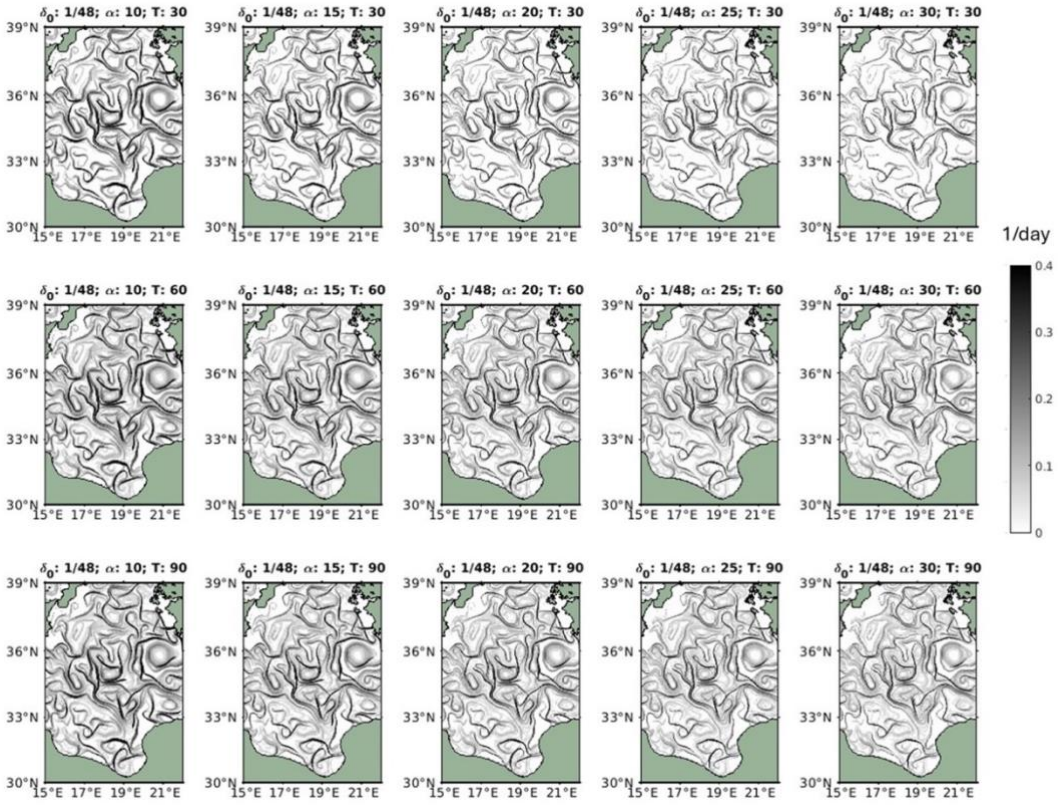


Figure 49: FSLE maps over the Ionian Sea at 01/04/2016 for different parameter configurations computed using currents from the L4 MIOST dataset for α = 10,15,20,25,30 (various columns) and T = 30, 60 and 90 days (various rows) and an initial distance separation d0 = 1/48 deg.

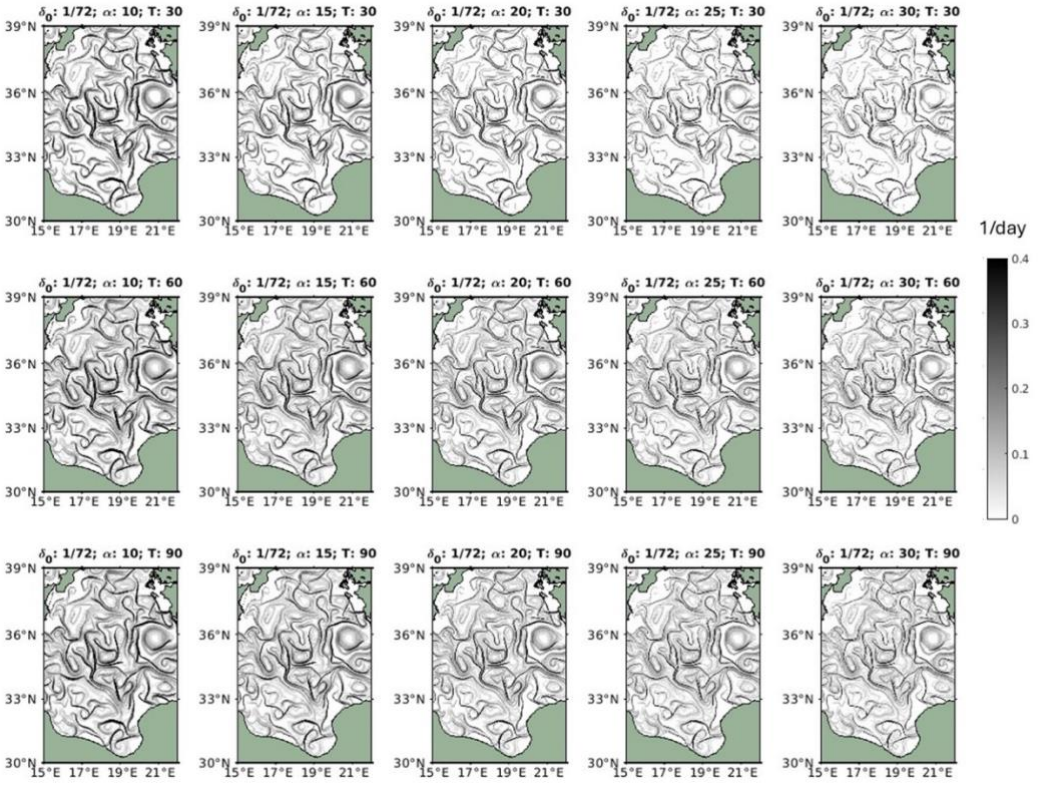


Figure 50: FSLE maps over the Ionian Sea at 01/04/2016 for different parameter configurations computed using currents from the L4 MIOST dataset for α = 10,15,20,25,30 (various columns) and T = 30, 60 and 90 days (various rows) and an initial distance separation d_0 = 1/72 deg.

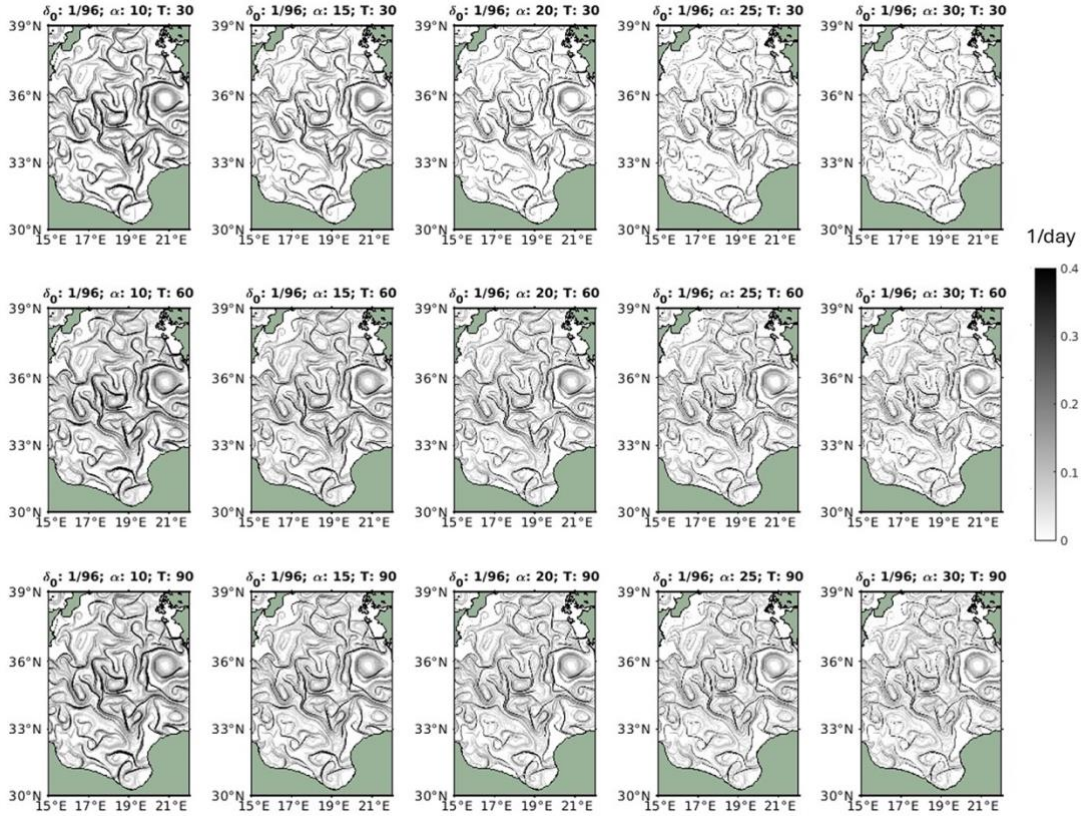


Figure 51: FSLE maps over the Ionian Sea at 01/04/2016 for different parameter configurations computed using currents from the L4 MIOST dataset for α = 10,15,20,25,30 (various columns) and T = 30, 60 and 90 days (various rows) and an initial distance separation d_0 = 1/96 deg.

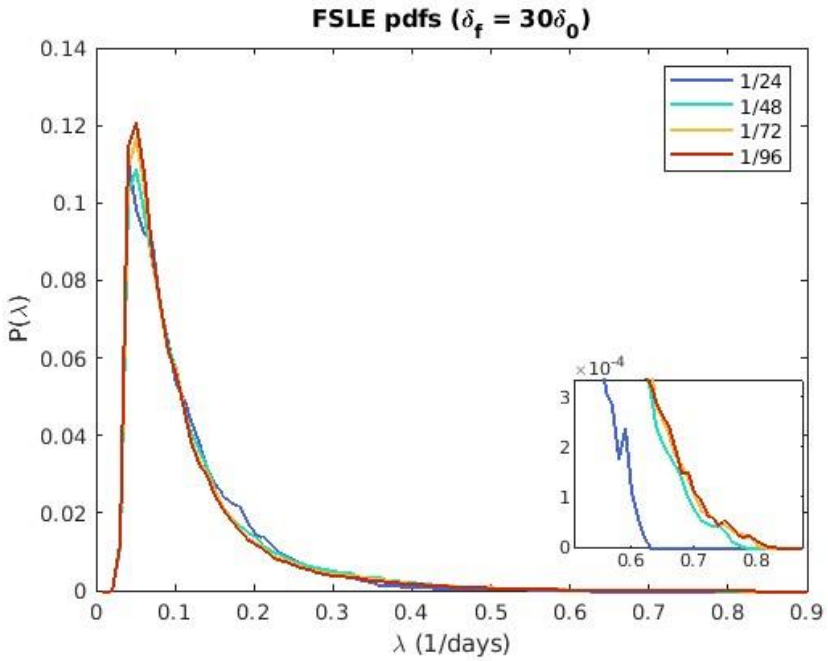


Figure 52: Probability distribution function of FSLE for MIOST dataset over the Ionian Sea for α = 30, T = 90 days and different initial distance separation d_0. The zoom shows the tail of the probability distribution density.

Increasing the value of the amplification factor α, for the same value of the integration period T, results in thinner but fewer filaments (upper row in Figure 48 to Figure 51), as the pairs of particles do not have sufficient time to separate fixed final separation distance δ_f given by the amplification factor α. On the other hand, using an amplification factor of α = 30 and an integration period of T = 90 days helps reduce the thickness of FSLE fronts, while providing enough time for more particles to achieve the defined final separation (note the greater filaments comparing upper and lower panels in Figure 48). Regarding the various initial separations (Figure 52), a separation of 1/96° allows for the reconstruction of finer filaments (indicated by a narrower peak in the probability density function) and results in a greater occurrence of large FSLE values (more values in the tail of the distribution). Moreover, the FSLE fields for 1/72° and 1/96° initial separations yield similar results, with the 1/96° calculation taking notably longer (80 minutes compared to 40 minutes for 1/72°).

Selecting 1/24° (1/72°) as initial separation, δ_0 , resulted in a computational cost of 1 (10) days for a standard computer and for one year of daily snapshots, producing 365 files of 2.25 (19.9) GB in size. For the seven years period 2016-2022, the computation was conducted in parallel, producing files with a total size of 16 (126) GB.

Parameter values selection:
FSLE fields generated in WP3000 will be computed using the following parameter values: **δ_0 = 1/24°** (~4 km) and **δ_0 = 1/72°** (~1 km); **α = 30**; **T= 90 days**, **dt = 1 hour**

2.5.1.4 Finite-time Lagrangian Vorticity (FTLV)

However, these FSLE-LCS give information about the evolving position of the transport barriers or boundaries separating regions with different dynamical properties, and don't provide a characterization of the specific coherent kinematic properties dominating in each of these dynamical regions. Here, in addition to the LCS computed from FSLE, we propose a new methodology based on the Lagrangian assessment of Eulerian flow functions (i.e. vorticity), which provides information on the cumulative effect of the Eulerian observable along trajectories. This approach has the advantage of distinguishing coherent vortex based on well-defined contours of Lagrangian-averaged vorticity deviation (Haller et al, 2016) or hyperbolic structures (i.e fronts, filaments, Hernandez-Carrasco et al., 2018) as well as to classify the flow in regions with different mixing activity (Mezić et al., 2010).

The expression of Finite-time Lagrangian Vorticity (FTLV) used in our computation is based on the objective (i.e. invariant with respect to all Euclidean frame changes) definition of relative dynamic rotation angle, for two-dimensional flows reported in Haller et al, 2016 as:

$$\psi(x,\, y,\, t_0,\, t) \;= \int_{t_0}^{t} (\omega 3(x(t),\, y(t),\, t) \;-\; < \omega 3(t) >) dt, \qquad \text{Eq. 2.5.5}$$

with ω3 the z-component of the vorticity vector and $< \omega 3(t) >$ its spatial mean, which is averaged over a domain that should be chosen large enough so that the averaged

vorticity is representative of the overall mean rotation of the flow under study. For geostrophic flows, the mean rotation is expected to be zero ($< \omega 3(t) >$ =0). While for any general flow the objective and dynamically consistent definition of rotation is given by Eq.2.5.5, the expression of FTLV that will be applied to the 4DMEDSea products is also objective but only for geophysical flows, since the spatial average of the vorticity in these flows is negligible and is expressed as follows:

$$\Omega_T(x, y, t, T) = \frac{1}{T}\int_t^{t+T} \omega 3(x(t), y(t), t)dt, \qquad \text{Eq. 2.5.6}$$

To obtain the FTLV we integrate the observed kinematic variable over the time-evolution of moving water parcels arriving to a specific point. We evaluate the relative vorticity along the particle trajectories computed using the Lagrangian model described above, for two different time periods of integration: T=40 days and T=80 days. Figure 53 shows an example of the FTLV for two different integration time T values.

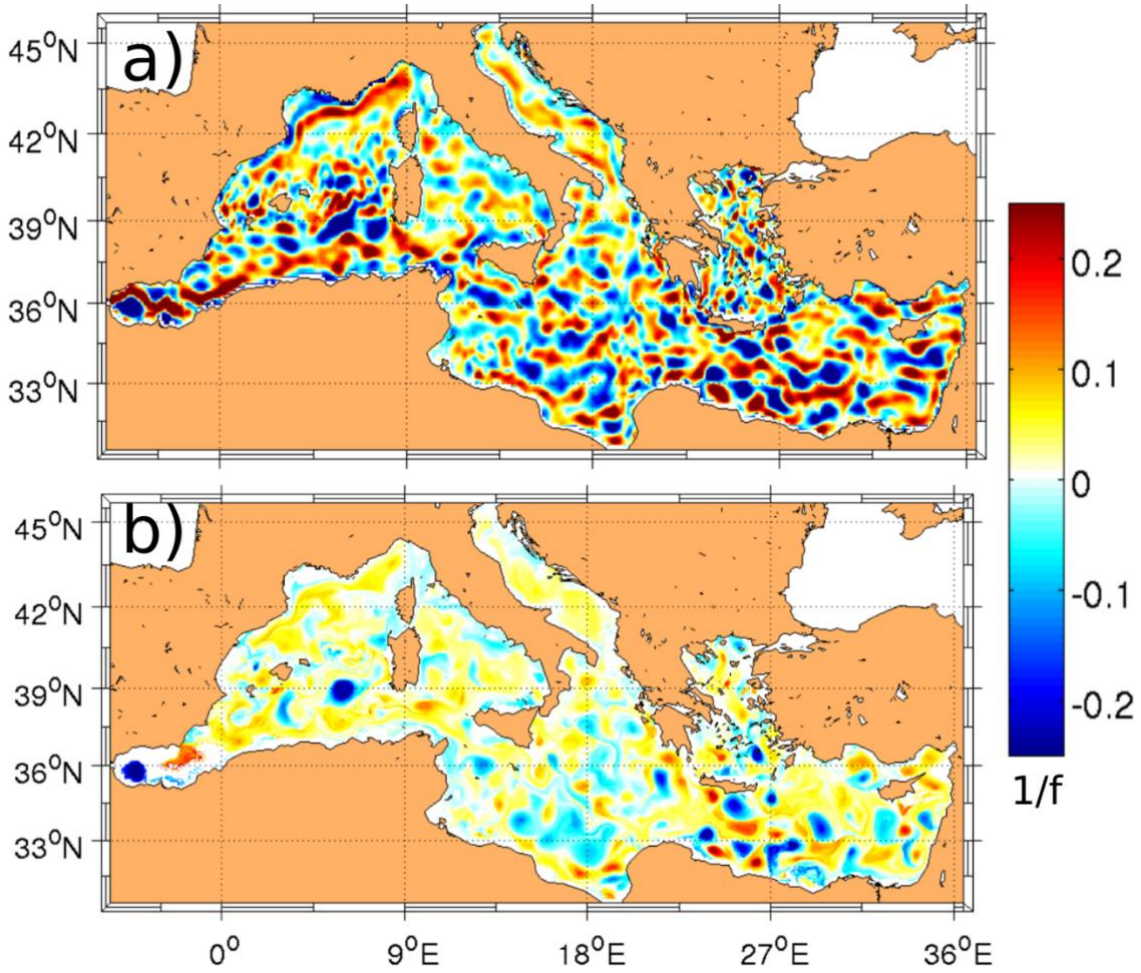


Figure 53: a) FTLV computed for T=0 days averaged over 40 days (Eulerian average from April 18, 2009 to May 28, 2010) b) FTLV computed for T=40 days (Lagrangian average from April 18, 2009 to May 28, 2010).

2.5.2 Lagrangian Flow Network model for connectivity analysis.

2.5.2.1 Lagrangian Flow Network's overview

Lagrangian Flow Network (LFN) is a Lagrangian modeling framework coded in C++ and parallelized (currently running on Lunix operating system) that simulates dispersion or transport in the ocean and the atmosphere (Rossi et al. 2014; Ser-Giacomi et al. 2015). Firstly, the domain defined by the user is subdivided into subareas called box or nodes or cells with the same area and particles are initialized in the domain. Secondly, a fourth order Runge Kutta (RK4) algorithm coupled to a spatiotemporal bi-linear interpolation are used to integrate particle's trajectories and compute their final position. Finally, a connectivity matrix file described in the next section quantifies the exchange between nodes.

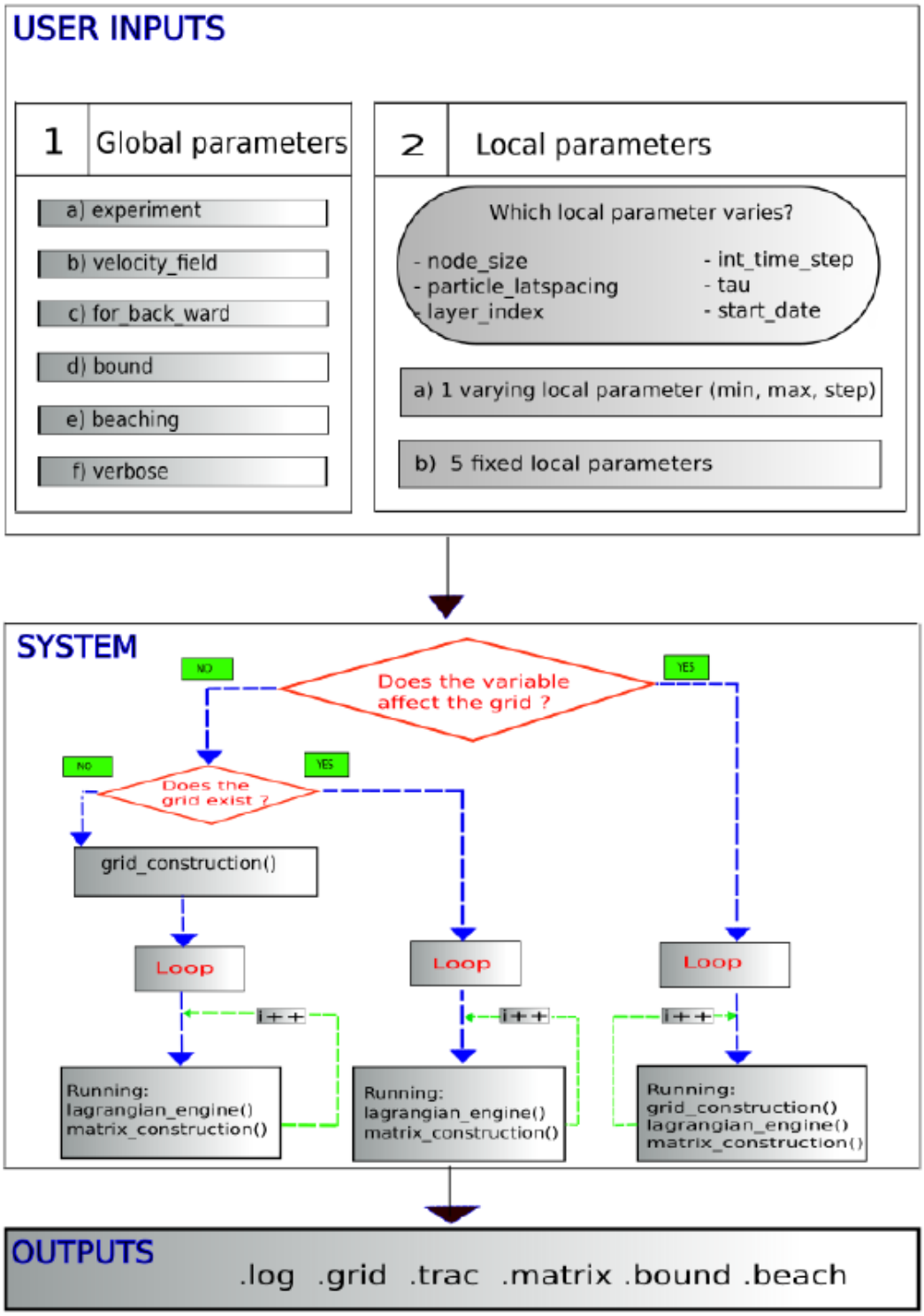


Figure 54: Schematic describing the main architecture of the LFN model.

AMU has been implementing and testing several important developments during the last few months, including (i) improved time precision; nowadays the LFN can be coupled to any flow fields whose temporal frequency range from minute to month (obtained after unifying and upgrading both daily and hourly versions) (ii) added options allowing the model to be ran forward- or backward-in-time.

2.5.2.2 Users input

LFN comprises a total of 12 parameters, which are categorized into two distinct sub-groups: global parameters and local parameters. These parameters are extracted from parameters.dat file within INPUTS subdirectory of the model's root. Additionally, one of these local parameters can vary within a range defined by a tuple in the following form (min, max, range).

The main functions (subroutines) of the LFN code are:

- grid_construction(): build regular grid and initialize particles. Particles must be in a node as well as in the grid of velocity field.
- lagrangian_engine(): read grid and initial positions of tracers before integrating trajectories of particles along the domain using RK4 algorithm and a spatiotemporal bi-linear interpolation of velocities in any point inside the domain.
- matrix_construction(): read initial & final positions of tracers and calculate the connectivity matrix characterizing the exchange between nodes.

2.5.2.3 Coupling of the LFN with the new 4DMED products (WP2520)

Following the delivery of the first sample product (covering years 2016-2022) entitled “ESA 4DMED-SEA - Mediterranean Sea Gridded L4 Sea Surface Heights And Derived Variables", at 1/24° last February 12, 2024 by CLS, AMU has been downloading the files and is currently working on coupling this new flow field with the LFN.

LFN model has the ability to detect automatically some of the main specifications of flow field, while other are entered manually. The main specifications concern:

- ncdir: velocity field’s path.
- name_file: file’s name format.
- cNc: netcdf temporal coverage in minutes.
- pNc: netcdf period in minutes.
- sorigin_date: epoch time origin (in format dd-mm-yyyy-hh-MM).
- name_time: time coordinate name.
- name_lat: latitude coordinate name.
- name_lon: longitude coordinate name.
- name_depth: depth coordinate name.
- name_u: magnitude u name.
- name_v: magnitude v name.
- name_missing_value: missing value name.
- name_scale_factor: scale factor name.

First outputs generated by the LFN fed by the ESA 4DMED-SEA - Mediterranean Sea Gridded L4 Sea Surface Heights And Derived Variables will be presented at the next

internal meeting planned on April 5, 2024. Some exemplary results are included in the deliverable D.3.2 leaded by CSIC.

2.5.3 Bio-Lagrangian backtracking and Kinematic Lagrangian Model for regional studies

Lagrangian simulations and the subsequent backtracking analysis of biological tracers will be computed by numerical integration of velocity fields defined on a 3D spatial grid. For ocean surface transport simulations, only the upper layer is considered. The general Lagrangian equation for a passive particle trajectory $\vec{r} = (x, y)$ is:

$$\frac{d\vec{r}}{dt} = \vec{v}_{LS} + \vec{v}_{SG} \qquad \text{Eq. 2.5.6}$$

The term $\vec{v}_{LS}$ is the large-scale velocity that will be obtained from 4DMED currents, with a finite spatial and temporal resolution (e.g., daily fields with spatial grid step l_{grid} of some kilometers). This term is subject to space and time interpolation at particle position: time interpolation consists of a simple linear scheme; spatial interpolation is carried out by means of a classic cubic-spline algorithm. Unresolved dynamics below l_{grid} is replaced by the sub-grid term $\vec{v}_{SG} = \vec{v}^{(KLM)}$. This term comes from a multi-scale kinematic Lagrangian model (KLM) which takes into account the effects of small-scale turbulent motions on particle trajectories. From the n-mode stream-function:

$$\Psi_{KLM}(x, y, t) = \sum_{i=1}^{n} (A_i / k_i)\ \sin(k_i(x - \alpha_i \sin(\omega_i t))) \sin(k_i(y - \beta_i \sin(\omega_i t + \varphi))),$$

Eq. 2.5.7

in which: $l_i = \frac{2\pi}{k_i}$ are the spatial wavelengths, A_i are the corresponding velocities, ω_i are the oscillation pulsations of the kinematic field, α_i and β_i are the oscillation amplitudes of the kinematic vortices, φ is a phase difference between the two oscillation directions, the KLM velocity components are derived according to the Hamiltonian formalism:

$$v_x^{(KLM)} = \frac{\partial \Psi_{KLM}}{\partial y}, \qquad v_y^{(KLM)} = -\frac{\partial \Psi_{KLM}}{\partial x} \qquad \text{Eq. 2.5.8}$$

The KLM set-up is calibrated on observational Lagrangian data (ocean surface drifters), i.e., the key parameters are the range of spatial scales (l_{min}, l_{grid}) that covers the unresolved velocity modes down to the minimum spatial scale available to observation, and the equivalent turbulent dissipation rate that defines the kinematic scale-dependent velocities $v^{(KLM)}(l) = (\varepsilon l)^{1/3}$, e.g., $\varepsilon \sim 10^{-9}\, m^2 s^{-3}$ is a typical value for Mediterranean sea surface turbulence directly measured from real drifting buoys relative dispersion. The two velocity fields $\vec{v}_{LS}$ and $\vec{v}_{SG}$ are decoupled, so that the two contributions to trajectory evolution can be computed separately, i.e. the KLM sub-grid increment is computed by means of a standard 4th order Runge-Kutta algorithm. All these properties hold for both forward and backward integrations. In most cases, small-sized bio-chemical tracers can be approximated to passive Lagrangian tracers. Parallel to (forward/backward) trajectory integration, additional indicators can be considered to characterize the type of tracer under analysis, e.g. the integral Chl content encountered by larvae trajectories along their pathway (Falcini et al., 2020).

For eventual, additional scientific studies that need to account for small-scale vertical mixing, a 3-D lattice of convective cells can be introduced. This model simulates chaotic Langmuir cell-like vertical convection and is adopted because often such a vertical mixing process overwhelms the action of the small-scale turbulence. We define an incompressible 3-D velocity field of components (*u*, *v*, *w*), as the curl of a potential vector Φ, of components (Φ_1, Φ_2, 0), given by:

$$\Phi_1 = A\hat{k}\sin\left[k\left(x - \varepsilon sin(\omega t)\right)\right]\sin\left[\hat{k}\left(z - \varepsilon sin(\omega t)\right)\right]$$

$$\Phi_2 = A\hat{k}\sin\left[k\left(y - \varepsilon sin(\omega t)\right)\right]\sin\left[\hat{k}\left(z - \varepsilon sin(\omega t)\right)\right] \qquad \text{Eq. 2.5.9}$$

where *A* is the velocity scale, $k = \frac{2\pi}{\lambda_0}$ is the horizontal wave number associated to the wavelength λ_0 of the flow, $\hat{k} = 2k$ is the vertical wave number assumed to be twice the horizontal wave number for isotropy, $t = \frac{\lambda_0}{A}$ is the convective time scale; ε and ω are amplitude and pulsation of the time-dependent oscillating terms. To account for the suppression of vertical convection below the mixed layer, both stream functions Φ_1 and Φ_2 can be multiplied by a damping term, e.g., $\Upsilon(z) = e^{-\frac{|z|}{\eta}}$, which is of the order of the mixed layer depth. By definition, the three components of the velocity field are: relaxed exponentially to zero at depths *z* much larger than η, where η is a length scale of

$$v_x^{(KLM\ 3D)} = \frac{\partial \Phi_1}{\partial z}, \qquad v_y^{(KLM\ 3D)} = -\frac{\partial \Phi_2}{\partial z}, \qquad v_z^{(KLM\ 3D)} = -\frac{\partial \Phi_1}{\partial x} - \frac{\partial \Phi_2}{\partial y}$$

Eq. 2.5.10

Finally, to mimic diurnal vertical migration (DVM) of anchovy larvae, in the general Lagrangian equation we will use the term:

$$\dot{z}(t) = -\gamma[z(t) - z_0], \qquad \text{Eq. 2.5.11}$$

where the value of the transfer rate γ is chosen in order to let the vertical positions of the larvae relax to a depth z_0 with reasonable velocities.

An important technical aspect to consider is that the kinematic field must be computed in a quasi-Lagrangian reference frame, i.e. the center of mass of a pair of trajectories evolving simultaneously. This is needed in order to avoid the well-known "sweeping effect" problem that otherwise would make the kinematic simulations ineffective (Palatella et al., 2014).

2.5.4 II-Kind FSLE analysis for velocity fields intercomparison

The FSLE (Finite-Scale Lyapunov Exponent) is a well-established dynamical system analysis technique to study relative dispersion of Lagrangian tracer trajectories. It is defined by:

$$\lambda(\delta) = \frac{\ln r}{\langle \tau(\delta) \rangle} \qquad \text{Eq. 2.5.12}$$

where $\langle \tau(\delta) \rangle$ is the average growth time of the separation between two trajectories from scale d to scale rd with r > 1, computed in a given scale range $[\delta_{min}, \delta_{max}]$. If d refers to

the separation between homogeneous trajectories, i.e. belonging to the same dynamical system, the FSLE is said of the I-kind. Several physical characteristics of the dispersion process can be obtained from a I-kind FSLE analysis: the scaling exponent in a function of the form $\lambda(\delta) \sim \delta^{-\nu}$ is related to the dominant dispersion mechanism (e.g., chaos, turbulence, diffusion, etc.) in a given scale sub-range; the best-fitting parameter of a given scaling law to the data provides information on the order of magnitude of the characteristic physical quantity related to the specific regime under consideration (e.g., Lagrangian Lyapunov exponent, turbulent dissipation rate, diffusion coefficient, etc). Moreover, in case of strongly anisotropic systems, the I-kind FSLE analysis can be split into components, to have a different characterization of the dispersion properties along each independent direction (e.g., zonal and meridional). If d, instead, refers to the separation between two heterogeneous trajectories, i.e. evolving in two different dynamical systems, the FSLE analysis is called of the II-kind. Generally, the two dynamical systems represent "real" vs "model" dynamics, or "model A" vs "model B" dynamics. For this reason, the II-kind FSLE analysis turns out to be a powerful and rigorous comparison tool between test and reference trajectories, e.g., in the framework of Lagrangian validation procedures (Lacorata et al., 2019). There is no special algorithm to be implemented for the numerical computation of I-kind and II-kind FSLE; it is only needed to follow the correct theoretical definition.

As an example, here we report results from the I- and II-kind FSLE analysis performed from a set of $\simeq$ 5 10^4 numerical trajectory pairs, generated from ESA GlobCurrent Geostrophic velocity fields (GCG) at the surface, for the period 2014–2015 in the Mediterranean basin (Figure 55). Numerical simulations were also performed by using the coupled GCG + KLM (kinematic Lagrangian model) model, with N_m = 6 kinematic modes distributed in the model inertial range $[l_{N_m}/2, l_1/2]$ with a density factor $a = \sqrt{2}$, turbulent dissipation rate $\epsilon \simeq$ 10–9 $m^2 s^{-3}$, maximum and minimum spatial wavelengths l_1 = 120 km and l_{N_m}=20 km, respectively. Let us recall that the kinematic eddy size is half the corresponding wavelength, and that the function of the KLM is to restore the missing contributions to the mean dispersion rates in the mesoscale range.

For this analysis, the "real" dynamics is represented by the product INSITU_MED_PHYBGCWAV_DISCRETE_MYNRT_013_035, from Copernicus Marine Service (https://doi.org/10.48670/moi-00044), where a total number of ~104 of simultaneous drifter pairs were identified, 330 of which were found with separation below a threshold Δ = 5 km (Lacorata et al., 2014). Time gaps in the data were compensated by cubic spline interpolation. Time sampling is 1 h.

From Figure 55 it is worth noting that the error threshold scale δ*, above which FSLE-I and FSLE-II overlap, is significantly reduced from $\simeq$ 80 km, in the GCG case, to $\simeq$ 40 km, in the GCG + KLM case (Lacorata et al., 2019).

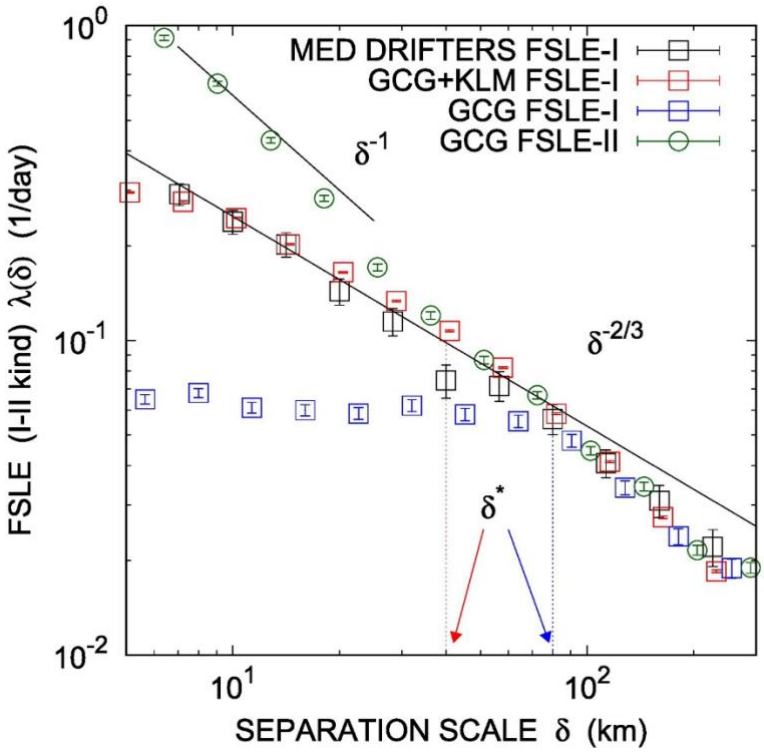


Figure 55: FSLE (I and II) analysis for the Mediterranean Sea. The scaling exponent (−2/3) corresponds to the Richardson's regime. Error threshold scale is $\delta^* \simeq 80$ km, for GCG, and $\delta^* \simeq 40$ km, for GCG + KLM. The scaling exponent (−1) corresponds to uniform linear error growth.

3 Product Validation Plan

3.1 2DMED ALT product validation

To assess the impact of the specific development choices in terms of input data and mapping solutions, we designed several Observing System Experiments (OSE) and Observing System Simulation Experiment (OSSE). An OSE or an OSSE serves as a method for evaluating the influence and efficacy of different observational systems or mapping solution in enhancing the accuracy of targeted products. In an OSE, data can be selectively included or excluded to analyze their impact on the system's performance. By comparing system reconstructions with and without particular observational data, we can pinpoint areas where mapping techniques can be improved or optimize the design of future observational networks or mapping methodologies.

The validation metrics are based on statistical and spectral analysis. One quantitative assessment is based on the comparison between SSH maps and independent SSH along-track data (Figure 56). This diagnostic follows three main steps: (1) the SSH gridded data are interpolated to the locations of the independent SSH along-track, geo-referenced by their longitude, latitude, and time; (2) the difference $SSH_{error} = SSH_{map} - SSH_{along\text{-}track}$ is calculated; and (3) a statistical analysis on the SSH_{error} is performed in 1° longitude × 1° latitude boxes. Prior to the statistical analysis, a filtering operation can be applied to isolate the spatial scales of interest. For example, the analysis can be performed over the spatial range [65–200 km] typically representative of the short length scale mesoscale ocean signal. In the studies, the validation metric is based on the error variance (σ_{err}) or root mean squared error (RMSE) scores in 1° longitude × 1° latitude boxes defined as:

$$\Delta SSH(i) = SSH_{obs}(i) - SSH_{reconstruction}(i)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\Delta SSH(i))^2}{N}} \qquad (M4, see \text{ Table } 10)$$

$$\sigma err = \frac{\sum_{i=1}^{N}(\Delta SSH(i) - \overline{\Delta SSH(i)})^2}{N} \qquad (M5, see \text{ Table } 10)$$

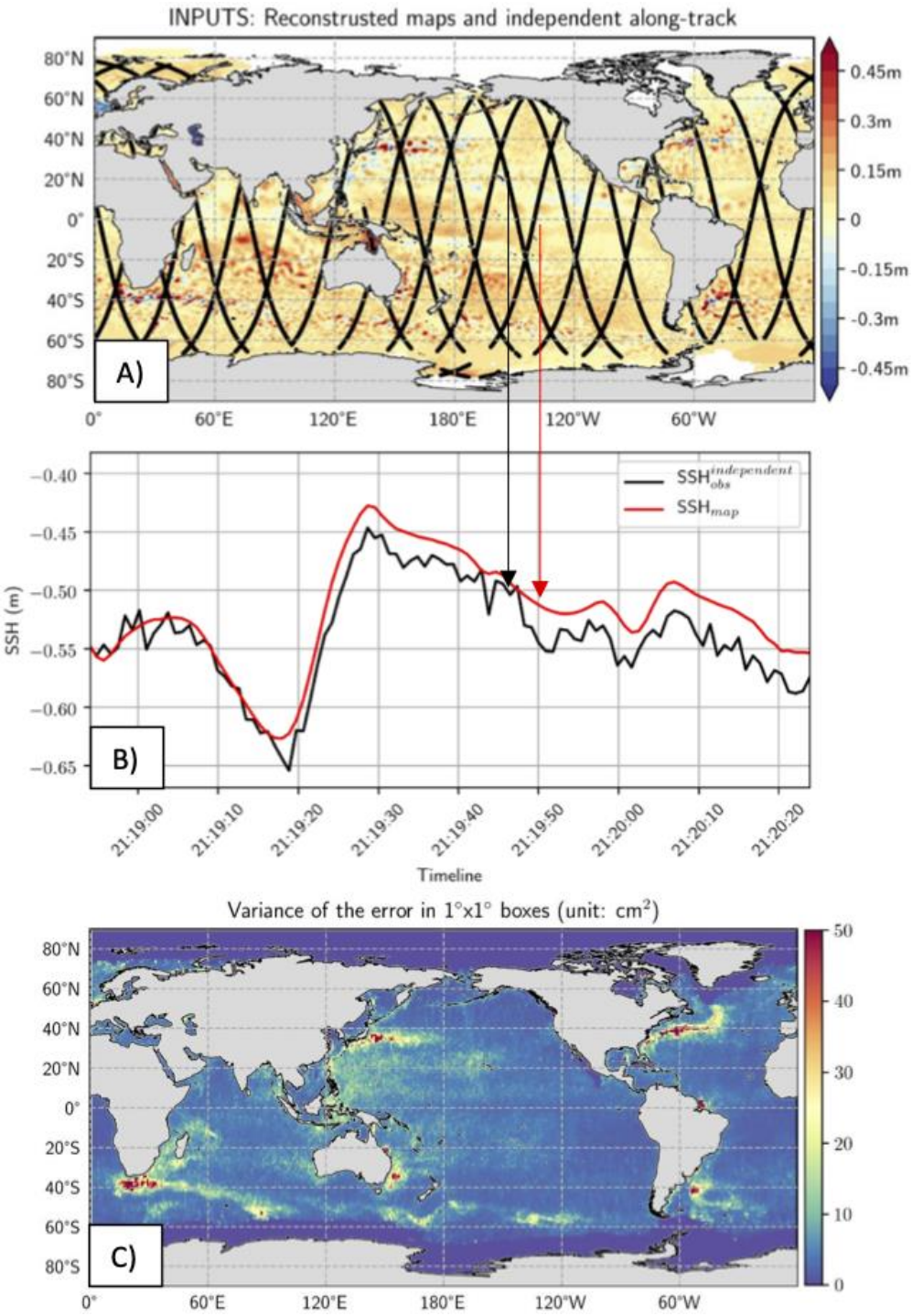


Figure 56: Illustration of the comparison maps vs independent along-track method: A) inputs SSH gridded and along-track fields, B) colocation gridded SSH and along-track SSH; and C) maps of statistical analysis

For a more descriptive assessment by wavelength and to avoid spatio-temporal filtering of independent and study datasets, diagnostics can be performed in frequency space, using spectral analysis of SSH altimetry and gridded datasets. More specifically, a spectral analysis can be applied to altimetry data to estimate the effective resolution of gridded SSH products (Figure 57). It is described, for example, in Ballarotta et al. (2019). It is based on the ratio between the spectral content of the mapping error and the spectral content of independent signals:

$$\mathrm{NSR}(\lambda s) = \frac{S_{\mathrm{diff}}(\lambda s)}{S_{\mathrm{obs}}(\lambda s)} = 0.5 \qquad (M6, see\ \text{Table } 10)$$

where $S_{\mathrm{diff}}(\lambda s)$ is the power spectral density of the mapping error and $S_{\mathrm{obs}}(\lambda s)$ is the power spectral density of the independent SLA signal. The effective resolution is then given by the wavelength λs where the NSR(λs) is 0.5.

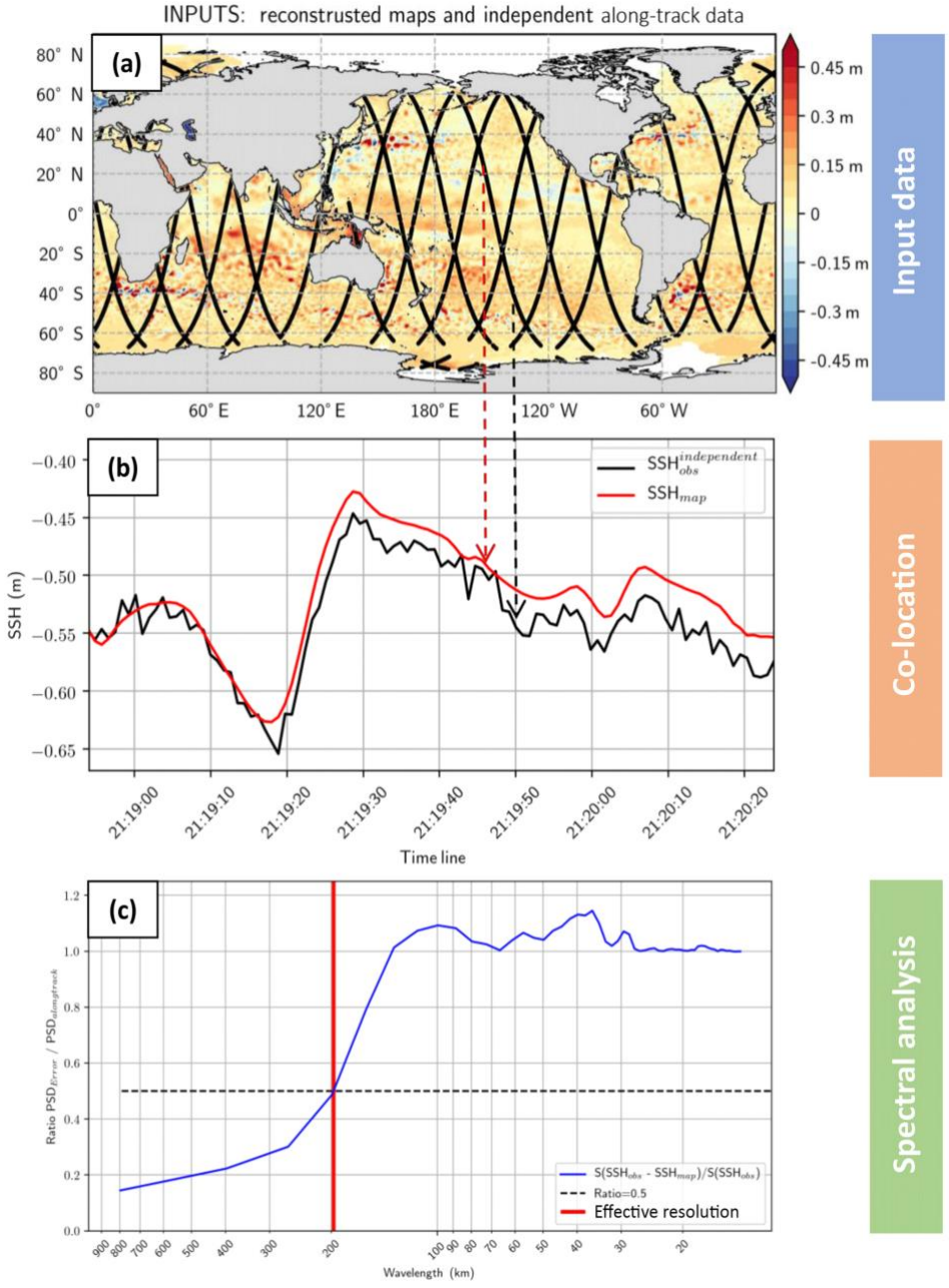


Figure 57: Illustration of the comparison maps vs independent along-track method: A) inputs SSH gridded and along-track fields, B) colocation gridded SSH and along-track SSH; and C) spectral analysis

In the previous diagnostic the validation is undertaken on the SLA fields. To complement our analysis, a comparison of reconstructed geostrophic current maps with independent drifter geostrophic velocities is undertaken and proposed in the data-challenge (Figure 58). Geostrophic currents are computed from the derivative of the SSH maps produced by the mapping method. These gridded zonal and meridional geostrophic currents can

then be compared to the total currents of the in-situ drifters' instruments. Like the comparison between maps and along-track, the gridded zonal and meridional components are interpolated onto the drifter's paths. Each component is then compared to the velocity component seen by the drifters. The statistical results are presented in box-averaging statistics maps or timeseries (https://2024c-dc-4dmedsea-esa.readthedocs.io/en/latest/2_eval_generic/overall_driftereval.html).

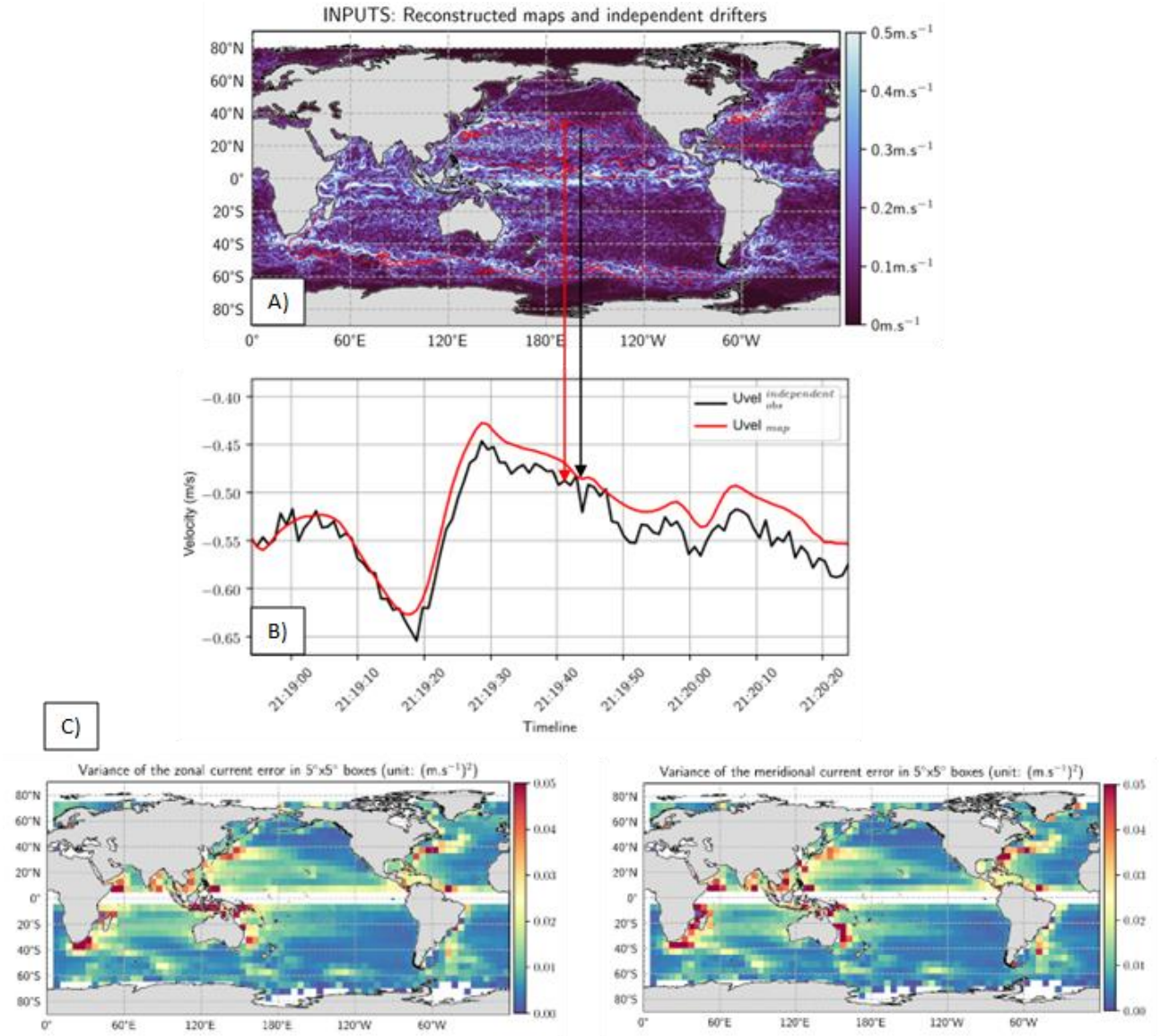


Figure 58: Illustration of geostrophic currents maps vs in-situ drifters geostrophic components comparison method: A) currents intensity maps and independent drifters' path (in red); B) colocation gridded currents and drifters' data and C) maps of variance errors

To complement the previous analysis, we develop a second performance assessment which involves a Lagrangian approach (see https://2023a-ssh-mapping-ose.readthedocs.io/en/latest/5_metrics_det/metrics_driftertraj.html). This approach involves comparing simulated drifter trajectories with their real counterparts. At each time point and for every drifter, we quantify the disparities (distance of separation) between its actual position and the positions projected using evaluated gridded velocity

fields. Consequently, we compute the distances between simulated and real drifter locations, providing insights across various time horizons. These distance metrics can be visualized through averaged grid box maps (Figure 59) or basin averages. This method is similar to the one implemented in Le Guillou et al. (2023).

The distance is computed using the great-circle distance formulation between two sets of coordinates (lon1, lat1) and (lon2, lat2):

$$\text{distance} = R\,.\,c \qquad (M7, see\ \text{Table 10})$$

with R the Earth radius and c:

$$c = 2 * arctan2\left(\sqrt{\left(\sin\left(\frac{lat1 - lat2}{2}\right)^2 + \cos(lat1) * \cos(lat2) * \sin\left(\frac{lon1 - lon2}{2}\right)^2\right)},\right.$$

$$\left.\sqrt{\left(1 - sin(\frac{lat1 - lat2}{2})^2 + cos(lat1) * cos(lat2) * sin(\frac{lon1 - lon2}{2})^2\right)}\right)$$

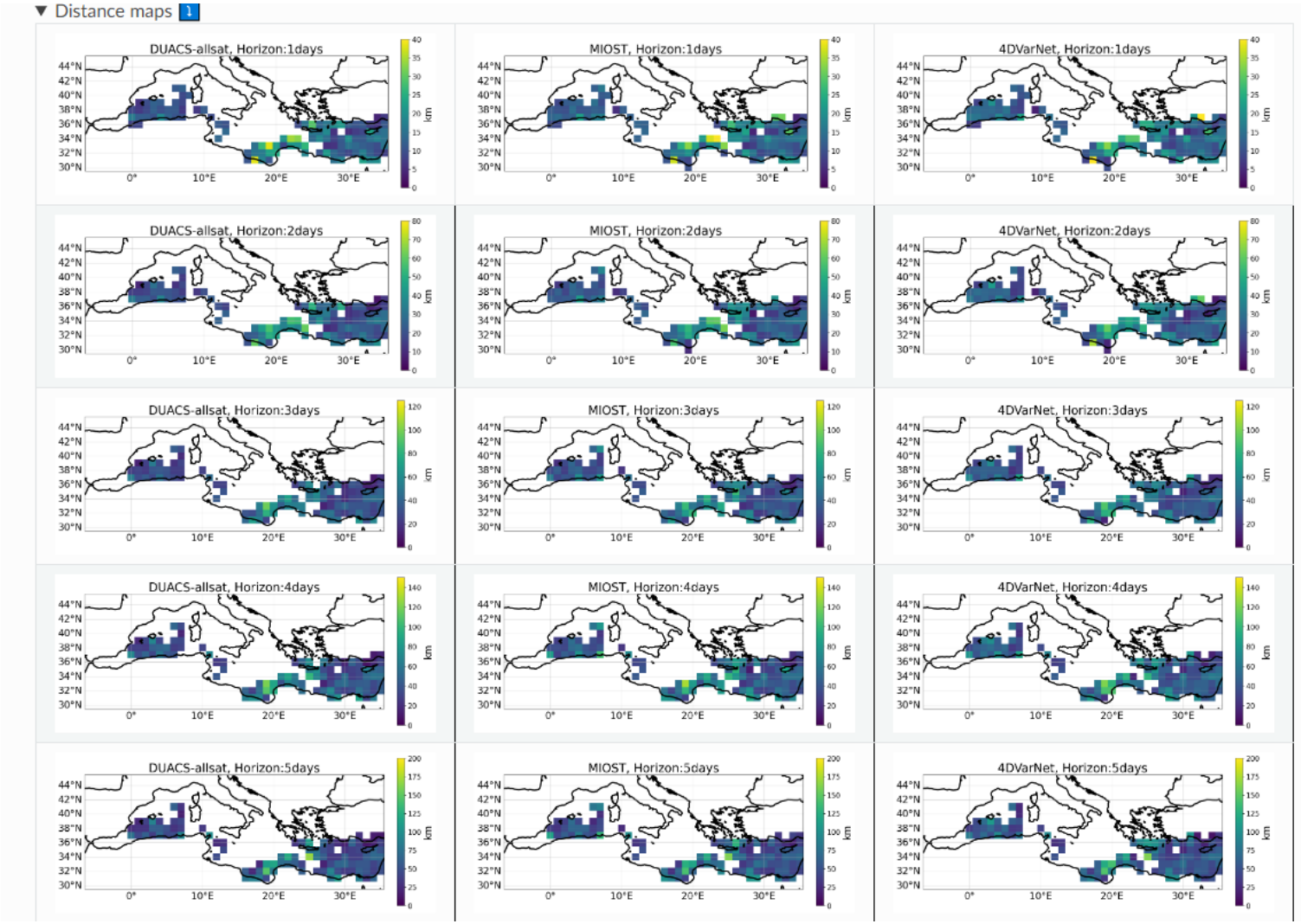


Figure 59: Distance maps between real drifters trajectories and simulated trajectories from DUACS (left), MIOST (middle) and 4dVArNET (right) and at different time horizons between 1 day (top) to 5 days (bottom).

3.2 Lagrangian metrics for 2DMEDSea velocity field validation

3.2.1 Averaged Lagrangian Separation distance (D)

We compare synthetic drifter trajectories integrated in the geostrophic velocity field and the real drifter trajectories by computing the distance (D) between each virtual drifter trajectory and the real one. The integration of virtual drifter trajectories is initialized at the same position of the real drifters and then the distance between real-virtual drifter pairs is computed each 1 day, stopping after 15 days of integration. We reinitialize the virtual drifter position at the position of the real drifter every 1 day. Synthetic trajectories are computed using the 4th order Runge-Kutta integration scheme and bilinear interpolation. Values of D as a function of t should be compared to the mean drift separation over the corresponding area of the drifter motion. This provides an idea about how energetic is the dynamic over the region.

3.2.2 Lagrangian Skill Score (SS)

Following the approach proposed by Liu and Weisberg (2011), the Lagrangian Skill Score (SS) is computed based on the cumulative Lagrangian separation distances between the endpoints of virtual and observed drifter trajectories, normalized by the associated cumulative trajectory lengths, to assess the performance of the experimental 4DMEDSea velocity field through particle trajectory evaluations (as schematized in Figure 60). This dimensionless index is used to define the trajectory model skill score (SS) as following:

$$SS = 1 - \left(\frac{\sum_{i=1}^{N} d_i}{\sum_{i=1}^{N} l_{oi}}\right)\left(\frac{1}{p}\right), \qquad \text{(M8, , } see \text{ Table 10)}$$

where d_i, is the separation distance between the virtual and observed endpoints of the Lagrangian trajectories at i time step after the initialization (virtual particle release), l_{oi} is the length of the observed trajectory and N is the total number of time steps. The higher skill score value, the better the model performance, with a value=1 implying a perfect fit between observation and simulation.

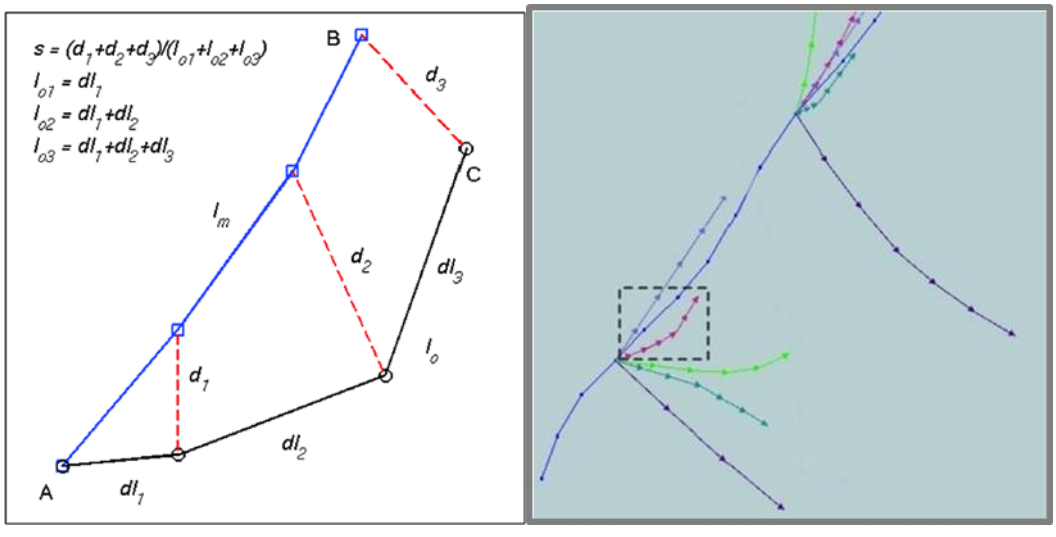


Figure 60: Illustration of the separation distances between observed drifter data (blue lines) and modeled (black lines) endpoints of Lagrangian trajectories (A-B and A-C, respectively), adapted from (Liu and Weisberg, 2011) -left panel- an application of the SS to different modeled trajectories over the observed one (blue line) -right panel-. In this case, the SS was calculated every six hours.

An example of application of the SS is shown in Figure 61. It results of the computation of the SS of the DUACS L4 multimission altimetry product with 29 drifters deployed in the Alboran sea in 2016. This figure of SS allows identifying over what geographical regions or under what dynamical conditions the new product performs better.

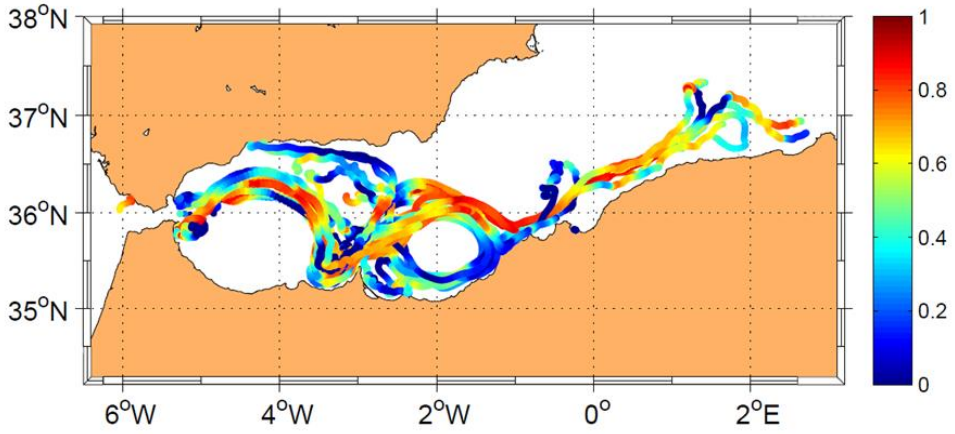


Figure 61: DUACS L4 – Altimetry skill score after 15 days of trajectories. A total of 29 real drifter trajectories have been used in this validation

3.2.3 I and II-Kind Finite-Scale Lyapunov Exponent

As explained in section 2.5, these two Lagrangian metrics can be used to evaluate the dynamical scales resolved by the new altimetry products. Both I and II – FSLE allow for separating nonlocal from local stirring and, in the case of the 4DMEDSea velocity fields, the scales of these new products that dominate the dispersion. Comparing the scaling laws exhibited in the FSLE spectrum for the different altimetry 4DMEDSea products (i.e. MIOST at 1/24°, 4DVarNet at two resolutions: 1/8° and 1/20°) and real drifters one can infer the size of the oceanic structures that are involved in the particle dispersion and the dynamical scales that can be resolved.

3.2.4 In situ data sets for validation collected from multiplatform experiments in the western Mediterranean

3.2.4.1 PRESWOT

The objective of this campaign was to carry out a multi-platform experiment south of the Balearic Islands to improve our understanding of mesoscale and sub-mesoscale dynamics and its impact on biogeochemical processes in an area of high oceanographic variability. The exact sampling area was determined according to the presence of filaments, whirlpools or ocean fronts from satellite images, covering a sampled area shown in Figure 62. The campaign was divided into 2 Legs, the first from May 6 to May 10, and the second from the 13th to the 16th May 2018. In situ systems, including 2

gliders, 12 drifters, CTD, ADCP, and more than 2000 water samples, were collected to determine physical and biochemical ocean variability.

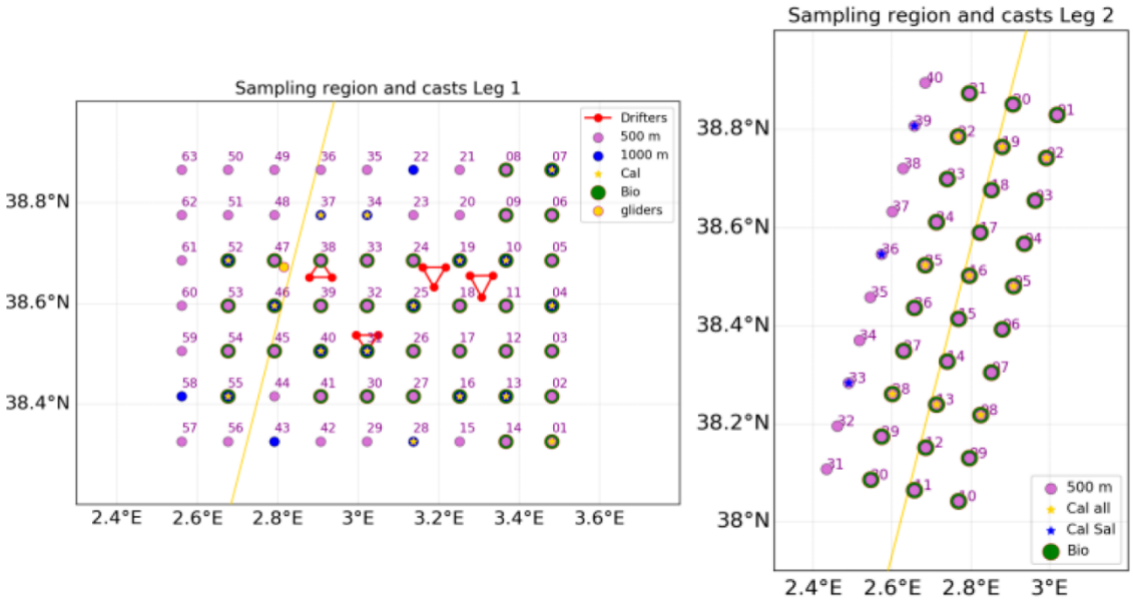


Figure 62: PRESWOT Leg 1 (left) and Leg 2 (right) rosette casts (Barcelo-Llull et al., 2021). Yellow track represents the track of Sentinel 3 on 13-May-2018

- **Salinity, temperature and density** from CTD data are distributed in two regular meshes of different resolution. Maximum depths of the CTD casts were 500 m. CTD fields were objectively interpolated onto a regular grid with a horizontal resolution of 2 km and a vertical resolution of 5 m with correlation scales Lx=Ly=20km.
- **Geostrophic velocities** derived from the CTD measurements will be used to validate the ALT 2DMED. Dynamic height was inferred using density data objectively interpolated with a correlation scale of 20 km (see Figure 63).
- **12 SVP drifters** with a drogue centered at 15 m were deployed based on the location of fronts. To remove the inertial signals, the drifter trajectories were filtered using a fifth-order Butterworth low-pass filter with a cutoff of 1.5 times the inertial period of the area, which corresponds to 19 hours.

During both Legs the **two gliders** equipped with CTD and **fluorimeter** followed the transect shown in Figure 63.

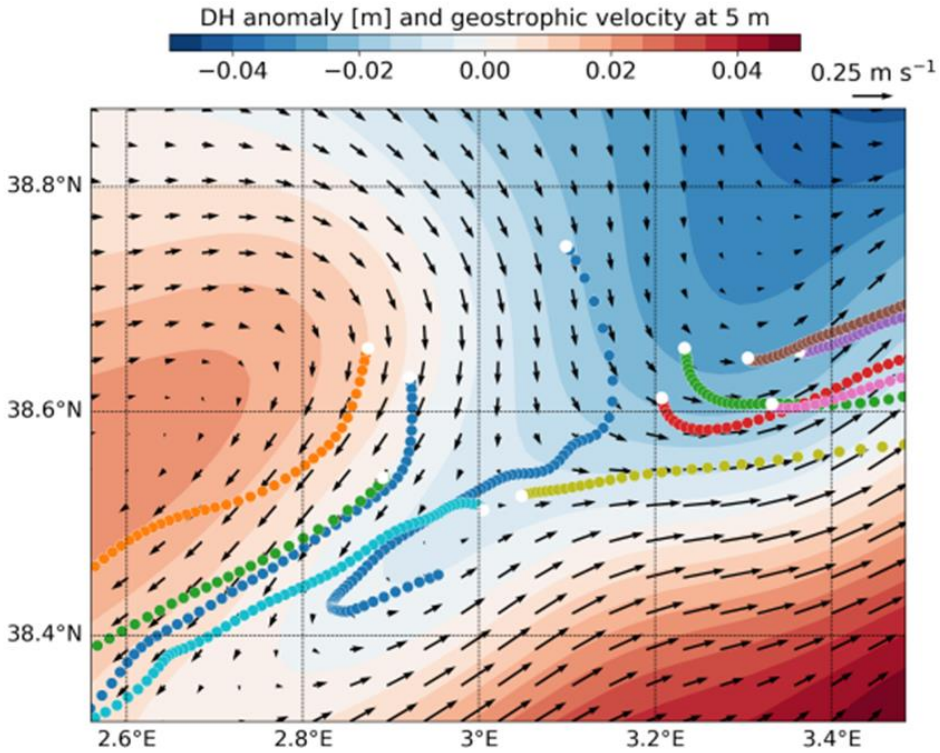


Figure 63: Dynamic height anomaly at 5 m depth inferred using Leg 1 density data objectively interpolated with Lx=Ly=20km. We assume a level of no motion at 1000 m depth. Geostrophic velocity vectors at 5 m depth are estimated through thermal wind balance. Dots show the trajectories of 12 SVP drifters deployed during the PRESWOT campaign

Raw data were post-processed and quality controlled. Further details of the CTD and ADCP data calibration and processing please can be found in Barcelo-Llull et al 2021. Data processed are available at: https://digital.csic.es/handle/10261/182615.

3.2.4.2 CALYPSO

We use a large drifter data set in the Alboran Sea as part of the ONR (Office of Naval Research) Departmental Research Initiative, "CALYPSO" (Coherent Lagrangian Pathways from the Surface Ocean to Interior) campaign whose goal is to better understand the three-dimensional pathways of Lagrangian particles in the ocean. A total of 82 drifters drogued at three different depths were deployed. Drifter data are divided into surface (0.6 and 1 m) and near-surface (15 m) drifters.

- 35 CARTHE drifters. The Consortium for Advanced Research on Transport of Hydrocarbon in the Environment (CARTHE) consists of a buoy attached to a drogue that extends 60 cm below the surface. A total of 35 CARTHE drifters were deployed and transmitted their GPS position every 5 min.
- 14 CODE drifters. The Coastal Ocean Dynamics Experiment (CODE) drifters follow the top 1-m average current. CODE drifters transmitted their position every 10 min.
- 33 SVP drifters. The Surface Velocity Program (SVP) drifters consist of a surface buoy attached to a drogue centered at 15 m depth. During the experiment transmitted their GPS position every 5 min.

The data set, including all three drifter types, was processed to remove spikes in velocity and acceleration, interpolated to 5 min, and filtered with a 1-h hamming window (Tarry et al.,2021). Data sets are available at https://zenodo.org/records/4592311

3.2.4.3 CANALES

The Canales missions at the Balearic Islands Coastal Ocean Observing and Forecasting System (SOCIB) started in 2011 with the objective to monitor the circulation of the Mallorca and Ibiza channels. The Canales project provides glider data of the Balearic Island channels with a maximum time gap between missions of 1 month. This long-term monitoring program has collected 55 transects of glider data across the Mallorca channel during 8 years of glider missions distributed equally by seasons and covering all months. Measured variables include pressure, temperature, and salinity. Raw glider data were processed with the SOCIB Glider Toolbox, which includes thermal lag correction and quality control. In the Mallorca channel there is a total of 5,216 profiles (maximum depth of 958 m) with a horizontal resolution that changes with bathymetry, and which is approximately 2 km in the deepest part of the 65-km-wide channel. Each transect took on average 2.8 days to be completed. Potential temperature and practical salinity data (hereinafter temperature and salinity, respectively) have been interpolated onto a vertical resolution of 5 m and a horizontal resolution along the glider track line of 2 km. Dynamic height profiles have been assuming a reference level of no motion at 800-m depth or the seafloor where profile depth is shallower. Only transects with at least one profile deeper than the reference level have been considered for this computation; this results in 50 transects from a total of 55 with the reference level at 800-m depth. To infer the geostrophic velocity perpendicular to the transect through thermal wind balance, the dynamic height was previously smoothed with a Loess filter considering a spatial scale of 15 km. Data are available at: https://www.socib.es/?seccion=observingFacilities&facility=glider

3.3 2DMED SSS product validation

The assessment of the new experimental 2DMED SSS product is carried out by using an independent dataset of Thermosalinograph (TSG) data covering the Mediterranean Sea from 2017 to 2018. In order to evaluate the improvement of the optimized algorithm, the performances of the new 2DMED SSS have been compared also with the previous Mediterranean SSS daily product developed in Sammartino et al. 2022 (here after MED LR) at 1/16° grid of resolution and with the Global weekly SSS product (hereafter GLO LR) at ¼° grid resolution (MULTIOBS_GLO_PHY_S_SURFACE_MYNRT_015_013, version 8.0).

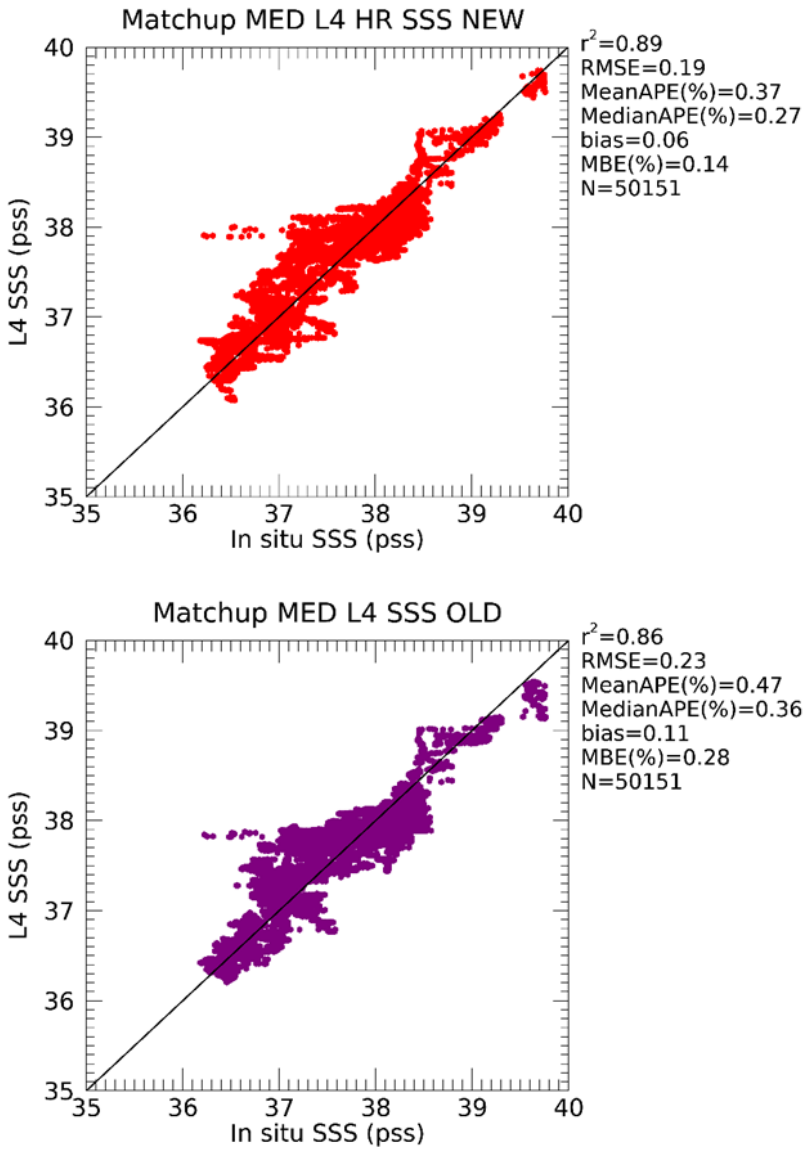
3.3.1 In situ data used for validation

The independent in situ dataset used for the assessment of the High Resolution 2DMED SSS product comprises the Global Ocean Surface Underway Data (GOSUD) thermo-salinometer (TSG) observations. It is based on TSG data collected by French research

vessels, since the early 2000 (Kolodziejczyk, et al. 2021). This dataset is carefully calibrated, updated annually, and is made freely available after a delayed-mode processing [R2]. For our analysis, only adjusted TSG data from 2017 to 2018 with quality flags = 1 have been extracted and matched up with the 2D salinity fields.

3.3.2 Scientific analysis

The first comparison of the 2DMED SSS product with the previous MED LR and GLO LR are shown in Figure 64. This figure shows the scatterplot of the in situ SSS vs. the 2DMED SSS, MED LR and GLO LR fields and the statistical results of this independent validation [M2]. Here, the data are mostly distributed along the 1:1 line, with some exceptions for low salinity concentrations. As highlighted from the statistics reported in each panel, the new 2DMED outperforms the other datasets. In the optimized model, the errors are reduced with a RMSE of 0.19 for the new 2DMED SSS with respect to the RMSE=0.23 and RMSE=0.25 of the MED LR and GLO LR, respectively.



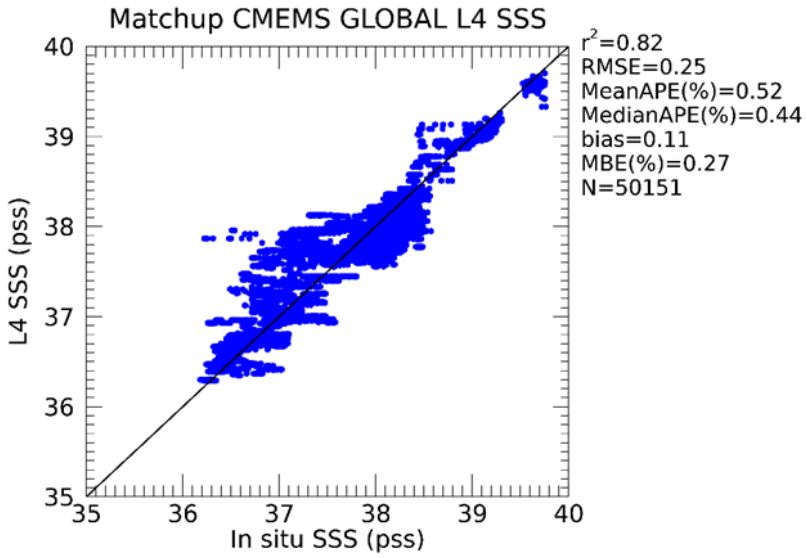
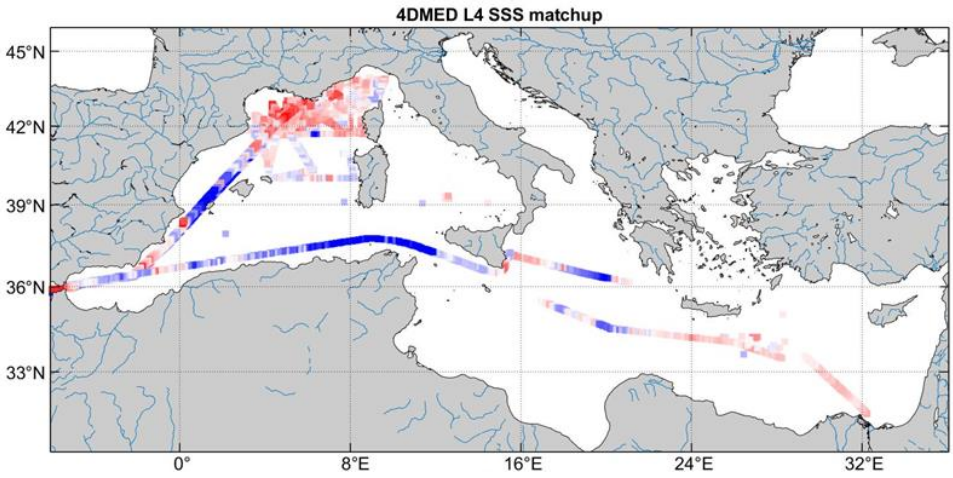


Figure 64: Scatterplots of the validation of multi-observation products against the independent in situ dataset for the years 2017-2018. Comparison of the new 2DMED SSS product (red dots), with MED LR of Sammartino et al. (2022) (purple dots) and GLO LR v8.0 (blue dots).

To spatially visualise the differences between in situ and each L4 SSS dataset, in all validation matchups, three different maps have been created and given in Figure 65.



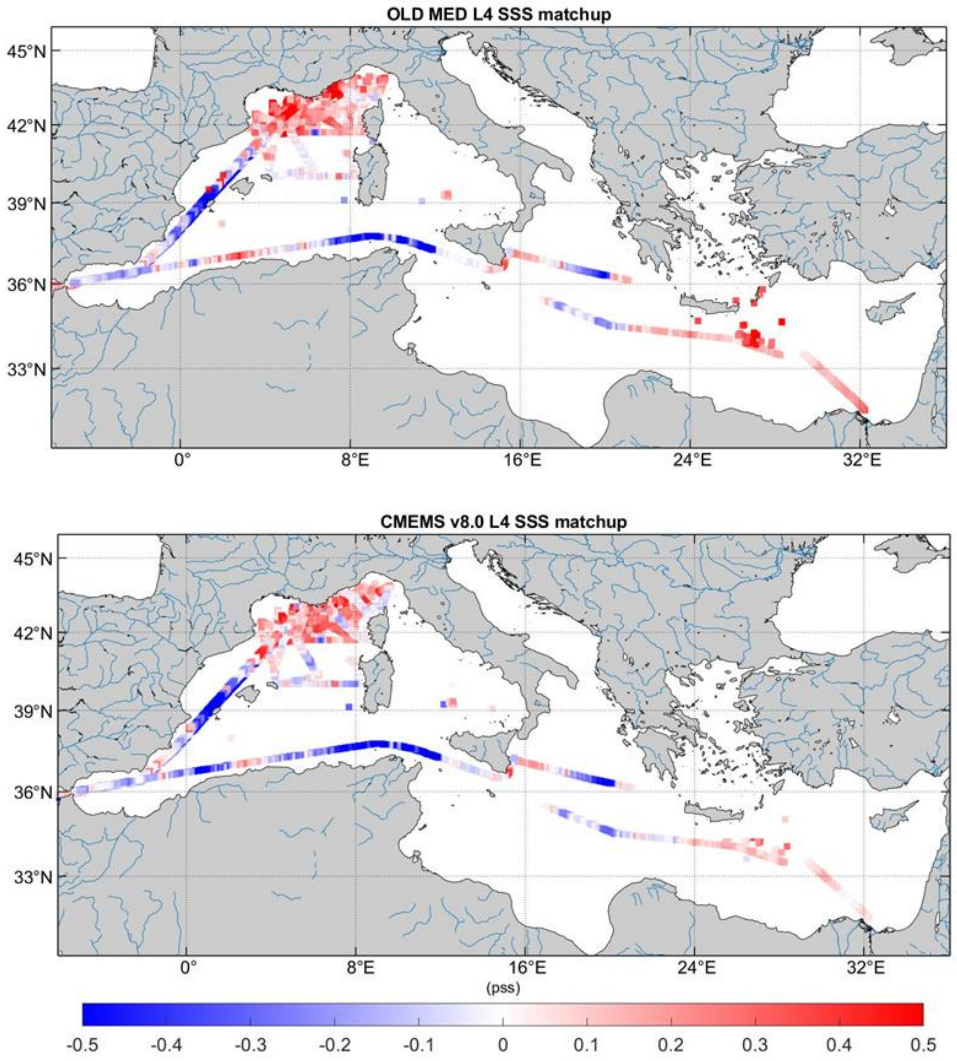


Figure 65: Spatial distribution of the validation matchups between the multi-observation L4 SSS datasets and in situ observations. The colour of the dots refers to the biases computed between in situ estimates and the corresponding values extracted from new 2DMED L4 SSS values (a), previous version MED LR (Sammartino et al., 2022) (b) and GLO LR v8.0 (c).

The maps accounts for two years of matchups (2017-2018). The color of the dots refers to the biases between in situ estimates and the corresponding extracted values. The new 2DMED SSS product shows reduced errors with respect to the previous version MED LR and GLO LR products in most areas of the basin. The improvement of the new 2DMED SSS product with respect to the MED LR is particularly evident in some high dynamic and coastal areas such as the Gulf of Lion and Ligurian Sea. The errors are also reduced in open ocean areas such as Levantine Sea.

Figure 66 shows the spatial power spectral densities (PSD) as a function of the frequency computed for each of four different L4 SSS datasets: the HR and LR 2DMED SSS products; in situ climatology (see section 2.2.1) the CMEMS GLO LR L4 SSS v8.0 dataset and the Mediterranean L4 SSS product from the Barcelona Expert Center (BEC).

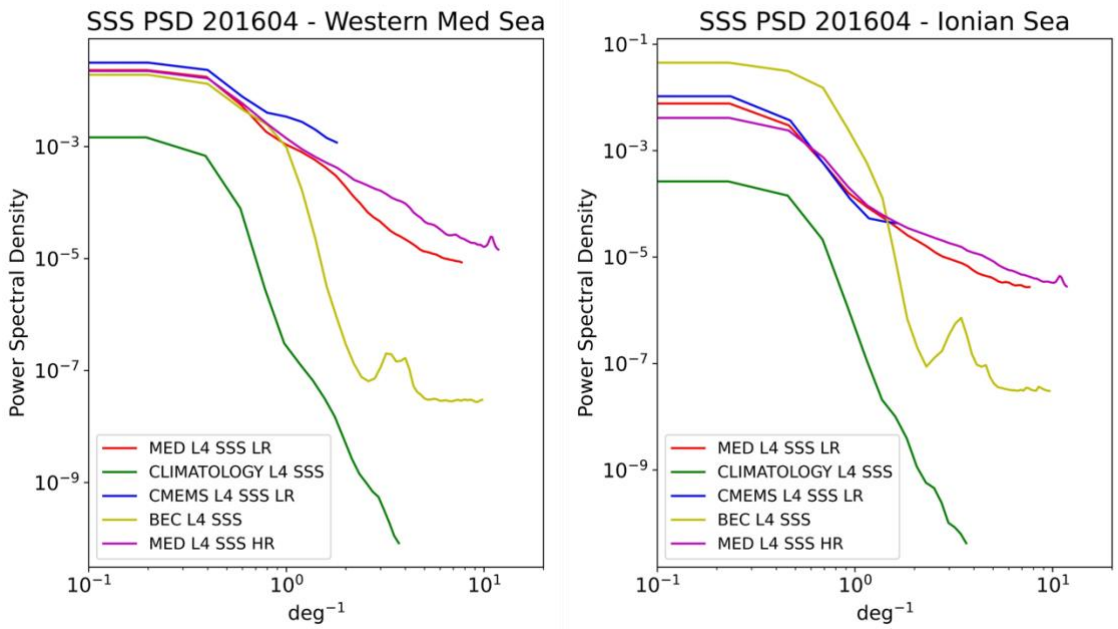


Figure 66: Spatial power spectral density computed from four L4 SSS products: previous version MED LR (Sammartino et al., 2022) (red line); in situ climatology (see section 2.2.1) (green line); CMEMS GLO LR L4 SSS v8.0 (blue line); Barcelona Expert Center L4 MED SSS (light green line) and the new 2DMED HR L4 SSS (purple line).

To simplify the comparison between the effective resolution resolved by each product, two different boxes (one in the western basin and the other in the Ionian Sea) have been considered for April 2016. The PSD spectra were computed on daily images following the latitudinal variations and averaging the results obtained for each longitude, to provide a single spectrum. Before computing the Fast Fourier Transform, the data were detrended (using a linear fit) and a Blackman-Harris windowing function was applied to diminish the spectral leakage (Sammartino et al., 2022).

As expected, the climatology shows the lowest variance for both areas with a drop already at 0.5 degree^{-1} (about 200 km). It is followed by the BEC dataset that shows an abrupt drop between 1–2 deg^{-1}, with a variance that still decreases until smaller scales (>4 deg-1). The 2DMED HR shows the highest effective spatial resolution with respect to the other datasets, followed by the 2DMED LR (Sammartino et al., 2022). In particular, in the Western Mediterranean Sea box, the comparison of the two datasets highlights that the 2DMED HR shows highest spatial variance in the mesoscale (>2 deg-1) than the MED LR (Figure 66).The improvement of the new 2DMED SSS product with respect to both MED LR and GLO LR datasets is mostly related to the optimized model, both in terms of the use of a new mask specifically focused only on big rivers and the inclusion in the OI algorithm of a Ultra-High resolution SST fields that implicitly accounts for smaller scale dynamics.

3.4 Combined physical-biological 4DMED experimental product validation

The choice of the best performing set up of the network has been assessed by using the 20% of the reference in situ database which was randomly extracted before the training and left for the validation of the algorithm on dataset never seen before. The results of this validation are given in Figure 67 Table 10in terms of error profiles. Here, the Root

Mean Square Error (RMSE) profiles (M3, see Table 7) are computed for each reconstructed variable between observed and predicted values. The CHL errors are computed on linear values and not on log10-transformed values.

Furthermore, to avoid edge effects related to network structure, before computing the RMSE, the surface data of each predicted profile was substituted by the observed data used as input, while the last bottom predicted value was substituted by that of nearest upper level.

From Figure 67, it is highlighted that the errors for T, S and CHL are comparable to those obtained in Buongiorno Nardelli (2020) and Sammartino et al. (2020).

However, the comparison with previous works shows evident improvements of the new 4DMED product, especially in the first surface layers, with RMSE (T<0.2, S<0.08 and CHL<0.07) lower than those obtained in the Buongiorno Nardelli, 2020 (T<~0.6, S<0.2) and Sammartino et al. 2020 (CHL >~0.25).

For all physical variables, the highest errors are observed between 20 and 60 meter of depth corresponding to the area of the highest and most complex dynamic for the Mediterranean basin, while for chlorophyll this area is wider reaching the 80 meters of depth. As expected, at deeper levels, the errors reduce until the bottom. This especially occurs for chlorophyll for which the reduction of the light deeply affects the photosynthetic activity at deeper layers.

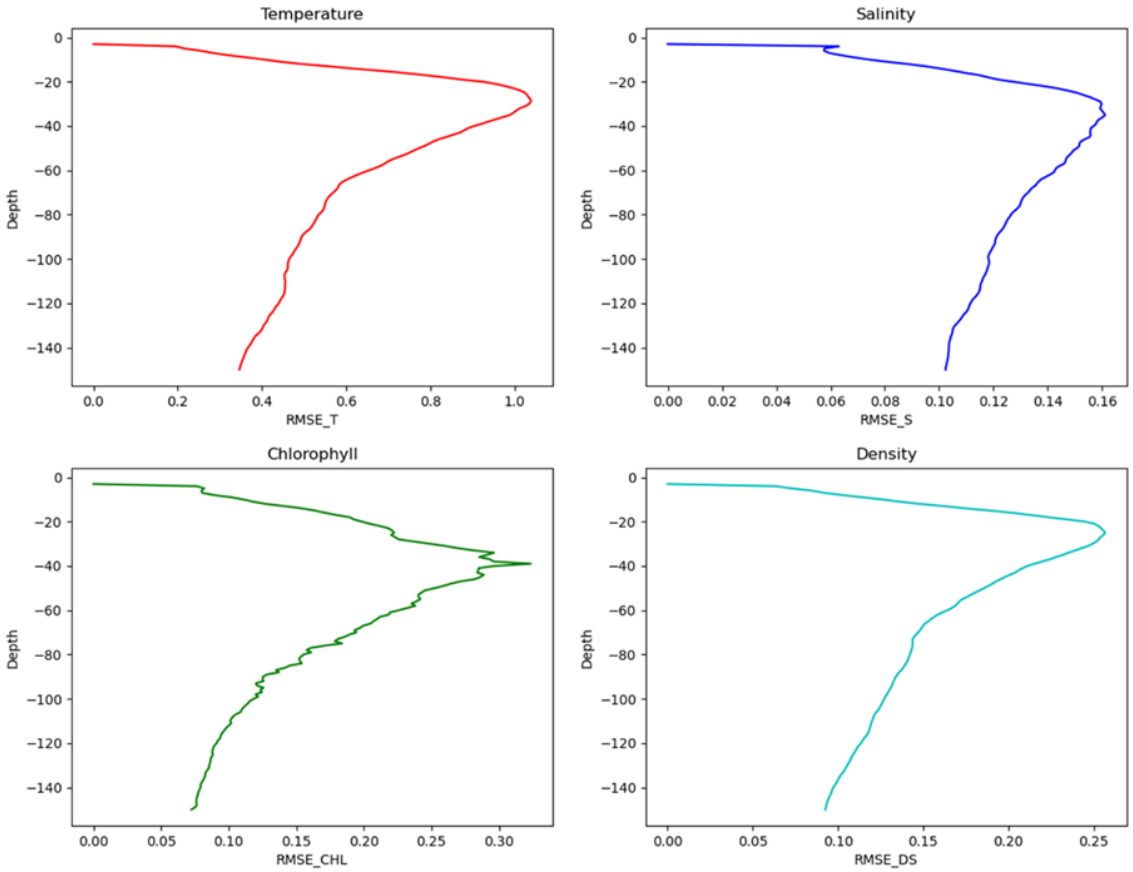


Figure 67: RMSE profile of the four reconstructed bio-physical variables computed on the test set as a function of depth.

Nevertheless, a deep validation of the combined physical-biological 4DMED experimental product will be carried out in the WP3700 and on a common in situ database between WP3300 and WP3400.

3.5 Experimental physical 4DMED product

The experimental physical 4DMED product inferred from machine learning trained model and with satellite observations will be validated with in situ observations (R1) before and after the merging with in situ Argo floats observations. The Root Mean Square Error profiles will be calculated to evaluate the skills of the product to represent the temperature and salinity observed variables.

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - Model_i)^2 \qquad \text{(M1, see Table 10)}$$

, with x_i the observed value, Model is the corresponding reconstruction and N is the number of matchup

Nevertheless, a deep validation of the combined physical 4DMED experimental product will be carried out in the WP3700 (with the in situ observations CANALES (R6) and PRESWOT (R7)).

The Phys 4DMED solution will also be compared to ARMOR3D product available at a 1/8° and daily resolutions (it will be delivered to the Copernicus Marine Service at the November 2024 Release).

Finally, the geostrophic velocities at 15 m will be validated against the velocities of the AOML SVP drifters, filtered at inertial frequency (~24h). Once the Phys 4DMED velocities have been interpolated to the positions of the drifters, mean and root mean square errors will be calculated to assess the product's ability to represent the zonal and meridional velocities calculated from the drifter trajectories. It should be noted that geostrophy represents only a part, but it is a significant one, of the velocity signal. This result is compared with the results of ARMOR3D over the same Mediterranean region.

Validation results of temperature and salinity fields

To validate the temperature and salinity fields, mean differences and RMSE between the in situ Argo temperature and salinity observations on one hand and, on a second hand, (1) ML solution, (2) ML + OI solution (final Phys 4DMED product), (3) ARMOR3D 1/8°- daily solution (Copernicus Marine service): step 1 (multi linear regression of satellite data on the vertical), (4) ARMOR3D 1/8° - daily solution (Copernicus Marine service): step 2 (merging with in situ observations using an Optimal Interpolation), (6) MEDSEA, (7) WOA climatology.

For salinity, the step 1 and step 2 of Phys 4DMED shows slightly better agreement with the observations than the steps 1 and 2 (resp.) of ARMOR3D and for temperature it's the opposite. For the two parameters, the Phys 4DMED does better than the initial MEDSEA target. On the final product the RMSE is maximum between 20 and 60 meters because of the high variability and complex physics of this layer. It reaches 0.5°C and

0.11-0.12 psu. At other depths the RMSE are around 0.2-0.3°C and 0.06-0.1 psu. Phys 4DMED is not biased when compared to the observations.

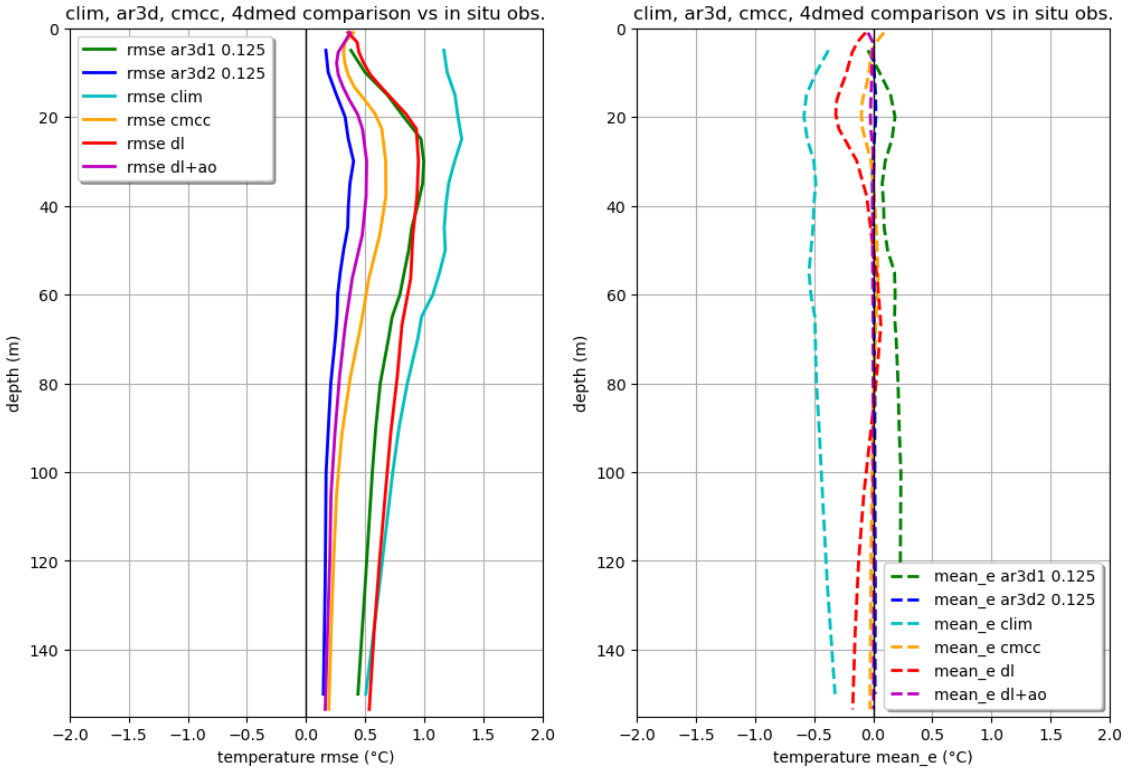


Figure 68: RMSE (left) and Mean of the differences (right) of the temperature fields between in situ observations and AR3D step 1 (green), AR3D step2 (dark blue), Climatology (cyan), MEDSEA (yellow), ML solution (red), ML + OI solution (Final PHY 4DMED product) (purple).

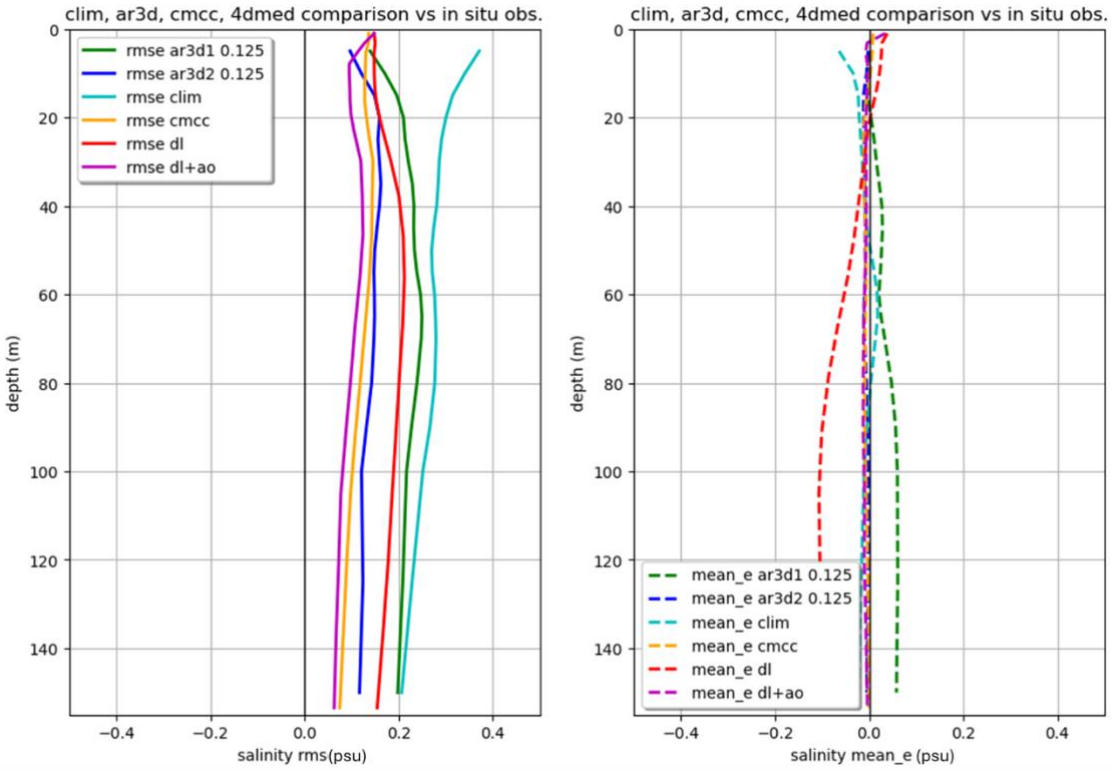


Figure 69: RMSE (left) and Mean of the differences (right) of the salinity fields between in situ observations and AR3D step 1 (green), AR3D step2 (dark blue), Climatology (cyan), MEDSEA (yellow), ML solution (red), ML + OI solution (Final PHY 4DMED product) (purple).

Validation results of 15-m velocities fields

The comparison of ARMOR3D and Phys 4DMED with the drifters' velocities at 15m-depth shows slightly agreement for Phys 4DMED. The velocities are not biased compared to the observations and the RMSE is reduced, mainly for the zonal component (Table 8).

Larger errors are found in the more energetic regions such as, for example, along the Algerian current and in the Alboran Sea (Figure 70).

Table 8: Mean and RMS of the differences between velocities from SVP drifters at 15-m depth (AOML) and (1) ARMOR3D [1/8°, daily] and Phys 4DMED

Data vs drifters 15m	**U**		**V**	
	Mean E (cm/s)	**RMSE (cm/s)**	**mean E (cm/s)**	**RMSE (cm/s)**
ARMOR3D	0.6	13.2	1.1	12.7
Phys 4DMED	-0.0	12.5	0.2	12.5

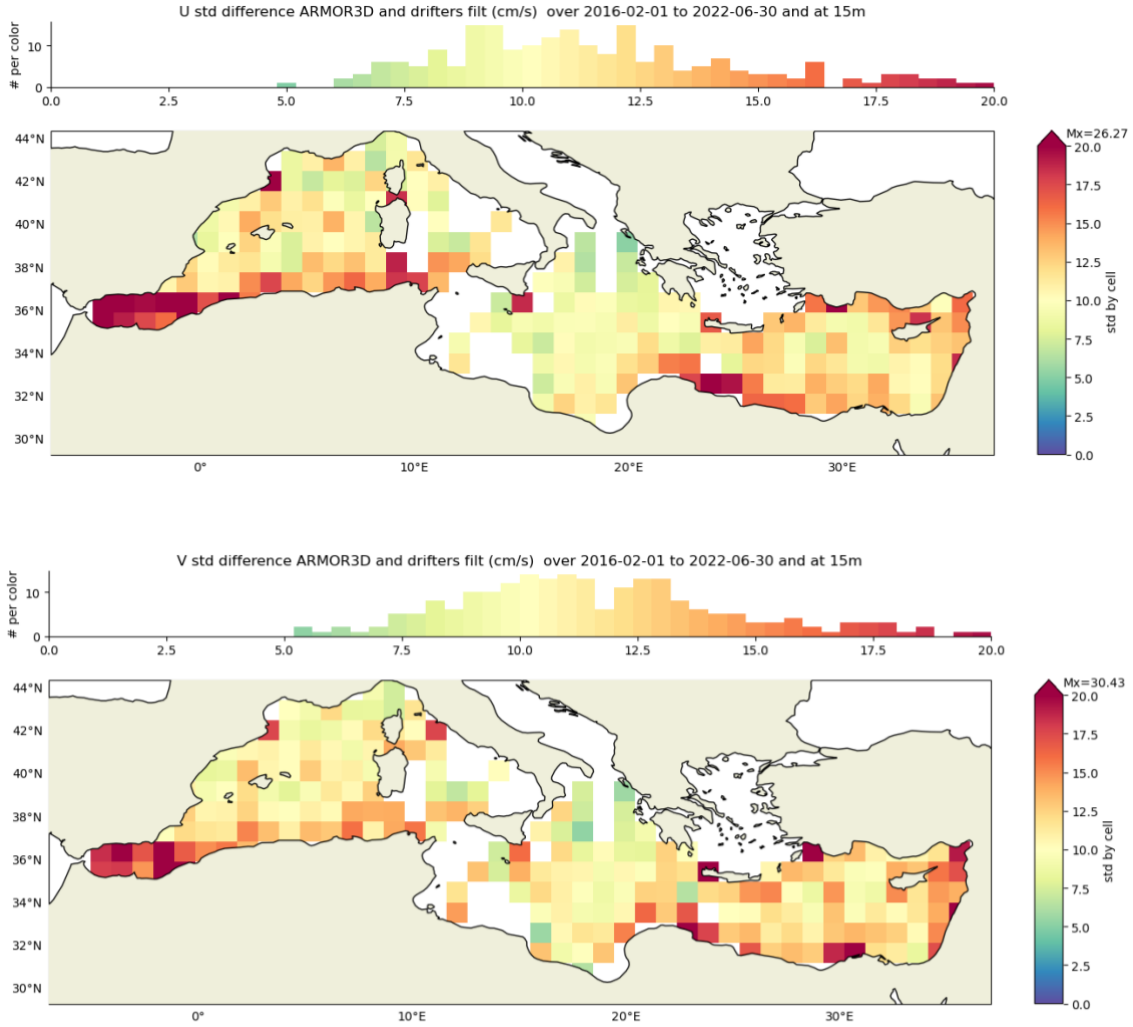


Figure 70: Standard deviation of Phys 4DMED minus 15-m depth velocities over period 2016-2022/06 (cm/s) for zonal component (top) and meridional component (bottom).

3.6 Summary of the observations and metrics used for validation of 4DMEDSea experimental products

Table 9 lists the reference observations data, in situ and from satellites, used to validate the 4DMED experimental products. Table 10 summarizes the metrics of validation.

Table 9 – List of reference data used for validation metrics

Number	Validation input and reference data	Description	Link/reference
R1	In situ observations	Copernicus Marine Service database of	https://doi.org/10.17882/46219

		temperature and salinity	
R2	In situ observations	Global Ocean Surface Underway Data (GOSUD) thermo-salinometer (TSG) observations	https://doi.org/10.17882/39475; Kolodziejczyk, et al., **2021**.
R3	In situ observations	Biogeochemical and hydrographic profiles from oceanographic cruises and Argo program	Sammartino et al. (2020) for oceanographic cruises ftp.ifremer.fr for Bio-Argo data
R4	Altimeter satellite	Copernicus Marine Service - European Seas Along Track L 3 Sea Surface Heights Reprocessed 1993 Ongoing Tailored For Data Assimilation	https://doi.org/10.48670/moi-00139
R5	In situ observation	Copernicus Marine Service - Global Ocean-Delayed Mode in-situ observations of ocean surface currents	https://doi.org/10.17882/86236
R6	In Situ Observations	CANALES glider temperature and salinity 2D transects	https://www.socib.es/?seccion=observingFacilities&facility=glider

R7	In Situ Observations	PRESWOT CTD profiles, glider transects (temperature, salinity, fluorimeter) Lagrangian drifters	https://digital.csic.es/handle/10261/182615
R8	In Situ Observations	CALYPSO drifters Lagrangian drifter trajectories	https://zenodo.org/records/4592311
R9	In Situ Observations	SVP AOML drifters	https://doi.org/10.17882/86236

Table 10 – List of metrics/methods used to validate the WP3 products

Number	Product name (the one that is validated)	Description	Reference data used (if relevant)
M1	Phys 4DMED	RMS and mean error vertical profiles over the MED	R1, R6, R7
M2	SSS 2DMED	RMS and mean error computed on surface values extracted from in situ observations	R2, R6, R7
M3	Phys-Bio 4DMED	RMS and mean error vertical profiles over the MED	R3, R6, R7
M4	ALT 2DMED	RMSE maps vs independent along-track	R4, R5
M5	ALT 2DMED	Variance of the error vs along-track independent	R4
M6	ALT 2DMED	Effective resolution	R4

M7	ALT 2DMED	Distances between simulated and real drifter locations	R5
M8	ALT 2DMED	Lagrangian skill score (S)	R7, R8
M9	ALT 2DMED	Normalized Lagrangian distance between synthetic and real drifters (D)	R7, R8
M10	ALT 2DMED	I and II – kind FSLE. Effective dynamical resolution	R7, R8

References

- Antoine, D., and Morel, A. (1996). Oceanic primary production: 1. Adaptation of a spectral light-photosynthesis model in view of application to satellite chlorophyll observations. Global biogeochemical cycles, 10(1), 43-55.
- Arbic, B. K., Scott, R. B., Chelton, D. B., Richman, J. G., and Shriver, J. F.: Effects of stencil width on surface ocean geostrophic velocity and vorticity estimation from gridded satellite altimeter data: OCEAN GEOSTROPHIC VELOCITY ESTIMATION, J. Geophys. Res. Oceans, 117, n/a-n/a, https://doi.org/10.1029/2011JC007367, 2012.
- Arhan, M. and Colin de Verdière, A.: Dynamics of eddy motions in the Eastern North Atlantic, J. Phys. Oceanogr., 15, 153–170, 1985.
- Aurell, E., Boffetta, G., Crisanti, A., Paladin, G. & Vulpiani, A. Predictability in the large: an extension of the Lyapunov exponent. *J. Phys. A* **30**, 1–26 (1997).
- Aydogdu et al., 2023: The dynamical role of upper layer salinity in the Mediterranean Sea, in: 7th editionnof the Copernicus Ocean State Report (OSR7), Copernicus Publications, State Planet, 1-osr7, 6, https://doi.org/10.5194/sp-1-osr7-6-2023.
- Ballarotta, M., & Verbrugge, N. (2024). ESA 4DMED-SEA - Mediterranean Sea Gridded L4 Sea Surface Heights And Derived Variables [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10648981
- 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.
- Ballarotta, M., Ubelmann, C., Veillard, P., Prandi, P., Etienne, H., Mulet, S., Faugère, Y., Dibarboure, G., Morrow, R., and Picot, N.: Improved global sea surface height and current maps from remote sensing and in situ observations, Earth Syst. Sci. Data, 15, 295–315, https://doi.org/10.5194/essd-15-295-2023, 2023.
- Barceló-Lull, B., Pascual, A., Sánchez-Román, A., Cutolo, E., d'Ovidio, F., Fifani, G., Ser-Giacomi, E., Ruiz, S., Mason, E., Cyr, F., Doglioli, A., Mourre, B., Allen, J. T., AlouFont, E., Casas, B., Díaz-Barroso, L., Dumas, F., GómezNavarro, L., and Muñoz, C.: Fine-Scale Ocean Currents Derived From in situ Observations in Anticipation of the Upcoming SWOT Altimetric Mission, Front. Mar. Sci., 8, 1070, https://doi.org/10.3389/fmars.2021.679844, 2021
- Boffetta, G., Lacorata, G., Redaelli, G. & Vulpiani, A. Detecting barriers to transport: A review of different techniques. *Physica D* **159**, 58–70 (2001).
- Böhme, L., & Send, U. (2005). Objective analyses of hydrographic data for referencing profiling float salinities in highly variable environments. *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*(3-4), 651-664.
- Bretherton, F., R. Davis, and C. Fandry: A technique for objective analysis and design of oceanographic experiments applied to MODE-73. Deep-Sea Res., 23, 559–582, 1976
- Buongiorno Nardelli, B. A Deep Learning Network to Retrieve Ocean Hydrographic Profiles from Combined Satellite and In Situ Measurements. Remote Sens. 2020, 12, 3151. Buongiorno et. al, 2020
- Buongiorno Nardelli, B. A Novel Approach for the High-Resolution Interpolation of In Situ Sea Surface Salinity. J. Atmos. Ocean Technol. 2012, 29, 867–879.
- Buongiorno Nardelli, B.; Droghei, R.; Santoleri, R. Multi-dimensional interpolation of SMOS sea surface salinity with surface temperature and in situ salinity data. Remote Sens. Environ. 2016, 180, 392–402.

- Cabanes, C., Thierry, V., & Lagadec, C. (2016). Improvement of bias detection in Argo float conductivity sensors and its application in the North Atlantic. Deep Sea Research Part I: Oceanographic Research Papers, 114, 128-136.
- Calafat, F. M., Frederikse, T., & Horsburgh, K. (2022). The sources of sea-level changes in the Mediterranean Sea since 1960. Journal of Geophysical Research: Oceans, 127, e2022JC019061. https://doi.org/10.1029/2022JC019061
- d'Ovidio, F., Fernández, V., Hernández-García, E. & López, C. Mixing structures in the Mediterranean sea from Finite-Size Lyapunov Exponents. *Geophys. Res. Lett.* **31**, L17203 (2004)
- d'Ovidio, F., J. Isern-Fontanet, C. López, E. Hernández-García, and E. García-Ladona (2009), Comparison between Eulerian diagnostics and finite-size Lyapunov exponents computed from altimetry in the Algerian basin, Deep Sea Res. Part I, 56, 15–31, doi:10.1016/j.dsr.2008.07.014
- Droghei, R.; Buongiorno Nardelli, B.; Santoleri, R. A New Global Sea Surface Salinity and Density Dataset from Multivariate Observations (1993–2016). Front. Mar. Sci. 2018, 5, 84.
- Droghei, R.; Nardelli, B.B.; Santoleri, R. Combining in Situ and Satellite Observations to Retrieve Salinity and Density at the Ocean Surface. J. Atmos. Ocean Technol. 2016, 33, 1211–1223.
- Eppley,R.W. (1972) Temperature and phytoplankton growth in the sea. Fisheries Bulletin, 70, 1063-1085.
- Fablet, R., Amar, M. M., Febvre, Q., Beauchamp, M. & Chapron, B. End-to-end physics-informed representation learning for satellite ocean remote sensing data: Applications to satellite altimetry and sea surface currents. ISPRS Annals Photogramm. Remote. Sens. & Spatial Inf. Sci. (2021).
- Falcini, F., Corrado, R., Torri, M., Mangano, M. C., Zarrad, R., Di Cintio, A., ... & Lacorata, G. (2020). Seascape connectivity of European anchovy in the Central Mediterranean Sea revealed by weighted Lagrangian backtracking and bio-energetic modelling. Scientific Reports, 10(1), 18630
- Gregg, W. W., & Carder, K. L. (1990). A simple spectral solar irradiance model for cloudless maritime atmospheres. Limnology and oceanography, 35(8), 1657-1675.
- Gregg, W. W., & Casey, N. W. (2009). Skill assessment of a spectral ocean–atmosphere radiative model. Journal of Marine Systems, 76(1-2), 49-63.
- Haller, G. (2016). Dynamic rotation and stretch tensors from a dynamic polar decomposition. Journal of the Mechanics and Physics of Solids, 86, 70–93. https://doi.org/10.1016/j.jmps.2015.10.002
- Hernández-Carrasco, I., López, C., Hernández-García, E. & Turiel, A. How reliable are finite-size Lyapunov exponents for the assessment of ocean dynamics? *Ocean Modelling* **36**, 208–218 (2011)
- Hernández-Carrasco, I., Orfila, A., Rossi, V., & Garçon, V. (2018). Effect of small-scale transport processes on phytoplankton distribution in coastal seas. Scientific Reports, 8(1), 8613. Hernandez-Carrasco et. al 2018
- Hollingsworth, A. and Lönnberg, P. (1986), The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus A, 38A: 111-136. https://doi.org/10.1111/j.1600-0870.1986.tb00460.x (HL86 method)
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
- Incarbona et al., 2016: Mediterranean circulation perturbations over the last five centuries: Relevance to past Eastern Mediterranean Transient-type events. Sci Rep 6, 29623. https://doi.org/10.1038/srep29623

- Iona, A.; Theodorou, A.; Watelet, S.; Troupin, C.; Beckers, J.-M.; Simoncelli, S. Mediterranean Sea Hydrographic Atlas: Towards optimal data analysis by including time-dependent statistical parameters. Earth Syst. Sci. 2018, 10, 1281–1300.
- Kar, S.; McKenna, J.R.; Anglada, G.; Sunkara, V.; Coniglione, R.; Stanic, S.; Bernard, L. Forecasting Vertical Profiles of Ocean Currents from Surface Characteristics: A Multivariate Multi-Head Convolutional Neural Network–Long Short-Term Memory Approach. *J. Mar. Sci. Eng.* **2023**, *11*, 1964. https://doi.org/10.3390/jmse11101964.
- Kolodziejczyk, N.; Diverres, D.; Jacquin, S.; Gouriou, Y.; Grelet, J.; Le Menn, M.; Tassel, J.; Reverdin, G.; Maes, C.; Gaillard, F. Sea Surface Salinity from French RESearcH Vessels: Delayed mode dataset, annual release. SEANOE 2021.
- Lacorata, G., Corrado, R., Falcini, F., & Santoleri, R. (2019). FSLE analysis and validation of Lagrangian simulations based on satellite-derived GlobCurrent velocity data. Remote sensing of environment, 221, 136-143.
- Lacorata, G., L. Palatella, and R. Santoleri (2014), Lagrangian predictability characteristics of an Ocean Model, J. Geophys. Res. Oceans, 119, 8029–8038, doi:10.1002/2014JC010313.
- Le Guillou, F., Gaultier, L., Ballarotta, M., Metref, S., Ubelmann, C., Cosme, E., and Rio, M.-H.: Regional mapping of energetic short mesoscale ocean dynamics from altimetry: performances from real observations, Ocean Sci., 19, 1517–1527, https://doi.org/10.5194/os-19-1517-2023, 2023.
- Le Traon P.-Y., F. Nadal, N. Ducet, An Improved Mapping Method of Multisatellite Altimeter Data, J. Atmos. Oceanic Technol. 15, 522-534, 1998
- Liu Yonggang & Robert H. Weisberg: Evaluation of trajectory modeling in different dynamic regions using normalized cumulative Lagrangian separation, JGR, vol 116, issue C9, https://doi.org/10.1029/2010JC006837
- Margirier, F., Testor, P., Heslop, E. et al. Abrupt warming and salinification of intermediate waters interplays with decline of deep convection in the Northwestern Mediterranean Sea. Sci Rep 10, 20923 (2020). https://doi.org/10.1038/s41598-020-77859-5.
- Méda, G., Ballarotta, M., Treboutte, A., Faugere, Y., Fablet, R., Ubelmann, C., Dibarboure, G. : 4DVarNet : Data-driven mapping of Ocean Surface Topography using SWOT altimetry, SWOT ST 2023, Toulouse, (https://swotst.aviso.altimetry.fr/fileadmin/user_upload/SWOTST2023/posters_earlycareer/meda_ocean_InversionAssimilation_EarlyCareer.pdf)
- Menemenlis, D., & Chechelnitsky, M., (2000). Error Estimates for an ocean general circulation model from altimeter and acoustic tomography data. Monthly Weather Review – March 2000, 128(3), 763,16p
- Mezić, I., Loire, S., Fonoberov, V. A., & Hogan, P. (2010). A new mixing diagnostic and gulf oil spill movement. Science, 330(6003), 486–489
- Millot et al., 2006: Large warming and salinification of the Mediterranean outflow due to changes in its composition, Deep Sea Research Part I: Oceanographic Research Papers, Volume 53, Issue 4, Pages 656-666, ISSN 0967-0637, https://doi.org/10.1016/j.dsr.2005.12.017.
- Morel, A. (1991). Light and marine photosynthesis: a spectral model with geochemical and climatological implications. Progress in oceanography, 26(3), 263-306.
- Morel, A.; Berthon, J.-F. Surface pigments, algal biomass profiles, and potential production of the euphotic layer: Relationships reinvestigated in view of remote-sensing applications. Limnol. Oceanogr. 1989, 34, 1545–1562.

- Owens, W. B., & Wong, A. P. (2009). An improved calibration method for the drift of the conductivity sensor on autonomous CTD profiling floats by θ–S climatology. *Deep Sea Research Part I: Oceanographic Research Papers*, *56*(3), 450-457.
- Palatella, L., Bignami, F., Falcini, F., Lacorata, G., Lanotte, A. S., & Santoleri, R. (2014). Lagrangian simulations and interannual variability of anchovy egg and larva dispersal in the Sicily Channel. Journal of Geophysical Research: Oceans, 119(2), 1306-1323.
- Pauthenet, E., Bachelot, L., Balem, K., Maze, G., Tréguier, A. M., Roquet, F., ... & Tandeo, P. (2022). Four-dimensional temperature, salinity and mixed-layer depth in the Gulf Stream, reconstructed from remote-sensing and in situ observations with neural networks. *Ocean Science*, *18*(4), 1221-1244.
- Pingyang & Toste, 2020: Recent Changes in Deep Ventilation of the Mediterranean Sea; Evidence From Long-Term Transient Tracer Observations, Frontiers in Marine Science, https://www.frontiersin.org/articles/10.3389/fmars.2020.00594, DOI=10.3389/fmars.2020.00594
- Pujol, M.-I., Faugère, Y., Taburet, G., Dupuy, S., Pelloquin, C., Ablain, M., and Picot, N.: DUACS DT2014: the new multi-mission altimeter data set reprocessed over 20 years, Ocean Sci., 12, 1067–1090, https://doi.org/10.5194/os-12-1067-2016, 2016
- Rio, M.H.; Pascual, A.; Poulain, P.-M.; Menna, M.; Barceló-Llull, B.; Tintore, J. Computation of a new mean dynamic topography for the Mediterranean Sea from model outputs, altimeter measurements ando ceanographic in situ data. Ocean Sci. 2014, 10, 731–744.
- Rossi, V., Ser-Giacomi, E., López, C., and Hernández-García, E. (2014), Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves, Geophys. Res. Lett., 41, 2883–2891,https://doi.org/10.1002/2014GL059540
- Sammartino, M., & Buongiorno Nardelli, B. (2024). ESA 4DMED-SEA - Mediterranean Multivariate Optimal Interpolated Salinity and Density fields [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10822461
- Sammartino, M., B. B. Nardelli, S. Marullo, and R. Santoleri (2020), An artificial neural network to infer the mediterranean 3D chlorophyll-a and temperature fields from remote sensing observations, Remote Sens., 12(24), 1–27, doi:10.3390/rs12244123.
- Sammartino, M., S. Aronica, R. Santoleri, and B. Buongiorno Nardelli (2022), Retrieving Mediterranean Sea Surface Salinity Distribution and Interannual Trends from Multi-Sensor Satellite and In Situ Data, Remote Sens., 14(10), 2502, doi:10.3390/rs14102502.
- Ser-Giacomi Enrico, Vincent Rossi, Cristóbal López, Emilio Hernández-García; Flow networks: A characterization of geophysical fluid transport. Chaos 1 March 2015; 25 (3): 036404. https://doi.org/10.1063/1.4908231
- Ser-Giacomi, E., Jordá-Sánchez, G., Sotto-Navarro, J., Thomsen, S.,Mignot, J., Sevault, F., & Rossi, V. (2020). Impact of climate changeon surface stirring and transport in the Mediterranean Sea.Geophysical Research Letters, 47, e2020GL089941.
- Slingo, A. (1989). A GCM parameterization for the shortwave radiative properties of water clouds. Journal of Atmospheric Sciences, 46(10), 1419-1427.
- Taburet, G., Sanchez-Roman, A., Ballarotta, M., Pujol, M.-I., Legeais, J.-F., Fournier, F., Faugere, Y., and Dibarboure, G.: DUACS DT2018: 25 years of reprocessed sea level altimetry products, Ocean Sci., 15, 1207–1224, https://doi.org/10.5194/os-15-1207-2019, 2019.
- Tanré, D., Herman, M., Deschamps, P. Y., and De Leffe, A. (1979). Atmospheric modeling for space measurements of ground reflectances, including bidirectional properties. Applied optics, 18(21), 3587-3594

- Tarry, D. R., Essink, S., Pascual, A., Ruiz, S., Poulain, P.-M., Özgökmen, T., et al. (2021). Frontal convergence and vertical velocity measured by drifters in the Alboran Sea. *Journal of Geophysical Research: Oceans*, 126, e2020JC016614. https://doi.org/10.1029/2020JC016614
- Ubelmann, C. et al. Reconstructing ocean surface current combining altimetry and future spaceborne doppler data. J. Geophys. Res. Ocean. 126, e2020JC016560 (2021).
- Volpe, G., Buongiorno Nardelli, B., Colella, S., Pisano, A. and Santoleri, R. (2018). An Operational Interpolated Ocean Colour Product in the Mediterranean Sea, in New Frontiers in Operational Oceanography, edited by E. P. Chassignet, A. Pascual, J. Tintorè, and J. Verron, pp. 227–244.
- Wong, A. P. S., G. C. Johnson, and W. B. Owens, 2003: Delayed-Mode Calibration of Autonomous CTD Profiling Float Salinity Data by θ–S Climatology. *J. Atmos. Oceanic Technol.*, **20**, 308–318, https://doi.org/10.1175/1520-0426(2003)020<0308:DMCOAC>2.0.CO;2.
- Xiao, C., Tong, X., Li, D., Chen, X., Yang, Q., Xv, X., ... & Huang, M. (2022). Prediction of long lead monthly three-dimensional ocean temperature using time series gridded Argo data and a deep learning method. International Journal of Applied Earth Observation and Geoinformation, 112, 102971.