Project information

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

Deliverable information

Deliverable number	D4.5
Deliverable title	**Synthesis of satellite validation results**
Description	This report presents the results of Task 4.4: Improving the use of in situ observations for the long-term validation of satellite observations
Work Package number	WP4
Work Package title	Data integration, Assimilation, and Forecasting
Lead beneficiary	CLS
Lead authors	Sandrine Mulet (CLS)
Contributors	Sandrine Mulet (CLS), Christine Boone (CLS), Bàrbara Barceló-Llull (CSIC), Nikolaos Zarokanellos (SOCIB), Ananda Pascual (CSIC), Ivan Manso-Narvarte (AZTI), Ainhoa Caballero (AZTI), Anna Rubio (AZTI), Marine Bretagnon (ACRI-ST), Antoine Mangin (ACRI-ST)
Due date	31.10.2022
Submission date	04.11.2022
Comments	

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

https://doi.org/10.3289/eurosea_d4.5

Table of contents

Executive summary

This report presents the results of Task 4.4 on “Improving the use of in situ observations for the long-term validation of satellite observations”.

In situ observations are essential for the calibration and validation of satellite observations. Calibration activities organized by satellite agencies rely on dedicated instrumentation, whereas validation activities rely on the global ocean observing system. Task 4.4 demonstrate the potential of emerging in-situ networks for satellite validation activities and there are important links with the other tasks of WP4.

Thus, the task considers the use of Argo, gliders, HF radars and surface drifters for the validation of S3A&B SAR altimeter observations in terms of sea level and currents and addressing regional and global ocean. Besides, it focuses on the role of the BGC Argo array for the validation of ocean colour satellite missions (in particular S3A&B) for chlorophyll-a concentration, particulate backscattering coefficient and diffuse attenuation coefficient.

Four groups are contributing to Task 4.4: CSIC, AZTI and CLS for the blue ocean and ACRI-ST for the green ocean. Although the four groups conduct separate activities using specific datasets in dedicated regions of interest, the methods used are similar. The triple collocation technique is applied to estimate the error associated with satellite measurement. The triple collocation analysis is a method for quantifying the random error standard deviation of 3 datasets of the same geophysical variable by combining the covariances between the datasets (Mignot et al. 2019). In addition to the triple collocation analysis, usual validation metrics (e.g., rms, bias, correlation) are also considered.

1. Validation of Sentinel-3 altimeter observations with SOCIB glider data in the Balearic Sea (CSIC)

1.1. Objective

Validate Sentinel-3A altimeter observations along track 57 with glider observations in the Balearic Sea (see example Figure 1).

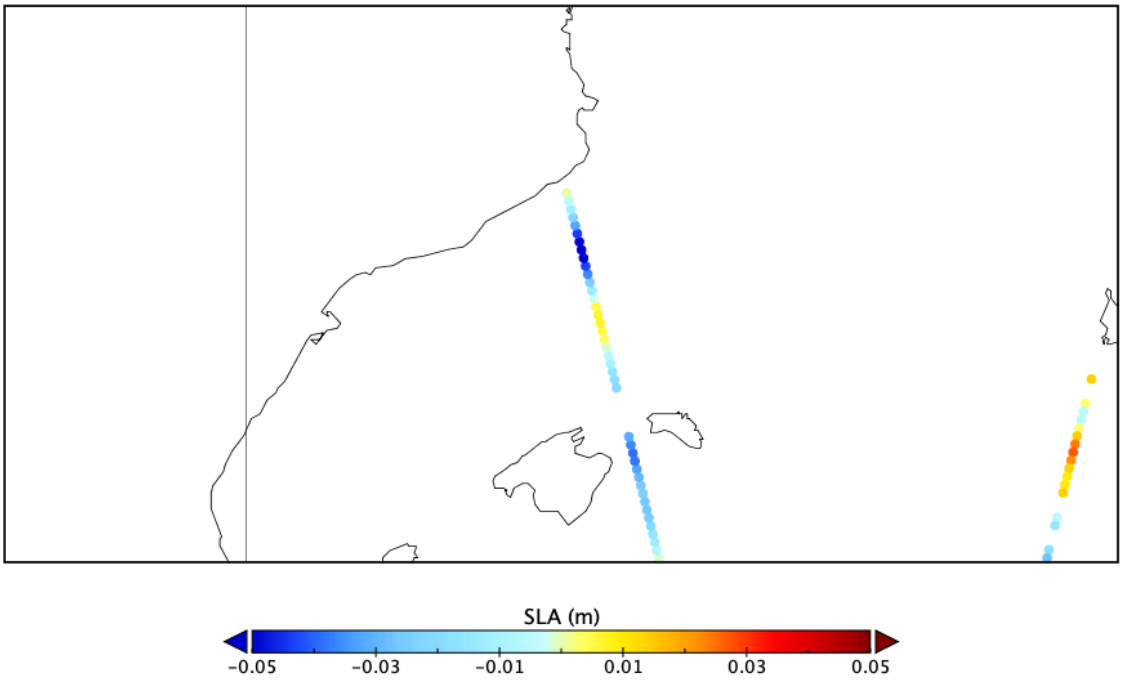


Figure 1. Sea level anomaly (SLA) observed from Sentinel-3A along track 57 on 24 March 2021.

1.2. Data and methods

Sentinel-3A observations

The altimetric product used for this study is the SSALTO/DUACS Near-Real-Time Level-3 sea surface height measured by Sentinel-3A over European Seas and distributed by the Copernicus Marine Environment Monitoring Service (CMEMS[1]). We use the 1Hz along-track data collected by S3A along the ground-track 57 with a spatial resolution of 7 km. This product provides the unfiltered and filtered (cut-off wavelength of 40 km2) sea level anomaly (SLA) fields. The absolute dynamic topography (ADT) is calculated from SLA observations plus the mean dynamic topography (MDT), also provided by this product.

Glider observations

Observations from a Slocum glider were obtained in the Balearic Sea (Figure 2) from the Balearic Islands Coastal Observing and Forecasting System (SOCIB). The glider took measurements between Mallorca and the Spanish peninsula along the Sentinel-3A (S3A) track 57 between 8 March 2021 and 28 May 2021 (Figure 2). During 81 days at sea, it covered 1156 km and sampled 1900 physical and biogeochemical profiles. The Slocum glider carried a suite of sensors that measured temperature, conductivity, and pressure (CTD), dissolved oxygen (oxygen optode), chlorophyll fluorescence and turbidity (FNLTU). The glider was programmed to profile from the surface up to 700 m at a nominal vertical speed of 0.18 ± 0.02 m/s and moved horizontally at approximately 20–24 km/day. A total of 10 glider transects along the S3A track were completed during the experiment. Two of these transects were conducted during S3A overpasses and have been used for the validation:

---

[1] https://doi.org/10.48670/moi-00140

[2] https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-SL-QUID-008-032-068.pdf

- Section 3 was conducted from 22-03-2021, 20:17h to 28-03-2021, 23:33h, while S3A overpassed the region on 24-03-2021, 21:16h. The glider transect was sampled from south to north during 6.1 days.
- Section 7 was conducted from 15-04-2021, 09:09h to 21-04-2021, 08:59h, while S3A overpassed the region on 20-04-2021, 21:16h. The glider transect was sampled from south to north during 6.0 days.

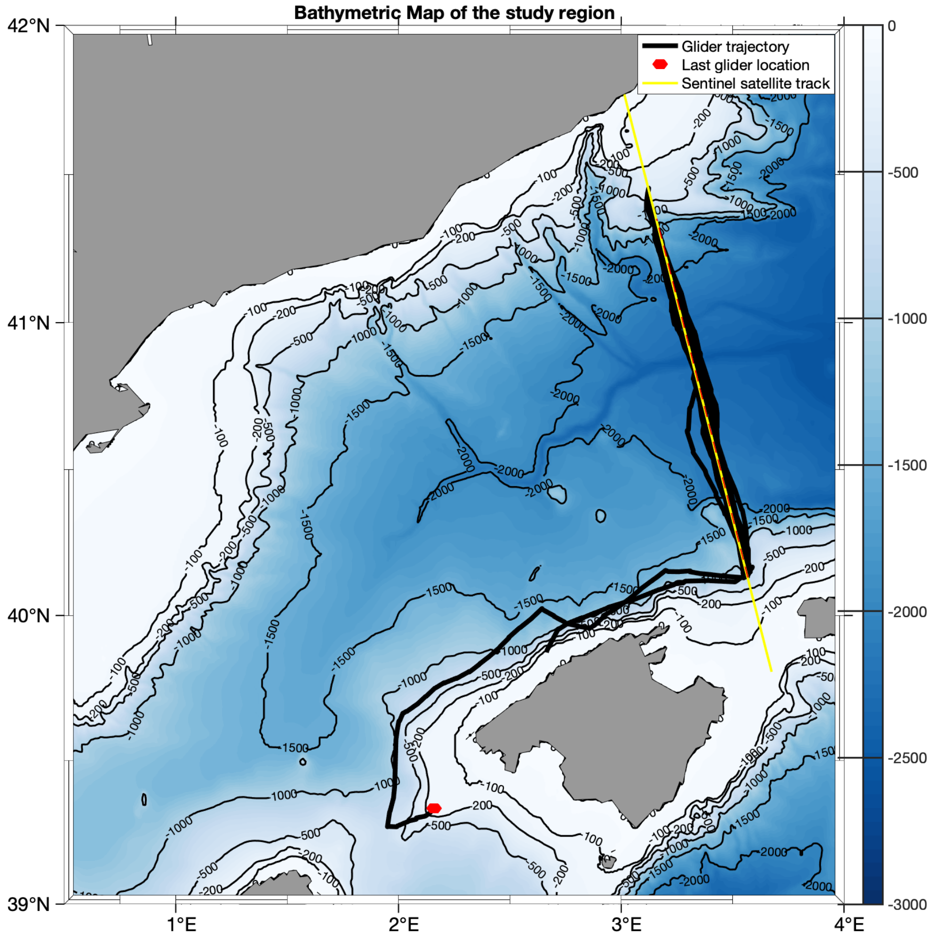


Figure 2. Trajectory of the SOCIB Slocum glider that sampled the Sentinel-3A track 57 between 8 March 2021 to 28 May 2021.

Glider observations of potential temperature and practical salinity along sections 3 and 7 were interpolated onto a regular two-dimensional (2D) grid with a resolution along the glider track of 4 km and a vertical resolution of 5 m between 10 and 690 m depth (Figure 3 and Figure 4). The original distance axis was calculated from the initial coordinates: (3.57ºE, 40.23ºN) for Section 3 and (3.55ºE, 40.19ºN) for Section 7.

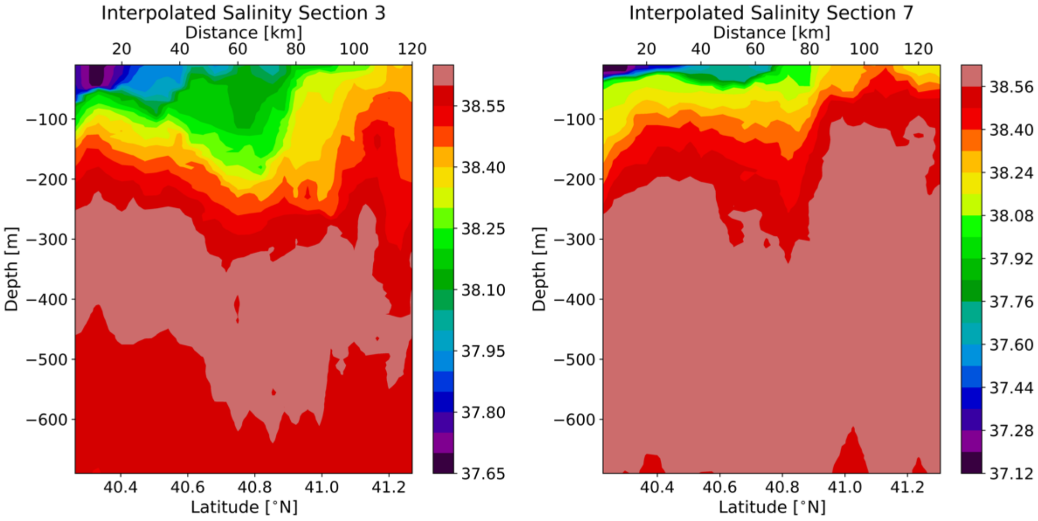


Figure 3. Interpolated salinity fields from glider observations along sections 3 and 7.

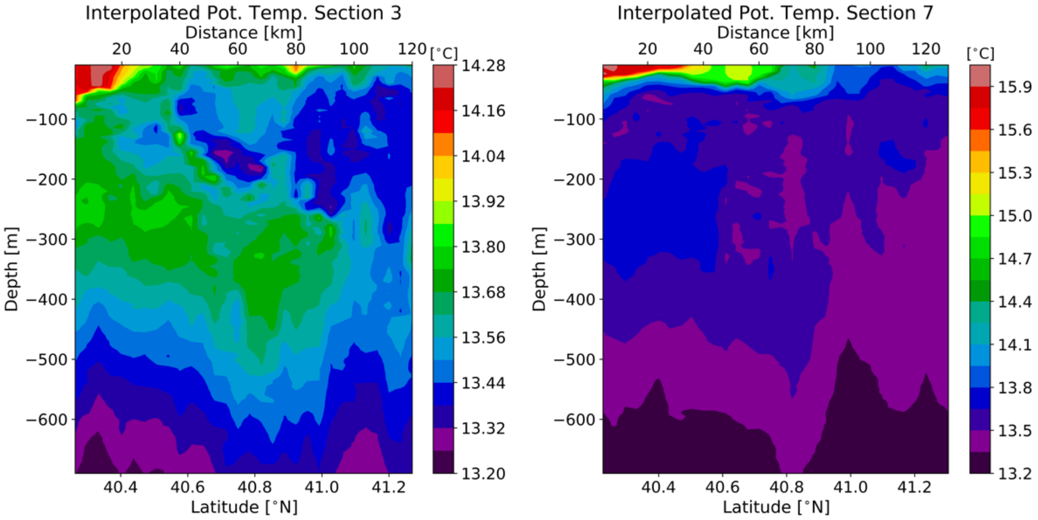


Figure 4. Interpolated temperature fields from glider observations along sections 3 and 7.

DH and geostrophic velocity calculation

With the interpolated temperature and salinity fields, we computed the dynamic height (DH) and across-track geostrophic velocity at the upper layer (10 m) considering a reference level of no motion at 690 m. These fields were compared with the S3A absolute dynamic topography (ADT) and across-track surface absolute geostrophic velocity, respectively. The following metrics were used for the comparison: correlation coefficient, root-mean-square-error (RMSE) and RMSE-based score (RMSEs), defined as:

RMSEs = 1 – [RMS(S3A_data - glider_data)/RMS(glider_data)]),

where a RMSEs of 1 would mean equal fields, and a value of 0 different fields.

To compare glider DH data to S3A ADT observations, we removed the mean of each field to compare anomalies (DHa and ADTa, respectively) and then the ADTa was linearly interpolated onto the latitude axis of the glider data to calculate statistics. We perform the comparison considering filtered and unfiltered fields.

The DH was filtered using a Loess smoother with a half-power filter cut-off of 40 km (as the cut-off wavelength used to filter Sentinel-3A data).

From DH and ADT fields, we calculated the corresponding across-track geostrophic velocities. From glider observations, the across-track geostrophic velocity at the upper layer (10 m) was calculated as:

$$v_{glider}^{g} = \frac{g}{f}\frac{\partial DH(z=-10m)}{\partial x},$$

where g is gravity, f is the Coriolis parameter and x is the distance axis along the glider track from the southeast to the northwest (Bouffard et al., 2010). The along-track gradient of DH at the upper depth layer (10 m) was calculated through forward difference for the first point, backward difference for the last point and centered differences in the inner points. For this computation, we used the filtered DH.

From S3A observations, the across-track absolute geostrophic velocity was calculated following the same procedure:

$$v_{S3A}^{g} = \frac{g}{f}\frac{\partial ADT}{\partial x}.$$

For this study we used the filtered ADT because it is less noisy than the unfiltered field. Note that the ADT is calculated as SLA (S3A observations) + MDT. The impact of the MDT on the results is evaluated by calculating the corresponding across-track geostrophic velocity:

$$v_{MDT}^{g} = \frac{g}{f}\frac{\partial MDT}{\partial x}.$$

With these definitions and considering the distance axis oriented from the southeast to the northwest of the track, the across-track geostrophic velocities will have a positive sign if they are oriented to the southwest, and a negative sign if they are oriented to the northeast. However, as currents to the northeast are typically considered positive and to the southwest negative, we change the sign of the velocity. With the new definition, the across-track velocities have a positive (negative) sign if they are oriented to the northeast (southwest).

The depth-averaged velocity (DAV) derived from glider data (Figure 5) is used to evaluate the impact of the reference level on the calculation of DH and to quantify the presence of barotropic motions (Bouffard et al., 2010). The DAV components are available along the zonal and meridional axes (DAV_u, DAV_v). To have the DAV along and across the glider track, we needed to rotate the original components. To do this, we used the angle between the glider track and the meridional axis. This angle was calculated from the initial and final positions of each trajectory ($tan(\theta) = \frac{lon_f - lon_i}{lat_f - lat_i}$) and is -16º for section 3 and -15º for section 7. Then, we used the following trigonometric equations to rotate the DAV components:

$$x' = x\,cos\theta + y\,sin\theta$$

$$y' = -x\,sin\theta + y\,cos\theta,$$

where x indicates zonal axis, y is meridional axis, $x'$is the across-glider-track direction and y' is the along-glider-track direction. For the DAV components:

$$DAV_{across} = DAV_u\,cos\theta + DAV_v\,sin\theta$$

$$DAV_{along} = -DAV_u \sin\theta + DAV_v \cos\theta$$

We compare the DAV_{across} (DAV across glider track) to the geostrophic velocity component calculated from glider observations.

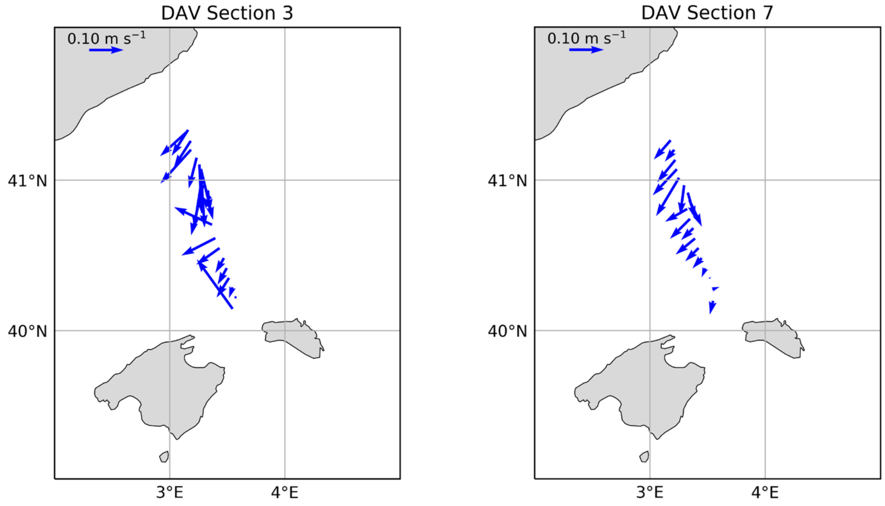


Figure 5. Depth-averaged-velocity (DAV) derived from glider data along sections 3 and 7.

1.3. Results

Glider DH vs. Sentinel-3A ADT

Figure 6 shows the filtered and unfiltered glider DHa and S3A ADTa along section 3. The unfiltered ADTa is noisier than the filtered field, but overall, there is a good correspondence between S3A and glider observations (Table 1). The higher differences are observed at ~40.8ºN, in a salinity front detected by glider data (Figure 3) but not captured by altimetry.

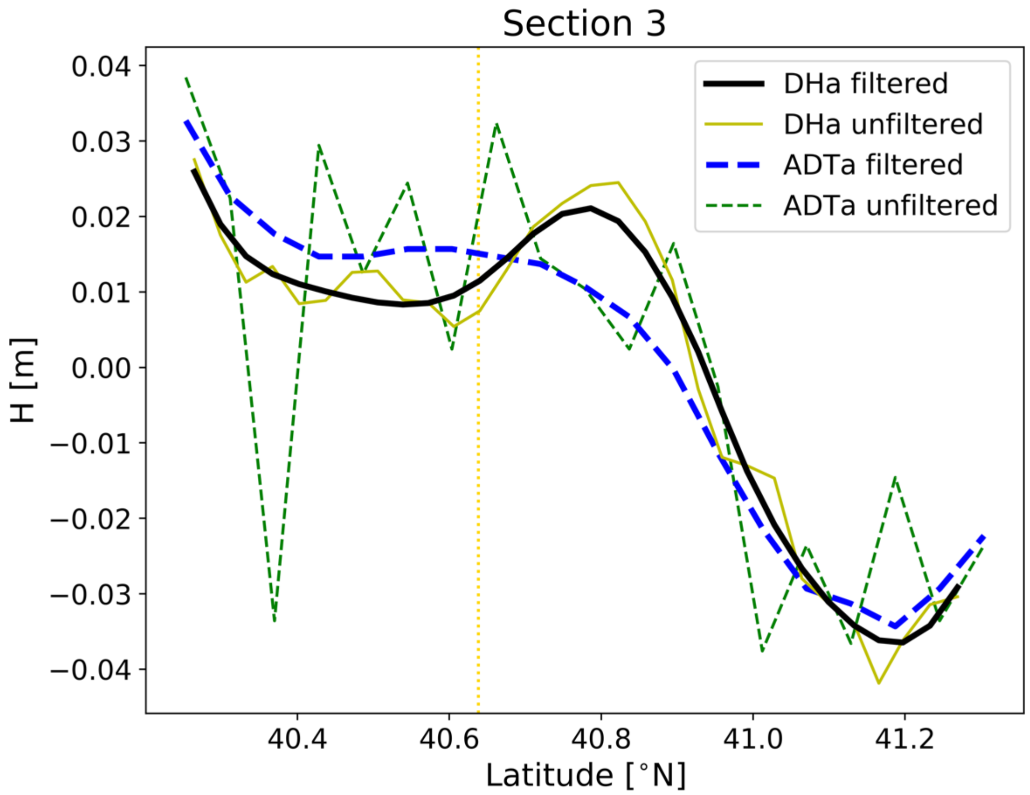


Figure 6. Filtered and unfiltered DH and ADT anomalies along section 3. The vertical yellow dotted line represents the glider position during the S3A overpass.

Table 1. Statistics between the glider DH anomaly and the S3A filtered and unfiltered ADT anomaly fields along section 3.

Glider data	**S3A Product**	**Correlation coefficient**	**RMSE [cm]**	**RMSEs**
Unfiltered DHa	Filtered ADTa	0.93	0.8	0.64
	Unfiltered ADTa	0.80	1.4	0.36
Filtered DHa	Filtered ADTa	0.95	0.6	0.70
	Unfiltered ADTa	0.83	1.2	0.39

For section 7, the unfiltered ADTa is less noisy than in section 3, and both ADTa fields have similar patterns than the filtered and unfiltered DHa (Figure 7). However, there are some dissimilarities that are translated in higher RMSE and lower values of the correlation coefficient and RMSEs (Table 2) than in section 3 (Table 1). Note that in this case S3A overpassed the region of study at the end of the glider sampling, hence, there is more difference in time between the data sampled at the beginning of the glider sampling and the observations of S3A than in section 3.

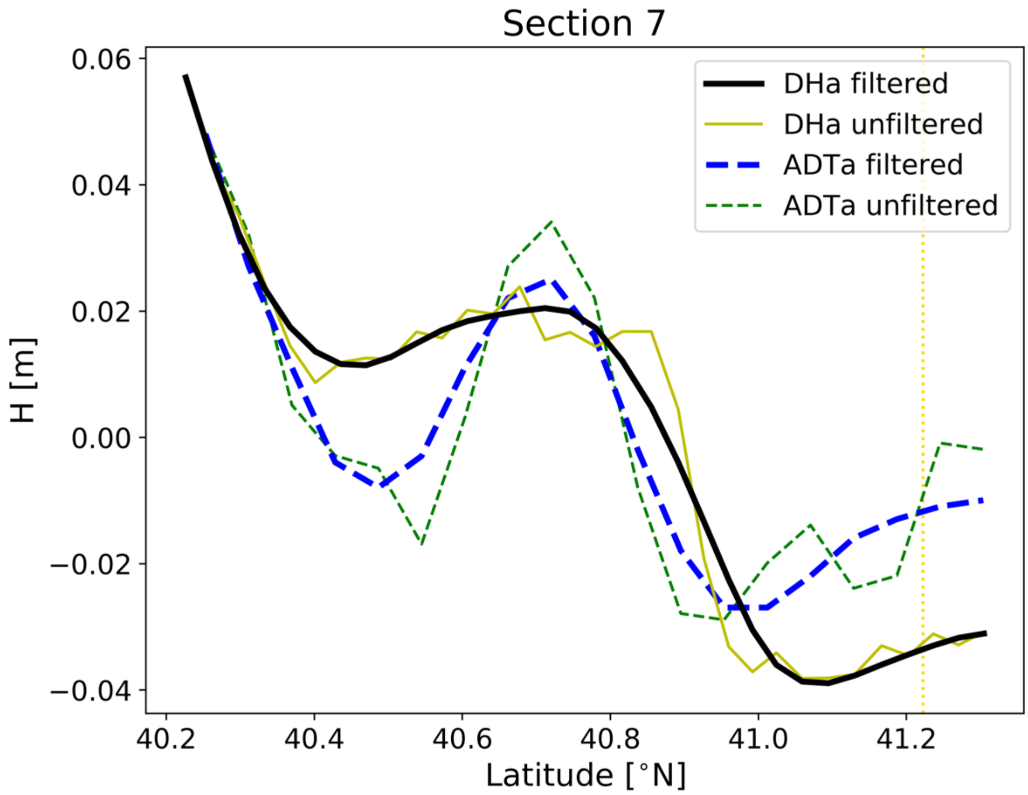


Figure 7. Filtered and unfiltered DH and ADT anomalies along section 7. The vertical yellow dotted line represents the glider position during the S3A overpass.

Table 2. Statistics between the glider DH anomaly and the S3A filtered and unfiltered ADT anomaly fields along section 7.

Glider data	**S3A Product**	**Correlation coefficient**	**RMSE [cm]**	**RMSEs**
Unfiltered DHa	Filtered ADTa	0.85	1.4	0.49
	Unfiltered ADTa	0.74	1.8	0.36
Filtered DHa	Filtered ADTa	0.86	1.3	0.51
	Unfiltered ADTa	0.76	1.7	0.39

Comparison of across-track geostrophic velocities

The analysis of the across-track geostrophic velocity derived from glider vs. Sentinel-3A observations along section 3 shows in general a good correspondence between both fields (Figure 8). However, a significant difference has been observed in the salinity front detected at ~40.8ºN (Figure 3, Δs ~ 0.3). This difference has also been identified in the comparison of the DH and ADT anomalies in Figure 6. The correlation coefficient, in this case, is 0.89, the RMSE is 4.4 cm/s and the RMSEs is 0.57 (Table 3).

The across-track geostrophic velocity derived from the MDT is smooth and does not capture the smaller-scale variability observed in the fields derived from DH and ADT (Figure 8). This is translated in a low correlation coefficient of -0.13, a higher RMSE of 12.3 cm/s and a lower RMSEs of -0.21.

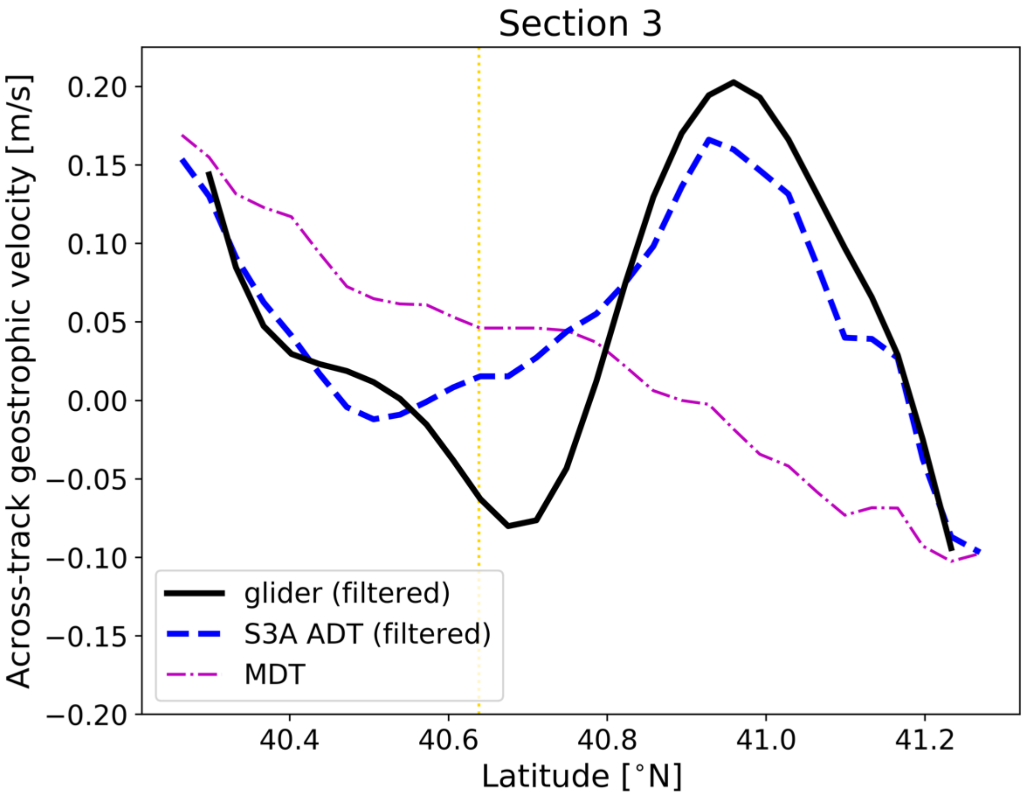


Figure 8. Across-track geostrophic velocity computed from glider filtered DH (solid black line), S3A filtered ADT (dashed blue line) and MDT (dotted-dashed magenta line) along section 3. Across-track velocities that are oriented to the northeast (southwest) have.

Table 3. Statistics between the across-track geostrophic velocity computed from glider filtered DH vs. Sentinel-3A filtered ADT and MDT along section 3.

Glider data	**Product**	**Correlation coefficient**	**RMSE [cm/s]**	**RMSEs**
Filtered DHa	Filtered ADTa	0.89	4.4	0.57
	MDT	-0.13	12.3	-0.21

Regarding section 7, the across-track geostrophic velocity calculated from glider DH has a pattern similar to the velocity estimated from S3A ADT (Figure 9). However, the DH-derived velocity pattern is displaced about 0.1º northwards along the latitude axis with respect to the ADT-derived velocity, being both velocities similar in the north, i.e., at the end of the glider transect, when the time of both observations is closer. These temporal and spatial discrepancies impact the statistical metrics used for the comparison (Table 4): the correlation coefficient is 0.79, the RMSE is 9.6 cm/s and the RMSEs is 0.27.

As in section 3, the across-track geostrophic velocity derived from the MDT is smooth and does not capture the smaller-scale variability observed in the velocities calculated from DH and ADT (Figure 9). This is translated in a low correlation coefficient of 0.40, a higher RMSE of 12.0 cm/s and a lower RMSEs of 0.09 (Table 4).

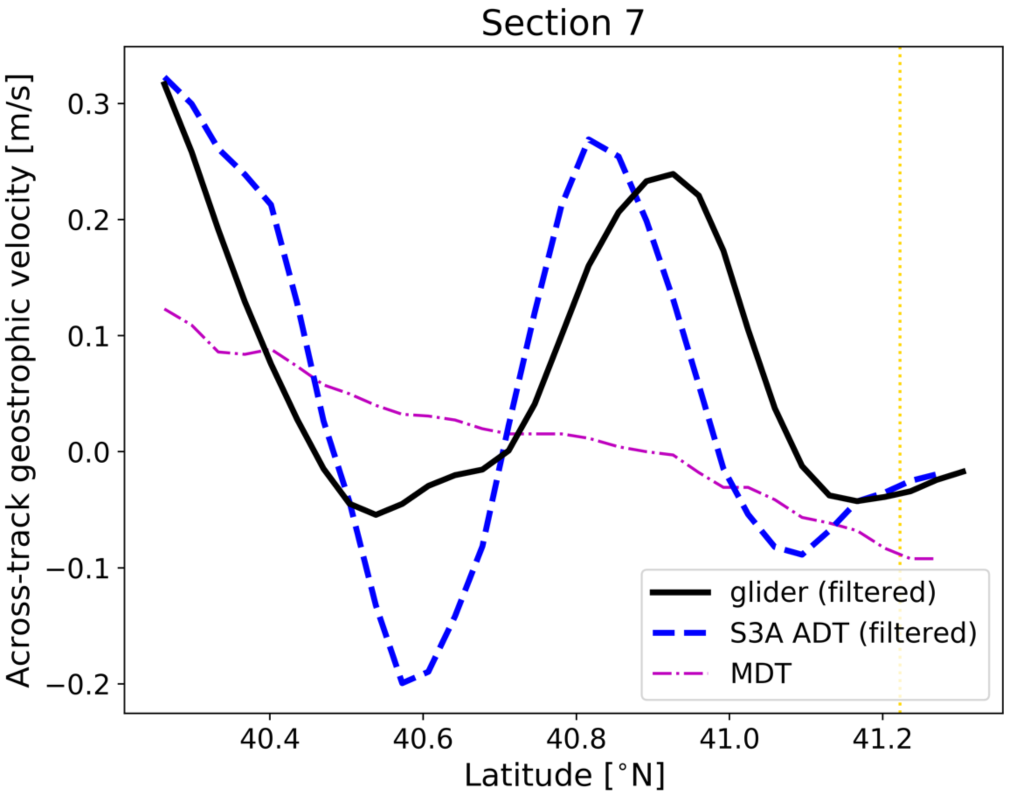


Figure 9. Across-track geostrophic velocity computed from glider filtered DH (solid black line), S3A filtered ADT (dashed blue line) and MDT (dotted-dashed magenta line) along section 7. Across-track velocities that are oriented to the northeast (southwest) have positive (negative) sign. The vertical yellow dotted line represents the glider position during the S3A overpass.

Table 4. Statistics between the across-track geostrophic velocity computed from glider filtered DH vs. Sentinel-3A filtered ADT and MDT along section 7.

Glider data	Product	Correlation coefficient	RMSE [cm/s]	RMSEs
Filtered DHa	Filtered ADTa	0.79	9.6	0.27
	MDT	0.40	12.0	0.09

Comparison of the glider-derived geostrophic velocity with the DAV

The DAV computed from the GPS glider positioning has some correspondence with the shape of the DH-derived geostrophic velocity for both sections (Figure 10 and Figure 11). However, there are important dissimilarities, especially in the magnitude of both fields. The DH-derived geostrophic velocity has a higher magnitude than the DAV. Note that for the calculation of the geostrophic velocity we have assumed a reference level of no motion at 690 m, and this velocity is the result of the baroclinic geostrophic contribution. On the other hand, the DAV has a different physical content that includes ageostrophy, high frequency barotropic signals, cyclostrophy and inertial currents (Bouffard et al., 2010). Also, the DAV estimates can have an associated error of 2-3 cm/s due to the attitude sensor of the platform and the angle of attachment (Merckelbach et al., 2008). The correlation coefficients between both fields are 0.82 and 0.72, and the RMSE are 9.7 cm/s and 13.2 cm/s for sections 3 and 7, respectively (Table 5 and Table 6).

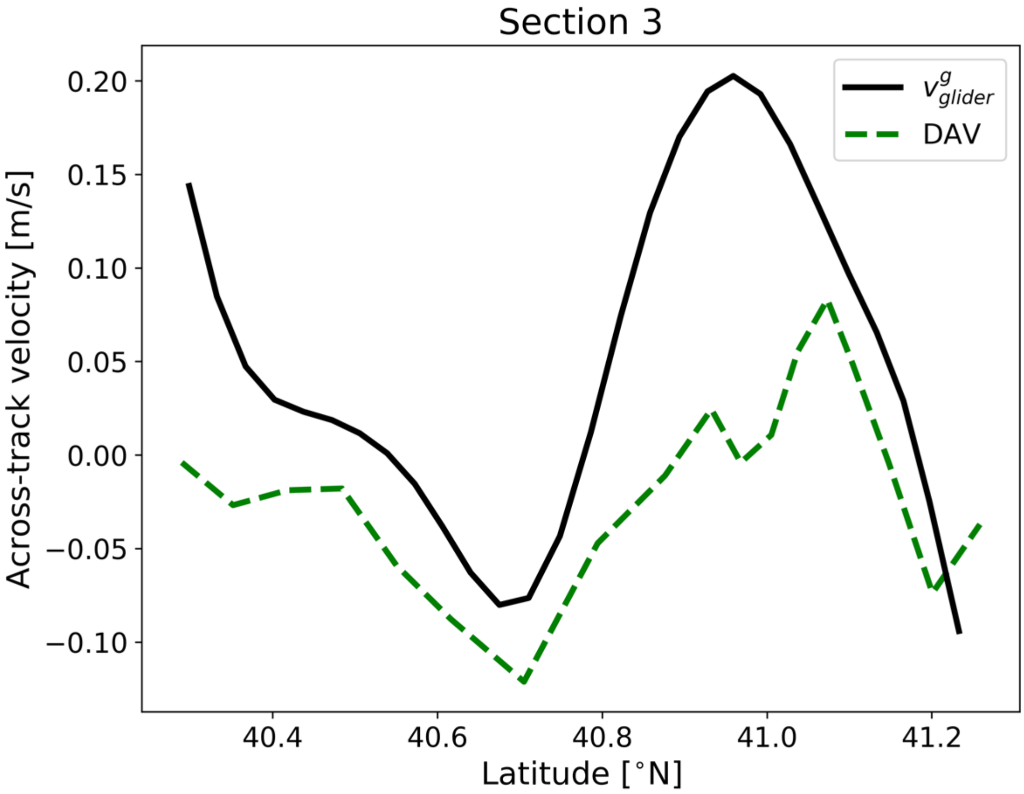


Figure 10. Across-track geostrophic velocity computed from glider filtered DH (solid black line) and across-track component of the glider DAV (dashed green line) for section 3.

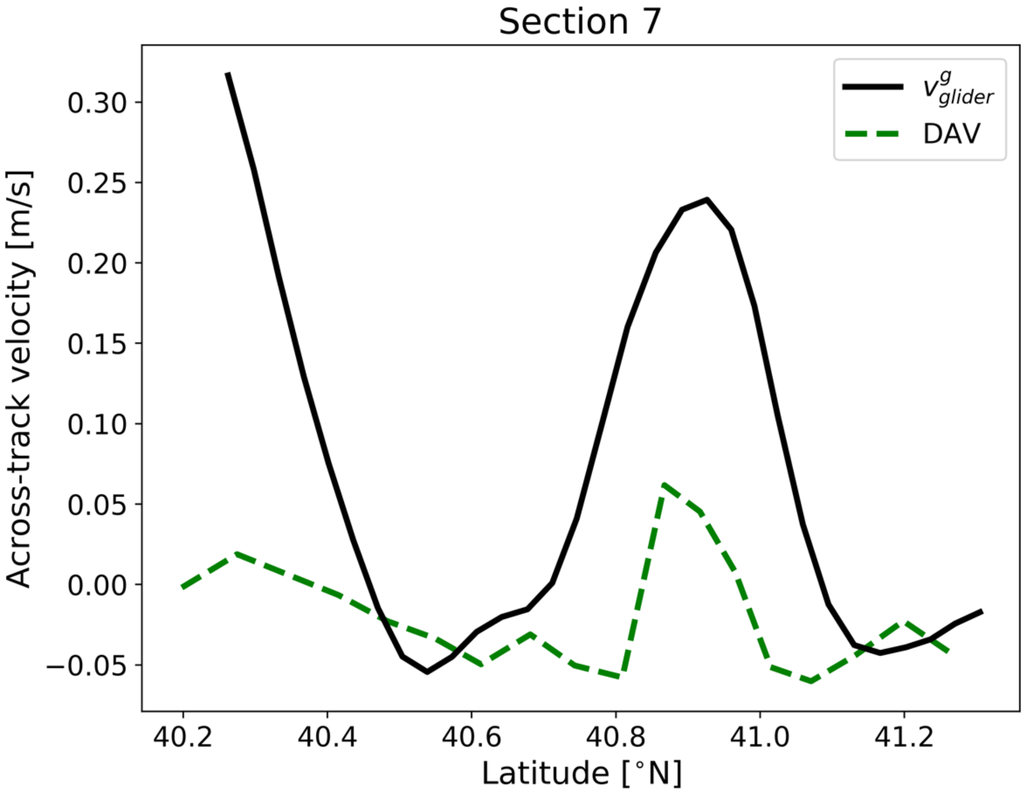


Figure 11. Across-track geostrophic velocity computed from glider filtered DH (solid black line) and across-track component of the glider DAV (dashed green line) for section 7.

Table 5. Statistics between the across-track geostrophic velocity computed from glider filtered DH vs. the across-track component of the DAV for section 3.

Correlation coefficient	**RMSE [cm/s]**
0.82	9.7

Table 6. Statistics between the across-track geostrophic velocity computed from glider filtered DH vs. the across-track component of the DAV for section 7.

Correlation coefficient	**RMSE [cm/s]**
0.72	13.2

1.4. Conclusions

The Sentinel-3A altimeter observations along track 57 have been validated with glider measurements conducted along the same track in March-April 2021. The DH calculated from glider observations shows good correspondence with the S3A ADT. In section 3, the correlation coefficient between filtered fields is 0.95 and the RMSE is 0.6 cm. The higher difference is detected in a salinity front not captured by altimetry. Along section 7 the two fields show higher dissimilarities related to the difference in time between observations. The correlation coefficient is 0.86 in this case, and the RMSE is 1.3 cm. The comparison of the across-track geostrophic velocities reveals similar results. In section 3 the correlation coefficient is 0.89, and the RMSE is

4.4 cm/s. In section 7 there is an offset between both fields of ~0.1º along the latitude axis, except at the end of the glider track, where the time of both observations is closer. The statistics in this case indicate a correlation coefficient of 0.79 and a RMSE of 9.6 cm/s.

The glider-derived geostrophic velocity has been compared to the DAV (depth-averaged-velocity) computed from the GPS glider positioning. The across-track components of both fields show a similar shape, while there are significant differences in their magnitudes. Dissimilarities between both fields are expected because they represent a different physical content. While the geostrophic velocity is the result of the baroclinic geostrophic contribution, the DAV also includes ageostrophy, barotropic signals and inertial currents (Bouffard et al., 2010). In addition, the DAV estimates can have errors of 2-3 cm/s. The statistics between both fields reveal a correlation coefficient of ~0.7-0.8 and a RMSE of the order of 10 cm/s.

2. Validation of Sentinel-3 altimeter observations with HF radar and ADCP in the Bay of Biscay (AZTI)

2.1. Introduction

Satellite measurements have multiple sources of errors such as errors in the orbit, range and/or the applied corrections. In addition, some of the errors can be amplified in coastal areas because of the degradation of the measurements, due to the contamination of the altimeter footprint area by the land and the inaccuracies in geophysical (atmospheric and tidal) corrections (Bouffard et al., 2010). Given the errors within the satellite measurements, coastal in-situ observations have great potential for the long-term validation of those measurements (Bonnefond et al., 2011).

For instance, recent studies on the evaluation of the performance of altimetry using high-frequency (HF) radars, concluded that HF radars offer a way to improve the validation and accuracy of altimetry products for coastal areas (Chavanne and Klein, 2010; Liu et al., 2012; Pascual et al., 2015; Troupin et al., 2015; Roesler et al., 2013). One of the most extended approaches found in the literature to study the synergy between altimetry and HF radar data consists of the comparison of the total across-track currents in the along-track direction (e.g., Morrow et al., 2017; Troupin et al., 2015; Pascual et al., 2015).

In the south-eastern Bay of Biscay (SE-BoB), altimeter measurements were already compared to HF radar measurements in terms of current velocities (Manso-Narvarte et al., 2018), observing that the best agreement was identified in areas and periods of high geostrophic signal with correlations up to 70%. The mesoscale surface circulation in this area is strongly affected by the slope current that flows poleward in winter along the Spanish and French slopes (see Figure 11) and in summer it is reversed and weakened (e.g., Charria et al., 2013; Solabarrieta et al., 2014). At the surface, the circulation is also driven by winds (e.g., González et al., 2004; Solabarrieta et al., 2015) with prevailing southwesterlies in autumn and winter that induce a northward and eastward drift over the shelf. During spring, the wind regime changes to northeasterlies inducing a westward and south-westward drift, with a similar situation during summer. However, weak, and variable winds make wind-driven currents more uncertain in spring and summer, challenging the correct combination with altimetry data. Also, in general (and despite the season), over the shelf, wind-driven currents prevail over tidal or density-driven currents in this area, due to the narrow shelves and the low influence of rivers that induce significant density currents (e.g., González et al., 2004; OSPAR, 2000). In Manso-Narvarte et al. (2018), the highest correlations and lowest errors were observed within the slope and for winter periods for the track 248 of the Jason-2 mission, due to the intense and persistent geostrophic signal of the slope current (see Figure 12a). Mesoscale eddies also provided a good qualitative agreement (see Figure 12b).

The current availability of the HF radar and moored ADCP observations and a higher number and wider coverage of altimetry observations in the area enables further analysis of the agreement between altimeter-derived and in-situ currents. Therefore, this contribution consists in identifying the most appropriate conditions for comparing along-track altimetry-derived currents with HF radars measurements and in quantifying the measurement errors of altimetry, HF radar and ADCP-derived currents by means of the Triple Collocation (TC) method. To this end, data from Sentinel 3A (S3A) and Sentinel 3B (S3B) satellites were compared with HF radar and moored ADCP data.

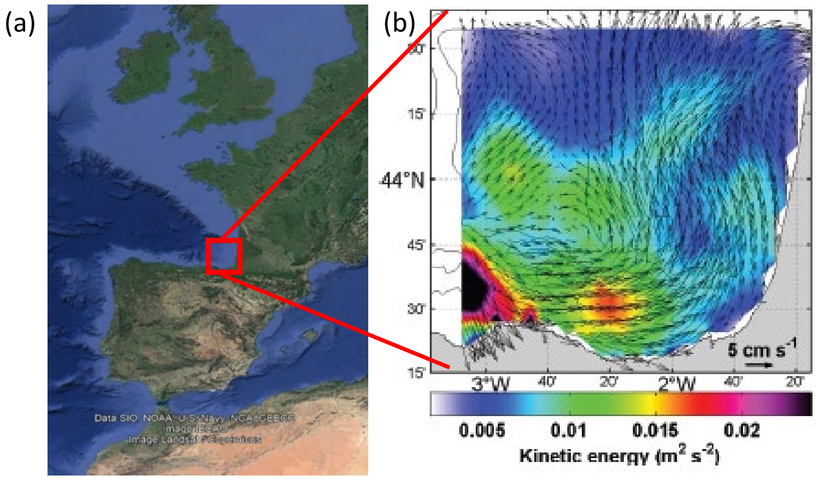


Figure 11. (a) Location of the study area. (b) Surface currents and kinetic energy per unit of mass (in m2·s-2) for winter from HF radar observations (Caballero et al., 2020).

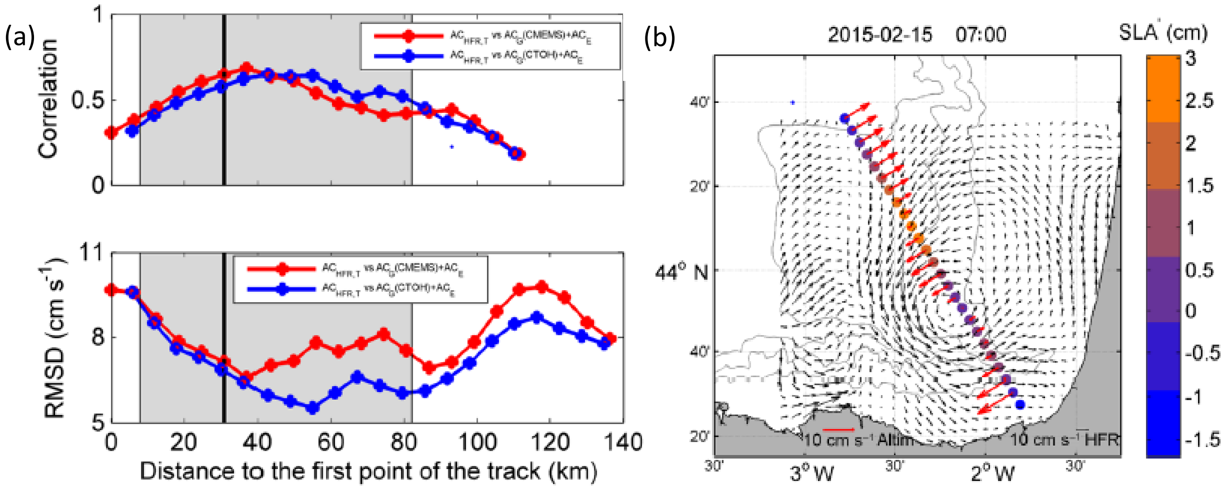


Figure 12. Figure extracted from Manso-Narvarte et al. (2018). (a) Correlation and root mean square difference (RMSD) between altimetry-derived and HF radar-derived currents along Jason-2 track 248 from the coast towards the open ocean for two different altimetry products (in red and blue). The grey-coloured area corresponds to the slope between the 200 and 1000 isobaths. (b) The dots show the points of track 248. SLA values are indicated in the colour scale. Black arrows depict the HFR current fields. Red arrows correspond to across-track altimetry-derived currents. Grey lines: 200, 1000, and 2000m isobaths. Note that the scale of each kind of arrow is not the same.

2.2. Data and Methods

Data

On the one hand, the in-situ data used as a reference for validating the satellite measurements are HF radar and ADCP data from the Basque Operational Oceanography System (EuskOOS[3],Figure 13). The HF radar system is composed of two sites, one in Cape Higher and another one in Cape Matxitxako and works at a central frequency of 4.46 MHz with an operational bandwidth of 30 kHz. It provides hourly surface current velocity fields grided onto a regular orthogonal mesh of 5 km resolution covering ~150 km off the coast (displayed in Figure 13). This system has provided data since 2009 with some interruptions mostly due to severe atmospheric conditions and its validity to study coastal processes and transport patterns has been widely demonstrated (e.g., Rubio et al., 2011, 2018; Solabarrieta et al., 2014, 2015, 2016). In order to perform the comparisons against altimetry data, the biggest spatial coverage possible is needed, therefore spatially gap-filled currents were used, filled by an open-mode analysis (OMA) (Kaplan and Lekien, 2007), which is a widely used method for filling gaps in HF radar current fields. The ADCP is co-located at the Donostia mooring, anchored in a water depth of 550 m at the Spanish slope (at 43.56°N−2.03°W, red point in Figure 13) and works at a central frequency of 150 kHz providing hourly current velocity data from −12.26 m extending down 200 m in the water column with bins every 8 meters, since 2007. However, in this study, only the data of the first bin, at −12.26 m, is going to be considered. The performance of the ADCP has been demonstrated in several works (e.g., Rubio et al., 2013a; Solabarrieta et al., 2016, 2014).

Wind data, used as input in the Ekman model (see Section 2.2), were obtained from the Weather Research and Forecasting model (WRF[4]) provided by the meteorological agency of Galicia (MeteoGalicia). This model, with a native resolution of 12 km, reproduces the offshore wind fields of the SE-BoB with reasonable accuracy (Ferrer et al., 2010).

On the other hand, details of the S3A&B altimetry sea level anomaly (SLA) data selected for the validation, are shown in Figure 13 and Table 7. These data were extracted from the reprocessed SEALEVEL_GLO_PHY_L3_MY_008_062 product of CMEMS.

---

[3] https://www.euskoos.eus/en/

[4] http://mandeo.meteogalicia.es/thredds/catalog/modelos/WRF_HIST/d02/catalog.html

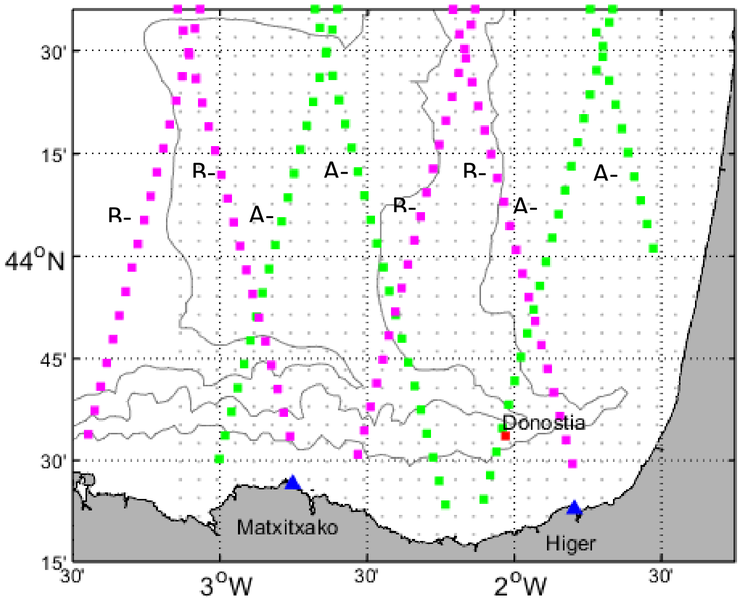


Figure 13. HF radar system (antennas represented by the triangles, the regular grid used for OMA currents represented by the grey dots) and ADCP (red square) location in the study area. Tracks 102, 216, 257 and 371 of S3A (green) and S3B (magenta).

Table 7. Attributes of the altimetry data used.

Altimeter	**Period**	**Tracks**	**Spatial resolution**	**Revisit period**
Sentinel 3A	From 16-03-2016 to 08-12-2020	102, 216, 257 and 371	7 km	27 days
Sentinel 3B	From 20-12-2018 to 18-12-2020	102, 216, 257 and 371		

Finally, mean dynamic topography (MDT) data were obtained from the COMBAT project (Caballero et al., 2020), which was improved by the inclusion of EusKOOS HF radar data in the MDT computation for the study area. These data were used to obtain absolute dynamic topography (ADT) data from the summation with SLA data.

Methods

The currents observed by in-situ platforms and the ones derived from altimetry capture different components of the flow. The altimetry-derived currents measure the geostrophic component, whereas the others measure both the geostrophic and ageostrophic components of the current. Therefore, data were processed making them closer to each other for the comparisons. In order to remove the HF and ageostrophic signals contained in the in-situ data, such datasets were low pass (LP) filtered. To that end, different filters were tested to find out the one that provides the best results as was done in the study area in Manso-Narvarte et al. (2018). The sensitivity tests showed that the 2-day running average filter provided the best agreement between altimetry and in-situ derived currents (not shown). Currents derived from the HF radar were interpolated to the altimetry along-track points and rotated in across-track direction, while the ADCP-derived ones were also rotated.

The altimetry-derived across-track currents (V_g) were obtained by finite-differentiating the ADT (obtained by the summation of the SLA and MDT). More precisely, such currents were computed by the geostrophic velocity equation using a three-point central difference operator:

$$V_g = \frac{g}{f}\frac{\Delta(ADT)}{\Delta x}$$

where g is the acceleration of gravity, f is the Coriolis parameter and x is the along-track distance. In order to obtain a similar spatial smoothing to that of the currents obtained from in-situ observations and reduce the noise provided by finite differentiating, the altimetry-derived currents were spatially averaged along-track with their 4 adjacent points. This spatial window provided the best results in Manso-Narvarte et al. (2018), where the spatial resolution of the altimeter data was the same as in this case.

In order to make altimetry-derived currents closer to in-situ LP filtered current data, the low-frequency wind-driven Ekman currents were added to the altimetry-derived ones as in Liu et al. (2012) and Manso-Narvarte et al. (2018). Ekman currents were obtained following the M1 model proposed by Rio and Hernandez (2003) that previously provided good results in the study area (Caballero et al. 2008, Manso-Narvarte et al. 2018). Once the Ekman currents were estimated they were LP filtered as the in-situ data to obtain the same low-frequency signal and then interpolated to the altimetry-derived current points and rotated in across-track direction. Note that sensitivity tests using and not-using Ekman currents were carried out (compare the results in Table 9 and Table 10) with a better agreement when they were used, as observed in previous studies (Liu et al., 2012; Manso-Narvarte et al., 2018).

In summary, 2-day LP filtered currents derived from in-situ observations were compared to altimetry-derived + LP filtered Ekman currents (hereinafter altimetry+Ekman). Table 8 shows the different data sources, their processing and the final variables used. These variables were compared by different approaches explained in the following subsections.

*Table 8. Summary of the data sources, the processing of the data and the final variables used. *Ekman, HF radar and ADCP currents were LP filtered to remove most of the ageostrophic signal.*

Source	**Processing**	**Final variable used**	**Used in the triple collocation (TC) and/or Along-track method (AT)**
Reprocessed SLA along the tracks of S3A&B (SEALEVEL_GLO_PHY_L3_MY_008_062 product)	Geostrophic across-track currents	Across-track geostrophic + Ekman* currents along the track's points (quantity already used in Liu et al. (2012) and Manso-Narvarte et al. (2018))	TC and AT
MDT from Combat project that assimilated HF radar data (Caballero et al., 2020)			
Wind data from Meteogalicia	Ekman currents interpolated to the altimetry along-track points and rotated in across-track direction		
HF radar surface total current fields	Currents interpolated to the altimetry along-track points and rotated in across-track direction	Across-track currents* along the track's points	TC and AT
ADCP total currents of the first bin (-12.26 m)	Currents rotated in across-track direction	Across-track currents* at the mooring location	TC

Along-Track

This approach compares the altimetry-derived and HF radar-derived across-track currents along S3A&B tracks by means of correlation (significance of 90%), root mean square difference (RMSD) and RMSD relative (RRMSD) to HF radar's root mean square (RMS, see Figure 14). The number of simultaneous HF radar and altimetry-derived current data pairs used in the comparisons changes depending on the along-track point, ranging between 33 and 46 for S3A and 14 and 18 for S3B. The study area was split into 4 different zones of different dynamical characteristics in order to analyse the sensitivity of the comparisons for different scenarios (see Figure 14):

- Zone 1: Spanish slope area affected by the slope current (intense RMS and mean KE values, see Figure 11b and Figure 14).
- Zone 2: Open ocean.
- Zone 3: French platform where currents are mainly driven by winds and tides.
- Zone 4: French slope area affected by the slope current (but with less intensity than in Zone 1).

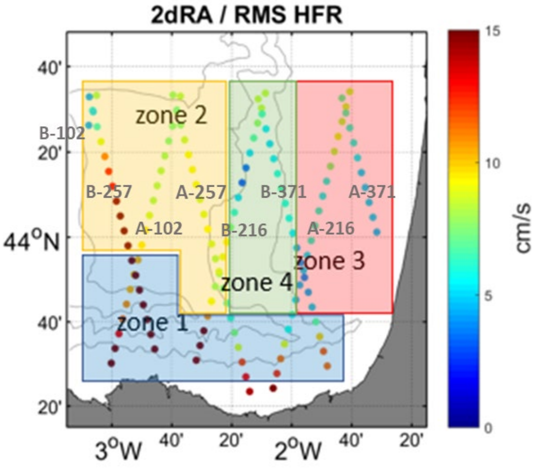


Figure 14. The RMS at the S3A&B altimeter tracks, estimated from 2-day running averaged HF radar observations. The area is split into 4 zones of different dynamics.

Triple Collocation

The triple Collocation (TC) quantifies the random error (hereinafter RMSE or ε_i) of three datasets of the same geophysical variable by combining the covariances between them (e.g., Mignot et al., 2019). Moreover, if one data set is assumed to be perfectly calibrated, the TC can provide the additive (hereinafter offset or α_i) and multiplicative (hereinafter gain or β_i) instrumental biases of the two other datasets and the RMSE rescaled (rRMSE) to the data space of the dataset assumed as perfectly calibrated (in order to compare the 3 platforms at the same level). The error model is as follows:

$$d_i = \alpha_i + \beta_i t + \varepsilon_i$$

where d_i are the three collocated datasets (i=1, 2, 3) and t is the unknown true value.

To apply the method, simultaneous altimetry+Ekman and HF radar-derived across-track velocities in the points inside black circles (Figure 15) and ADCP-derived currents at the ADCP location (red point in Figure 15) were gathered as data triplets. Note that the two different points highlighted by black circles were in areas of very similar dynamics and were jointly considered for increasing the number of triplets. However, since still only 51 simultaneous data triplets were obtained (few for estimating robust covariances) 1,000 bootstrap simulations were run (as in Mignot et al., 2019) and mean RMSE, rRMSE, gain and offsets were obtained assuming that the measurements of the ADCP were perfectly calibrated (since it is the platform that provides the most in-situ measurements).

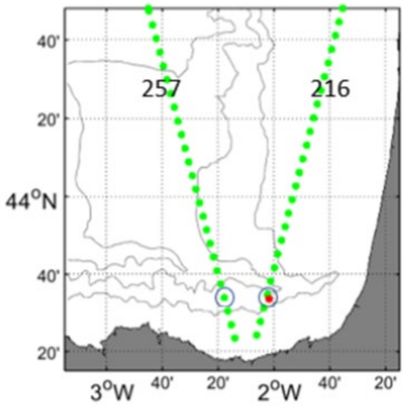


Figure 15. Points of tracks 257 and 216 of S3A (green). ADCP location (red point). The altimetry points used for the TC are depicted inside black circles.

2.3. Results and Discussion

Along-Track

First, if we compare the results using (i.e., altimetry+Ekman) and not-using Ekman (i.e., altimetry) we observe that the addition of the Ekman currents improves the results (compare the results in the general case in Table 9 and Table 10). The correlation increases by 19% whereas the RMSD and RRMSD slightly decrease. This effect was already observed in the study area where an increase in the correlation of up to 10% was obtained (Manso-Narvarte et al., 2018). Since the along-track comparison covers a wide area and the addition of Ekman currents improves the results, the altimetry+Ekman dataset was used in all the following comparisons of this study.

By using the along-track method for the validation, different results were obtained depending on the geographical area (Table 10 and Figure 16). In zone 1, strongly affected by the slope current, high correlations and relatively low RRMSDs were obtained. This behaviour was already observed in the study area by Manso-Narvarte et al. (2018), obtaining similar maximum correlation values. In zone 2, where the geostrophic signal is lower than in zone 1, medium correlations and RRMSDs were observed. However, for the track 257 of S3B high correlations and low RRMSDs were obtained (points inside the black circle in Figure 16) where an intense current regime was observed (see Figure 16). In zone 3, an area of low geostrophic signal, correlations and RRMSDs were low and high, respectively. Finally, in zone 4, very low correlations and high RRMSDs were observed. This is somehow surprising since this is an area potentially affected by the slope current where a good agreement was observed in a previous study (Manso-Narvarte et al., 2018; see Figure 12). This unexpected result might be because the altimeter tracks did not sample the places or dates where the core of the slope current was flowing as shown by the weak RMS values observed by the HF radar (Figure 14). Note that the pointwise correlations shown in the maps and the correlations shown in the table for each zone might be derived from a slightly different number of data since the zonal results of the table comprise all the data of the points of each zone that might be discarded in the maps due to low significance.

In general, the goodness of the results shows a marked sensitivity to the dynamics of the area, that is, the persistence of the currents and the strength of the geostrophic component provide better results. This behaviour was already observed in the study area (Manso-Narvarte et al., 2018), and in this study the wider spatial coverage corroborates it. Therefore, when comparing or validating altimetry-derived currents with HF radar-derived currents the dynamics of the area should be considered, and periods or areas of a strong geostrophic signal are recommended.

Table 9. Correlation, RMSD and RRMSD between altimetry and HF radar-derived currents along S3 A&B tracks.

	RMSD (cm/s)	RRMSD	Correlation
General	9.34	1.00	0.32

Table 10. Correlation, RMSD and RRMSD between altimetry+Ekman and HF radar-derived currents along S3 A&B tracks for 4 different zones.

f	RMSD (cm/s)	RRMSD	Correlation
Zone 1	8.99	0.70	0.68
Zone 2	9.28	1.04	0.46
Zone 3	7.96	1.29	0.37
Zone 4	9.79	1.61	0.17
General	8.91	0.95	0.51

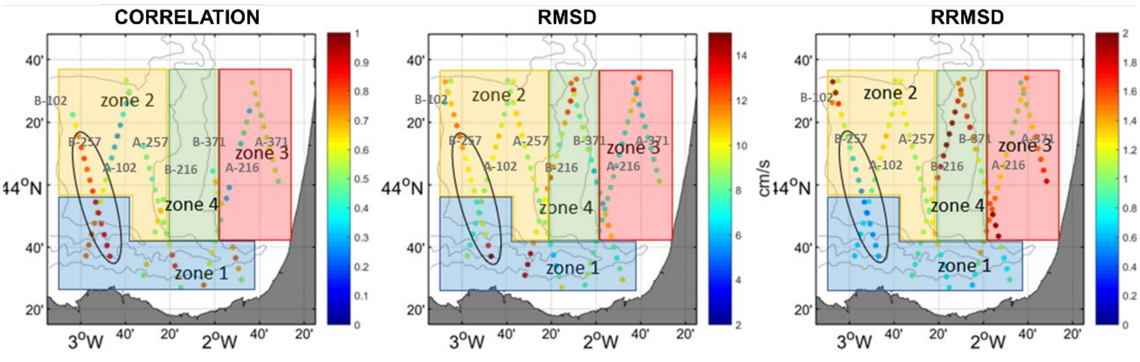


Figure 16. Correlation, RMSD and RRMSD between altimetry+Ekman and HF radar-derived currents along S3A&B tracks in the SE-BoB. The study area is split into 4 zones of different dynamics.

Triple Collocation

As stated before, the TC method allows estimating the random and instrumental errors of the collocated observations based on the covariances between them. Despite quality controlling the data, given the very low covariances between altimetry+Ekman and ADCP observations in some of the bootstrap simulations, mean RMSE, gain and offset values of the HF radar were very high, providing a high standard deviation (STD) of the mean values that made them not representative. In order to avoid the effects of those outlying values, the iterations where data triplets provided covariances between altimetry+Ekman and ADCP under a threshold value were repeated. This threshold value was set as a compromise between considering a high number of times each of the 51 triplets (over 800 times out of 1000 iterations) and providing lower STD values than the mean values of the errors, making those mean values representative (as can be seen in Table 11).

Once this modification was made, the results show that the ADCP has a mean RMSE of 3.4 cm/s, whereas the altimetry+Ekman provides slightly higher values when rescaled to the ADCP data space, suggesting that both datasets are equally precise in terms of random error. The HF radar provides even lower rRMSE values of 1.13 cm/s and therefore a higher precision (see Table 11). With respect to the range of each dataset, these values represent a relative random error of 14.5%, 7.6% and 13.2% for ADCP, HF radar and altim+Ekman, respectively. The altimetry+Ekman tend to overestimate the currents (Figure 17a), whereas the HF radar tend to underestimate them (Figure 17b), with respect to the ADCP.

Concerning the mean gains and offsets, the ADCP and HF radar differ with a gain of 1.49 and an offset of -0.96 cm/s, but less than between ADCP and altimetry+Ekman with a gain of 1.98 and an offset of 3.85 cm/s. Therefore, as expected altimetry+Ekman performs worse than the HF radar in estimating the velocities with respect to the ADCP in terms of instrumental error.

Table 11. TC estimates of gain, offset, RMSE, and Rescaled RMSE (rRMSE) for the ADCP, HF radar and altimetry+Ekman datasets. Note that the mean and standard deviation (given in parentheses) of the 1,000 bootstrap estimates are shown for each estimate, except for

	ADCP	HF radar	altim+Ekman
<Gain>	1	1.49(0.19)	1.98(0.37)
<Offset> (cm/s)	0	-0.96(0.74)	3.85(1.55)
<RMSE> (cm/s)	3.40(0.53)	1.69(0.75)	6.67(1.19)
<rRMSE> (cm/s)		1.13(0.47)	3.49(0.90)

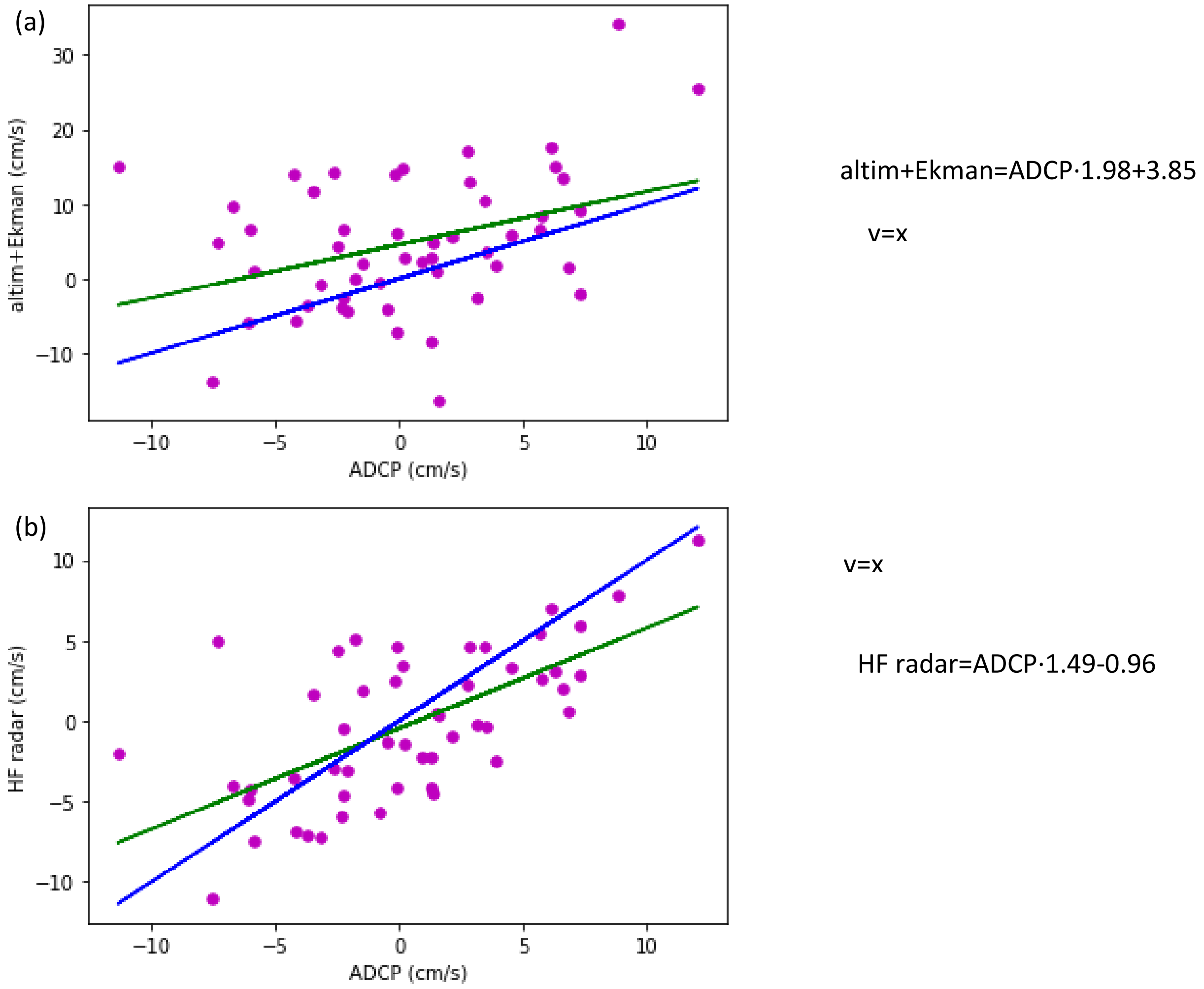


Figure 17. Scatter plots of (a) altimetry+Ekman and (b) HF radar currents as a function of ADCP data. Blue line represents line y=x. The green line is the linear regression between the two datasets.

2.4. Conclusions

In order to detect the most appropriate conditions for comparing altimetry data against in-situ HF radar and ADCP measurements for validation exercises of altimeters and to quantify the measurement errors, after the processing of the datasets, two approaches were used. Concerning the along-track comparison of the currents, it was observed that the dynamics of the area should be considered and that isolating, for the comparison, periods or areas of a strong geostrophic signal is recommended. With regard to the TC, a similar precision was observed for the ADCP and the altimeter in terms of random error (around 3.45 cm/s) whereas the altimeter measurements differed from the ADCP in terms of instrumental error with a gain (i.e., multiplicative error) of 1.98 and an offset (i.e., additive error) of 3.85 cm/s. It was also observed that the addition of the Ekman currents into the analysis improved the agreement between in-situ and altimetry-derived currents, as in previous studies. This suggests that the sensitivity to the addition of the wind effects should be considered in future altimeter validation comparisons, at least in terms of current velocities.

3. Global validation of altimetric data (Sentinel-3) with in-situ observations (CLS)

3.1. Sea level anomalies validation

Introduction

The main objective of this sub-task is to validate altimetry data, here along track Sentinel3B (S3B) SAR altimeter observations in terms of sea level and current using two global ocean observing systems: Argo and surface drifters. The triple collocation analysis (TCA) will be applied to estimate the random error associated with the altimeter measurement. In addition to the triple collocation analysis, usual validation metrics (e.g. rms, bias, correlation) will be also retrieved. The study will be performed at a global scale for the open ocean, as well as on a regional scale.

TCA is a method for quantifying the standard deviation of the random error of 3 datasets of the same geophysical variable by combining the covariances between the datasets (Mignot et al. 2019). The first output of the study will be the observational errors (=systematic error (bias/offset) + random error) associated to each of the 3 datasets. The second output will take the form of recommendations for the design of the in-situ observing system for the validation of satellite altimetry & observational error analysis.

The following datasets for sea level and currents will be used:

- In-situ: Argo, surface drifters (acting as the reference)
- Along-track altimetry: S3B
- Altimeter L4 products: C3S mapped sea level anomalies (specific L4 maps, without S3B)

The tasks to be carried out include:

- Set up of the method
- Sensitivity analysis:
    - Using different L4 products (L4 satellite, model outputs)
    - Using different configuration of the in-situ observing system (Argo/2, drifters/2)
- Analysis over at least one year (2019 or 2020) + seasonal cycle
- Analysis at global scale for the open ocean + the North Atlantic Ocean, Balearic Sea and South East Bay of Biscay

Data

In situ data

We use the temperature and salinity profiles from Argo floats and delivered by the INSitu Thematic Assembly Center (INSTAC) of the Copernicus Marine Service. From these profiles, we calculate the dynamic height referenced to 900m. Then we subtract a 2003-2019 synthetic reference climatology.

Combined gridded sea level anomalies

We use the climatic time series of Sea Level Anomaly gridded Maps (MSLA) computed by the Copernicus Climate Change Service (C3S[5]) and distributed by the Copernicus Marine Service[6]. These maps are computed using only 2 satellites, over our time period of study (2019-2020) the maps used Jason-3 (J3) and Sentinel-3A (S3A) data, thus Sentinel-3B (S3B) is an independent dataset. The anomaly in this dataset is referenced to the 1993-2012 period.

Along-track sea level anomalies from altimeters

We use the along-track sea level anomalies (SLA) from the S3B satellite computed by the Copernicus Marine Service Sea Level Thematic Assembly Center (SLTAC). There are used as an independent dataset from the dynamic height and from the gridded C3S product. The SLA is re-referenced to the same period as the dynamic height.

In order to see the impact of the SLA, three processing-level steps of S3B will be considered here:

- **L2P**: These SLA are not corrected for the orbit error and long wavelength error.
- L3: level 3 data corrected for the orbit error and long wavelength error (Hereafter **L3U**)
- Filtered L3: level 3 data corrected for the orbit error, long wavelength error and filtered (Hereafter **L3F**)

Temporal mean of MSLA

Some ancillary data are used in order to reference all the data sets over the same time period. A mean sea-level anomaly over 2003-2019 is calculated from the MSLA maps to subtract from the gridded and along-track data such that the anomalies are referenced to the same time period as the Argo (i.e. 2003-2019).

Methods

In order to evaluate the altimetric performances using in-situ data, we use different metrics:

Absolute difference mean error from three datasets

Select collocated data within 75% of the ocean scales and estimate an error based on:

Mean (1/3 (|X-Y| + |X-Z| + |Z-Y|)) per grid cell.

Triple collocation analysis

The TCA as describe in (Janssen, 2007) is a method to quantify the random error standard deviation RMSE using the covariances of three independent datasets (thus assuming they have uncorrelated random error). The systematic errors are supposed eliminated via data quality control.

The three different datasets need to be collocated.

- Argo Insitu dynamic height estimated from the temperature and salinity profiles (reference depth is 900m here).
- MSLA are interpolated on in-situ locations

---

[5] https://climate.copernicus.eu/

[6] https://doi.org/10.48670/moi-00145

- Along track, SLA are nearest neighbours (time and space) of in-situ locations using KDTree with data “non-dimensionalized” by the ocean scales (see 4.2). Here we start with 1 degree in longitude, and latitude, and 6 days in time, and then it is refined later.

Thus, we have 3 collocated datasets.

The method is very sensitive to outliers, thus data lower or larger than 1.5 times the first and last quarters are eliminated, see (Tukey, 1977) as referenced in (Mignot et al, 2019).

The in-situ data are assumed perfectly calibrated, with quality control indicators QC = 1 (good) taken for temperature and salinity.

If the amount of data is deemed insufficient for robust covariance estimation, or extra statistics are needed, a bootstrapping method can be used to assess the impact of the number of available data. The bootstrap method is a statistical technique to estimate quantities about a population by averaging estimates from multiple small data samples. Importantly, samples are constructed by drawing observations from a large data sample one at a time and returning them to the data sample after they have been chosen.

Oceanic correlation scales and oceanic signal variance impact

The selection of matchups from the 3 different datasets is based on the oceanic time and space scales to ensure that the data from the matchups observes the same phenomenon. Also, in areas of large oceanic signal variance, a slight difference in time and space between the data from a matchup can involve huge differences. For example, in Figure 18, one can see for the L3U along-track data (center picture), that the random error from the TCA analysis increases when data are farther away, and when the signal variance increase. It also shows on the density and repartition of the matchups over the globe. The left and right center pictures are from less processed (L2P) to better processed (L3F) along track data. Note that the better processed data have a lower random error than the less processed data since a better adjustment between the three independent dataset (dynamic height, along track and gridded data) are in better agreement.

The time and space ocean correlation scales from DUACS (see Figure 18) are used to determine the “closeness” of the ARGO/S3B pair. Pairs are formed within a generic 1-degree longitude by 1-degree latitude by 6-day. Then selected pairs are refined locally with the time and space distances allowed to form a pair within 75% of the correlation scales, to ensure that the pair represents a large part of the same oceanic signal.

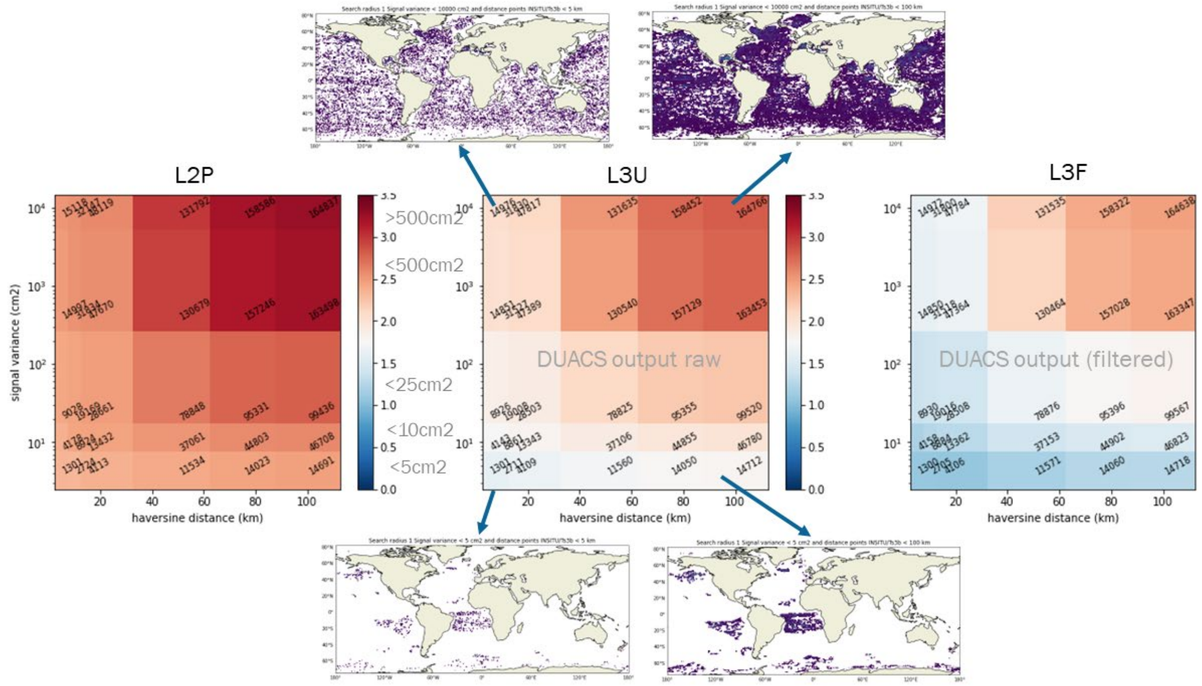


Figure 18. impact of the spatial distance and/or the signal variance on the number of selected matchups and RMSE output from the TCA.

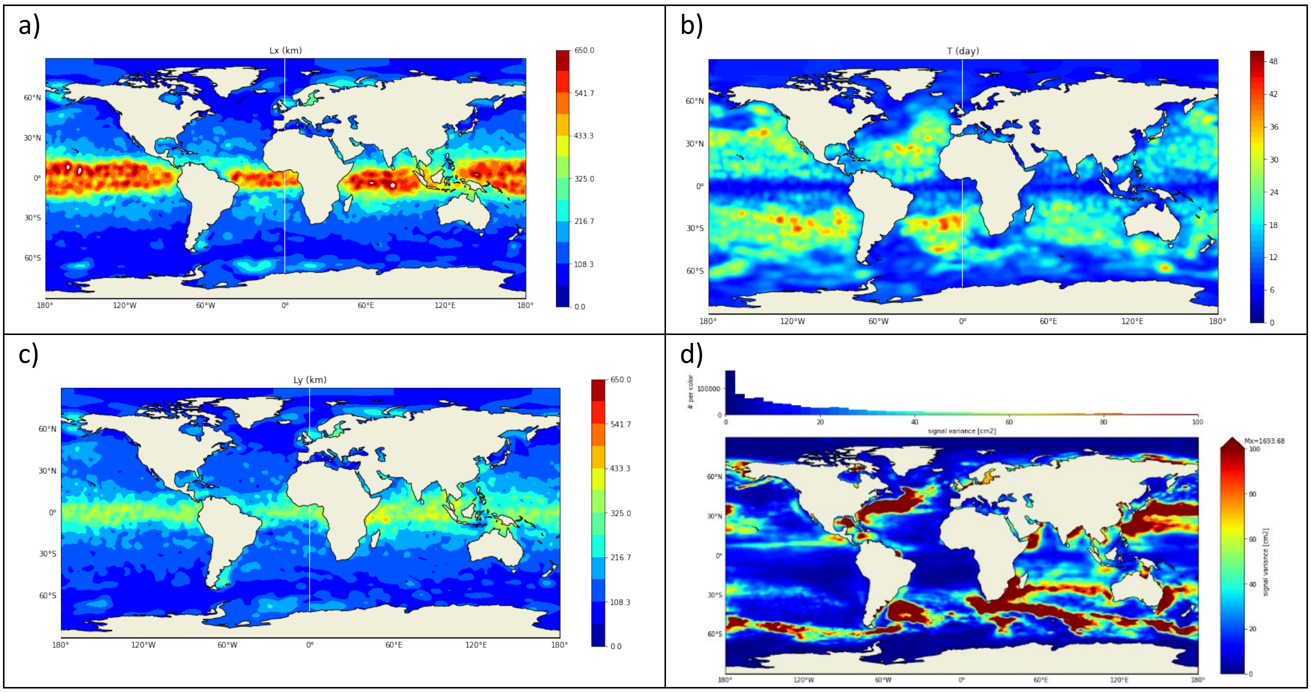


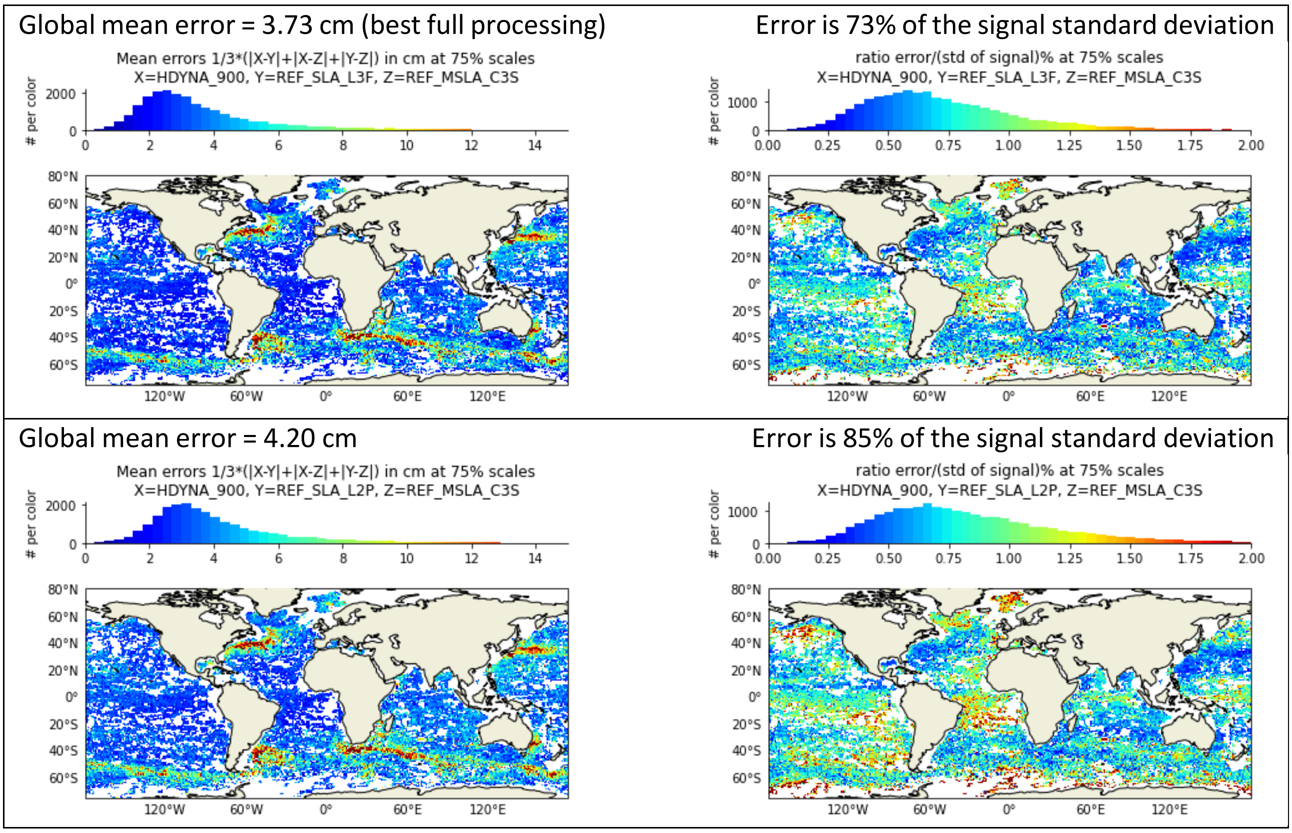
Figure 19. DUACS 2021 correlation (a) zonal, (b) meridional and (c) temporal ocean scales. (d) Oceanic signal variance from DUACS-2018

Results

Absolute difference mean error from three datasets

We select collocated data within 75% of the ocean scales (Figure 19) and estimate an error based on:

Mean (1/3 (|X-Y| + |X-Z| + |Z-Y|)) per grid cell. Where X represents the dynamic height anomaly from *in situ* profiles dataset, Y the along track sea level anomaly from S3B and Z the gridded C3S altimetric sea level anomaly.



Global mean error = 3.73 cm (best full processing)

Error is 73% of the signal standard deviation

Global mean error = 4.20 cm

Error is 85% of the signal standard deviation

Figure 20. Mean error from the combination of 3 dataset: Dynamic height anomalies, gridded combined C3S sea level anomalies, and different along-track sea level anomalies: top) L3F best, bot) L2P least corrected.

One can see from Figure 20 that the L2P is improved by the L3F processing. The global error drops from 4.2cm to 3.73 cm, and the ratio error to signal drops from 85 to 73%. The error in high variability region also includes the representativity with a slight difference in time/space location of the collocated data can lead to a large difference in the value (for example one data set measuring inside and eddy and another dataset just the outside of an eddy), thereby increasing the error. However, when compared with signal standard deviation (square root of the variance), the error represents less than 50 per cent of the signal in the Gulf Stream, Kuroshio, Antarctic circumpolar current, and Agulhas current.

Triple collocation analysis and along-track random error

The triple collocation analysis is described in §3.2.2. To check the method, two versions of along-track S3B data are considered. In the first case (L2P), the data used contains an error due to a correction not applied correctly (it has been corrected since then). The second case is the best-qualified level 3 product: The L2P corrected by DUACS (orbit correction, long wavelength error correction). The TCA applied for the sets dynamic heights, gridded combined sea level anomalies and either 1) S3B-L2P or 2) S3B-L3F along track data.

The selected data are chosen within 75% of the ocean scales such that each matchup from the datasets will be representing a similar signal.

If the TCA is applied over the globe, the results over the 211950 matchups (after removal of the 21882 outliers) are presented Table 12.

Table 12. TCA results when applied over the whole globe. Values in parenthesis are the standard deviation (cm) obtained from the bootstraps estimates.

TCA parameters	**Dynamic height anomalies from insitu**	**Along track sea level anomaly for S3B, with L3F processing**	**Gridded C3S sea level anomalies**
Gain	1.0	1.267 (0.004)	1.354 (0.004)
Offset (cm)	0	0.941 (0.013)	0.764 (0.011)
RMSE (cm)	2.960 (0.007)	3.602 (0.014)	1.742 (0.018)
RRMSE (cm)	-	2.841 (0.016)	1.286 (0.015)
Correlation with unknown truth	0.820 (0.001)	0.830 (0.001)	0.957 (0.001)

Over the globe, the rescaled random error from the filtered (L2P) along track S3B satellite is 2.84 +/- 0.04 cm in the rescaled dynamic height data space (if assumed perfectly calibrated) and is of the same order as the RMSE from the height anomalies (2.96cm) showing an equally precise random error. C3S however has a slightly lower random error with and RRMSE 1.29 cm.

Now it would be interesting to estimate how this error might be spread geographically. The TCA is now applied on a 10 by 10 degree grid. The grid cells containing less than 10 matchups or where the mean (RMSE) < 3 *std (RMSE) are not taken into account because of poor representativeness and unreliable statistics. The standard variation to the RMSE is obtained from bootstrap estimates (method applied here on 100 replicates per each grid cell to enable statistics). If the bootstrapped sample mean of the RMSEs per grid cell is less than the error (3*std (RMSE) per grid cell) then the results are not reliable because there is too much variability between the bootstraps outputs. On average 600 matchups per grid points are used for the TCA (see Figure 21) and 23 matchups discarded on average (Figure 22)

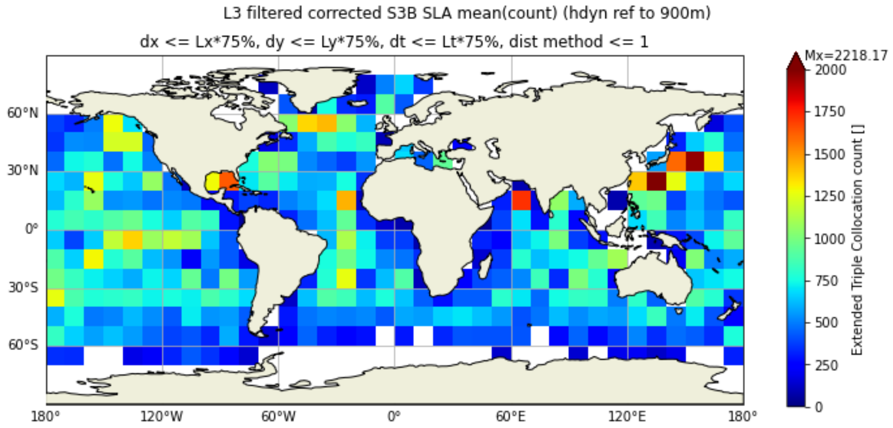


Figure 21. Number of matchups per cells used for TCA.

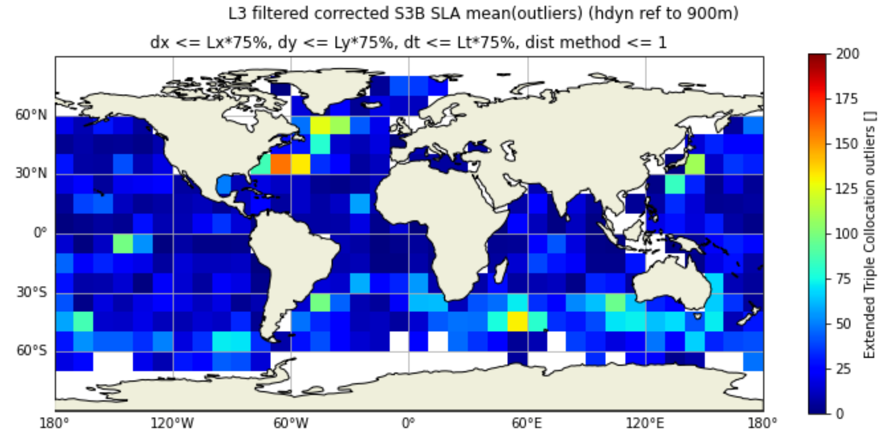


Figure 22. Number of matchups containing at least one outlier. These matchups will not be used in TCA (a mean of ~4 % of the input matchups are removed per cell)

Results on in-situ dynamic heights

The results shown below are the results with the three datasets: dynamic heights anomalies referenced at 900 meters, S3B L3 Filtered (best) along track sea level anomalies, and gridded C3S level anomalies.

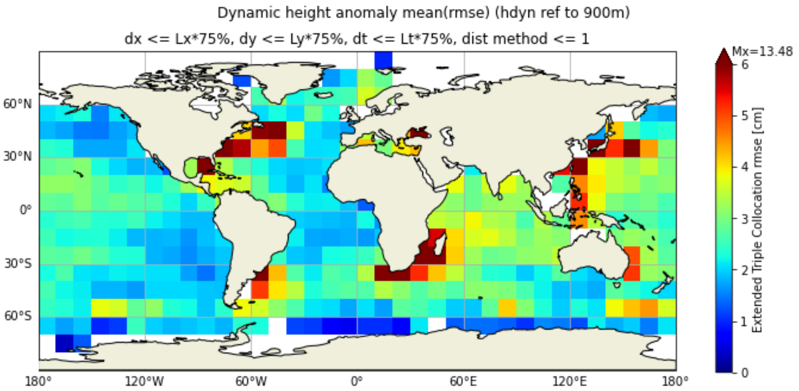


Figure 23. Random error results (RMSE) of the in-situ dynamic heights on a 10x10 degree grid. Results are in cm. Mean RMSE=2.75 cm.

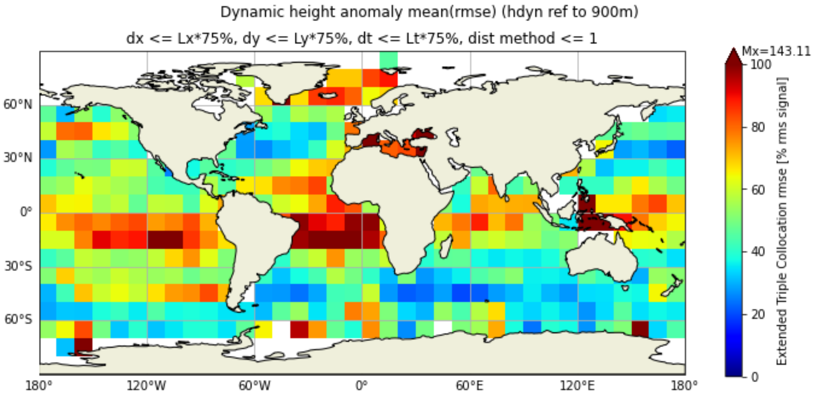


Figure 24. ratio of the in-situ dynamic height random errors (as in Figure 6) and the standard deviation of the oceanic signal (Figure 2d) in percent. The mean percentage is 55.5%.

The random error (RMSE) obtained from the TCA is 2.75 cm over the globe (Figure 23) and shows the highest error in the strong oceanic currents (RMSE above 5 cm). The ratio of the HDYN RMSE with respect to the oceanic signal is 55.5% on average over the globe, with very large values around the equator (see Figure 24). In all the strong current (where the highest errors lie see Figure 23), this ratio drops dramatically.

The correlation as defined by McColl et al, 2014 between the dynamic height anomalies and the unknown truth is globally good with a global mean of 0.8 (Figure 25).

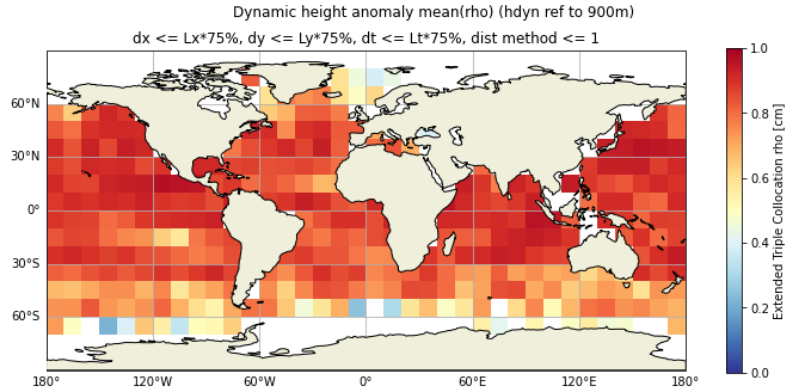


Figure 25. Correlation between the dynamic height anomalies from in situ profiles and the unknown truth. The global mean correlation is 0.8.

Results on along-track error estimations and impact of the processing

The results shown below are the results with the three datasets: dynamic heights anomalies referenced at 900 meters, gridded C3S level anomalies, and along-track S3B sea level anomalies.

To quantify the impact of the added processing on the S3B along-track data starting from L2P data to L3 unfiltered (hereafter L3U) to L3 Filtered (best hereafter L3F), the random error on the along-track data are displayed on a 10x10 degree map on Figure 26. The S3B random error varies from 3.62cm (L2P) to 3.19cm (L3U) to reach 2.84cm for the best processing (L3F Figure 26, top). Thus, processing from L2P to L3F lowers the error by 0.78cm (27.5%) (Figure 26, bottom) and processing from L3U to L3F lowers the error by 0.35 cm (12.3%) (see Figure 26, centre).

The correlation between the along track S3B dataset and the unknown truth increases from 0.75 to 0.83 after DUACS processing (see difference between Figure 27 S3B-L2P (bot) and S3B-L3F (top)).

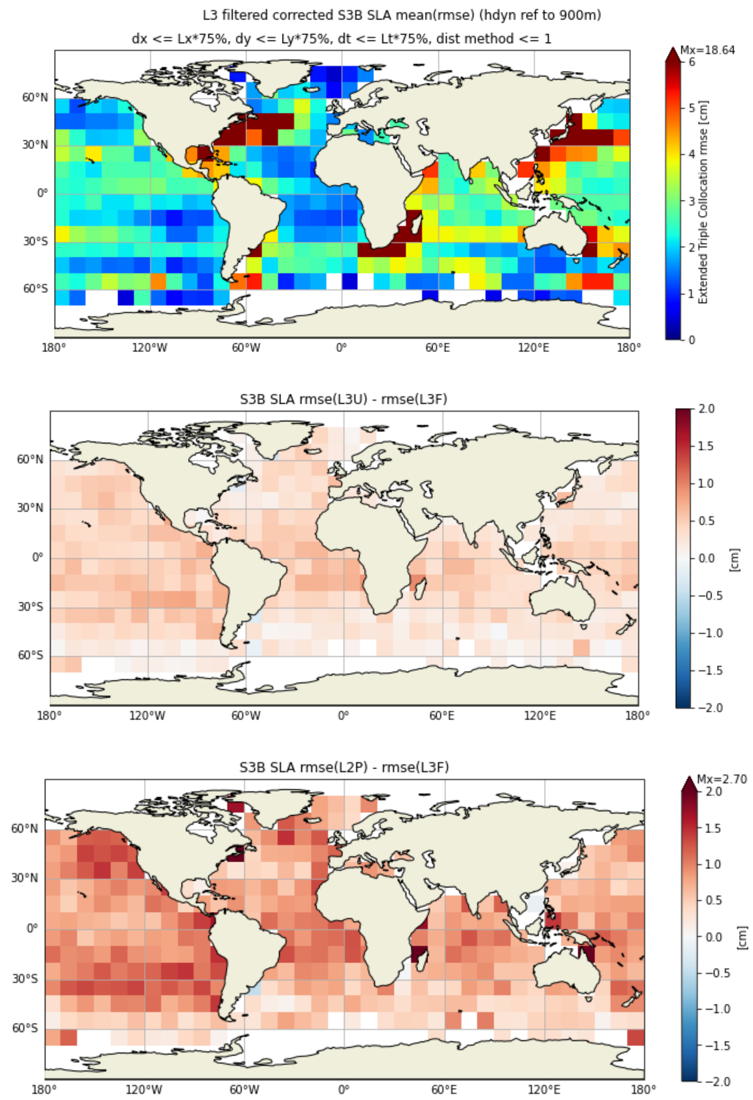


Figure 26. Result on along-track random error (RMSE) in cm from TCA using in-situ dynamic heights, C3S gridded altimetric sea level anomalies and along track S3B sea level anomaly of difference quality. Top: best quality (L3F), RMSE=2.84cm. middle: RMSE difference between medium L3U and best L3F quality, (global difference of 0.35cm) and bottom: RMSE difference between L2P low and best L3F quality (global difference of 0.78cm). The RMSE increases as the data quality decreases.

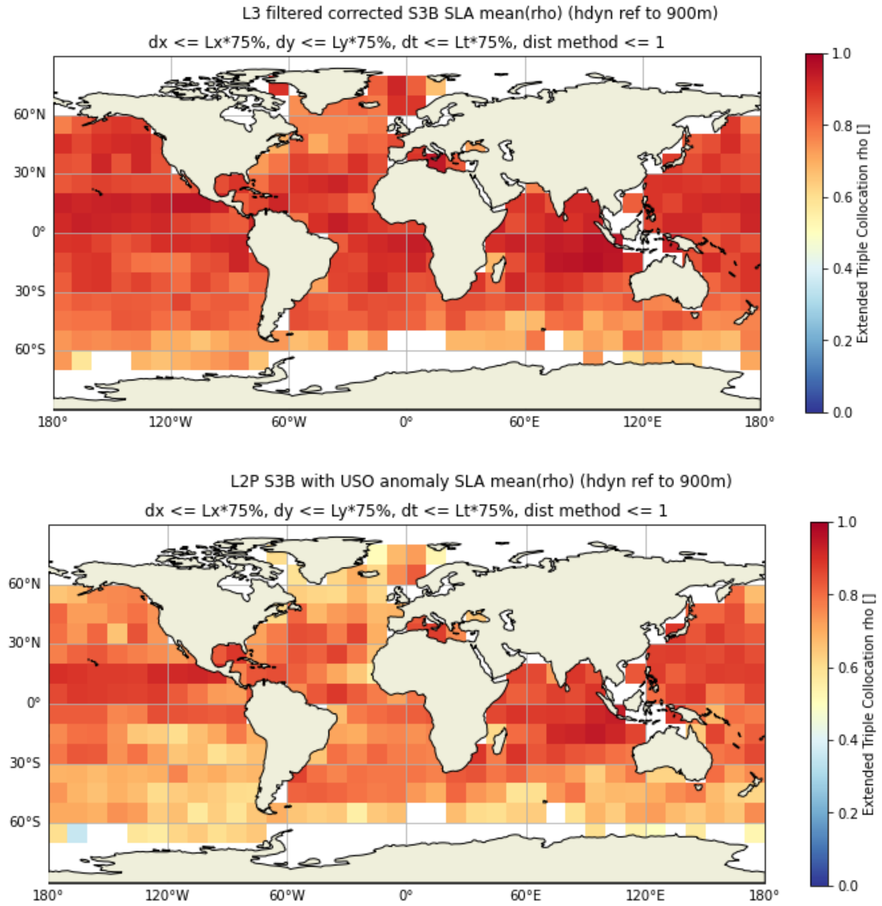


Figure 27. Result on along-track correlation and the unknown truth from the extended TCA using in-situ dynamic heights, C3S gridded altimetric sea level anomalies and top) S3B-L3 (best) and bottom) S3B-L2P sea level anomalies. The results are on a 10x10 degrees grid. The grid cells containing less than 10 measurements are not ignored. The global S3B-L3F correlation is 0.83 and down to 0.75 for S3B-L2P.

Sea level anomalies conclusion

A large part of the work is to select the good matchups to use in the different analyses. The choice is based on time and distance, with a pre-required knowledge of the oceanic correlation scales and signal variance. Many different choices can be made, leading to different results.

Three primary independent datasets are used, a dynamic height calculated from *in situ* temperature/salinity profiles, a gridded sea level anomaly field (C3S) and along-track data from the S3B satellite. For the latter, we will check the improvements from two consecutive validation steps (L2P and L3F). L3F is the one distributed by SSALTO/DUACS in CMEMS.

The two technics shown herein give an indication on the error between three independent datasets. An average of the two-by-two differences between the three datasets show an improvement with the latest processing, reducing the overall error from 4.2cm to 3.73 cm. As expected, the errors are higher in the high variability regions reaching over 12 cm. The second method used in this study gives a random error and other statistics for each dataset, with the dynamic height anomalies from *in situ* data as a reference. The globe is broken into 10 by 10 degrees squares, and the TAC is applied in each square containing from 200 to over 1000 matchups. The overall random error is lowered by 27.5% from 3.62cm for L2P processing to 2.84 +/-

0.04 cm peaking at 18 cm for the best processed along track S3B sea level anomalies. The correlation between the SLA and the unknown truth is increased from 0.75 to 0.83. Both methods show the improvement with the L3F processing, giving an estimate of the error depending on the geophysical location.

3.2. Geostrophic velocity anomalies validation

Data processing

Geostrophic velocities from along track S3B sea level anomalies

Data used for along-track altimetry data is the same as 2.3, using the filtered version.

Before the across-track velocity calculation, the SLA signal is filtered to remove any high-frequency noise. A 4-lobe Lanczos filter with a 45 km cutoff corresponding to half window lengths of 180 km. Since the filter is applied over the time domain rather than the spatial domain, assuming a constant satellite ground speed of 6.64 km s−1, the filter is really applied with a ~27 s half-window length.

The velocity is then calculated using a finite difference to estimate the geostrophic current which results from the balance between the pressure gradient and the Coriolis effect. The two 2 points before and 2 points after the location for the estimated velocity are used.

$$u = -\frac{1}{f\rho}\frac{\delta p}{\delta y}$$

$$v = \frac{1}{f\rho}\frac{\delta p}{\delta x}$$

$$f = 2\Omega \sin(\phi)$$

$$p = \rho \eta g$$

u and v are the zonal and meridional component velocities

ρ the sea water density

f the Coriolis parameter

Ω is the earth rotation rate

ϕ the latitude

p the pressure

η the surface elevation

g the acceleration due to gravity

Note that at the equator, ϕ=0 so is f is null and therefore u and v are infinite. Thus, the data within the equatorial band (at +/- 10 degrees of latitude) will be removed from the error estimations.

Another limitation besides the above approximations is that only the along-track η gradient is known at any given observation point, and its direction is determined by the angle of the satellite track. There is no information on the across-track gradient so only a part of the ocean dynamics can be estimated.

The angle between the eastward direction and the across track velocity x_{s3b} at any given point is noted α_{s3b} (depends on the local s3b track orientation).

In situ drifters

Drifter velocities are extracted from the 6-hour AOML drifter database, using a drifter with a drog at 15m. 1565 drifters are present in 2019, representing 1 105 182 velocities. The anomaly of geostrophic velocity is calculated as the total drifter current at 15m minus the Ekman current (deduced for ERA5 winds) and minus the mean dynamic topography (MDTCLS18).

The currents are filtered in time with a Lanczos filter of half window size of 3 days and 0 lobe to produce velocities without tides and inertial signal, and smooth enough to compare to the altimetry signal. The resulting fields are $u_{drifter}$ and $v_{drifter}$.

After altimetry and drifter collocation using a KDtree method (see Triple collocation analysis), with longitude and latitude scales of 1 degree and a time scale of 6 days. The +/- 10-degree equatorial band is removed from the data since the classical geostrophic balance does not apply where the Coriolis parameter is null.

The $u_{drifter}$ and $v_{drifter}$ drifter components are then projected on the collocated across the track to compare with the across track geostrophic velocity from altimetry.

across track drifter velocity $x_{drifter}$= $u_{drifter}\cos(\alpha_{s3b}) + v_{drifter}\sin(\alpha_{s3b})$

Geostrophic velocity anomalies from C3S gridded fields

The gridded C3S geostrophic velocities anomalies are interpolated on AOML drifter location. They are projected across-track the same way as the drifter velocities. Note: by using interpolated η along data on a grid, longitudinal and latitudinal gradients can be calculated thus lifting the constraint of a unidirectional gradient. The across-track geostrophic velocity anomaly is labelled x_{c3s}. These data are only used in the triple collocation analysis but results pertaining to this dataset will not be discussed here.

Triple collocation

The gridded altimetric C3S data are already interpolated on drifter location. The only collocation needed is then between drifter location and along-track altimetric data (thus using the term matchups in relation to time/space location). This is done by matching each drifter location with the closest altimetric location in time and space using a KDtree method with a longitude and latitude scaling of 1 degree, and a 6-day time scaling to set the corresponding importance of the longitude, latitude, and time. For example, a matchup separated by 6 days has the same weight as a matchup at 1-degree longitude or latitude. This was done by checking for the best results with a few tests, therefore it can be improved.

An example in the eastern part of the Gulf stream is shown in Figure 28, where the altimetric along-track SLA is shown (see colour bar) with the derived geostrophic velocity as grey arrows. The drifter locations are marked with a black + sign, and the black lines connecting drifter location and along track location show the matchups used for the collocation (closest in time and space).

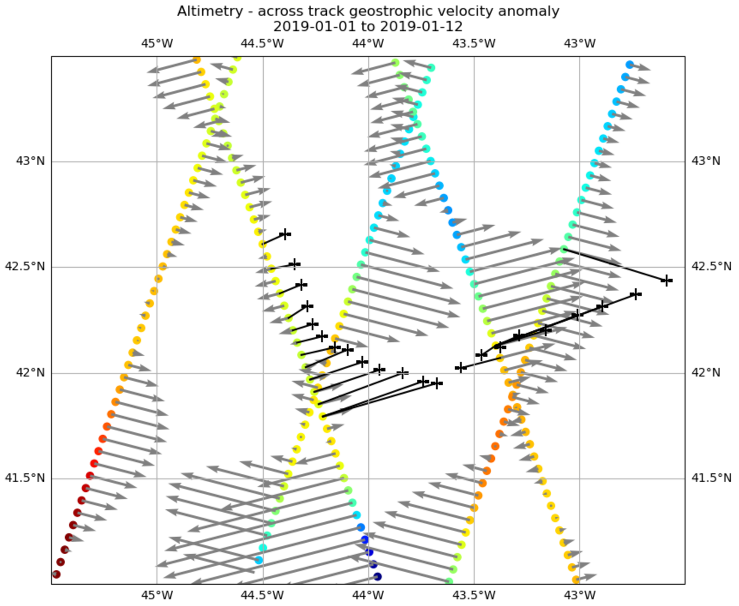


Figure 28. Drifter 64502410 over 12 days in January 2019 (black +) with S3B filtered along track Sea level anomaly (colour dots) and associated calculated across track geostrophic velocity anomaly (grey arrows). The black lines show the related alti-drifter matchup closest in time and space.

The data within the equatorial +/-10 degrees latitude band are removed since the geostrophic balance does not hold. 984 039 matchups are found for the year 2019 outside the equatorial band. From these, only the closest matchups that are where the distance is less than 20km and the time difference is less than 3 days, grid cells where enough data is available (>10 matchups) and the rms difference of matchups velocities is sufficiently low compared to the signal (see Figure 29). The importance of closeness in the matchups of the different sources is important since they are supposed to measure the same quantity even in regions of strong gradient. Thus, the interest in keeping only the closest matchups.

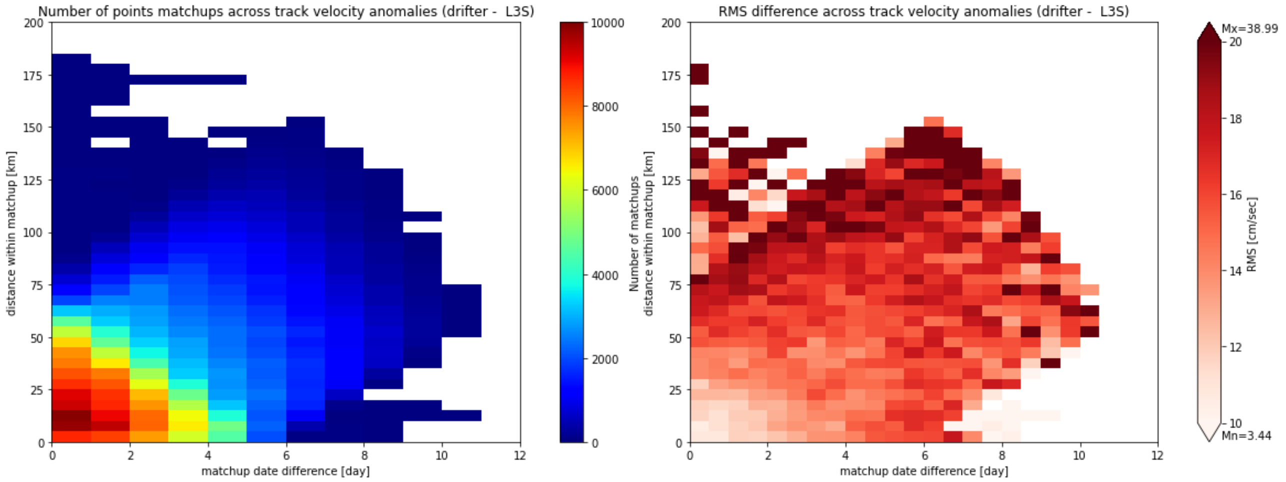


Figure 29. drifter and S3B altimetry matchup distribution in function of matchup distance in time (abscissae) and kilometres (ordinates). The closer in time and space the matchup is, the more data available (left) and the lower the rms difference in velocity (right). Note than at least 10 data points per grid were required to calculate the statistics.

210 235 matchups are left (21%). The dispersions of the velocities with the full matchups or the most pertinent matchups are shown in Figure 30 (top left and right side respectively), showing a better fit and correlation (top right plot) when the matchups are chosen closer together. However, this is not the case between filtered drifters and C3S data. Is it because C3S data are first interpolated on drifter location and then collocated thus accumulating positioning errors? Or maybe scales representations are too different in the two datasets. C3S data tends to underestimate the velocities compared to drifter data (with a gain at 0.64, see the equation of the fit in bottom right figure). The geographical data distribution still has a similar spread as shown in Figure 31, but both show a lack of data in the southern Pacific Ocean.

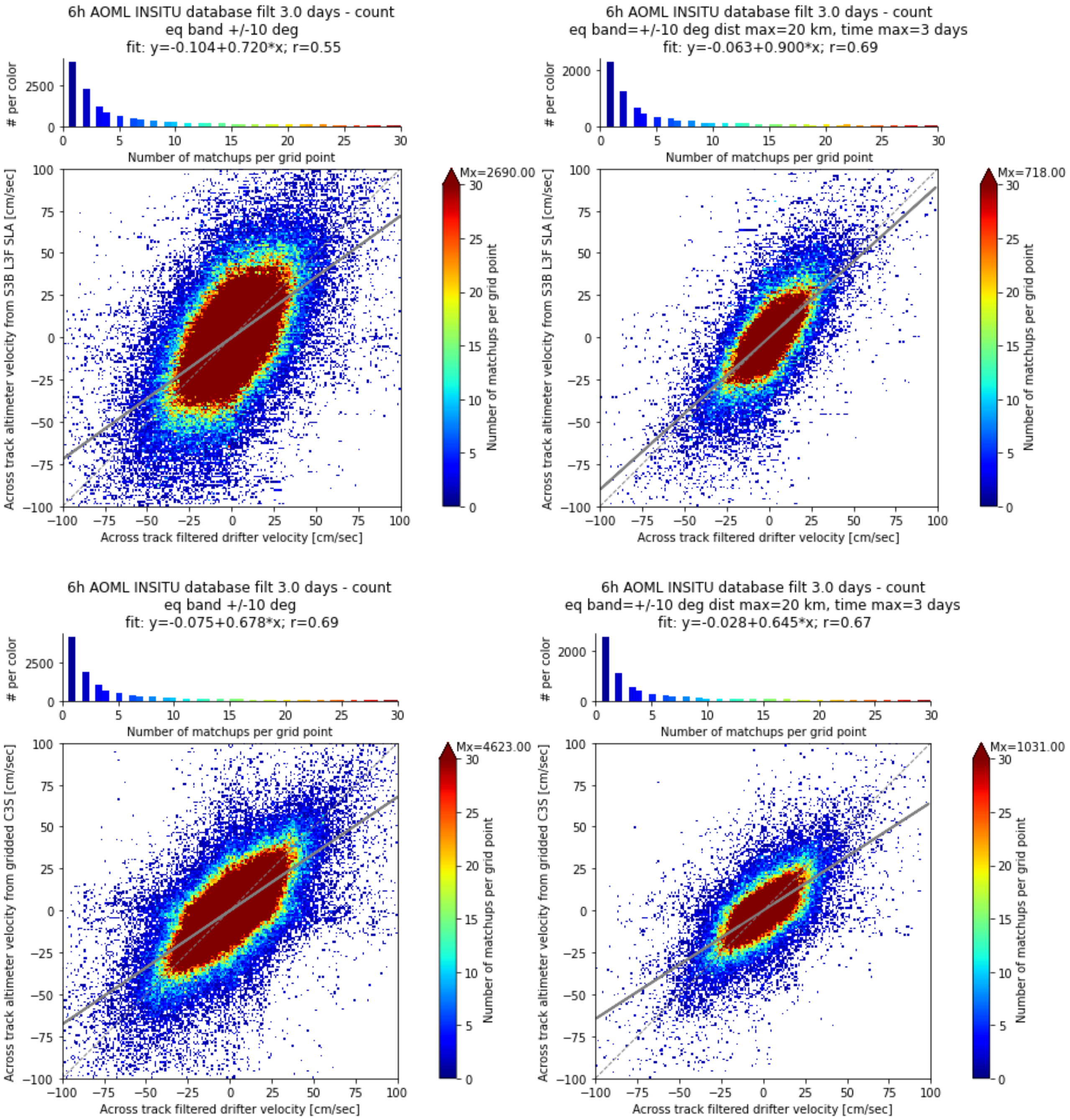


Figure 30. velocity dispersion between across track drifter and S3B altimeter velocity with full matchups (top left) and matchups less than 20 km and 3 days apart (top right). Only 21% of the data from left figure are present on right figure. The dash line represents the identity line, and the solid grey line the linear regression between drifter and altimetry derived geostrophic velocity anomalies. The bottom figures are the same as the top one, except for S3B replaced by C3S.

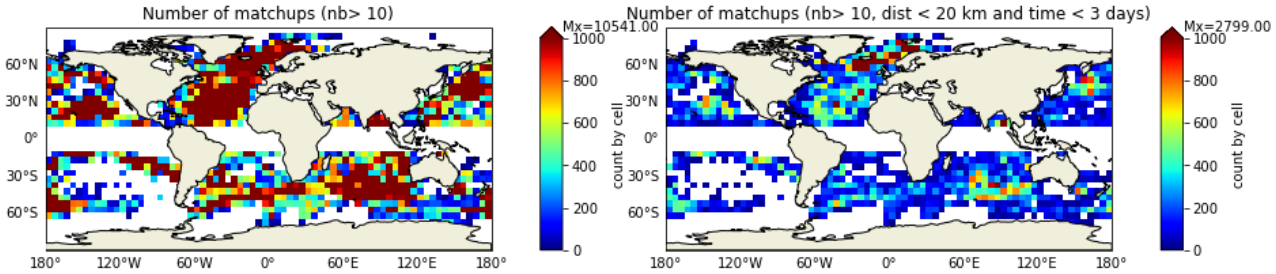


Figure 31. Number of matchups available per 5x5 degree with all matchups (left) and matchups such that the distance between points in a matchup is less than 20 km and the time difference less than 3 days (right). The grid box with less than 10 matchups are not shown.

Random error standard deviation estimation

The global triple collocation analysis is performed on values selected within the 1.5 times interquartile range to remove extra outliers. A bootstrap method gives access to the standard deviation of the results, thus ensuring robust estimated if the standard deviation is lower (here by a factor 3) than the mean of the results. The offset between the dataset is very low, and the gain show, as seen in §7.2 that C3S underestimates the geostrophic velocity anomaly. The mean of the dataset is low, and S3B shows more variability with a standard deviation of 10.78 cm.s^{-1} than drifters with 8.67 cm.s^{-1} and C3S with 7.28 cm.s^{-1}.

The random error based on the covariance between drifters, S3B and C3S have a mean RMSE=4.76 cm.s^{-1}, 7.82 cm.s^{-1} and 5.41 cm.s^{-1} respectively. The associated standard deviations show that, within the frame of the method, these estimations are robust (Table 13). Assuming the drifter perfectly calibrated the rescaled RMSE (rrmse) in the drifter plane for S3B and C3S is larger, showing a higher precision in altimetry than drifter. This can be a side effect of all the processing done, and data filtering might need adjustment to estimate the matchups.

Table 13. Results from the tri-collocation analysis of across track geostrophic velocity anomalies. The +/- 10 degree latitude band is removed, and matchups less than 20km and 3 days apart are taken into account. The random errors (RMSE and rescaled RMSE) are in are in cm.s^{-1}.

Across track geostrophic velocities anomalies in cm.s-1 (equatorial band +/-10 degrees removed)			
Estimates on 500 bootstraps with 10000 samples each without replacement TC selection=1.5*[75,25] quantiles			
Geostrophic velocities anomalies	$x_{drifter}$	x_{s3b}	x_{c3s}
Nb outliers	1372.41 (34.35)	1372.41 (34.35)	1372.41 (34.35)
Nb used	8627.59 (34.35)	8627.59 (34.35)	8627.59 (34.35)
mean	-0.01 (0.11)	-0.12 (0.13)	-0.05 (0.09)
std	8.67 (0.10)	10.78 (0.12)	7.28 (0.10)
gain	1.00	1.03 (0.02)	0.67 (0.02)
offset	0.00	-0.11 (0.10)	-0.05 (0.07)
Correlation with unknown truth	0.84 (0.01)	0.69 (0.01)	0.67 (0.01)
rmse	4.76 (0.12)	7.82 (0.12)	5.41 (0.08)
rrmse	-	7.63 (0.23)	8.05 (0.26)

The same analysis is carried out locally on a 5-degree longitude by 5-degree latitude grid to estimate the regional deviation between the three datasets, and the results are shown on Figure 32.

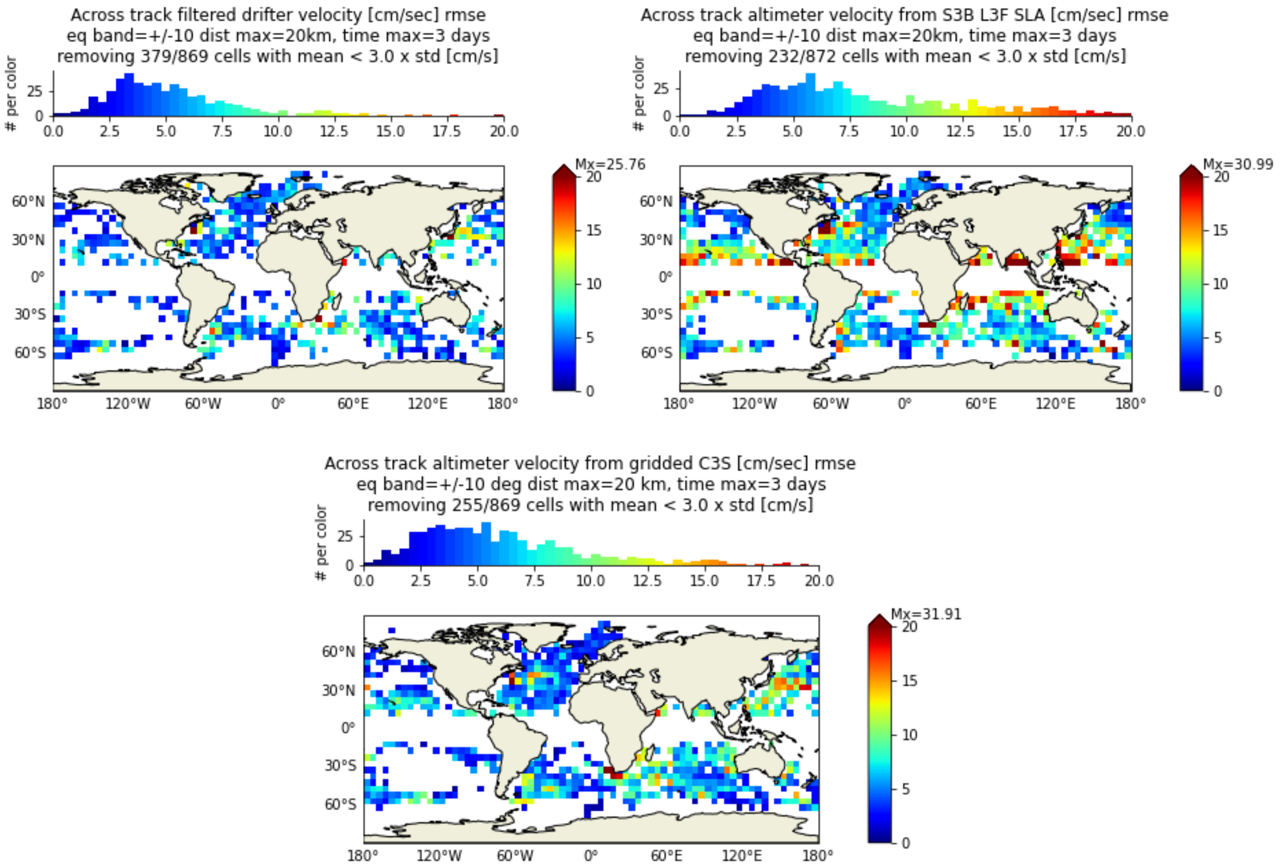


Figure 32. Results from the tri-collocation method between drifters, S3B along track altimetry, and gridded C3S across track geostrophic anomaly after removing the +/-10 degrees equatorial band (geostrophic equation not valid), keeping matchups where distance is less or equal to 20km, time difference is less than 3 days. Only RMSE results where bootstrapping results are reliable (the variability of the result is less than the result) are plotted.

The tri-collocation analysis gives information about the random error associated with the three datasets. These results are obtained after 500 draws of 10000 samples each per 5x5 degrees grid cells. RMSE errors area is still high around the equator (over 20 cm.s^{-1}) and in areas of strong current (e.g., the Gulf Stream, Kuroshio, etc.).

3.3. Conclusions

Two types of error estimation techniques were presented, one additive and the other one (the triple collocation analysis, TCA) to estimate a random bias between three independent global data sets. They show similar results where errors are higher in high variability regions, which is partly linked to errors in data collocation, especially where gradients are strong. The methods were applied to sea level anomalies and geostrophic velocities anomalies. The method showed the improvement of the S3B along-track data by DUACS processing, lowering the error with respect to Argo drifter and C3S sea level anomaly maps. Some areas were not very well represented in space, especially the southern Pacific Ocean for the drifter/altimetry velocity validation.

More precise processing such as filtering, scaling, and matching should probably be carried out:

- Ensure that the data filtering on drifter and altimetry products are more coherent and best for comparison
- The projection on the across-track and colocation between altimetry-derived velocity and drifter velocity leads to errors, and so does the interpolation of C3S onto drifter locations
- This method will probably be more adapted to altimetry data such as SWOT, where 2-D gradients can be estimated, thus enabling access to the whole signal
- Use regional scale representativeness to select the matchups with the drifter as was done for dynamic high processing.

However, these validations showed the importance of in-situ data to validate altimetry and a homogeneous time and space distribution is important to compare the datasets.

4. BGC Argo array for the validation of ocean colour satellite missions, in particular S3A&B (ACRI-ST)

4.1. Introduction

Remote sensing observations of the ocean colour has allowed to monitor the surface ocean and the biological activity at the global scale for more than 24 years now. Satellite observation of the ocean colour is a key information for a sustainable ocean. Indeed, chlorophyll surface concentration can be derived from the ocean colour and, as phytoplankton is at the basis of the marine food web, an index of the productivity of the system can be derived from the estimated chlorophyll concentration. What is a precious help for instance for fisheries, as the fish catch potential can be linked to the chlorophyll concentration. In addition, ocean colour data are also important for climate, as the productivity of the system is an important parameter to estimate the efficiency of the biological carbon pump. Ocean colour can be also used to monitor the sea-water quality. For instance, thanks to specific algorithms, these observations can help for fight against harmful algal bloom.

Nevertheless, remote sensing data should be regularly validated as the information derived from the ocean colour (e.g., chlorophyll concentration) is retrieved with some assumptions, for instance, one should consider that the atmospheric corrections are properly applied, and the algorithm used to derive the variable is suitable for the region of interest. Also, for sensors with relatively long time series, sensor drift can potentially occur, and in such cases, the validation of the ocean colour data can be useful to point it out and to trigger remediation (e.g., new calibration).

Satellite ocean colour data are validated against in situ data. Historically, these data were from bottle sampling and HPLC analysis in a laboratory or derived from direct fluorescence measurements. However, these measurements are costly and time-consuming, not to mention that some parts of the ocean are difficult to reach. Also, there is a delay which can be important between the acquisition time and the time when the data is publicly available.

BGC-Argo autonomous platforms increase the number of in situ data for the validation of ocean colour. These platforms now cover a wide part of the global ocean, with more than 100 000 of chlorophyll profiles acquired since the beginning of the program in 2005. The use of autonomous platforms takes on its full meaning for the quality control and the validation of some ocean colour data as data from these platforms are available soon after their acquisition and data can be available everywhere, regardless the sea state.

The aim of this study is to demonstrate how efficient are the BGC-Argo for the validation of the satellite ocean colour data from Sentinel3 A & B.

4.2. Data description

BGC-Argo data are freely accessible on the two Global Data Assembly Centres (US-Godae and Coriolis). In the frame of this study, data were collected from the FTP access of the Coriolis centre. There were 42924 profiles coincident with the Sentinel-3 period (April 2016 – present).

Once the access to the BGC-Argo has been established, one should extract the data of interest, in the present case, the chlorophyll concentration (CHL), the particulate backscattering (BBP) and the coefficient of diffuse attenuation (Kd). To be fully comparable with satellite observation, in situ data should be reported on the same scale than satellite data. Indeed, satellite can "see" only the upper layer of the ocean, while BGC-Argo float document the water column, from the surface up to 1000 or 2000 m depth.

The location of the BG-Argo profiles used for this study are presented on the map Figure 33.

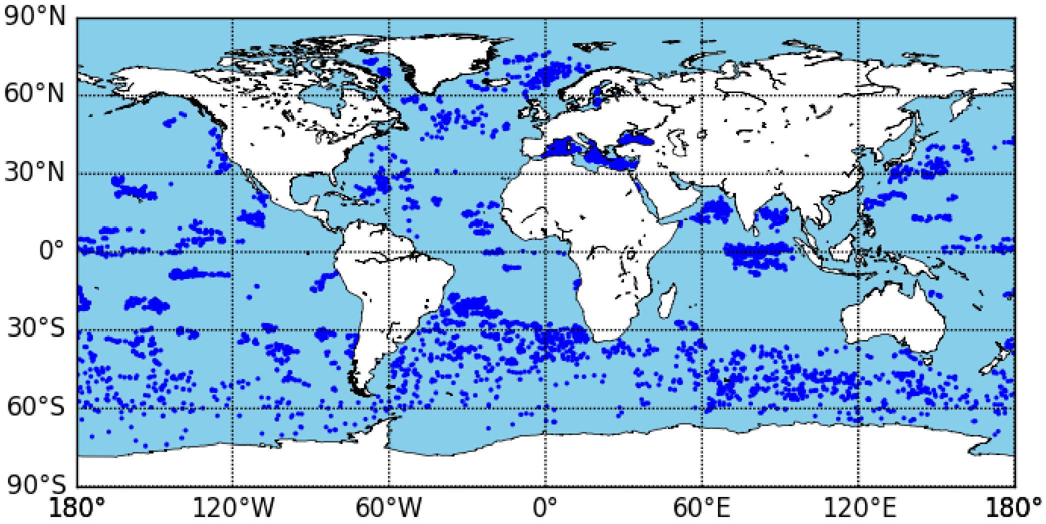


Figure 33. Location of the matchups between BGC-Argo profiles and Sentinel-3 observations (profiles acquired from April 2016 up to December 2021).

BGC-Argo are associated with a wide network allowing to document different environment. This presents a good advantage, notably for the validation of regional algorithm.

Chlorophyll a in situ processing

To make the *in situ* CHL comparable with satellite observations, the chlorophyll concentration should be first averaged over the first optical depth (Zpd) which represents roughly the depth at which 90% of the solar light is absorbed due to the seawater constituents (water, particles, dissolved matter). This depth is derived from the Morel et al. (2007) formula:

$$Kd = 0.166 + 0.077298 * CHL_0^{0.67155} \qquad (1)$$

And then

$$Z_{pd} = \frac{1}{Kd} \qquad (2)$$

Where CHL_0 correspond to the chlorophyll surface concentration. In the case of BGC-Argo floats, this chlorophyll surface concentration is taken as the median value of the chlorophyll measured by BGC-Argo float between the surface and 10 m. This allows computing Zpd (eq. 1 and 2 above).

The chlorophyll concentration measured by BGC-Argo is then averaged between the surface and Zpd.

For validation purposes, the “adjusted” chlorophyll from the BGC-Argo files was considered. These profiles have been processed in delayed mode, with more accurate correction factors.

Indeed, to convert fluorescence measurement into chlorophyll concentration, some corrections should be applied.

1. Dark correction (which corresponds to the removal of the background)
2. Slope correction (which corresponds to the correction to apply for the from fluorescence to chlorophyll concentration
3. Quenching correction

Using the adjusted chlorophyll profile for satellite validation is expected to improve the accuracy and reliability of the data.

In the case of this study, two ocean colour algorithms were compared with BGC-Argo, the GlobColour product based on OC5 and Hu algorithm (Gohin et al., 2002, Hu et al., 2012), and the GSM product (Maritorena and Siegel, 2005).

Particulate back scattering (BBP)

The particulate backscattering corresponds to an index for the sea concentration in particles. In the open ocean, the BBP covary with the concentration in particulate organic carbon as the essential source of particles is mainly the biological activity, and especially the phytoplankton productivity (Bishop, 2009, Loisel et al., 2002).

As for the chlorophyll concentration, the BBP is averaged over the first optical depth determined by the Morel et al. 2007 methodology (eq. 1 and 2).

From satellite observations, the BBP is derived from remote sensing thanks to the GSM algorithm (Maritorena and Siegel, 2005).

The BBP measured by the BGC-Argo floats are not given at the same wavelength as the one derived from satellite observations (520 and 700 nm for BGC-Argo and 443 nm for satellite observations). To be comparable, satellite observations should be converted to BBP of the same wavelength than BGC-Argo.

To perform this conversion, we applied the Quasi Analytical Algorithm (QAA v6) originally developed by Lee et al. (2002) to derive the absorption and backscattering coefficients by analytically inverting the spectral remote-sensing reflectance (Rrs (λ)).

The conversion is based on the following:

$$b_{bp}(\lambda) = b_{bp}(\lambda_0) * \left(\frac{\lambda_0}{\lambda}\right)^{\eta} \qquad (3)$$

Where λ_0 correspond to the wavelength of reference and η the BBP exponent. This exponent is estimated thanks to remote sensing reflectance as follows:

$$\eta = 2 * \left(1 - 1.2 * \exp\left(-0.9 * \frac{r_{rs}(443)}{r_{rs}(560)}\right)\right) \quad (4)$$

With:

$$r_{rs}(\lambda) = \frac{R_{rs}(\lambda)}{0.52+1.17*R_{rs}(\lambda)} \quad (5)$$

Then, the BBP at 443 nm of satellite observations is converted in BBP at 532 and 700 nm to be compared with BGC-Argo float.

Coefficient of diffuse attenuation (Kd)

The coefficient of diffuse attenuation is an indicator of the turbidity of the water column. From satellite observation, this coefficient is estimated from the Morel et al. (2007) algorithm, based on the chlorophyll concentration.

BGC-Argo floats do not directly measure this parameter, but it is derived from the downwelling irradiance (Ed) following the methodology defined by Organelli et al. (2017a). The method used is briefly described here below. First, Ed measurements from BGC-Argo are extrapolated toward the surface thanks to a polynomial function. Then, Ed data are binned to 1 m of resolution within the first optical depth. The coefficient Kd is estimated as the slope of a linear regression between the natural logarithm of the Ed and depth (Mueller et al., 2003). The coefficient Kd is considered as valid when the regression is estimated with, at least, 3 points and the coefficient of determination (R^2) of the regression Ed vs depth is higher than 0.9.

In the frame of this study, two satellite algorithms have been compared with BGC-Argo data. The algorithm defined by Morel et al. (2007) and the one developed by Lee et al. (2005).

Matchup procedure

Matchup data correspond to BGC-Argo data and satellite observation acquired coincidently at the same location. Here, a macro-pixel of 3x3 pixels were considered. A matchup is considered as valid when at least 50% of the pixels within the macro-pixel are valid, and the coefficient of variation is lower than 0.15. Once the matchup is considered as valid, satellite observation and BGC-Argo data can be compared. Satellite data which is used for comparison is the median of the values within the macro-pixel (to filter out outliers).

Note that the number of matchups for chlorophyll data is more important than the number of matchups for the other parameters as the fluorescence sensor equipped more often BGC-Argo than radiometer or backscattering sensor.

4.3. Global validation

Benefit of the BGC-Argo floats on the number of matchups

After the pre-processing of the data, chlorophyll concentration, BBP and Kd estimated from satellite observations can be compared with data measured by BGC-Argo.

As previously mentioned, one of the advantages of the BGC-Argo in the validation of ocean colour products is the fast availability of the *in situ* data, especially for OLCI as its observations start in April 2016. Indeed, the number of matchups with BGC-Argo available per year, and per region, is independent of the availability of human resources. For instance, the number of matchups remains elevated in 2020, despite the pandemic and the lockdown (Figure 34).

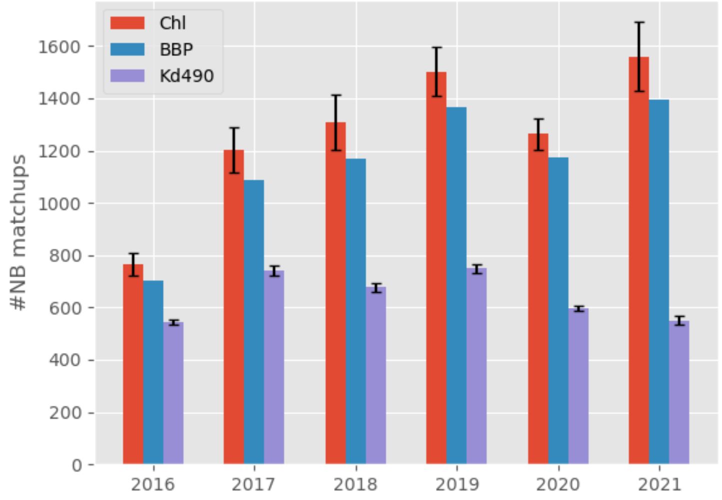


Figure 34. Number of matchups between OLCI observations and BGC-Argo data. The red bars are for the chlorophyll, the blue bars for the BBP and the violet bars for the Kd. For the chlorophyll and the Kd, the bar is representative of the mean between two considered

The number of in situ data used for the validation is considerably improved by the use of BGC-Argo floats. If one use only data from oceanographic cruises, the number of matchups will be reduced. For example, the Ocean Colour Thematic Assembly Centre identified, for the Copernicus Marine Service, less than 600 matchups[7] for the OLCI period (2016-2020). The validations of the Kd and the BBP with cruises data are even less meaningful than the chlorophyll ones.

Note that the term "validation" should be interpreted as an attempt at quality control of ocean colour algorithms. Indeed, the chlorophyll concentration is not directly measured by the floats but is derived from their fluorescence measurement.

Chlorophyll validation

The high number of matchups and the relatively global coverage of BGC-Argo allow for comparing different chlorophyll algorithms. In the present case, the usefulness of BGC-Argo is proved for the comparison of OC5 and GSM chlorophyll algorithms.

There are more matchups for the OC5 algorithm than GSM algorithm. This can be explained by the flagging strategy which is not equivalent to the two methods (Garnesson et al. 2019). BGC-Argo tend to demonstrate that there is no significant difference between both algorithms, but the validation results present a slight better correlation between BGC-Argo floats with GSM data than with OC5 data. Indeed, the slope of the regression line is closer than 1 with GSM and the RMSE is relatively lower for this last algorithm than with OC5 (0.47 mg.m-3 for OC5 and 0.42 mg.m-3 for GSM).

---

[7] http://octac.acri.fr/

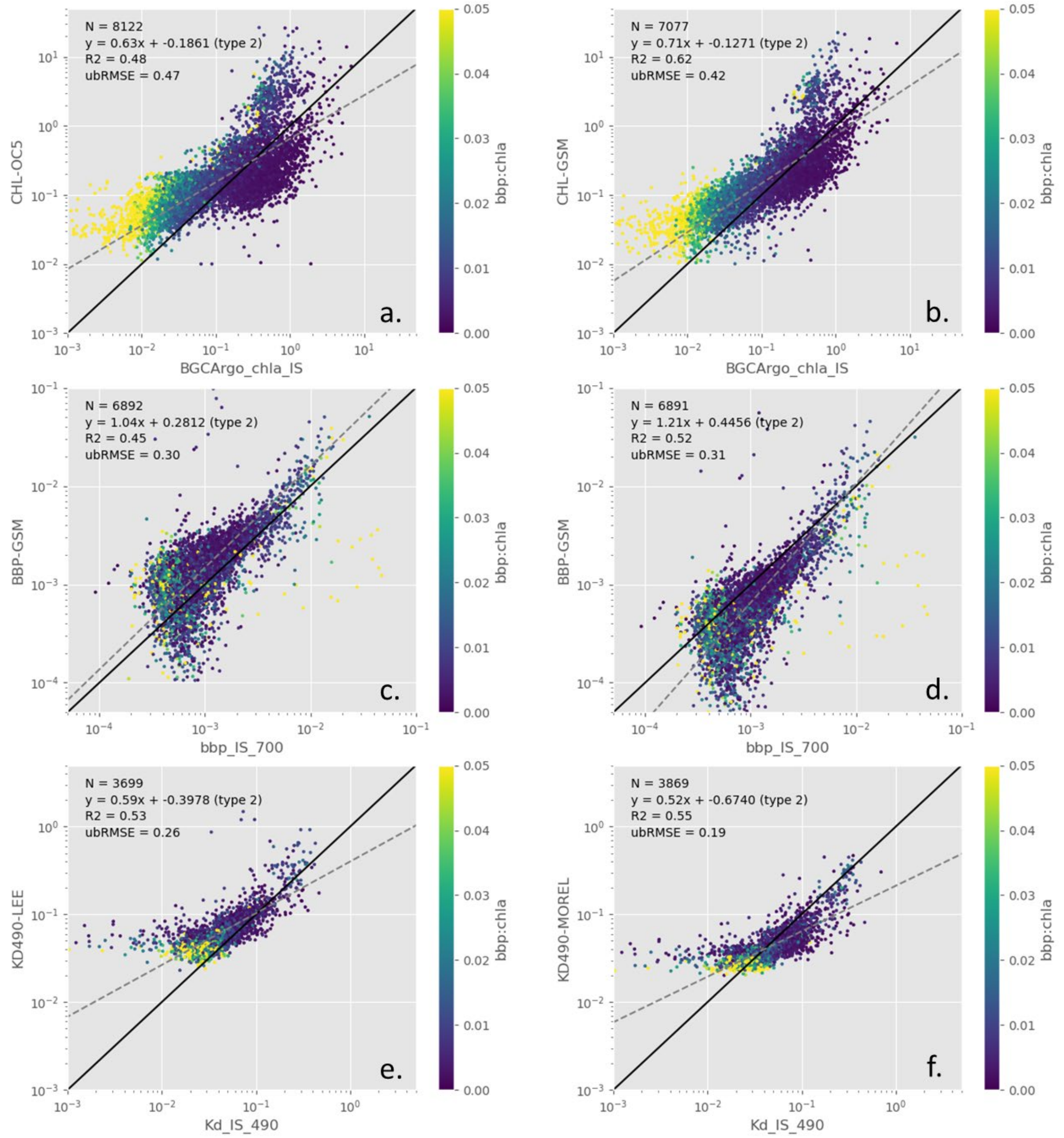


Figure 35. Scatter plot of the chlorophyll concentration (a. for OC5, b. for GSM), BBP without and with modification of the wavelength (c., d.) and Kd (e. for Lee, f. for Morel) derived from ocean colour against the in situ measurements from BGC-Argo floats. The

As BGC-Argo floats samples different environment of the global ocean, it is possible to assess the performance of the algorithm as a function of the range of chlorophyll concentration. The scatter plot of validation for OC5 data seems to overestimate the low chlorophyll concentrations compared to Argo floats. This overestimate is less pronounced with the GSM algorithm.

To understand the deviation between satellite and BGC-Argo, the set of parameters measured by the floats can be used as they are precious information on the surrounding environment. For instance, the ratio BBP/CHL measured by the BGC-Argo can be used to get more information on the origin of the particles.

Indeed, the ratio BBP/CHL is known to be an optical index of the phytoplankton community composition (Cetinić et al., 2015; Lacour et al., 2017). For instance, high BBP/CHL ratio might be due to the presence of viruses and/or bacteria or even a relatively important concentration in non-living material, such as detritus or inorganic particles (Bricaud et al., 2004; Claustre et al., 1999; Oubelkheir et al., 2005). Conversely, large phytoplankton cells (such as diatoms) are generally characterised by a low BBP/CHL ratio.

The scatter plot in Figure 35a-b presents a relatively high BBP/CHL ratio for low chlorophyll concentration, while the discrepancy between satellite observations and floats measurement increases. The discrepancy observed for the low range of chlorophyll concentration might be associated with a dominance of small phytoplanktonic cells.

The validation plot also presents a deviation for higher concentrations. Indeed, for some points of the dataset, the concentration of the chlorophyll *in situ* is comprised of between 0.5 and 1 mg.m^{-3} while satellite observations estimate the concentration between 1 and 100 mg.m-3, about one order of magnitude higher. This may occur in coastal regions, or in regions affected by mineral particles (e.g., dust, riverine discharges), as inorganic particles are known to be more refractive than organic ones. This deviation is less pronounced for GSM algorithm than OC5 indicating the potential weakness of this last algorithm in such environments. However, a BBP/CHL ratio, which would argue for the impact of mineral material, is not clearly noticeable for this range of chlorophyll concentration and does not allow to fully explain this deviation.

Validation of the BBP product

The validation of satellite products of BBP and Kd benefit also the availability of the BGC-Argo floats. For the BBP, the wavelength correction does not appear to improve the validation. Actually, the wavelength correction appears to degrade the validation results, with a slope of 1.21 for the corrected data while the slope is 1.04 for non-corrected data. The error (RMSE) is equal for both corrected and non-corrected data (0.3, Figure 35c-d). This seems to indicate that the algorithm used for the wavelength correction is not well adapted. For the BBP validation using the BGC-Argo float, the satellite observations can be directly compared to *in situ* measurement at 700 nm. Note that for some floats, the BBP is also measured at 532 nm. However, the number of matchups with BGC-Argo measurement at this last wavelength is noticeably reduced compared with measurements at 700 nm and presents weaker validation results. The discrepancy between satellite and BGC-Argo being reduced at 532 nm, one should preferentially consider BBP measurement at 700 nm for satellite validation purposes.

The *in situ* ratio BBP/CHL does not appear to be linked with the deviation between *in situ* and satellite BBP. This indicator is not as useful as for the chlorophyll validation in explaining deviation.

Validation of the coefficient of diffuse attenuation

For the coefficient of diffuse attenuation (Kd), and as for the chlorophyll, two algorithms are compared with BGC-Argo data (Lee: Figure 35e, and Morel: Figure 35f). The number of matchups for the Kd validation is the weakest compared to BBP and Chl. Indeed, at the moment, only few BGC-Argo floats are equipped with radiometer. At the moment of writing of this report, 13 floats are actively measuring irradiance data, while 62 floats measure the chlorophyll. Even if the number of floats measuring irradiance is significantly lower than the one measuring chlorophyll, there are 3699 matchups available for the validation of the Lee algorithm and 3869 matchups for the Morel validation. This number of matchups give confidence in the results. Performance of both algorithms appear to be relatively similar. First as the number of matchups with OLCI data is roughly equivalent, then because the slopes of the regression line are close to each other (0.59 for Lee and 0.52 for Morel). The RMSE are also relatively equivalent (0.3 for Lee and 0.2 for Morel).

As the network of the BGC-Argo is spread over the global ocean, the validity of the ocean colour products can be studied in function of the surrounding environment. Indeed, due to the presence of diverse particles (size, origin etc…), global ocean colour products might be less accurate in coastal waters than in the open ocean, where the particles are mainly from the degradation of the phytoplankton. Then, the validation of ocean colour products can be carried out in function of the biomes as defined by Longhurst et al., 1996. The results for the polar, coastal, trade winds and westerlies biomes can be compared between them and compared with the statistical results for the global ocean.

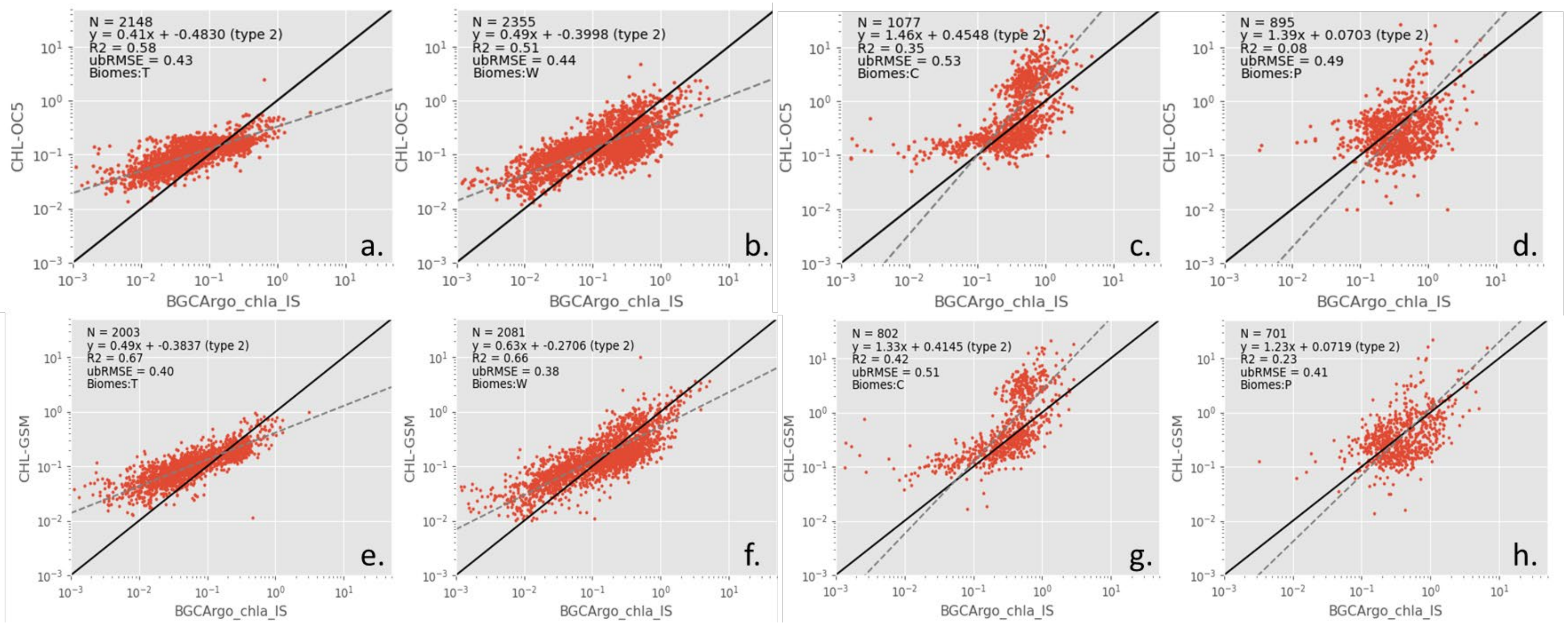


Figure 36. Scatter plot of the chlorophyll concentration derived from ocean colour against the chlorophyll concentration measured in situ with BGC-Argo for OC5 algorithm (upper panel) and GSM algorithm (bottom panel). The scatter plots are given for the four biomes as defined by Longhurst et al., 1995, the trade wind (a, e), westerlies (b, f), coastal (c, g) and polar (d, h).

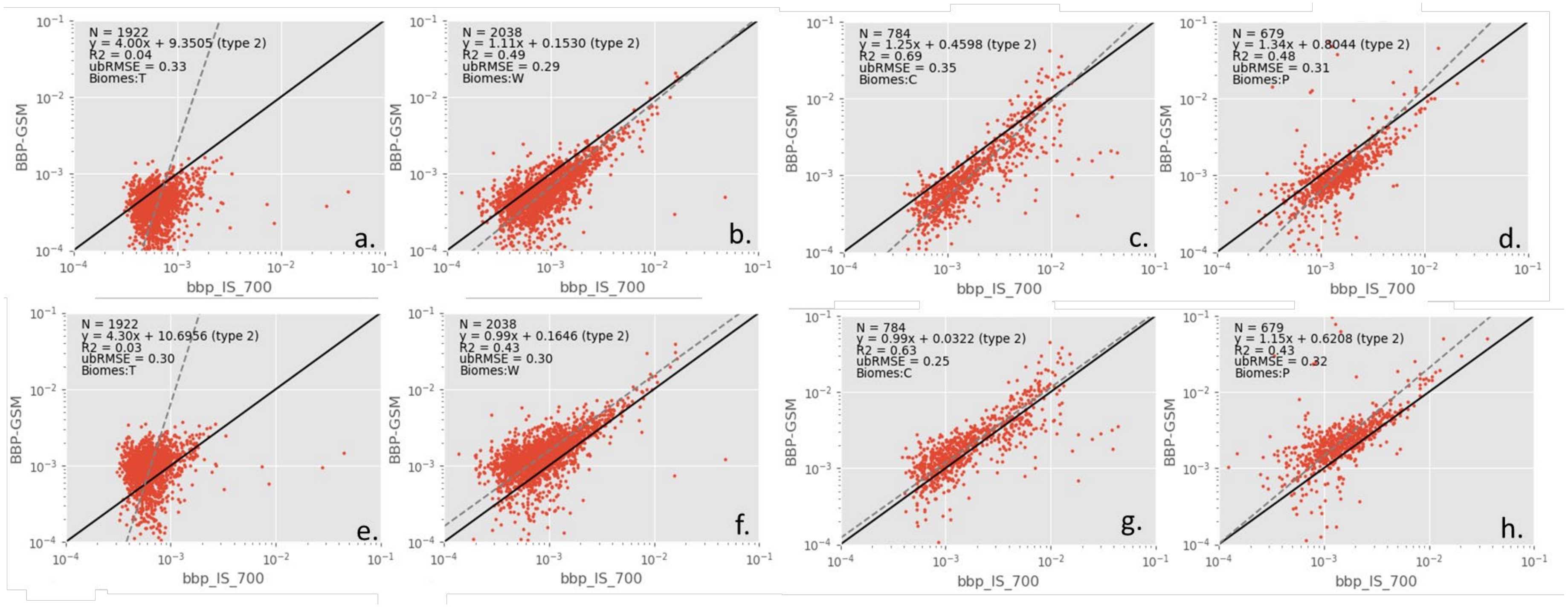


Figure 37. Scatter plot of the BBP derived from ocean colour against the BBP measured in situ with BGC-Argo for wavelength corrected (upper panel) and uncorrected wavelength (bottom panel). The scatter plots are given for the four biomes as defined by Longhurst et al., 1995, the trade wind (a, e), westerlies (b, f), coastal (c, g) and polar (d, h).

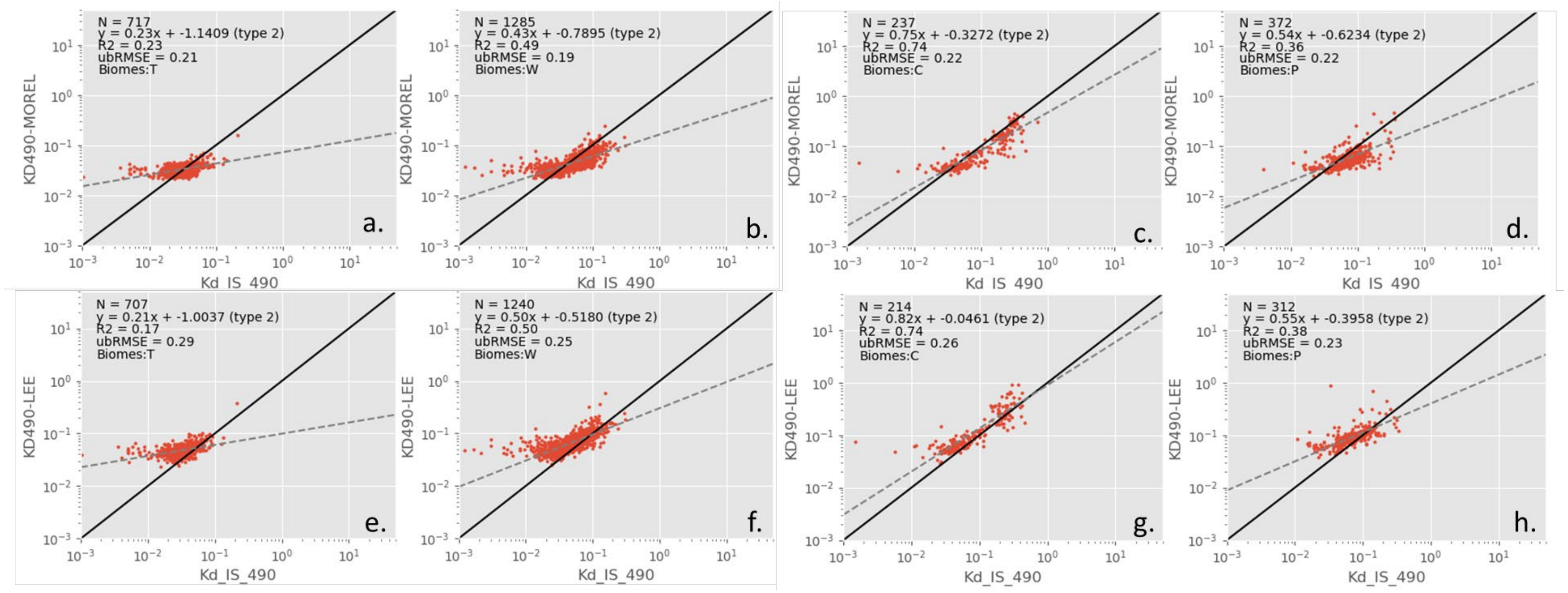


Figure 38. Scatter plot of the Kd derived from ocean colour against the Kd measured in situ with BGC-Argo for Morel algorithm (upper panel) and Lee algorithm (bottom panel). The scatter plots are given for the four biomes as defined by Longhurst et al., 1995, the trade wind (a, e), westerlies (b, f), coastal (c, g) and polar (d, h).

4.4. Biomes analysis

The performance of the ocean colour products appears to be relatively equivalent regardless of the considered biomes (Figure 36). However, better results are assessed for the trade wind and the westerly biomes, mainly representative of open waters. Polar and coastal areas are special regions, due to the physical environment, with strong wind for instance, or a wide diversity of particles close to the coast. The global algorithm might potentially be associated with stronger uncertainties in such regions (both on BGC floats measurements – due to waves agitation conditions and on Ocean Colour derived which is sensing surface with a lot of foam). One should also consider an increased uncertainty for BGC-Argo measurement in the polar biome. Indeed, in such an environment, the phytoplankton cell does not respond as the other part of the globe, as it lives in a highly dynamic environment, and photo-acclimatisation occurs. The ratio between the fluorescence and the chlorophyll concentration is a function of different parameters such as the growth phase and photophysiology (Cullen, 1982). This ratio is expected to be affected in the polar region, where the light availability might be limited due to mixing induced by the wind. Note that the BGC-Argo floats are corrected by HPLC measurement concomitant with the BGC-Argo deployment. However, the ratio of fluorescence-chlorophyll may vary with the region and the season and therefore with the float life. In the polar environment, the ratio used for the sensor correction may be not as accurate as for the beginning of the mission and may lead to the degradation of the quality of the Argo chlorophyll product (Organelli et al., 2017b).

The satellite observations of the BBP present a relatively good correlation with BGC-Argo data in all biomes except for the trade winds biome for which there is no correlation (regression slope of 4 and a coefficient of determination of 0.04, Figure 37). The weakest validation values are observed for the trade wind biomes, where the concentration in organic particles is the lowest, as in the oligotrophic gyres, and where therefore BBP values are the lowest. The bad results in this biome might be explained by the uncertainties associated with low BBP values. Satellite BBP present an underestimation within this biome.

For the Kd, as for the BBP, the weakest validation results are assessed for the trade winds biome (with a slope of the regression line of 0.23, Figure 38). The strongest validation results are assessed for the coastal biome, with the slope of the regression line higher than 0.75. Note however that for this biome, the number of matchups is limited and might be not as significant as for the westerly biome.

4.5. Estimate of the error

As previously demonstrated for the Euro-Argo RISE project, the triple collocation technique would allow to highlight regional deviation between BGC-Argo measurement and satellite observations of ocean colour. Indeed, the triple collocation estimates the error associated with three concomitant sources of data. Note that at the time of this study, the third source of data is missing to apply the triple collocation and estimate the error associated with BBP and Kd.

For the chlorophyll, thanks to the estimate of the error associated with each method, it is possible to flag pixels associated with each matchup in order to observe potential spatial trends between satellite observations and floats measurements.

For each point of matchup, the difference in concentration between satellite and floats is compared to x time the error of the matchups (σ).

$$\sigma = \sqrt{\sigma_{argo}^2 + \sigma_{satellite}^2}$$

In the present case, in order to benefit from the most accurate error estimate, the error is estimated at the global scale, and then σ equals 0.09.

When these deviations are plotted on a map (Figure 39), regional patterns seem to appear, notably for the Southern Ocean. Indeed, for the austral ocean, satellite observations are globally lower than BGC-Argo floats measurements, with a deviation between these two sources of data being equivalent to more than 2 times the error of the matchup. It should be notice that there is no consensus on the origin of these deviations. One can argue that the physical condition limits the quality of satellite measurements. Indeed, the austral ocean is a dynamic environment, with highly windy conditions potentially affecting satellite measurements due to the formation, for instance, of whitecaps at the sea surface.

Other can claim for BGC-Argo error is the chlorophyll retrieval. Indeed, in the austral ocean, the phytoplankton growth might be considered as co-limited by the iron and the light availability. Phytoplankton should adapt their photosynthetic apparatus to the environmental conditions, with a potential influence on the fluorescence to chlorophyll ratio (Petit et al., 2022). Some refinement of the relation between fluorescence and chlorophyll, considering the composition of the phytoplankton communities may improve the accuracy of the sensors.

Therefore, in the particular region of the Southern Ocean, the validation of chlorophyll products with BGC-Argo floats should be considered with caution, keeping in mind that both satellite and floats might be associated with bias and error.

For the other region (from mid-to-low latitude), no clear trend appears on the map. For this range of latitude, the deviation between satellite and BGC-Argo floats appears rather low, with a deviation comprised between 0 to 1 σ. This low deviation tends to indicate the suitability of BGC-Argo for the validation of satellite estimations of the chlorophyll concentration, as the pattern is relatively similar for both tested chlorophyll algorithms.

There is nevertheless an exception for the black sea, where satellite observations appear to overestimate the chlorophyll concentration. This might be due to the relatively complex water in this sea, for which the usual chlorophyll algorithms are not performing well.

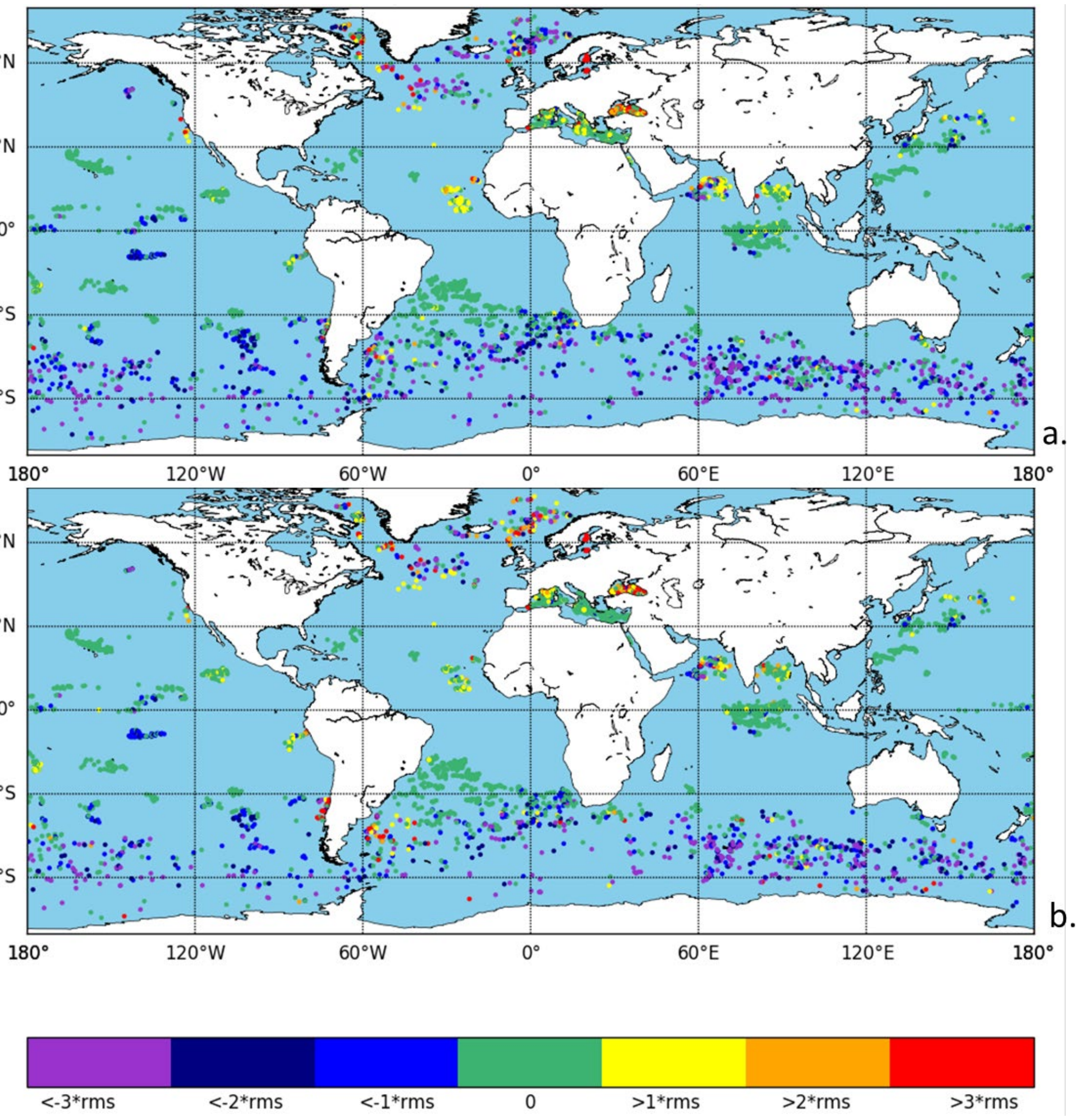


Figure 39. Map of the deviation between chlorophyll concentration estimated by satellite or derived from BGC-Argo floats measurements. The colour of the dots indicated how satellite and BGC-Argo deviate considering the error (sigma) associated to the source of data.

4.6. Conclusion

For the validation of the OLCI data, the BGC-Argo present strong advantages notably thanks to the important number of in situ measurement each year, and the fast availability of the data. Also, BGC-Argo allows documenting the global ocean, whatever the season, without engaging any human resources in difficult sampling conditions. Indeed, the validation of OLCI data, and especially for OLCI data from Sentinel-3B, might be challenging considering only data acquired during oceanographic cruises. First, data from cruises often require processing at the laboratory, which may delay the availability of the data. Secondly, because in some cases, the data shows properties of the organisation which did the measure, limiting the number of available data.

Another advantage of BGC-Argo floats in the validation of ocean colour data is the set of parameters measured by these platforms. The validation is not limited to chlorophyll algorithms, but also particular backscattering or coefficient of diffuse attenuation. The validation of these parameters is as important as chlorophyll validation since Kd and BBP are proxies for the particle concentration. Also, as BBP is considered proportional to the particulate organic carbon in the open ocean, it might be used for the estimation of the efficiency of the biological carbon pump.

When satellite observations and BGC-Argo floats measurements are associated with a third source of data, the error of each source of data can be determined thanks to the triple collocation. The error hence determined can be used as an index to detect trends in particular regions, and potentially propose some improvement of the algorithm.

The number of available matchups thanks to BGC-Argo floats may help to compare the performance of different ocean colour algorithms at a global and/or regional scale. For instance, in the present study, BGC-Argo pointed out that the accuracy of the GSM algorithm appears to be better than the OC5 for the chlorophyll, except in the polar biome, or both algorithms present strong discrepancies with Argo floats. In that biome, some uncertainties remain between fluorescence and chlorophyll and the systematic deviation observed in the austral ocean call for an improvement in the calibration of the fluorescence sensor.

Even if BGC-Argo floats should not be considered golden sensors, the expansion of the BGC-Argo network, together with the diversity of sensors on each float is undoubtedly a strength for the validation of ocean colour data.

General conclusion

The main objective of this task is to validate data from Sentinel-3A (S3A) and Sentinel-3B (S3B) satellite and remote sensing algorithms with *in situ* observations, such as sea level height from gliders and Argo drifters, velocities from gliders, drifters, HF radars and ADCP, and ocean colour from BGC Argo.

Different methods are applied, from traditional rms differences to triple collocation to estimate random errors from the correlations between three datasets.

A glider in the Balearic Sea sampling the ocean underneath S3A tracks shows good agreement in dynamic height with good correlation greater than 0.85 and RMS of 1.3cm, and the across-track geostrophic velocity RMS vary from 4.4cm/sec to 9.6 cm/s, the later when the satellite track and glider are farther apart. The comparison with depth average velocities (DAV) shows a ~0.75 correlation and RMS of 10cm/s due to a difference in the signal content between GPS-derived DAV and altimetry-derived signal.

In the Bay of Biscay, altimetry data is compared to in-situ HF radar and ADCP measurements. The region is broken into different sub-region with different current regimes. For TC comparison, periods or areas of a strong geostrophic signal are recommended. A similar precision was observed for the ADCP and the altimeter in terms of random error (around 3.45 cm/s) whereas the altimeter measurements differed from the ADCP in terms of instrumental error with a gain (i.e., multiplicative error) of 1.98 and an offset (i.e., additive error) of 3.85 cm/s. The addition of the Ekman currents into the analysis improved, as shown in previous studies, the agreement between in-situ and altimetry-derived currents. This suggests that the sensitivity to the addition of the wind effects should be considered in future altimeter validation comparisons, at least in terms of current velocities.

Globally TC applied to Argo floats and sea level anomalies from different altimeter shows improvements in altimeter data processing (between L2 and L3 products) with errors dropping from 3.62 cm to 2.84 cm and an increase in correlation. TC applied on longitude/latitude gridded locations shows higher errors in strong ocean variability regions. The TC is also applied to AOML drifters and along geostrophic velocities, and along-track S3B velocities have a less precise random error (7.8 cm/s) than the drifters (4.8 cm/s). As previously, the errors are higher in strong gradient areas. These methods show a sensitivity to the post-treatment of the data and should be carefully set up. They are a good tool to assess data quality.

Finally, S3A ocean colour for the OLCI sensor is validated against Argo BGC, and remote sensing OC5 and GSM are tested. The BGC-Argo present strong advantages notably a global and important number of *in situ* measurement, a large set of parameters and the fast availability of the data. Thus, in addition to chlorophyll algorithms validation, backscattering or coefficient of diffuse attenuation is also validated. These latter parameters are as important as chlorophyll since Kd and BBP are proxies for the concentration of particles. Also, as BBP is considered proportional to the particulate organic carbon in the open ocean, it might be used for the estimation of the efficiency of the biological carbon pump. With the addition of a third source of data, the error of each source of data can be determined by triple collocation analysis, and the estimated error can be used as an index to detect trends in particular regions, and potentially propose some improvement of the algorithm such as in the calibration of fluorescence sensor.

All these above studies show that *in situ* data are invaluable for the validation and calibration of satellite measured/derived ocean parameters. The diversity of sensors, their density in time and space and their rapid availability is a key to the improvement of present and future missions, and our knowledge of the ocean.

References

Bishop, J. K. (2009). Autonomous observations of the ocean biological carbon pump. *Oceanography*, *22*(2), 182-193.

Bonnefond, P., Haines, B.J., Watson, C., In situ absolute calibration and validation: a link from coastal to open-ocean altimetry, Coastal altimetry, 2011 – Springer, 259–296.

Bouffard, J., A. Pascual, S. Ruiz, Y. Faugère, and J. Tintoré (2010), Coastal and mesoscale dynamics characterization using altimetry and gliders: A case study in the Balearic Sea, J. Geophys. Res., 115, C10029, doi: doi.org/10.1029/2009JC006087.

Bricaud, A., Claustre, H., Ras, J., & Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. *Journal of Geophysical Research: Oceans*, *109*(C11). doi: doi.org/10.1029/2004JC002419

Caballero, A., Mulet, S., Ayoub, N., Manso-Narvarte, I., Davila, X., Boone, C., Toublanc, F., & Rubio, A. (2020), Integration of HF Radar Observations for an Enhanced Coastal Mean Dynamic Topography, Frontiers in Marine Science, doi: doi.org/10.3389/fmars.2020.588713.

Cetinić, I., Perry, M. J., D'asaro, E., Briggs, N., Poulton, N., Sieracki, M. E., & Lee, C. M. (2015). A simple optical index shows spatial and temporal heterogeneity in phytoplankton community composition during the 2008 North Atlantic Bloom Experiment. *Biogeosciences*, *12*(7), 2179-2194.

Charria, G., Lazure, P., Le Cann, B. Serpette, A., Reverdin, G., Louazel, S., Batifoulier, F., Dumas, F., Pichon, A., & Morel, Y. (2013), Surface layer circulation derived from Lagrangian drifters in the Bay of Biscay, Journal of Marine Systems, 109-110, supplement, S60-S76.

Chavanne, C.P., Klein, P. (2010), Can oceanic submesoscale processes be observed with satellite altimetry?, Geophysical Research Letters, 37,22, , doi: 10.1029/2010GL045057

Claustre, H., Morel, A.,Babin, M ., Cailliau, C., Marie, D., Marty, J.C., Tailliez, D., & Vaulot, D. (1999) Variability in particle attenuation and chlorophyll fluorescence in the tropical Pacific: Scales, patterns, and biogeochemical implications. Journal of Geophysical Research, 104, C2, 3401-3422.

Cullen, J. J. (1982). The deep chlorophyll maximum: comparing vertical profiles of chlorophyll a. Canadian Journal of Fisheries and Aquatic Sciences, 39(5), 791-803.

Ferrer, L., Liria, P., Bolanos, R., Balseiro, C., Carracedo, P., González-Marco, D., González, M., Fontán, A., Mader, J., & Hernández, C. (2010). Reliability of coupled meteorological and wave models to estimate wave energy resource in the Bay of Biscay. 3 rd International Conference on Ocean Energy, 6 October, Bilbao.

Garnesson, P., Mangin, A., Fanton d'Andon, O., Demaria, J., & Bretagnon, M. (2019). The CMEMS GlobColour chlorophyll a product based on satellite observation: Multi-sensor merging and flagging strategies. Ocean Science, 15(3), 819-830.

Gohin, F., Druon, J. N., Lampert, L. (2002). A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. International journal of remote sensing, 23(8), 1639-1661

González, M., Uriarte, A., Fontán, A., Mader, J., Gyssels, P. (2004), Oceanography and marine environment of the Basque Country, Marine Dynamics, 70, 133-157.

Hu, C., Lee, Z., Franz, B. (2012). Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. Journal of Geophysical Research, 117(C1). doi: 10.1029/2011jc007395

Janssen, P. A.-R. (2007). Error estimation of buoy, satellite, and model wave height data. *Journal of Atmospheric and Oceanic Technology, 24(9)*, 1665–1677. doi:doi.org/10.1175/JTECH2069.1

Kaplan, D., & Lekien, F. (2007), Spatial interpolation and filtering of surface current data based on open-boundary modal analysis, Journal of geophysical research, 112, C12, doi: doi.org/10.1029/2006JC003984

Lacour, L., Ardyna, M., Stec, K. F., Claustre, H., Prieur, L., Poteau, A., ... & Iudicone, D. (2017). Unexpected winter phytoplankton blooms in the North Atlantic subpolar gyre. *Nature Geoscience*, *10*(11), 836-839. doi: doi.org/10.1038/ngeo3035

Lee, Z. P., et al. (2002), Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters, Applied Optics, 41, 5755- 5772

Lee, Z. P., Du, K. P., & Arnone, R. (2005). A model for the diffuse attenuation coefficient of downwelling irradiance. Journal of Geophysical Research: Oceans, 110(C2).

Liu, Y., Weisberg, R.H., Vignudelli, S., L Roblou, L., & Merz, C., Comparison of the X-TRACK altimetry estimated currents with moored ADCP and HF radar observations on the West Florida Shelf, Advances in Space Research, 50, 8, 2012, 1085-1098

Longhurst, A.R et al. (1995). An estimate of global primary production in the ocean from satellite radiometer data. J. Plankton Res. 17, 1245-1271

Longhurst, A., Sathyendranath, S., Platt, T., & Caverhill, C. (1996). An estimate of global primary production in the ocean from satellite radiometer data. Oceanographic Literature Review, 2(43), 203.

Loisel, H., Nicolas, J. M., Deschamps, P. Y., & Frouin, R. (2002). Seasonal and inter-annual variability of particulate organic matter in the global ocean. *Geophysical Research Letters*, *29*(24), 49-1.

Manso-Narvarte, I., Caballero, A., Rubio, A., Dufau, C., & Birol, F. (2018), Joint analysis of coastal altimetry and high-frequency (HF) radar data: observability of seasonal and mesoscale ocean dynamics in the Bay of Biscay, 14, 5, 1265–1281.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., & Piles, M. a. (2014). Extended Triple Collocation: Estimating Errors and Correlation Coefficients with Respect to an Unknown Target. *Geophysical Research Letters 41*, 6229–6236.

Merckelbach, L. M., R. D. Briggs, D. A. Smeed, and G. Griffiths (2008), Current measurements from autonomous underwater gliders, paper pre- sented at Current Measurement Technology, 2008, 9th Working Confer- ence on Current Measurement Technology, Inst. of Electr. and Electr. Eng., New York.

Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., & Franz, B. A. (2007). Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. Remote Sensing of Environment, 111(1), 69-88.

Morrow, R., Fu,L-L., Farrar, J.T., Seo, H., & Le Traon,P-Y. (2017), Ocean Eddies and Mesoscale Variability, Satellite Altimetry Over Oceans and Land Surfaces, CRC Press.

Mueller, J. L. (2003). Ocean optics protocols for satellite ocean color sensor validation, revision 4: radiometric measurements and data analysis protocols (Vol. 3). Goddard Space Flight Center.

Maritorena, S. and Siegel, D.A. 2005. Consistent Merging of Satellite Ocean Colour Data Sets Using a Bio-Optical Model. Remote Sensing of Environment, 94, 4, 429-440.

Mignot, A., d'Ortenzio, F., Taillandier, V., Cossarini, G., & Salon, S. (2019). Quantifying Observational Errors in Biogeochemical-Argo Oxygen, Nitrate, and Chlorophyll a Concentrations. Geophysical Research Letters, 46, 8, 4330-4337.

Organelli, E., Claustre, H., Bricaud, A., Barbieux, M., Uitz, J., d'Ortenzio, F., & Dall'Olmo, G. (2017). (a) Bio-optical anomalies in the world's oceans: An investigation on the diffuse attenuation coefficients for downward irradiance derived from Biogeochemical Argo float measurements. Journal of Geophysical Research: Oceans, 122(5), 3543-3564.

Organelli, E., Barbieux, M., Claustre, H., Schmechtig, C., Poteau, A., Bricaud, A., ... & Xing, X. (2017). (b) Two databases derived from BGC-Argo float measurements for marine biogeochemical and bio-optical applications. Earth System Science Data, 9(2), 861-880.

Oubelkheir, K., Claustre, H., Sciandra, A., & Babin, M. (2005). Bio-optical and biogeochemical properties of different trophic regimes in oceanic waters. *Limnology and oceanography*, *50*(6), 1795-1809.

Pascual, A., Lana, A., Troupin, C., Ruiz, S., Faugère, Y., Escudier, R., & Tintoré, J. (2015), Assessing SARAL/AltiKa data in the coastal zone: comparisons with HF radar observations, Marine Geodesy, 38, Issue sup1: The SARAL/AltiKa Satellite Altimetry Mission, 260-276, doi: 10.1080/01490419.2015.1019656

Petit, F., Uitz, J., Schmechtig, C., Dimier, C., Ras, J., Poteau, A., ... & Claustre, H. (2022) Influence of the phytoplankton community composition on the in situ fluorescence signal: Implication for an improved estimation of the chlorophyll a concentration from BioGeoChemical-Argo profiling floats. *Frontiers in Marine Science*, 1837.

Rubio, A., Reverdin, G., Fontán, A., González, M., & Mader, J. (2011), Mapping near-inertial variability in the SE Bay of Biscay from HF radar data and two offshore moored buoys, Geophysical Research Letters, 38, 19, doi: doi.org/10.1029/2011GL048783

Roesler, C. J., Emery, W.J., & Kim, S.Y. (2013), Evaluating the use of high-frequency radar coastal currents to correct satellite altimetry, Journal of Geophysical Research, 118, 7, 3240-3259.

Solabarrieta, L., Rubio, A., Castanedo, S. Medina, R., Charria, G., & Hernández, C. (2014), Surface water circulation patterns in the southeastern Bay of Biscay: New evidences from HF radar data, Contiental Shelf Research, 74, 60-76.

Solabarrieta, L., Rubio, A., Cárdenas, M., Castanedo, S., Esnaola, G., Méndez, F., Medina, R., & Ferrer, L. (2015), Probabilistic relationships between wind and surface water circulation patterns in the SE Bay of Biscay, Ocean Dynamics, 65, pages1289–1303.

Solabarrieta, L., Frolov, S., Cook, M., Paduan, J., Rubio, A., González, M., Mader, J., &Charria, G. (2016), Skill Assessment of HF Radar–Derived Products for Lagrangian Simulations in the Bay of Biscay, Am. Meteo. Soc. 2585–2597, doi: doi.org/10.1175/JTECH-D-16-0045.1.

Troupin, C., Pascual, A., Valladeau, G., Pujol, I., Lana, A., Heslop, E., Ruiz, S., Torner, M., Picot, N., &Tintoré, J. (2015), Illustration of the emerging capabilities of SARAL/AltiKa in the coastal zone using a multi-platform approach, Advances in Space Research, 55, 1, 51-59.

Tukey, J. (1977). Exploratory Data Analysis. Addison-Wesley, Reading, MA.