HYDROCOASTAL

SAR/SARin Radar Altimetry for Coastal Zone and Inland Water Level

Product Validation Report

Deliverable D2.5

Sentinel-3 and Cryosat SAR/SARin Radar Altimetry for Coastal Zone and Inland Water

ESA Contract 4000129872/20/I-DT

Project reference: HYDROCOASTAL_ESA_PVR_D2.5
Issue: 2.0

25/07/2022

This page has been intentionally left blank

Change Record

Date	Issue	Section	Page	Comment
14/04/2022	1.0			First version
25/07/2022	2.0			Includes content from validation studies

Control Document

Process	Name	Date
Written by:	Mathilde Cancet, Luciana Fenoglio-Marc, Christine Gommenginger, Francisco Calafat, Jesús Gómez-Enri, Andrew Shaw, Denise Dettmering, Daniel Scherer, Pierre Fabry, Nicolas Bercher, Karina Nielsen, Elena Zakharova, Joana Fernandes, Angelica Tarpanelli	25/07/2022
Checked by	David Cotton	
Approved by:		

Subject	Radar Altimetry for Coastal Zone and Inland Water Level	**Project**	HYDROCOASTAL

Author	**Organisation**	**Internal references**
Mathilde Cancet	NOVELTIS	HYDROCOASTAL_ESA_PVR_D2.5
Luciana Fenoglio-Marc	U Bonn	
Christine Gommenginger	NOC	
Francisco Calafat	NOC	
Jesus Gómez-Enri	U Cadiz	
Andrew Shaw	Skymat	
Denise Dettmering, Daniel Scherer	TUM	

Pierre Fabry	Along Track	
Nicolas Bercher	AltHydroLab.fr	
Karina Nielsen	DTU	
Elena Zakharova	NUIM	
Joana Fernandes	U Porto	
Angelica Tarpanelli	CNR-IRPI	

	Signature	Date
For HYDROCOASTAL team		
For ESA		

Table of Contents

1 Introduction

1.1 The HYDROCOASTAL Project

The HYDROCOASTAL project is a project funded under the ESA EO Science for Society Programme, and aims to maximise the exploitation of SAR and SARin altimeter measurements in the coastal zone and inland waters, by evaluating and implementing new approaches to process SAR and SARin data from CryoSat-2, and SAR altimeter data from Sentinel-3A and Sentinel-3B.

One of the key objectives is to link together and better understand the interactions processes between river discharge and coastal sea level. Key outputs are global coastal zone and river discharge data sets, and assessments of these products in terms of their scientific impact.

1.2 Scope of this Report

This document is the Product Validation Report (PVR) report for HYDROCOASTAL and it corresponds to the deliverable D2.5. of the project. The scope of this report is to compile the main findings of the validation activities performed on the HYDROCOASTAL test datasets.

1.3 Document Organisation

This document is organised in four main sections:

- Section 1: A short introduction defining the scope of this report.
- Section 2: The results of the validation activities for the L2 products in the coastal zones (CZ).
- Section 3: The results of the validation activities for the L2, L3 and L4 products in the inland waters (IW).
- Section 4: The results of the validation activities for the new Wet and Dry Troposphere corrections.

1.4 Reference documents

HYDROCOASTAL Proposal: SAR/SARin Radar Altimetry for Coastal Zone and Inland Water Level. Proposal, January 2020.

HYDROCOASTAL Product Validation Plan (PVP – D2.4)

2 Validation of the Test Datasets in different Coastal Zone Scenarios

2.1 Introduction

In this section, we present the results of the validation of the Test Dataset Geophysical parameters against other satellites and in situ data, in different Coastal Zone Scenarios.

The validation activities include the analysis of the influence of land proximity and ground-track orientation on SAR/SARin, the analyses of the different algorithms proposed to produce the final dataset, and validation against independent observations.

A common methodology to define the validation metrics for the Coastal Zones has been described in the PVP document.

2.2 Validation in the German Bight/Baltic Sea region (U Bonn)

In this section the validation of S3A/S3B/CS2 sea level anomaly/SWH/U10 data is presented in the German Bight/Baltic Sea area (Figure 2.2.1). The analysis includes the five retrackers BONN, DTU, ESA, ISR and TUM presenting results for SLA and SWH. Missing is S3A/S3B and CS2 U10.

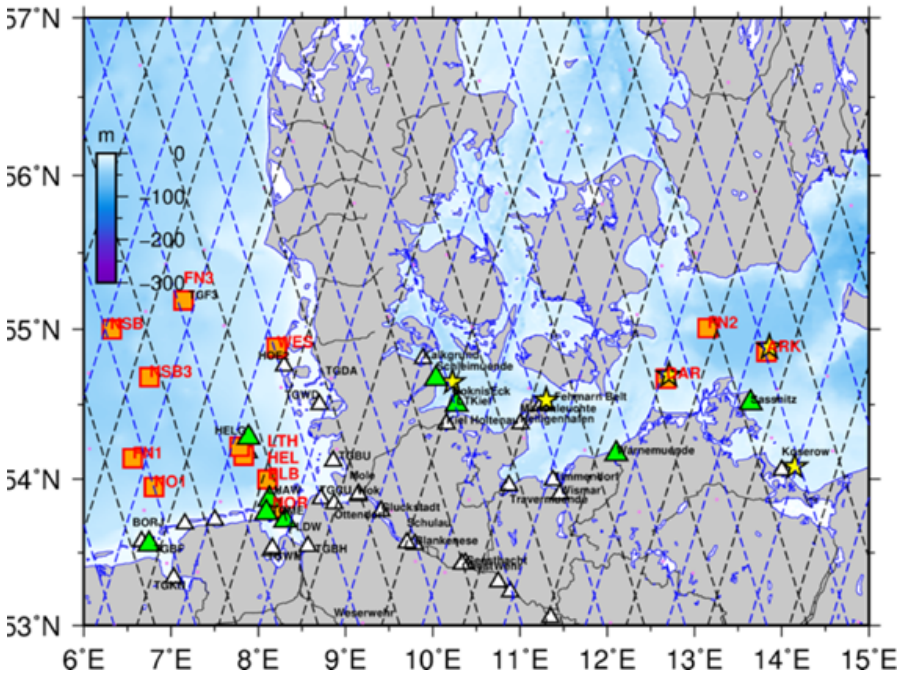


Figure 2.2.1. Tide gauge locations with S3A/B ground-tracks (green triangle), 9 TG and 8 buoys.

2.2.1 Sea Level Anomaly (SLA): S3A and S3B

From the s3_l2e_reduced files the quantities used are:

- Altitude: 'alt',
- Range: 'retracked_range_ESA',
- Range corrections: 'GIM_iono', 'mod_dry_tropo_cor_zero_altitude', 'mod_wet_tropo_cor_zero_altitude', 'upt_dry_tropo', 'gpd_wet_tropo',
- Geophysical corrections: 'ocean_tide_fes'/'ocean_tide_got', 'ocean_tide_non_eq', 'solid_earth_tide', and 'geocentric_polar_tide',
- Mean Sea Surface: 'MSS',
- Significant Wave Height: 'swh_ESA' to compute the SSB correction.

From the output files the range is extracted for each retracker: 'retracked_range_BONN', 'retracked_range_dtu', 'retracked_range_isr', 'retracked_range_tum'.

From the 's3_l2e_reduced_upt_geo' files the University of Porto Dry and Wet tropospheric corrections are used: 'upt_dry_tropo' and 'gpd_wet_tropo'. The following results are first shown with standard corrections. For the ocean tide correction, the t-tide program is used to correct the tide gauge by the tides.

The precision is first evaluated considering the along-track difference between SLA at less than 20 kilometers from the coast (Figure 2.2.2). The spectra indicate a higher noise in the ISR data (Figure 2.2.3).

With the processing "overpass" used in Fenoglio et al. (2020), 20 Hz give larger STDD than for 1Hz. In this case the DAC is applied and the same ocean model is used to correct both altimetry and tide gauges. The smallest STDD is 9.1 cm and corresponds to the ESA retracker, the largest STDD corresponds to the DTU retracker. Finally results of overpass at nine tide gauges are merged to build a boxplot, which shows a mean standard deviation of about 10 cm. The number of retained points is consistent, maximum expected is 27 for S3A.

The along-track data are binned in 20 Hz time-series. DAC correction is not applied, ocean tide is removed from the altimeter using the FES ocean model. We apply a selection criterion over all the binned data eliminating the sea level anomaly uncorrected for the DAC larger than +/- 1.5 meters for all the tracks. In extreme conditions, neglecting the DAC correction can generate unwanted effects. In the overpass procedure we apply a 3-sigma criterion on the absolute median deviation of all the sea level anomalies uncorrected for DAC. The statistics of the difference between binned 20 Hz time-series and in-situ data (stdd and correlation) is first computed at each location, no 3-sigma selection criteria applied. Finally, we average along track over the points with distance between 2 and 20 km from the coast, only points within 20 km from the station are accepted (Figure 2.2.4). The averaged statistics at Helgoland for the Sentinel-3 SLA along-track is shown in Table 2.2.1 for SLA uncorrected for SSB and in Table 2.2.2 for SLA corrected for SSB. The stdd is about 10 cm for both S3A and S3B. The DTU retracker is the noisiest, with stdd of 15 cm. The statistics of differences of altimetric and in-situ SLA for HELG for different choices of SSB is given in the tables below. Table 2.2.3 gives the results for LTKI for SLA with SSB not applied. Table 2.2.4 gives the STDD of SLA for 9 stations.

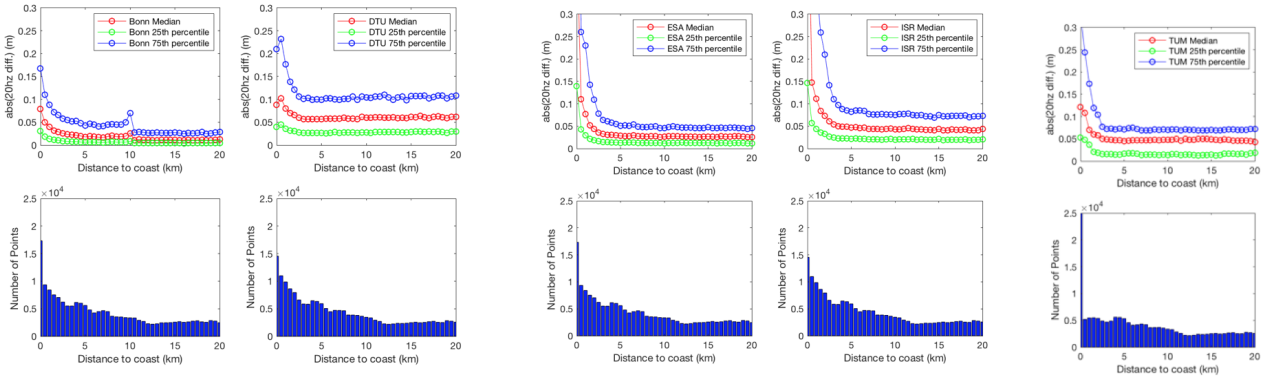


Figure 2.2.2. Noise and number of observations for 20 Hz Sentinel-3A (from left to right in alphabetic order Bonn-DTU-ESA-ISR-TUM). Time interval is from 2018-06-01 to 2020-06-30.

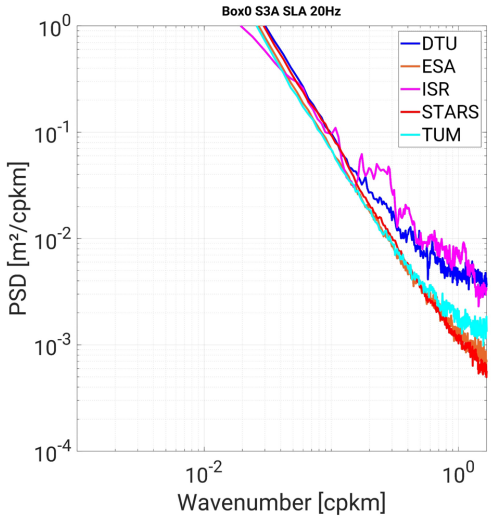


Figure 2.2.3. SLA spectra of S3A corresponding to standard ESA and to the retrackers in the German Bight.

Table 2.2.1. Statistics (STDD and correlation) of S3A and S3B SLA at Helgoland TG station with SSB not applied

In-situ	S3A stdd	S3A cor	S3B stdd	S3B cor
Bonn	0.109	0.944	0.112	0.929
DTU	0.158	0.887	0.102	0.937
ESA	0.106	0.946	0.109	0.928
ISR	0.104	0.947	0.110	0.927
TUM	0.111	0.943	0.108	0.930

Table 2.2.2. Statistics (STDD and correlation) for S3A SLA at Helgoland TG station using different types of SSB

In-situ	S3A stdd 5% SWH_ESA	S3A cor 5% SWH_ESA	S3A std 5% SWH or own	S3A cor 5% SWH or own
Bonn	0.114	0.941	0.114	0.940
DTU	0.174	0.874	0.174	0.873
ESA	0.113	0.941	0.115	0.941
ISR	0.113	0.939	0.117	0.938
TUM	0.116	0.939	0.127	0.930

Table 2.2.3. Statistics (STDD and correlation) of SLA for S3A and S3B LTKI TG station with SSB not applied

In-situ	S3A stdd	S3A cor	S3B stdd	S3B cor
Bonn	0.066	0.950	0.062	0.961
DTU	0.052	0.967	0.030	0.992
ESA	0.052	0.941	0.044	0.981
ISR	0.049	0.970	0.068	0.954
TUM	0.050	0.970	0.034	0.990

Table 2.2.4. Statistics (STDD) of SLA for S3A for 9 stations with SSB applied

SLA S3A SSB Each					
STDD	**BONN**	**DTU**	**ESA**	**ISR**	**TUM**
BORJ	0.1752	0.1733	0.1785	0.1901	0.1928
FLDW	0.2824	0.2462	0.3340	0.3463	0.3056
HELG	0.1160	0.1747	0.1143	0.1132	0.1161
LHAW	0.1230	0.1670	0.1738	0.1776	0.1664
LTKI	0.0691	0.0520	0.0468	0.0492	0.0489

SASS	0.0632	0.0580	0.0607	0.0570	0.0608
SCHL	0.0555	0.0478	0.0506	0.0537	0.0501
TGME	0.1481	0.2579	0.3219	0.3270	0.3092
WARN	0.0723	0.1445	0.0692	0.0891	0.0556
STDEV	0.0728	0.0797	0.1124	0.1146	0.1059
MEAN	0.1228	0.1468	0.1500	0.1559	0.1450

The methodology we have selected for the validation is slightly different from the methodology “overpass” used in Fenoglio et al. (2020), where one single measurement, the nearest to the in-situ, is retained per each pass and a time series is built with the selected data. Figure 2.2.5 shows the boxplots for both. We notice that the BONN retracker is the best for both methodologies in terms of deviation from the median.

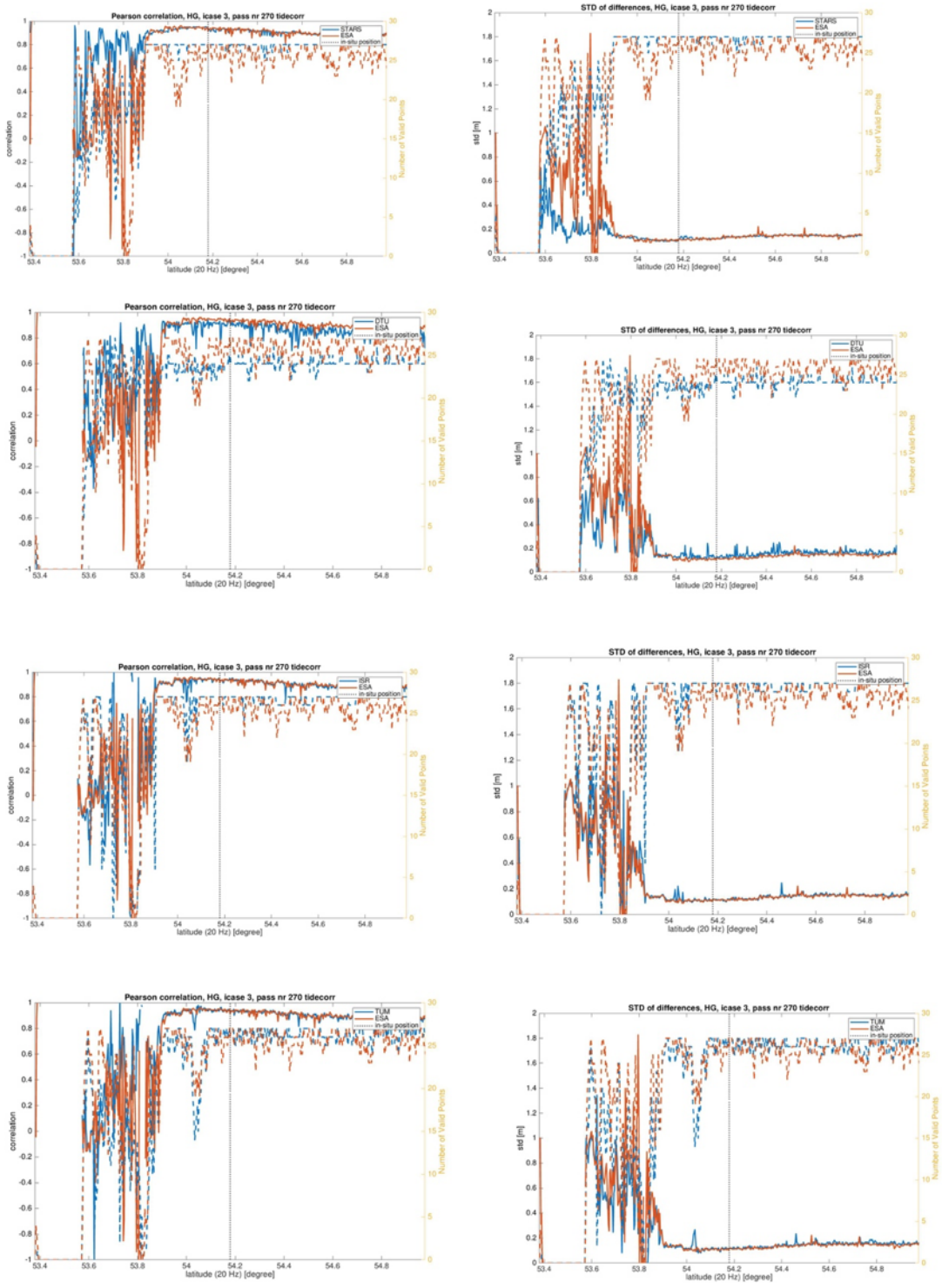


Figure 2.2.4. Correlation (left) and standard deviation (right) between in-situ and along-track binned time-series (here ISR und TUM). Number of retained points is on the right axis and presented with dashed lines of the same color as for the retracker. A black dashed vertical line gives the track point closest to the in-situ.

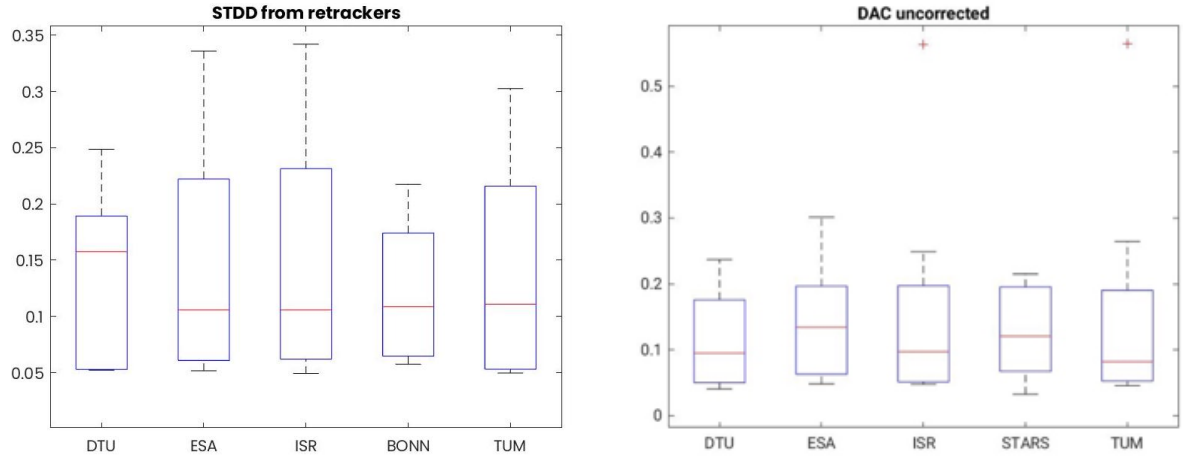


Figure 2.2.5. S3A without DAC applied. Boxplot of selected methodology (left) and boxplot of overpass (right).

The Validation is performed in the Elbe Estuary. Figure 2.2.6 shows the water height anomaly for Sentinel-3. The procedure "virtualpass" (Fenoglio et al., 2020), used for in-land water, defines a virtual point (VP) and a polygon for each in-situ station. The altimeter measurements falling inside the polygon within a chosen distance from the VP are averaged. The error assigned to the averaged altimeter height is the standard deviation of the averaged data, or the standard deviation read from the data products, if one measurement is used. The smallest stdd is found by the DTU retracker (stdd 58 cm), Fenoglio et al. 2020 report a stdd of 40 cm with SAMOSA+ and SAMOSA++ over 2016-2018.

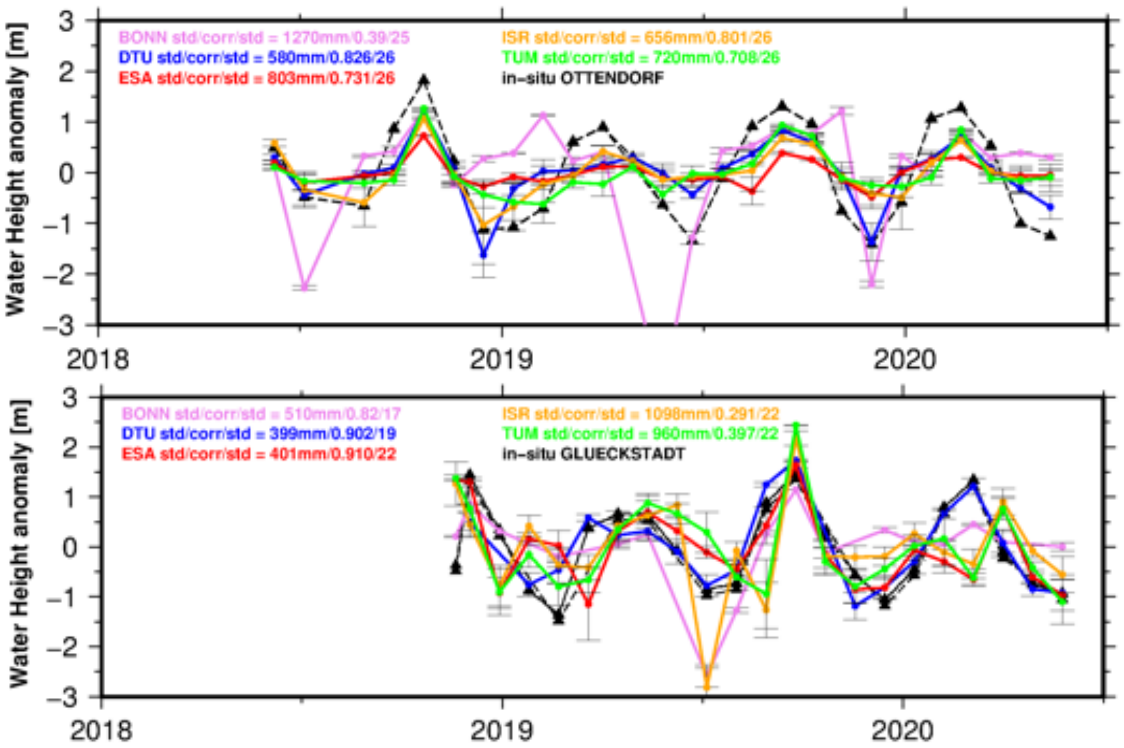


Figure 2.2.6. Water height anomaly from the 5 retrackers from S3A in Ottendorf (top) and from S3B in Glückstadt (bottom). The virtualpass method is used. DAC is not applied.

2.2.2 Significant Wave Height (SWH): S3A and S3B

Figure 2.2.7 gives the distribution of the SWH for each retracker. STARS has the least number of points, in ISR many values are set to zero. It appears that many SWH which are set to NaN in STARS are set to zero in the other retrackers. They correspond to small SWH. Figure 2.2.8 confirms the conclusion above. Tables 2.2.5 and 2.2.6 give the statistics (RMS, slope and intercept) of the comparison of SWH at 8 and 6 buoys for S3A and S3B respectively obtained by averaging between 2 and 20 km the statistics obtained for each 20Hz time-series. At most of the stations RMS is similar for S3A and S3B, the highest differences between S3A and S3B RMS are seen in LTH and are probably due to the different location of the tracks. From the final tables we see that the averaged RMSE is around 50 cm for all retrackers (Table 2.2.11). See the boxplot of RMS for S3A in Figure 2.2.9.

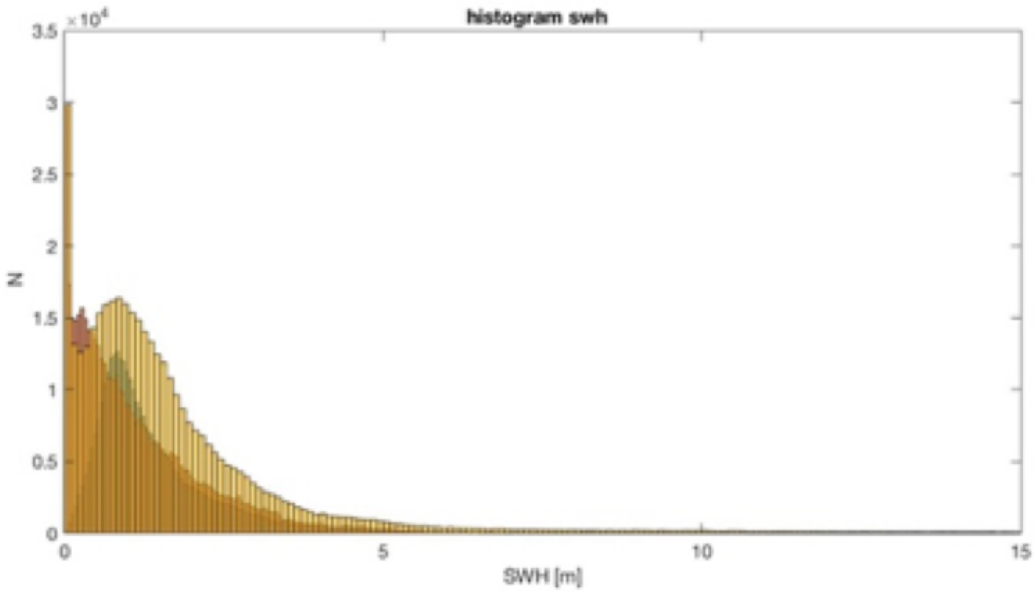


Figure 2.2.7. Scatterplot for SWH in ROI Bight-Baltic: STARS (green), ESA (light brown), ISR (brown)

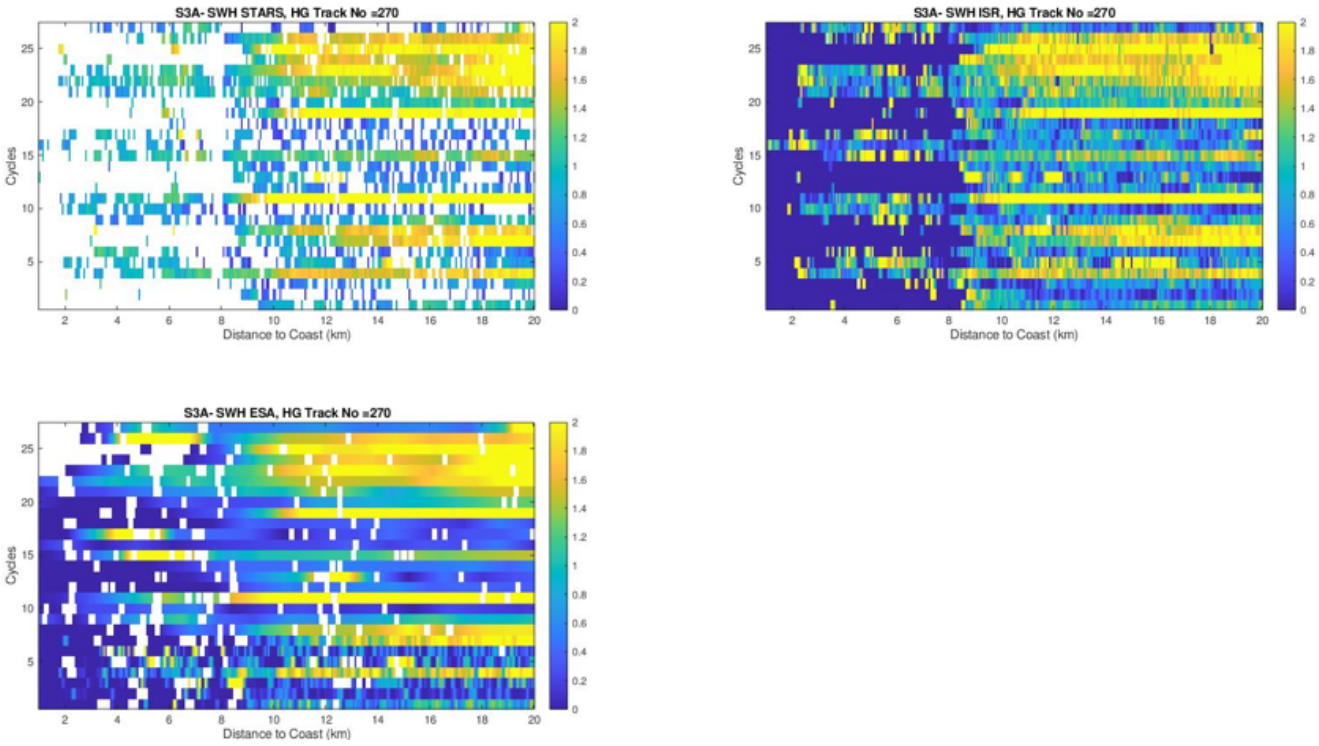


Figure 2.2.8. SWH values with distance to coast smaller than 20 km

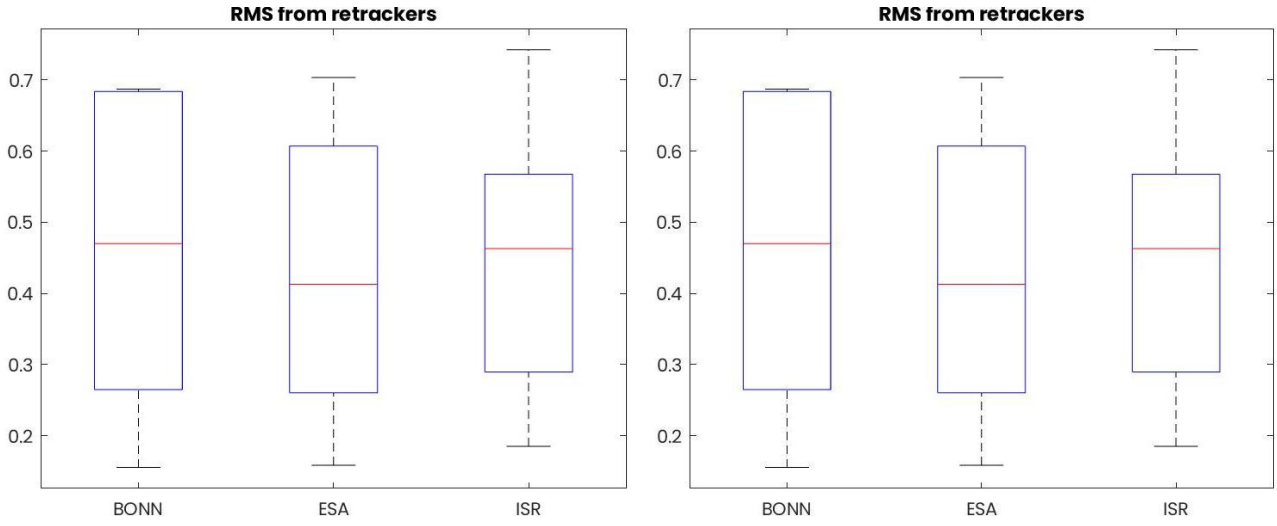


Figure 2.2.9. Boxplots of SWH RMS from S3A at 8 buoys (left) and S3B at 6 buoys (right).

Table 2.2.5. RMS, slope and intercept of SWH for S3A at 8 buoys.

SWH S3A								
ARK	RMS	SLOPE	Intercept		FN3	RMS	SLOPE	Intercept
ESA	0.8339	0.1541	0.7674		ESA	0.7561	0.5045	0.7258
ISR	0.4893	0.9177	0.4390		ISR	0.7367	0.9538	0.7831
BONN	0.7193	0.9593	0.7053		BONN	0.6994	0.6776	0.7647
DAR	RMS	SLOPE	Intercept		FNO	RMS	SLOPE	Intercept
ESA	0.6354	0.1443	0.7262		ESA	0.4629	1.1134	0.1322
ISR	0.7664	-0.0208	0.8316		ISR	0.4863	1.1849	0.0933
BONN	0.5627	0.2123	0.6702		BONN	0.4359	1.2072	-0.0451
ELB	RMS	SLOPE	Intercept		HELG	RMS	SLOPE	Intercept
ESA	0.3996	1.2936	0.1808		ESA	0.3299	1.0806	0.1104
ISR	0.3914	1.4152	-0.2620		ISR	0.3712	1.1715	-0.1506
BONN	0.3159	1.2110	-0.1161		BONN	0.3254	1.2003	-0.1178
FN2	RMS	SLOPE	Intercept		LTH	RMS	SLOPE	Intercept
ESA	0.6085	0.4620	0.6824		ESA	0.4651	1.0326	0.3128
ISR	0.6135	0.4387	0.6818		ISR	0.4700	1.1050	0.0783
BONN	0.5662	0.4494	0.6666		BONN	0.4201	1.2024	0.0664

Table 2.2.6. RMS, slope and intercept of SWH for S3B at 6 buoys.

SWH S3B							
ARK	RMS	SLOPE	Intercept	FN3	RMS	SLOPE	Intercept
ESA	0.6072	0.6462	0.2754	ESA	0.7035	0.4889	0.5758
ISR	0.5676	0.6863	0.2898	ISR	0.7426	0.4808	0.5661
BONN	0.5318	0.7830	0.2076	BONN	0.6839	0.5059	0.5612
DAR	RMS	SLOPE	Intercept	FNO	RMS	SLOPE	Intercept
ESA	0.4362	0.8771	0.3146	ESA	0.3896	0.9751	0.2835
ISR	0.5270	-0.2906	0.5230	ISR	0.3994	0.9756	0.3430
BONN	0.6874	-0.2131	0.6616	BONN	0.4085	1.0527	0.2769
ELB				HELG	RMS	SLOPE	Intercept
ESA				ESA	0.2605	0.8524	0.0993
ISR				ISR	0.2896	0.8917	-0.0343
BONN				BONN	0.2648	0.9015	-0.0145
FN2				LTH	RMS	SLOPE	Intercept
ESA				ESA	0.1585	0.9041	0.0896
ISR				ISR	0.1852	1.0168	-0.0890
BONN				BONN	0.1555	1.0490	-0.0584

2.2.3 Wind Speed (U10): S3A

Wind speed data measurements at the three platforms FINO1, FINO2 and FINO3 have been used to evaluate the accuracy of the altimeter products. We observe in Figure 2.2.10 that the median of the RMS is very similar for the three products and is about 2 m/sec. The spread is also smaller for ESA and larger for UBonn.

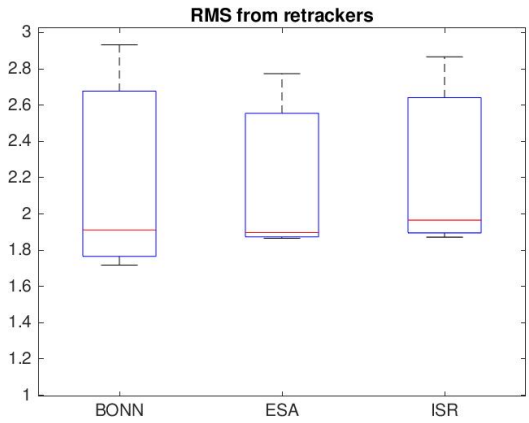


Figure 2.2.10. Boxplots of U10 RMS from S3A at 3 buoys

2.2.4 Sea Level Anomaly (SLA): CS2

The precision is first evaluated considering the SLA along-track difference at less than 20 km from the coast (Figure 2.2.11). Near coast, ISR has the highest noise and TUM retracker has the smallest number of data.

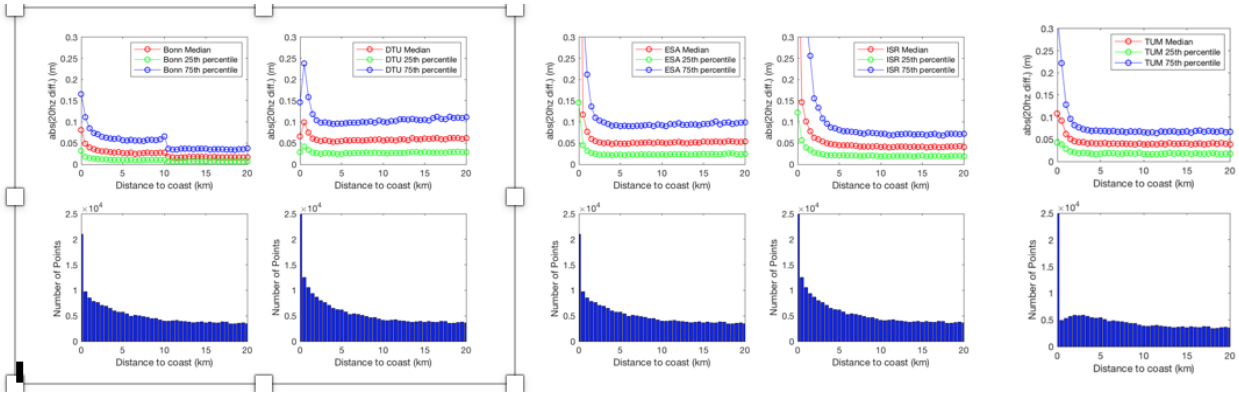


Figure 2.2.11. Noise (top) and number of observations (bottom) for 20 Hz CryoSat-2 (left to right in alphabetic order: Bonn-DTU-ESA-ISR-TUM). Time interval is from 2018-06-01 to 2020-06-30.

We select for CS2 using the following procedure: we have built time-series by averaging in the region between 2 and 20 km from coast the SLA uncorrected for DAC and the SWH. We then extract the corresponding point from the in-situ records and finally correlate the two time-series. Table 2.2.8 shows results for HELG without SSB correction, Table 2.2.9 shows results for the 8 stations with SSB corrections applied. The final performance of the retrackers are summarized in section 2.2.7.

Table 2.2.8. Statistics (STDD and correlation) of SLA for CS2 at HELG using no SSB correction

Retrackers	CS2 stdd	CS2 cor
Bonn	0.160	0.854
DTU	0.212	0.775
ESA	0.172	0.858
ISR	0.170	0.844
TUM	0.157	0.876

Table 2.2.9. Statistics (STDD and correlation) of SLA for CS2 at 8 tide gauge stations with SSB correction applied

SLA CS2 SSB Each					
BORJ	STD	CorrCoef	SASS	STD	CorrCoef
DTU	0.2815	0.7027	DTU	0.0581	0.9437
ESA	0.6078	0.7713	ESA	0.0812	0.9010
ISR	0.8592	0.3201	ISR	0.2118	0.1244
BONN	0.3470	0.4368	BONN	0.0561	0.9422
TUM	0.4223	0.6522	TUM	0.0589	0.9358
FLDW	STD	CorrCoef	SCHL	STD	CorrCoef
DTU	0.5753	0.6231	DTU	0.1611	0.7607
ESA	0.9091	0.0356	ESA	0.1346	0.7859
ISR	1.4177	-0.0970	ISR	0.1703	0.6571
BONN	0.6777	0.4460	BONN	0.1717	0.6061
TUM	0.6342	0.3220	TUM	0.1596	0.7268

HELG	STD	CorrCoef
DTU	0.1934	0.8397
ESA	0.1674	0.8925
ISR	0.1795	0.8656
BONN	0.1618	0.8805
TUM	0.1622	0.8846

TGME	STD	CorrCoef
DTU	0.4192	0.4947
ESA	0.7011	-0.1827
ISR	0.7769	-0.1072
BONN	0.4306	0.4440
TUM	0.5451	0.1846

LHAW	STD	CorrCoef
DTU	0.3885	0.5572
ESA	0.7473	-0.1080
ISR	0.7581	-0.3311
BONN	0.3735	0.5370
TUM	0.5569	0.2324

WARN	STD	CorrCoef
DTU	0.0643	0.8445
ESA	0.0738	0.7972
ISR	0.1206	0.5846
BONN	0.0829	0.8336
TUM	0.0606	0.8620

LTKI	STD	CorrCoef
DTU	0.1145	0.5924
ESA	0.1419	0.6345
ISR	0.1183	0.4886
BONN	0.1159	0.6862
TUM	0.1080	0.5408

We have also run the processing “overpass” used in Fenoglio et al. (2020). Figure 2.2.12 below shows the boxplots for both methodologies. BONN retrackers show in both methods the smallest deviation from the median.

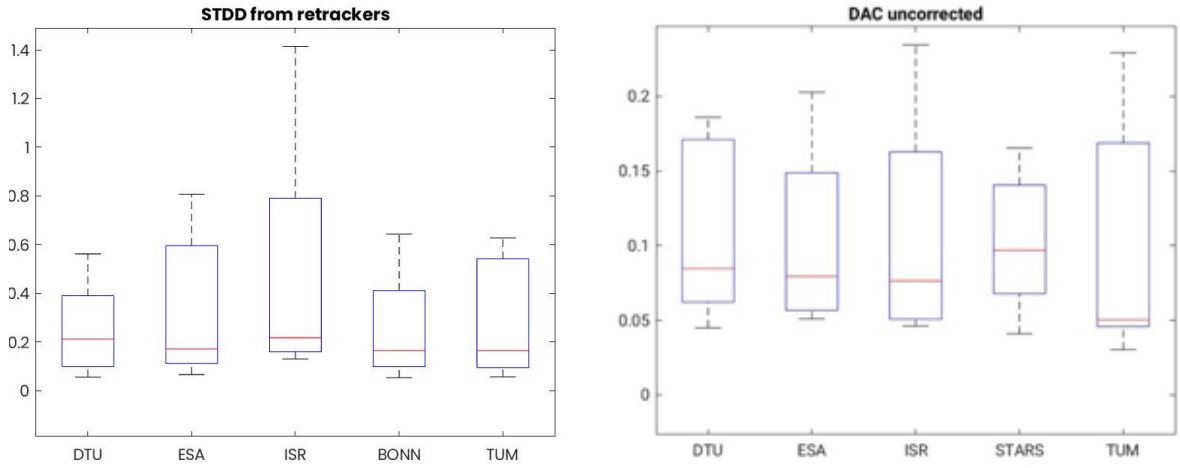


Figure 2.2.12. CS2 without DAC applied. Boxplot of selected methodology (left) and boxplot of overpass (right).

2.2.5 Significant Wave Height (SWH): CS2

Table 2.2.10 gives the statistics of the differences of SWH at Helgoland for all retrackers. The time series have been built by collecting data between 2 and 20 km and we correlate this time-series with the buoy. The performance of the three retrackers is different. The RMS (Figure 2.2.13) is higher for ESA and ISR (66 and 65 cm respectively) than for Bonn (37 cm).

Table 2.2.10. RMS, slope and intercept of SWH for CS2 at 8 buoys.

SWH CS2							
ARK	RMS	SLOPE	Intercept	FN3	RMS	SLOPE	Intercept
ESA	0.7590	-0.1842	1.1359	ESA	0.9096	-0.1362	1.4310
ISR	0.7478	0.8082	-0.6949	ISR	0.8355	0.5821	0.3485
BONN	0.3278	1.0261	-0.0675	BONN	0.8118	1.0035	0.7181
DAR	RMS	SLOPE	Intercept	FNO	RMS	SLOPE	Intercept
ESA	0.5124	0.1060	0.6597	ESA	0.5434	0.7946	0.0068
ISR	0.5715	0.8545	-0.2918	ISR	0.5212	0.8991	-0.2155
BONN	0.2416	1.0895	-0.0096	BONN	0.2981	1.0486	0.0209

ELB	RMS	SLOPE	Intercept
ESA	0.6670	0.5871	-0.0062
ISR	0.7822	0.3068	0.1691
BONN	0.4228	1.1706	-0.0426

HELG	RMS	SLOPE	Intercept
ESA	0.7238	0.4230	0.6368
ISR	0.5371	0.9192	-0.1472
BONN	0.2051	1.0959	-0.0978

FN2	RMS	SLOPE	Intercept
ESA	0.5376	0.2159	0.7495
ISR	0.6051	0.4398	0.1175
BONN	0.4009	1.0868	0.1630

LTH	RMS	SLOPE	Intercept
ESA	0.6448	0.8278	0.0554
ISR	0.5657	0.9600	-0.1931
BONN	0.2382	1.2480	-0.1355

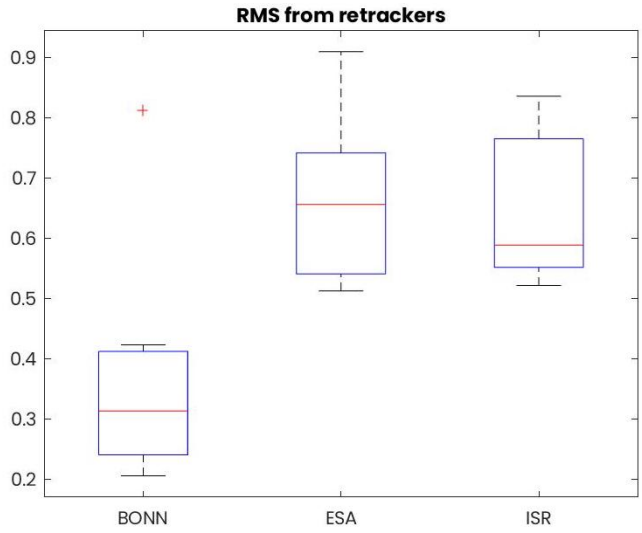


Figure 2.2.13. Boxplot of comparison of SWH RMS between CS2 and data at eight buoys

2.2.6 Wind Speed (U10): CS2

As for S3A, the wind speed at the same three platforms FINO1, FINO2 and FINO3 has been considered in evaluating the accuracy of the CS2 wind speed products. Also in this case the median of the RMS is similar for the three products and is between 6 and 7 m/s, which is larger than the 2.5 m/s obtained for S3A. The ISR results have smaller RMS, as seen in Figure 2.2.14.

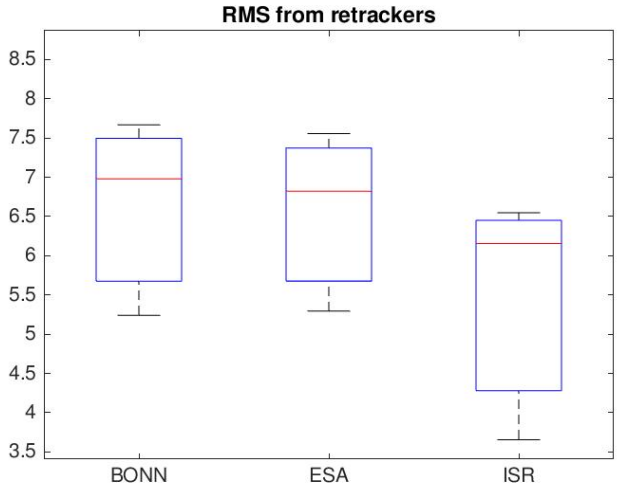


Figure 2.2.14. Boxplot of comparison of U10 RMS between CS2 and data at three in-situ stations

2.2.7 Final Summary Tables

Tables 2.2.11 to 2.2.18 summarize the performance according to the agreed metrics. The SSB applied is the 5% of SWH if the retracker has its own SWH (BONN and ISR) or a dedicated SSB (TUM).

- The SLA validation for S3A, S3B and CS2 is based on 9 tide gauges, S3A and S3B give similar results.
- For S3A SLA, BONN performs best with highest correlation and lowest STDD (12 cm), DTU follows (15 cm).
- For S3B, DTU and BONN have the best same STDD (13 cm).
- For CS2, DTU performs at best (25 cm), followed by Bonn (27 cm). SSB and no-SSB SLA show little difference.
- For S3A/S3B SWH is compared at 8/6 buoys. For S3A the lowest RMS is for UBonn, then for ISR and finally ESA. For S3B the smallest RMS is from ESA, ISR and Bonn follow.
- For CryoSat-2 SWH comparison at 8 buoys shows that U Bonn has the lowest RMS then ISR and ESA.
- For S3A/S3B U10, comparison at 3 stations gives a similar median (2 m/s) and small spread of RMS.
- For CryoSat-2 U10, comparison at 3 stations gives similar median (6 m/s) and larger spread of RMS.

Table 2.2.11. S3A SLA Altimeter v Tide Gauges, GB and EBaltic, Coast – 2-20km – with SSB

SLA S3A SSB own					
Order	Name	Mean of Correlation	Mean of STDD	STD of STDD	N stations Averaged
1	BONN	0.89	0.12	0.07	9
2	DTU	0.86	0.15	0.08	9
3	TUM	0.85	0.15	0.11	9
4	ESA	0.81	0.15	0.11	9
5	ISR	0.78	0.16	0.11	9

Table 2.2.12. S3B SLA Altimeter v Tide Gauges, GB and EBaltic, Coast – 2-20km – with SSB

SLA S3B SSB own					
Order	Name	Mean of Correlation	Mean of STDD	STD of STDD	N stations Averaged
1	DTU	0.86	0.13	0.10	8
2	BONN	0.83	0.13	0.08	8
3	TUM	0.76	0.16	0.18	8
4	ESA	0.75	0.17	0.19	8
5	ISR	0.73	0.17	0.17	8

Table 2.2.13. CS2 SLA Altimeter v Tide Gauges, GB and EBaltic, Coast – 2-20km – with SSB

SLA CS2 SSB own					
Order	Name	Mean of Correlation	Mean of STDD	STD of STDD	N stations Averaged
1	DTU	0.71	0.25	0.18	9
2	BONN	0.65	0.27	0.20	9
3	TUM	0.59	0.30	0.24	9
4	ESA	0.50	0.40	0.34	9
5	ISR	0.28	0.51	0.46	9

Table 2.2.14. S3A SWH Altimeter v buoys, GB and EBaltic, Coast – 2-20km

SWH S3A					
Order	Name	Mean of SLOPE	Mean of RMSE	STD(RMSE)	N stations Averaged
1	BONN	0.89	0.51	0.16	8
2	ISR	0.90	0.54	0.15	8
3	ESA	0.72	0.56	0.18	8

Table 2.2.15. S3B SWH Altimeter v buoys, GB and EBaltic, Coast – 2-20km

SWH S3B					
Order	Name	Mean of SLOPE	Mean of RMSE	STD(RMSE)	N stations Averaged
1	ESA	0.79	0.43	0.21	6
2	ISR	0.63	0.45	0.20	6
3	BONN	0.68	0.46	0.22	6

Table 2.2.16. CS2 SWH Altimeter v buoys, GB and EBaltic, Coast – 2-20km

SWH CS2					
Order	Name	Mean of SLOPE	Mean of RMSE	STD(RMSE)	N stations Averaged
1	BONN	1.10	0.37	0.20	8
2	ISR	0.72	0.65	0.12	8
3	ESA	0.33	0.66	0.13	8

Table 2.2.17. S3A U10 Altimeter v anemometers, GB and EBaltic, Coast – 2-20km

U10 S3A					
Order	Name	Mean of SLOPE	Mean of RMSE	STD(RMSE)	N stations Averaged
1	ESA	1.10	2.18	0.52	3
2	ISR	1.04	2.19	0.65	3
3	BONN	1.01	2.23	0.55	3

Table 2.2.18. CS2 U10 Altimeter v anemometers, GB and EBaltic, Coast – 2-20km

U10 CS2					
Order	Name	Mean of SLOPE	Mean of RMSE	STD(RMSE)	N stations Averaged
1	BONN	0.69	5.45	1.57	3
2	ISR	0.79	6.56	1.15	3
3	ESA	0.82	6.63	1.25	3

2.3 Validation in the Harvest region (NOC)

In this section, we present the validation of S3A/S3B/CS2 sea level anomaly/SWH/U10 data for the Harvest region against tide gauge and buoy observations. The analysis includes altimetry data from five retrackers, namely BONN, DTU, ESA, ISR and TUM (alphabetical order) and, in the case of the SLA, involves comparisons using different range and geophysical corrections.

2.3.1 Comparison of SLA against tide gauge data: S3A and S3B

The fields used to compute the along-track SLA were obtained from the input files (generated by IsardSat and available in the MEGA folder: s3_l2e_reduced):

- Altitude: 'alt'.
- Range: 'retracked_range_ESA'.
- Range corrections: 'GIM_iono', 'mod_dry_tropo_cor_zero_altitude', 'mod_wet_tropo_cor_zero_altitude' (these will be referred to as the DTC and WTC from Model 2). Also, 'upt_dry_tropo', 'gpd_wet_tropo' to test UPorto corrections.
- Geophysical corrections: 'ocean_tide_fes'/'ocean_tide_got', 'ocean_tide_non_eq', 'solid_earth_tide', and 'geocentric_polar_tide'.
- Mean Sea Surface: ‘MSS’.

- Significant Wave Height: 'swh_ESA' (ESA and DTU retrackers), ‘swh_BONN’ and ‘swh_isr’ to compute the SSB correction as -0.05*SWH. For the TUM retracker, we use the SSB correction supplied with the TUM files.

From the output files available in the MEGA directories corresponding to each retracker.

- Range: 'retracked_range_BONN', 'retracked_range_dtu', 'retracked_range_isr', 'retracked_range_tum'.

From the University of Porto outputs available in the MEGA directory: ‘s3_l2e_reduced_upt_geo’.

- Dry and Wet tropospheric corrections from the University of Porto: 'upt_dry_tropo' and 'gpd_wet_tropo'.

For this validation, we use data from four tide gauges, namely Crescent City, San Francisco, La Jolla, and San Diego. The tide gauge data have been obtained from the University of Hawaii Sea Level Center (https://uhslc.soest.hawaii.edu/datainfo/) and are provided as hourly values of relative sea level. The location of the tide gauges is shown in Fig. 2.3.1.

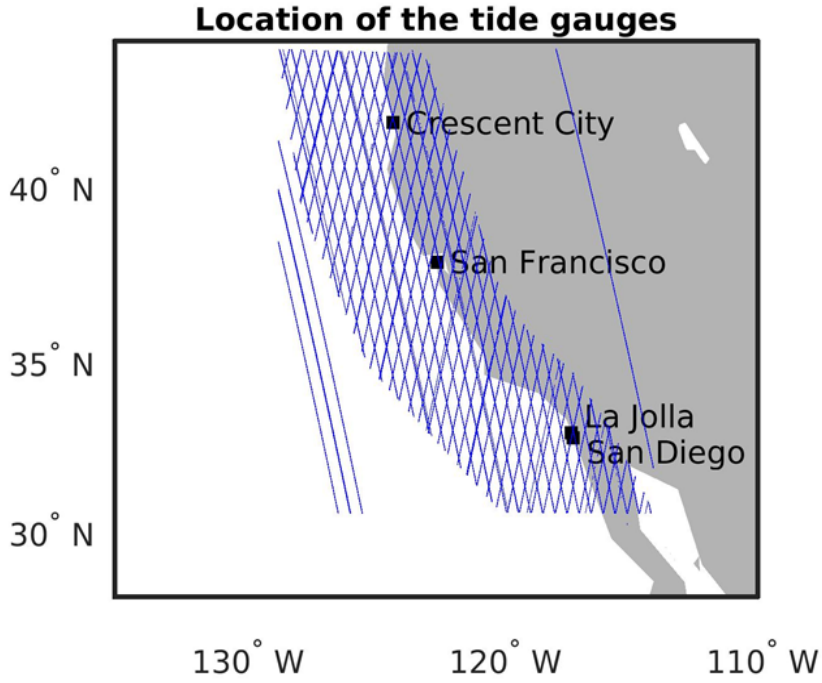


Figure 2.3.1. Location of the tide gauge stations used in the validation. The tracks for S3A/S3B are also shown in blue.

We begin by comparing the SLAs from S3A against the tide gauge data as a function of distance to the coast. We restrict this particular analysis to SLAs computed using the tidal correction from FES2014, the wet and dry tropospheric corrections (WTC and DTC) from the University of Porto and no SSB correction. We bin the altimetry data according to distances to the coast at intervals of one kilometer, thus generating one altimetry time series for each bin. Correlations with the tide gauge records as a function of distance to the coast are shown in Fig. 2.3.2. We note that, except for Crescent City, the agreement between the altimetry data and tide gauge measurements does not degrade in any significant way as the altimeter measurements get closer to the coast and even in the 0-1 km band the correlations are as high as at other farther distances. In Crescent City, the correlations show a decline from about 4 km of the coast, but otherwise they are comparable to those at other tide gauge stations. The

correlations show some variation across the distance bands, but they generally lie in the range (0.6,0.9). In the first 10 km from the coast, the data from the DTU and TUM retrackers tend to show higher correlations than the other retrackers, whereas the BONN retracker shows the worst performance, particularly at San Francisco and La Jolla.

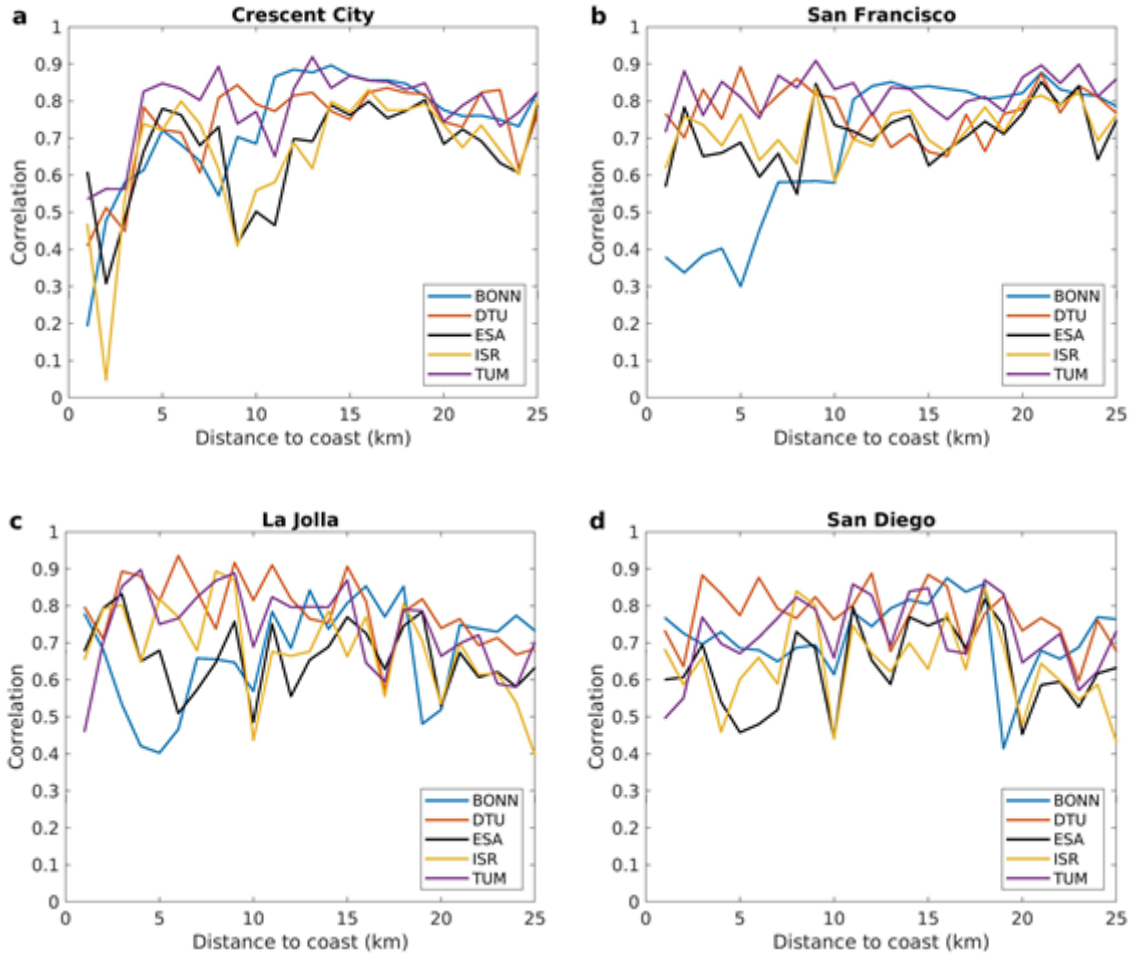


Figure 2.3.2. Correlation of SLAs from S3A (FES2014, DTC/WTC from UPorto, and no SSB correction) with tide gauge data as a function of distance to the coast at (a) Crescent City, (b) San Francisco, (c) La Jolla and (d) San Diego. The numbers along the x-axis denote the right edge of the distance bins (e.g., 5 km corresponds to the bin 4-5 km).

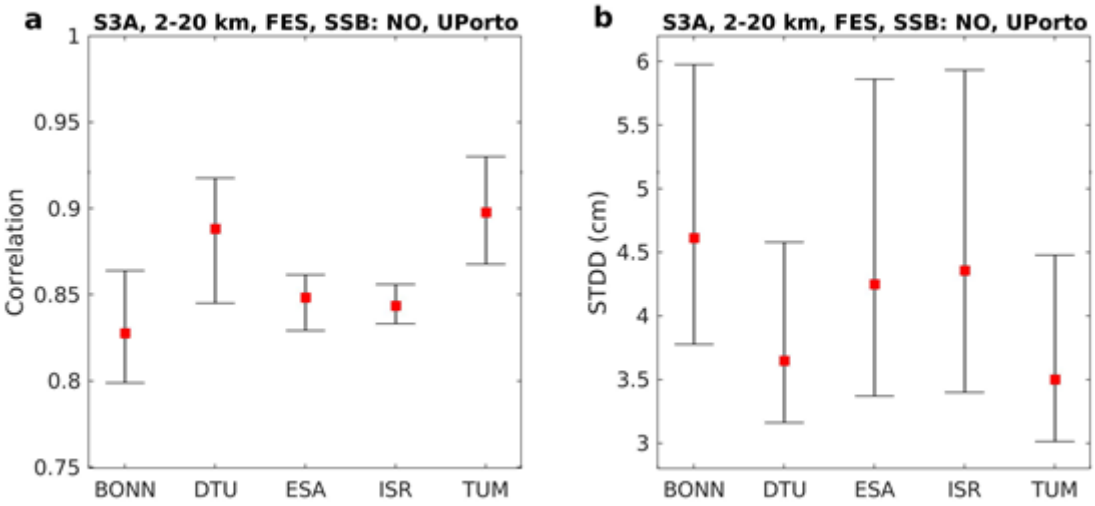
The relatively large spread of correlations across the 1-km bins (Fig. 2.3.2) primarily reflects the influence of measurement error and small-scale noise on the altimetry data. Such spread makes it challenging to assess the actual performance of the altimetry data and how such performance varies across retrackers and geophysical corrections. This challenge can be addressed, at least partly, by averaging the altimetry data over wider distance bins since this allows for a reduction in the overall noise of the altimetry measurements. To this aim, we choose to average the altimetry data within a distance of 2-20 km from the coast. The resulting altimetry time series are then compared with the corresponding tide gauge data.

We start by discussing the results for the Sentinel-3A data processed using the tidal correction from FES2014, the WTC/DTC from the University of Porto and no SSB correction (Table 2.3.1). Correlations range from 0.80 to 0.93 while the STDDs range from 3.0 to 6.0 cm across tide gauge stations and retrackers, indicating a very good agreement with the tide gauge data. The best-performing retrackers are TUM and DTU, as indicated by the highest

correlations and lowest STDDs achieved by those retrackers, followed by the ESA and ISR retrackers, which both show a similar performance. The BONN retracker still shows a good agreement with the tide gauge data, but overall such agreement is not as good as for the other retrackers. The average correlation (STDD) over the four tide gauges for BONN, DTU, ESA, ISR and TUM is (Fig. 2.3.3a,b), respectively, 0.83 (4.6 cm), 0.89 (3.7 cm), 0.85 (4.4 cm), 0.84 (4.3 cm) and 0.90 (3.5 cm).

Table 2.3.1. Correlation and standard deviation of the differences (STDD) between the SLAs from S3A (FES2014, DTC/WTC from UPorto, and no SSB correction) averaged over a distance of 2-20 km to the coast with tide gauge data. The highest correlation(s) at each tide gauge are highlighted in red.

	SENTINEL 3A **Distance to coast: 2-20 km** **Tidal model: FES; WTC/DTC: UPorto; SSB: No** Correlation (STDD in cm)				
Tide gauge	BONN	DTU	ESA	ISR	TUM
Crescent City	0.86 (6.0)	0.92 (4.6)	0.86 (5.9)	0.86 (5.9)	0.93 (4.5)
San Francisco	0.85 (4.6)	0.91 (3.6)	0.85 (4.5)	0.85 (4.3)	0.91 (3.5)
La Jolla	0.80 (3.8)	0.85 (3.3)	0.83 (3.4)	0.83 (3.4)	0.87 (3.0)
San Diego	0.80 (4.1)	0.88 (3.2)	0.84 (3.6)	0.83 (3.4)	0.88 (3.0)



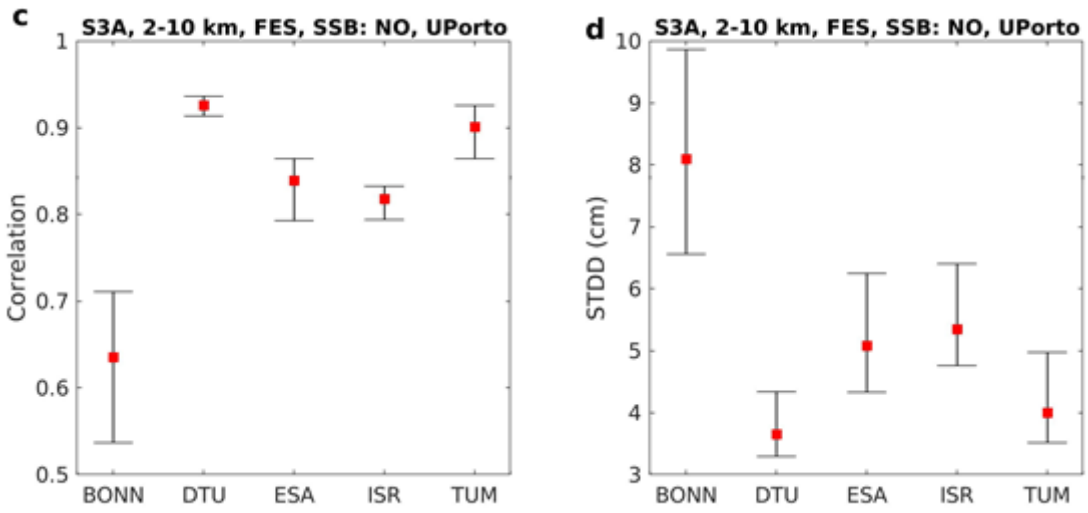


Figure 2.3.3. Correlation and STDD between SLAs from S3A (FES2014, DTC/WTC from UPorto, and no SSB correction) and tide gauge data for SLAs averaged over (a,b) 2-20 km and (c,d) 2-10 km from the coast. Red squares denote average over the four tide gauges while the whiskers represent the range of values.

For completeness, we have repeated the same comparison as in the preceding paragraph but for altimetry data averaged within a distance of 2-10 km from the coast. This is again done using the Sentinel-3A data processed using the tidal correction from FES2014, the WTC/DTC from the University of Porto and no SSB correction. The correlations and STDDs at each tide gauge are summarized in Table 2.3.2 while the diagnostics averaged over all tide gauges are shown in Fig. 2.3.3c,d. In this case, DTU is the best-performing retracker at all tide gauge stations, with correlations ranging from 0.91 to 0.94 and STDDs ranging from 3.3 cm to 4.3 cm. DTU is the only retracker that gives higher correlations for the 2-10 km distance range than for the 2-20 km range at all stations, indicating that DTU is the best retracker close to the coast and also that its performance does not degrade closer to the coast, at least between 2 and 10 km of the coast. The TUM retracker performs almost as well as DTU, with correlations ranging from 0.86 to 0.93. The ESA and ISR retrackers also show a good performance while the BONN retracker shows a moderate performance. The average correlation (STDD) over the four tide gauges for BONN, DTU, ESA, ISR and TUM is (Fig. 2.3.3c,d), respectively, 0.64 (8.1 cm), 0.93 (3.7 cm), 0.84 (5.1 cm), 0.82 (5.3 cm) and 0.90 (4.0 cm).

Table 2.3.2. Correlation and standard deviation of the differences (STDD) between the SLAs from S3A (FES2014, DTC/WTC from UPorto, and no SSB correction) averaged over a distance of 2-10 km to the coast with tide gauge data. The highest correlation(s) at each tide gauge are highlighted in red.

	SENTINEL 3A **Distance to coast: 2-10 km** **Tidal model: FES; WTC/DTC: UPorto; SSB: No** Correlation (STDD in cm)				
Tide gauge	BONN	DTU	ESA	ISR	TUM
Crescent City	0.68 (9.9)	0.93 (4.3)	0.84 (6.3)	0.83 (6.4)	0.91 (5.0)
San Francisco	0.54 (9.0)	0.93 (3.5)	0.86 (4.3)	0.83 (4.8)	0.91 (3.7)
La Jolla	0.61 (6.6)	0.94 (3.5)	0.86 (4.7)	0.81 (5.4)	0.93 (3.5)
San Diego	0.71 (7.0)	0.91 (3.3)	0.79 (5.0)	0.79 (4.8)	0.86 (3.8)

Next, we present the results for the Sentinel-3B data processed using the tidal correction from FES2014, the WTC/DTC from the University of Porto and no SSB correction, and averaged over the 2-20 km distance range (Table 2.3.3 and Fig. 2.3.4). The first thing we note is that, overall, the agreement with tide gauge data is worse than for Sentinel-3A, with average correlations (STDDs) over the four tide gauges (Fig. 2.3.4a,b) of 0.73 (6.5 cm), 0.73 (6.8 cm), 0.71 (7.9 cm), 0.69 (7.4 cm) and 0.62 (8.8 cm) for the BONN, DTU, ESA, ISR and TUM retrackers, respectively. In this case, the BONN and DTU retrackers achieve the best agreement with the tide gauge data, although ESA and ISR show nearly the same performance, while TUM shows the worst agreement. In the case of TUM, the agreement with tide gauge data is particularly poor at Crescent City and San Francisco (See Table 2.3.3).

Table 2.3.3. Correlation and standard deviation of the differences (STDD) between the SLAs from S3B (FES2014, DTC/WTC from UPorto, and no SSB correction) averaged over a distance of 2-20 km to the coast with tide gauge data. The highest correlation(s) at each tide gauge are highlighted in red.

	SENTINEL 3B **Distance to coast: 2-20 km** **Tidal model: FES; WTC/DTC: UPorto; SSB: No** Correlation (STDD in cm)				
Tide gauge	BONN	DTU	ESA	ISR	TUM
Crescent City	0.91 (5.4)	0.78 (10.2)	0.89 (5.4)	0.89 (5.4)	0.64 (12.6)
San Francisco	0.61 (10.5)	0.62 (8.8)	0.36 (18.0)	0.34 (15.6)	0.28 (14.6)
La Jolla	0.76 (4.5)	0.81 (3.6)	0.85 (3.4)	0.82 (3.6)	0.84 (3.3)
San Diego	0.64 (5.6)	0.73 (4.6)	0.74 (4.7)	0.70 (4.9)	0.73 (4.6)

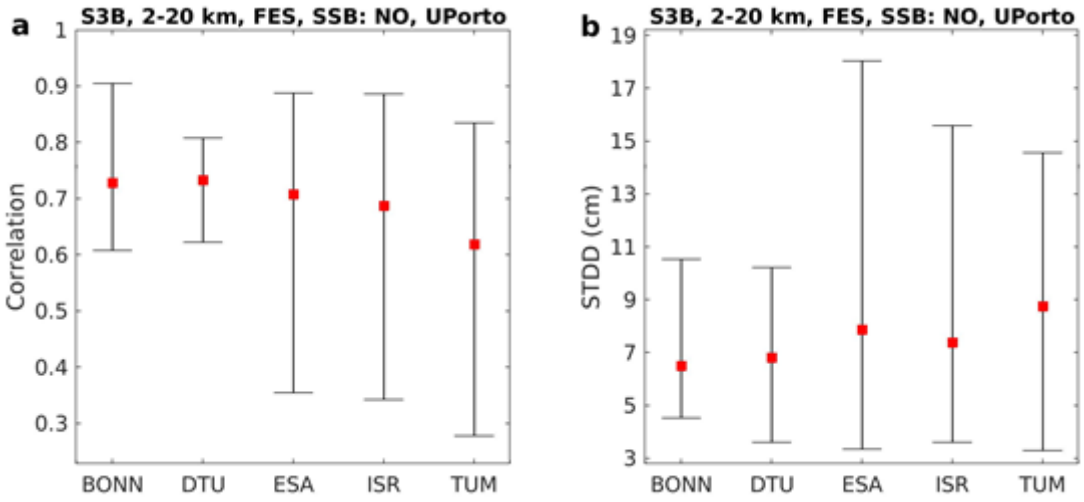


Figure 2.3.4. (a) Correlation and (b) STDD between SLAs from S3B (FES2014, DTC/WTC from UPorto, and no SSB correction) and tide gauge data for SLAs averaged over 2-20 km. Red squares denote average over the four tide gauges while the whiskers represent the range of values.

In the following, we extend the comparison with tide gauge data to altimetry data with different geophysical and range corrections applied. In particular, we repeat the comparison for the following cases: 1) SSB correction applied (along with FES2014, and the WTC/DTC from the University of Porto); 2) WTC/DTC from Model 2 (along with FES2014 and no SSB correction applied); and 3) GOT4.10 tidal correction (along with WTC/DTC from the University of Porto and no SSB correction applied). This analysis is only done for the Sentinel-3A data since the aim is simply to compare the performance of the different corrections by comparison with the results presented earlier in this section (tidal correction from FES2014, the WTC/DTC from the University of Porto and no SSB correction).

The results of this comparison are presented in Fig. 2.3.5. Our results show that applying the SSB correction leads to lower STDDs for all retrackers except DTU (Fig. 2.3.5a), suggesting that the SSB correction is well estimated and applying it improves the quality of the data. The average correlation (STDD) over the four tide gauges for BONN, DTU, ESA, ISR and TUM when applying the SSB correction is, respectively, 0.86 (4.3 cm), 0.85 (4.6 cm), 0.89 (3.4 cm), 0.89 (3.6 cm) and 0.90 (3.4 cm). For the ESA and ISR retrackers the STDD is reduced by nearly 1 cm when applying the SSB correction, indicating a significant improvement. The DTU retracker is the only retracker for which the SSB correction was estimated based on the SWH from a different retracker (i.e. from the ESA retracker), which is probably the reason that the DTU retracker showed no improvement after applying the SSB correction, contrary to the other retrackers. In short, applying the SSB correction generally improves the quality of the data, but such correction needs to be derived using SWH derived from the same retracker.

Regarding the WTC/DTC (Fig. 2.3.5b), we find that using the corrections from the University of Porto leads to lower STDDs for three of the retrackers (BONN, DTU and ESA) and the same STDDs for the other two retrackers (ISR and TUM). However, the differences in terms of STDD are small (~ 0.1 cm), indicating that the WTC/DTC from the University of Porto only slightly outperform those from Model 2. Finally, we find that applying the tidal correction from FES2014 leads to lower STDDs for all retrackers, indicating a better performance than GOT4.10. FES2014 reduces the STDD by as much as 0.3 cm.

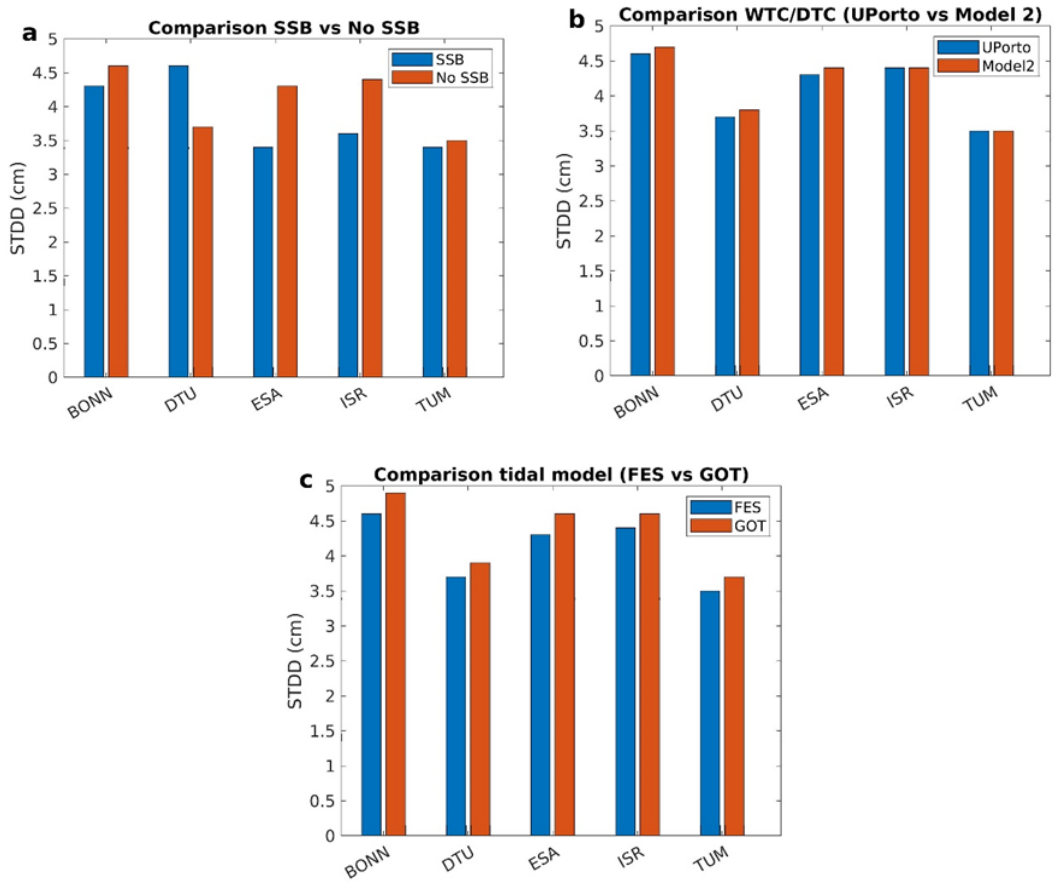


Figure 2.3.5. Average STDD over the four tide gauges for Sentinel-3A data for the cases: (a) SSB correction vs no SSB correction (both corrected using FES2014 and the WTC/DTC from the University of Porto); (b) WTC/DTC from the University of Porto vs that from Model 2 (both corrected using FES2014 and without the SSB correction applied); and (c) FES2014 vs GOT4.10 (both corrected using the WTC/DTC from the University of Porto and without the SSB correction applied).

2.3.2 Comparison of SLA against tide gauge data: CryoSat-2

In this section, we present the results of the comparison of data from CryoSat-2 with tide gauge observations. This comparison is only possible for the San Francisco tide gauge since the CryoSat-2 data are not available at the other tide gauge stations. The procedure to derive SLAs from CryoSat-2 data is the same as described in section 2.3.1 for Sentinel-3A/3B. We note that there is no valid data from the TUM retracker, and thus results for this retracker are not included here.

We begin by discussing the results for the data processed using the tidal correction from FES2014, the WTC/DTC from the University of Porto and no SSB correction. The correlation (STDD) of CryoSat-2 data with observations from the San Francisco tide gauge for the BONN, DTU, ESA, and ISR is, respectively, 0.71 (7.3 cm), 0.75 (6.6 cm), 0.80 (6.1 cm), 0.71 (6.8 cm). Hence, the ESA retracker gives the best agreement with the tide gauge data, followed by the DTU retracker. We also note that the agreement is lower than for Sentinel-3A for all retrackers, with both lower correlations and higher STDDs, but better than for Sentinel-3B.

For completeness, we have repeated the comparison using the WTC/DTC from Model 2 (using the tidal correction from FES2014 and no SSB correction). In this case, we find that the correlation (STDD) with data from the San

Francisco tide gauge for the BONN, DTU, ESA, and ISR is, respectively, 0.70 (7.2 cm), 0.74 (6.6 cm), 0.79 (6.2 cm), 0.70 (6.7 cm). Correlations are slightly lower than when using the WTC/DTC from the University of Porto for the BONN, DTU and ESA retrackers and the STDDs are only higher for the ESA retracker. These inconclusive results suggest that there is no significant difference in performance between the two sets of WTC/DTCs, which is consistent with the results presented in section 2.3.1 for the Sentinel-3 data.

2.3.3 Comparison of SWH against buoy data

In this section we present the results of the comparison of Sentinel-3A/3B and CryoSat-2 SWH data with observations from buoys. The location of the buoys used for this comparison is shown in Fig. 2.3.6 (only buoys farther than 2 km from the coast are used). This comparison is restricted to the BONN, ESA and ISR retrackers since no SWH data is provided by the DTU and TUM retrackers.

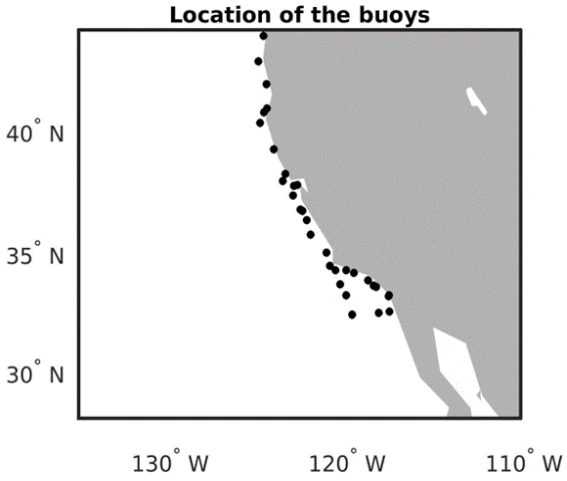
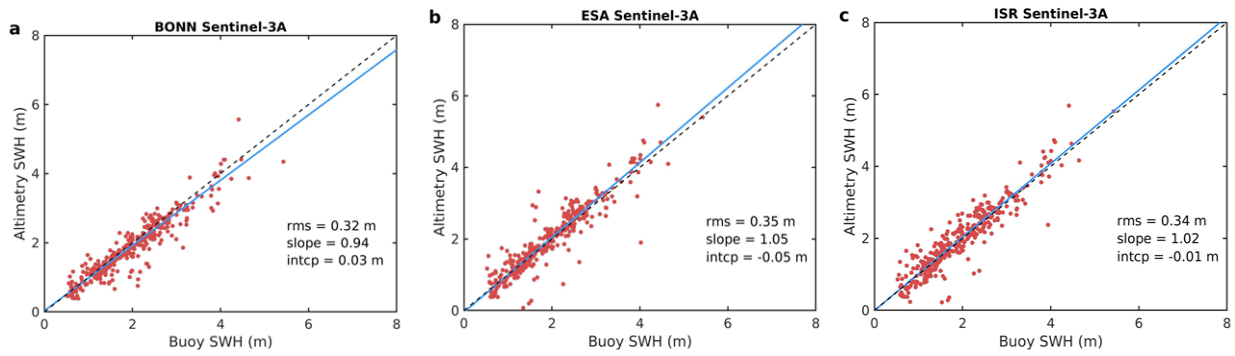


Figure 2.3.6. Location of buoys with available SWH/U10 data in our study domain.

The scatter plots (Fig. 2.3.7) for the three altimetry data sets (i.e., Sentinel-3A, Sentinel-3B and CryoSat-2) and for the three retrackers (i.e., BONN, ESA and ISR) show similar patterns, with most points falling along the 45-degree line, indicating a good agreement between altimetric and buoy observations and small offset or bias. The RMS error ranges from 0.27 m (CryoSat-2 BONN) to 0.47 (CryoSat-2 ISR) whereas the slope ranges from 0.93 (BONN Sentinel-3B and CryoSat-2) to 1.05 (Sentinel-3A ESA). While the agreement can be regarded as very good, all altimetry data sets and retrackers (except for Sentinel-3A ESA) show a tendency to underestimate the SWH relative to the buoy observations, though such underestimation is small.



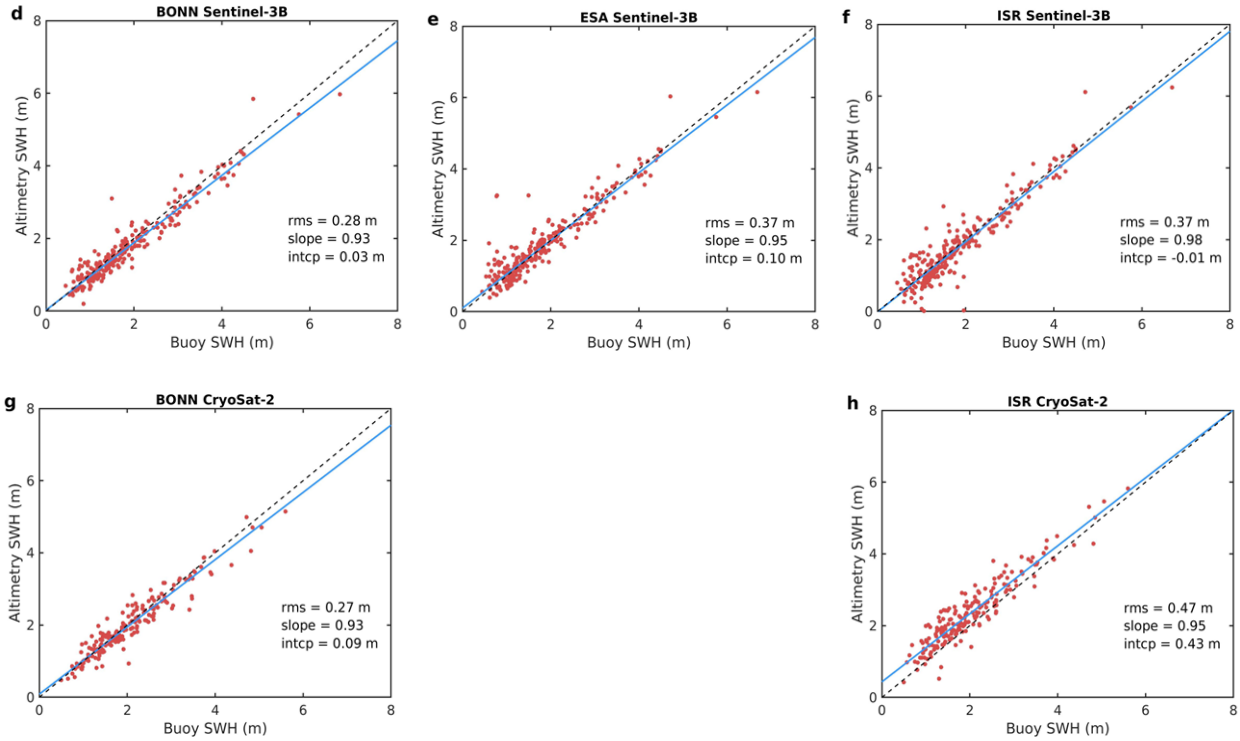


Figure 2.3.7. Scatter plot of the SWH from Sentinel-3A (a, b, c), Sentinel-3B (d, e, f) and CryoSat-2 (g, h) for the BONN, ESA and ISR retrackers against buoy observations. The blue line represents the result of a linear regression of the data (using robust regression) whereas the inset values denote the RMS error (rms), the slope, and the intercept (intcp) derived from the regression.

2.3.4 Comparison of 10-m wind (U10) against buoy data

In this section we present the results of the comparison of Sentinel-3A/3B U10 data with observations from buoys (no data from CryoSat-2 are available to compute wind). The output files of the retrackers do not include U10 as a field. Instead, they give sigma0. In order to compute U10, we add the atmospheric attenuation to the uncorrected sigma0 and then we estimate U10 following Abdalla's algorithm. The location of the buoys used for this comparison is the same as for the SWH comparison (Fig. 2.3.6). This comparison is restricted to the BONN, ESA and ISR retrackers since no U10 data is provided by the DTU and TUM retrackers.

The scatter plots (Fig. 2.3.8) for the BONN and ESA retrackers show similar patterns for Sentinel-3A and Sentinel-3B, with most points falling along the 45-degree line, indicating a good agreement between the altimetric and buoy observations and small offset or bias. The RMS error for these two retrackers ranges from 1.29 m/s (Sentinel-3B ESA) to 1.67 m/s (Sentinel-3A BONN) whereas the slope ranges from 0.84 (Sentinel-3B ESA) to 0.90 (Sentinel-3A BONN). The ISR retracker shows a good performance for Sentinel-3A, with a slope of 1.01 and an RMS error of 1.83 m/s, however for Sentinel-3B the ISR retracker shows a severe underestimation of U10 compared to the buoy data, with a slope 0.36 and an RMS error of 3.70 m/s. While the agreement can be regarded as good (except for Sentinel-3B ISR), all altimetry data sets and retrackers (except for Sentinel-3A ISR) show a tendency to underestimate the U10 relative to the buoy observations, though such underestimation is small.

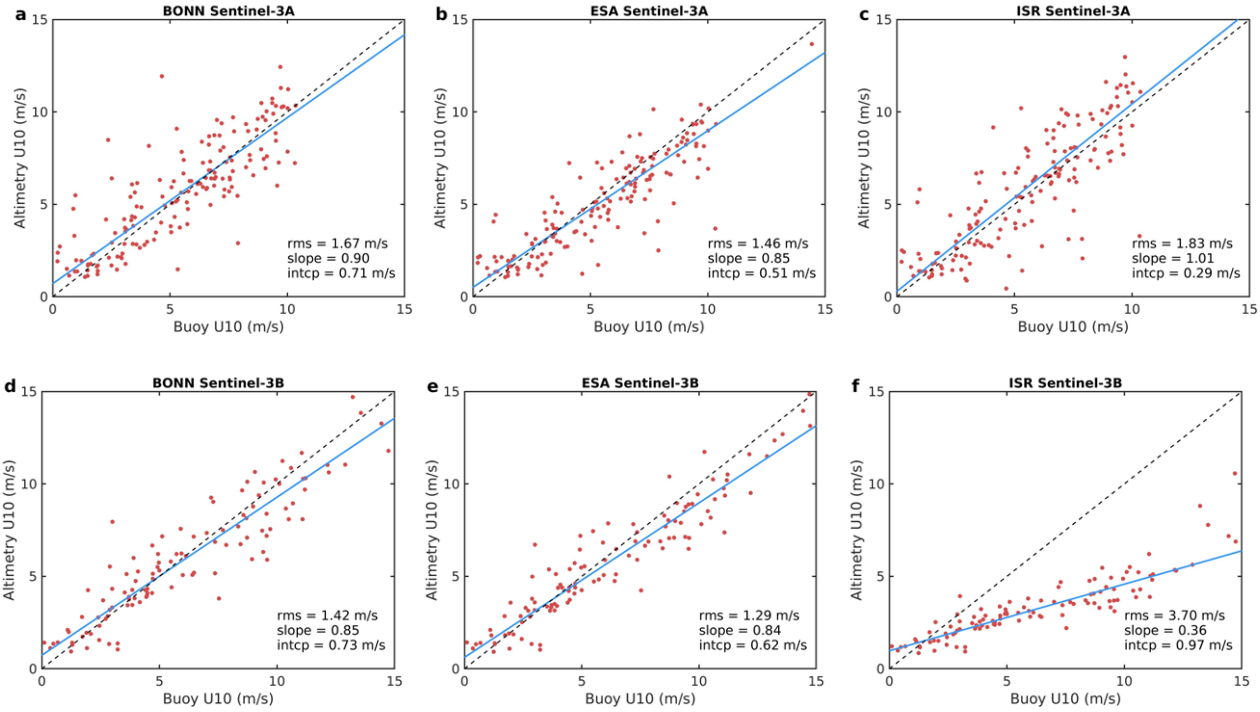


Figure 2.3.8. Scatter plot of the U10 from Sentinel-3A (a, b, c) and Sentinel-3B (d, e, f) for the BONN, ESA and ISR retrackers against buoy observations. The blue line represents the result of a linear regression of the data (using robust regression) whereas the inset values denote the RMS error (rms), the slope, and the intercept (intcp) derived from the regression.

2.3.5 Final Summary Tables

Below we give summary tables giving order of performances according to the agreed metrics.

Table 2.3.4. Sentinel 3A SLA Altimeter v Tide Gauges, California Coast – 2-20km – SSB applied**

Order	Name	Correlation	STDD* (cm)
1	TUM	0.90	3.4±0.4
2	ESA	0.89	3.6±0.3
3	ISR	0.89	3.6±0.3
4	DTU	0.89	3.7±0.4
5	U Bonn	0.85	4.6±1.1

* The ± ranges represent ±1 Mean Absolute Deviation
** The SSB has not been applied to DTU

Table 2.3.5. Sentinel 3B SLA Altimeter v Tide Gauges, California Coast – 2-20km – SSB applied**

Order	Name	Correlation	STDD* (cm)
1	U Bonn	0.74	6.2±1.7
2	DTU	0.73	6.8±2.7
3	ESA	0.72	7.9±4.2
4	ISR	0.69	7.5±2.6
5	TUM	0.62	8.9±5.0

* The ± ranges represent ±1 Mean Absolute Deviation
** The SSB has not been applied to DTU

Table 2.3.6. Cryosat-2 SLA Altimeter v San Francisco TG – 2-20km – SSB not applied

Order	Name	Correlation	STDD* (cm)
1	ESA	0.80	6.1
2	DTU	0.75	6.6
3	ISR	0.71	6.8
4	U Bonn	0.71	7.3
5	TUM	NA	7.2

* MAD is not provided since results for C2 are based on only 1 tide gauge.

Table 2.3.7. Sentinel 3A / 3B SWH, altimeter v buoys

Order	Name	RMS (m)	Slope	intcp
1	U Bonn	0.32 / 0.28	0.94 /0.93	0.03 / 0.03
2	ISR	0.34 / 0.37	1.02 / 0.98	-0.01 / -0.01
3	ESA	0.35/ 0.37	1.05 / 0.95	-0.05 / 0.10

Table 2.3.8. Cryosat-2 SWH, altimeter v buoys

Order	Name	RMS (m)	Slope	intcp
1	U Bonn	0.27	0.93	0.09
2	ISR	0.47	0.95	0.43

Table 2.3.9. Sentinel 3A / 3B U10, altimeter v buoys

Order	Name	RMS (m/s)	Slope	intcp
1	ESA	1.46 / 1.29	0.85 / 0.84	0.51 / 0.62
2	U Bonn	1.67 / 1.42	0.90 /0.85	0.71 / 0.73
3	ISR	1.83 / 3.70	1.01 / 0.36	0.29 / 0.97

- The SLA validation for S3A and S3B is based on 4 tide gauges.
- Different results were found for SLA for S3A and S3B, with the SLA from S3B data showing lower correlations and higher STDD. The reason for this is not yet understood.
- For S3A SLA, TUM showed the “best” performance with the highest correlation and lowest STDD. The performance of ISR, DTU and the original ESA data was very similar with STDDs within the range of the Mean Absolute Differences. U Bonn showed the lowest correlation and highest STDD, but with a larger range of Mean Absolute Differences (no doubt due to the lower correlation with TG data from San Francisco and La Jolla).
- For S3B SLA, U Bonn showed the “best” performance with the highest correlation and lowest STDD. The performance of DTU, the original ESA data, and isardSAT was similar with STDDs within the range of the Mean Absolute Differences. This time, in contrast to S3A, TUM showed the lowest correlation and highest STDD, but with a larger range of Mean Absolute Differences.
- For Cryosat-2 SLA, with comparisons against only 1 tide gauge (so no MAD), the ESA data showed the highest correlation and lowest STDD, followed by DTU, ISR, and U Bonn. No results were retrieved for TUM.
- For S3A and S3B SWH (comparison against 21 buoys), U Bonn showed the lowest RMS, followed by ISR and the original ESA data.
- For CryoSat-2 SWH (comparison against 18 buoys), U Bonn showed a significantly lower RMS than ISR. SWH values were not available from the ESA data
- For S3A and S3B U10 (comparison against 11 buoys), the original ESA data showed the lowest RMS, followed by U Bonn and ISR

In conclusion, there was no clear pattern of one retracker performing better than the others in comparing Sea Level Anomaly data against tide gauge data, with a different order of results across the three satellite data sets (Sentinel-3A, Sentinel-3B and CryoSat-2).

For S3A and S3B SWH data, U Bonn data were found to give lower RMS values than ISR or ESA. Similarly for CryoSat-SWH, U Bonn data showed a lower RMS than ISR.

For S3A and S3B U10 data, the ESA data were found to give lower RMS values than UBonn or ISR, with ISR showing much higher RMS values for S3B.

2.4 Validation in the Gulf of Cadiz and Strait of Gibraltar regions (U Cadiz)

In this section the validation of S3A/S3B/CS2 sea level anomaly/SWH/U10 data is presented in two areas of the Iberian Peninsula: Gulf of Cadiz (Huelva and Bonanza TGs) and Strait of Gibraltar (Tarifa TG). We divided the analysis made with the 5 retrackers: BONN, DTU, ESA, ISR, and TUM (following an alphabetical order) in six subsections, three for S3A/S3B (SLA, SWH, and U10) and three for CS2.

2.4.1 Sea Level Anomaly (SLA): S3A and S3B

The fields used to compute the along-track SLA were obtained from the input files (generated by IsardSat and available in the MEGA folder: s3_l2e_reduced):

- Altitude: 'alt'.
- Range: 'retracked_range_ESA'.
- Range corrections: 'GIM_iono', 'mod_dry_tropo_cor_zero_altitude', 'mod_wet_tropo_cor_zero_altitude'. Also, 'upt_dry_tropo', 'gpd_wet_tropo' to test UPorto corrections.
- Geophysical corrections: 'ocean_tide_fes'/'ocean_tide_got', 'ocean_tide_non_eq', 'solid_earth_tide', and 'geocentric_polar_tide'.
- Mean Sea Surface: 'MSS'.
- Significant Wave Height: 'swh_ESA' to compute the SSB correction (DTU, ESA retrackers).

From the output files available in the MEGA directories corresponding to each retracker.

- Range: 'retracked_range_BONN', 'retracked_range_dtu', 'retracked_range_isr', 'retracked_range_tum'.

From the University of Porto outputs available in the MEGA directory: 's3_l2e_reduced_upt_geo'.

- Dry and Wet tropospheric corrections from the University of Porto: 'upt_dry_tropo' and 'gpd_wet_tropo'.

2.4.1.1 Selection of tidal model

Two global tidal models were analysed: FES2014 and GOT4.10. The tidal elevation to be removed to the tide gauge data were computed using the constituents as in the global tidal models. To do this, the t_tide software was used.

For FES2014 two analyses were made:

- Tidal elevation (TGs) using all the constituents available in FES2014: K1, M2, M4, N2, O1, P1, Q1, S1, S2, K2, 2N2, E2 (EPS2), J1, L2, La2 (LDA2), M3, M6, M8, Mf, MKS2, Mm, MN4, MS4, MSf, MSqm, Mtm, Mu2, Nu2, R2, S4, Sa, Ssa, T2. N4 was not available.

- Tidal elevation (TGs) using the main constituents (amplitudes > 1 cm). 19 constituents for Huelva and Bonanza and 20 for Tarifa TG stations.

For GOT4.10:

- Tidal elevation: 10 constituents were used: Q1, O1, S1, K1, N2, M2, S2, K2, M4, P1.

We obtained time series from S3A / S3B (SLA_Sat) and from the tide gauges (SLA_TG). The along-track standard deviation of the differences (sdd) and Pearson's correlation coefficient (r) were estimated for all the tracks (S3A:

#114, #322, #051, #171; S3B: #114, #379, #051, #171), using three tide gauges: Huelva, Bonanza, and Tarifa. The location of the tide gauges (red dot) and track segments ([0 – 20] km) are given in Figure 2.4.1.

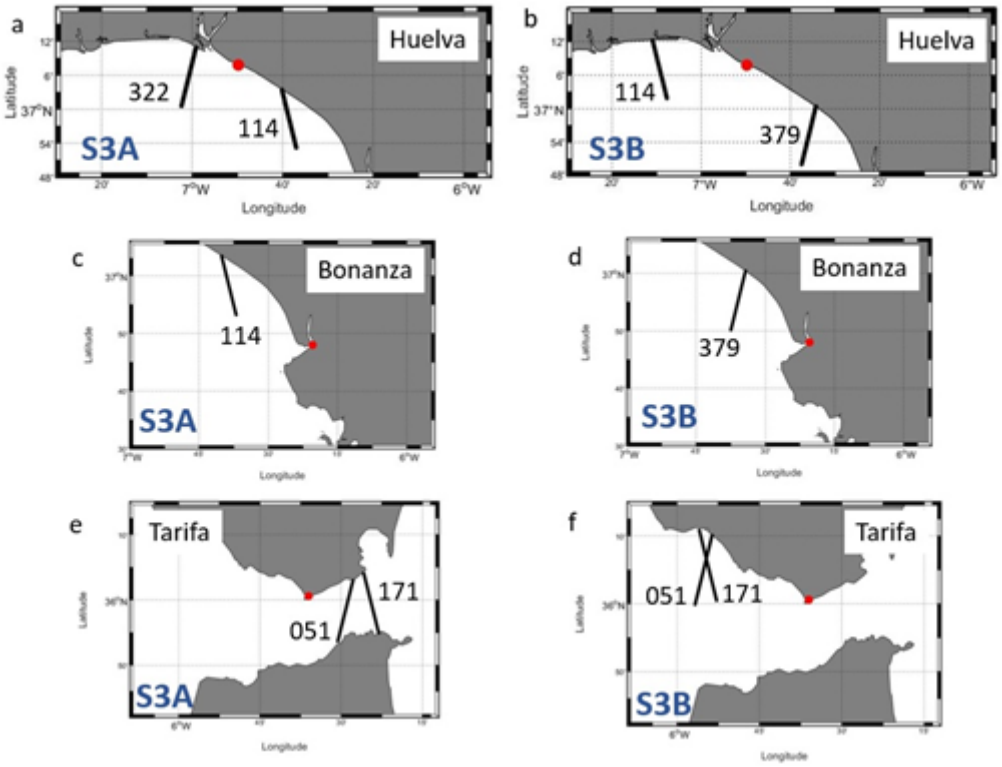


Figure 2.4.1. Red dots: Tide gauge locations: Huelva (2.4.1.a, 2.4.1.b), Bonanza (2.4.1.c, 2.4.1.d) and Tarifa (2.4.1.e, 2.4.1.f). The track segments of 20 km long are also shown for S3A and S3B.

The along-track percentage of valid data (%VD) (after screening), sdd and r is shown in Figure 2.4.2 for track #114 (S3A) using BONN, DTU, ESA (Figure 2.4.2.a, 2.4.2.c, 2.4.2.e) and ISR, TUM (Figure 2.4.2.b, 2.4.2.d, 2.4.2.f). The tide gauge used for comparison was Huelva. The tidal model applied to altimeter data was FES2014, no SSB was applied and dry/wet tropospheric corrections were those available in the input files. The tidal elevation from the tide gauge was removed using the same constituents as in the tidal model (all). The %VD is quite high in all the retrackers analysed (with the exception of TUM in the [2 – 4] km). The standard deviation of the difference is below 10 cm in the [2 -20] km segment and quite similar in all the cases. The same can be said for r, with high values in that segment of the track.

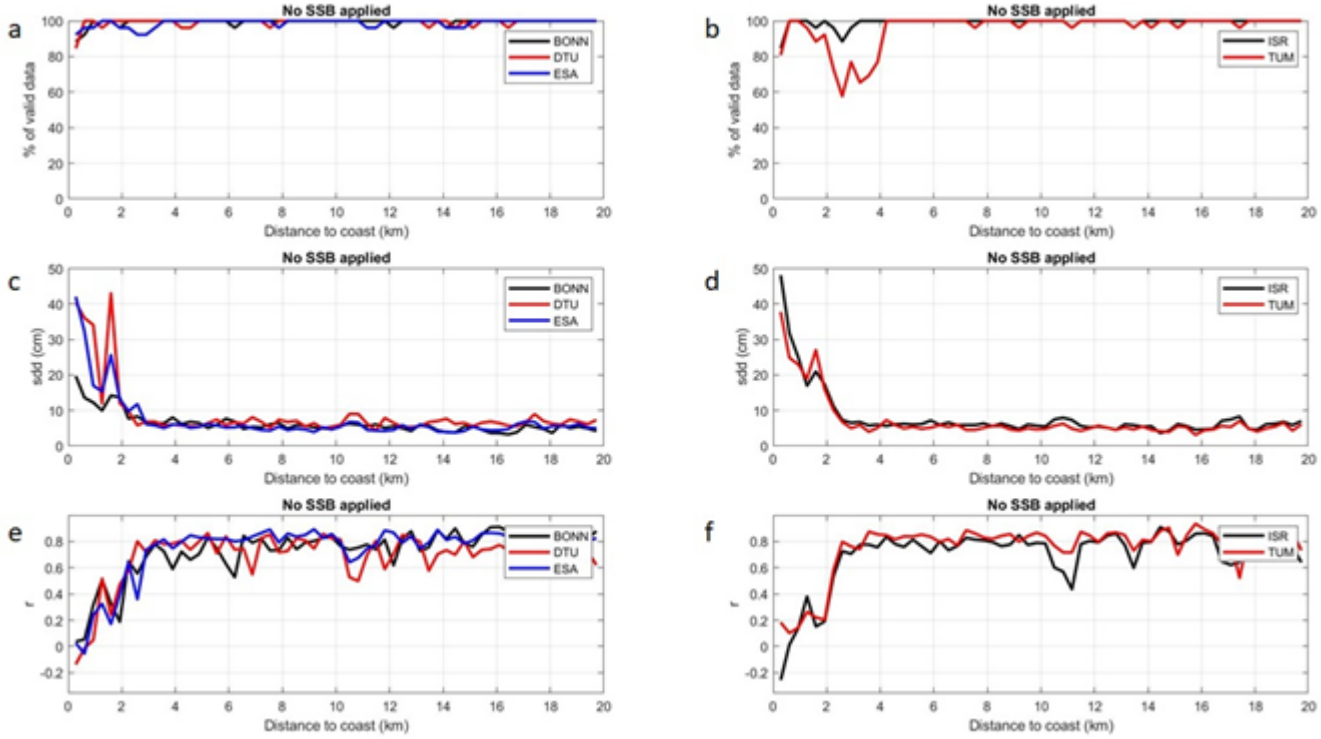


Figure 2.4.2. Along-track percentage of valid data (S3A: #114) for BONN, DTU and ESA retrackers (2.4.2.a), ISR and TUM (2.4.2.b). Same for sdd (2.4.2.c and 2.4.2.d) and r (2.4.2.e and 2.4.2.f).

In order to analyse the effect of the solutions adopted for removing the tidal elevation from the tide gauges, we computed the mean sdd in the [2 -20] km segment for all the retrackers. Figure 2.4.3 shows the average sdd for S3A: #114 (Figure 2.4.3.a), #322 (Figure 2.4.3.b); S3B: #114 (Figure 2.4.3.c), and #379 (Figure 2.4.3.d), using the 5 retrackers analysed. The same analysis was made using Bonanza (Figure 2.4.4) and Tarifa (Figure 2.4.5) TGs. fes (all) stands for tidal elevation removed from TGs using the same constituents as in FES2014; fes (main) stands for using only the main constituents; and got (all) stands for using all the constituents as in GOT4.10.

FES2014 gives the lower sdd in all the tracks using Huelva and Bonanza TGs. That is not the case for some tracks using Tarifa TG, as similar results are observed for FES2014 (all and main) and GOT4.10. No big differences are observed between the two solutions adopted for FES2014.

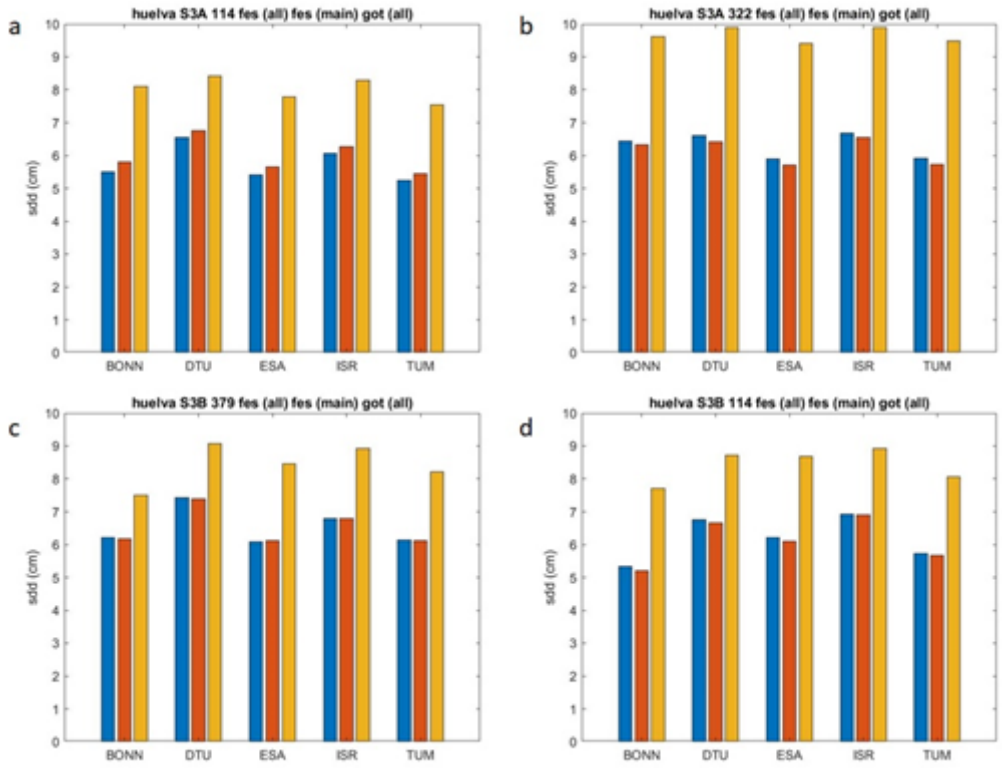


Figure 2.4.3. Average of sdd after removing tides from the Huelva TG (Gulf of Cadiz) using the same constituents as in FES2014 (all: blue bar), only the main constituents (main: red) and the same as in GOT4.10 (all: orange). The analysis was made for S3A, track #114 (2.4.3.a), #322 (2.4.3.b), and for S3B, track #114 (2.4.3.c), and #379 (2.4.3.d).

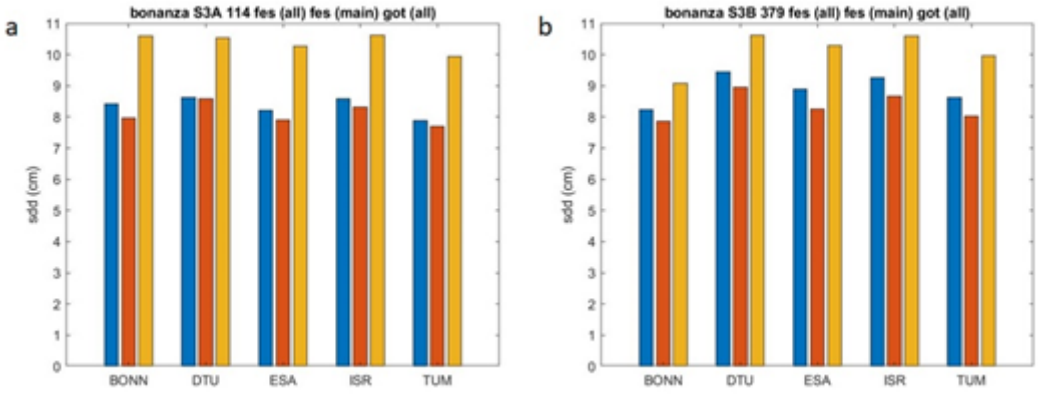


Figure 2.4.4. Same as in Figure 2.4.3 using Bonanza TG (Gulf of Cadiz).

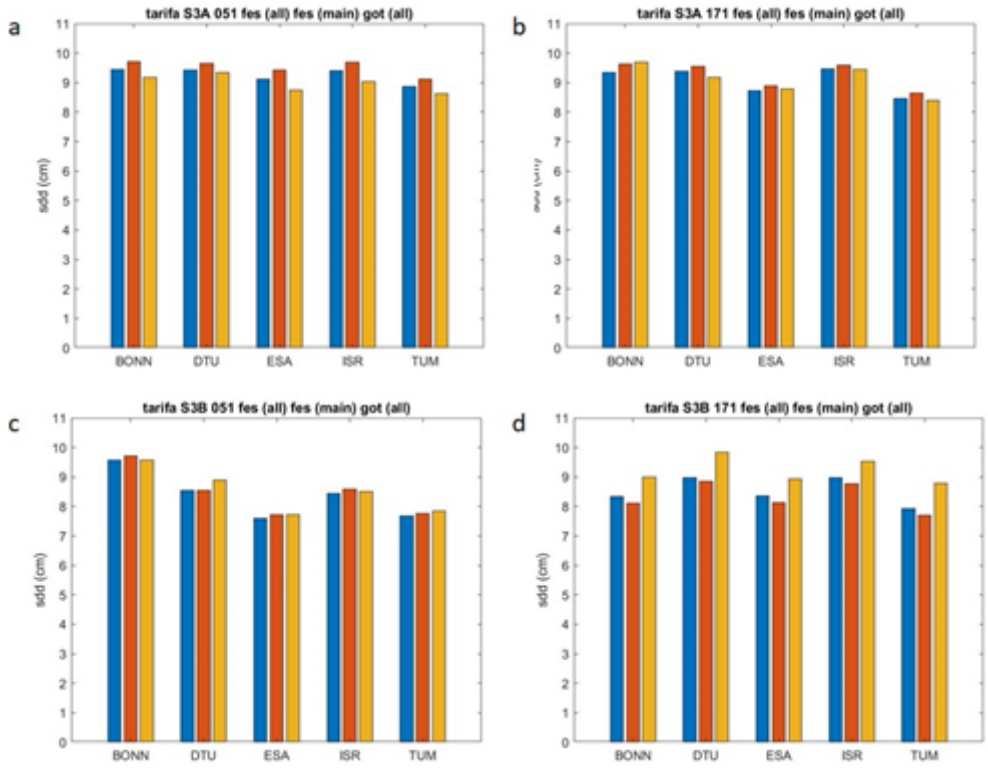


Figure 2.4.5. Same as in Figure 2.4.3 using Tarifa TG (Strait of Gibraltar).

FES2014 scores better in most of the cases with respect to GOT4.10. FES2014 (main) and (all) have similar scores most of the time with some discrepancies in some cases (especially in Bonanza). It is decided to use FES2014 (all) for the next analysis.

2.4.1.2 Analysis of SSB correction

Time series of SLA_Sat were created using different solutions of the SSB correction:

- DTU, ESA: 5% of SWH (from ESA retracker available in the input files).
- BONN: 5% of SWH (output from the retracker).
- ISR: 5% of SWH (output from the retracker).
- TUM: its own solution for SSB (no SWH is available in this product).

Figure 2.4.6 gives the results using Huelva TG for all the retrackers for S3A track #114 (Fig. 2.4.6.a), #322 (Fig. 2.4.6.b), S3B, track #114 (Fig. 2.4.6.c) and #379 (Fig. 2.4.6.d). The use of the SSB correction has a positive effect in all the cases (smaller sdd), with the exception of DTU retracker.

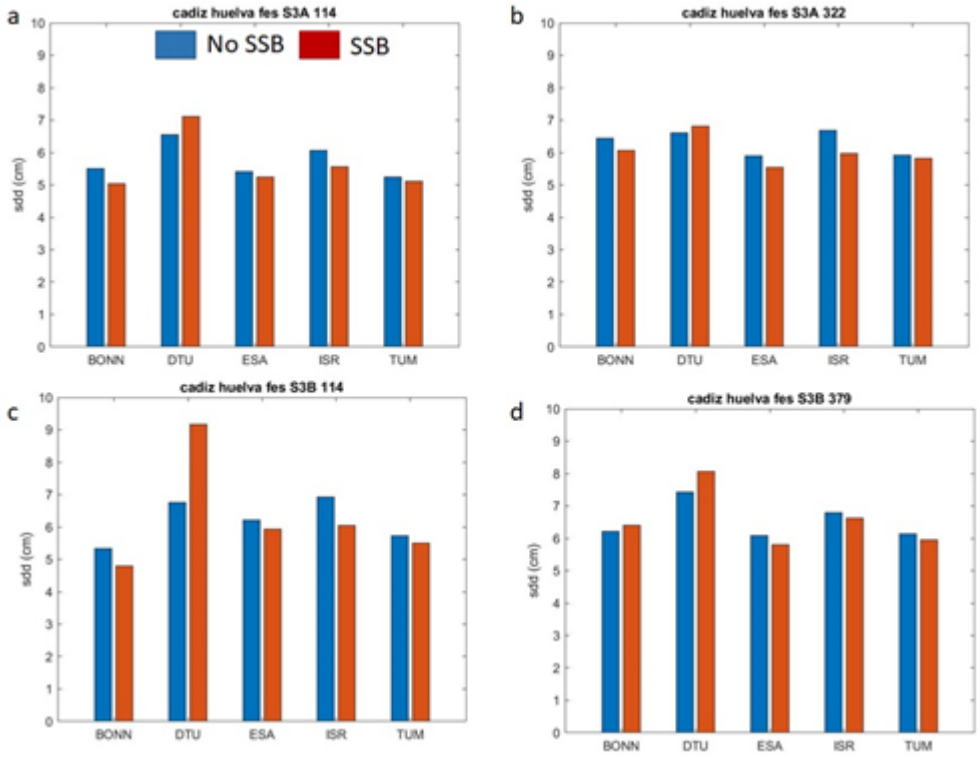


Figure 2.4.6. Average sdd for SSB = 0 (blue bars) and SSB correction applied (red) for all the retrackers using Huelva TG station. The analysis was made for S3A, tracks #114 (Fig. 2.4.6.a), #322 (Fig. 2.4.6.b), S3B, tracks #114 (Fig. 2.4.6.c) and #322 (Fig. 2.4.6.d).

The results obtained using Bonanza TG (Figure 2.4.7) indicates that the use of SSB reduces sdd to all the retrackers for S3A, track #114 (Fig. 2.4.7.a) and for S3B, track #379 (Fig. 2.4.7.b) (except for DTU).

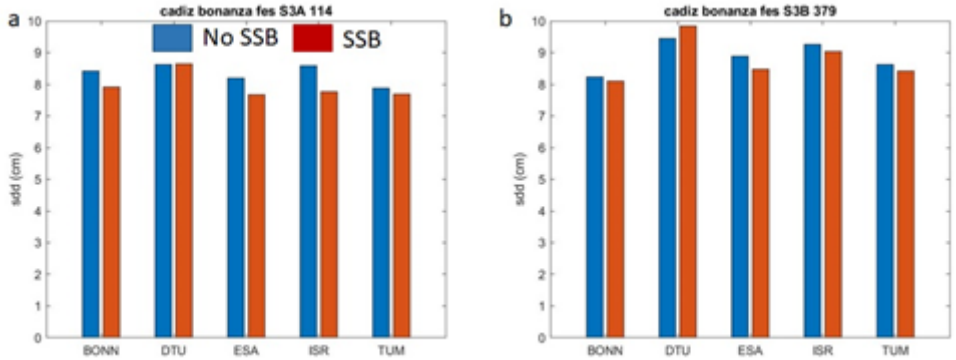


Figure 2.4.7. Same as in Figure 2.4.6 using Bonanza TG for S3A, track #114 (Fig. 2.4.7.a) and S3B, track #379 (Fig. 2.4.7.b).

The results using Tarifa TG station are given in Figure 2.4.8. The standard deviation of the differences is reduced when using SSB (all the retrackers) for S3A, track #051 (Fig. 2.4.8.a) and track #151 (Fig. 2.4.8.b) (except DTU). This is also the case for S3B, track #051 (Fig. 2.4.8.c) and track #171 (Fig. 2.4.8.d).

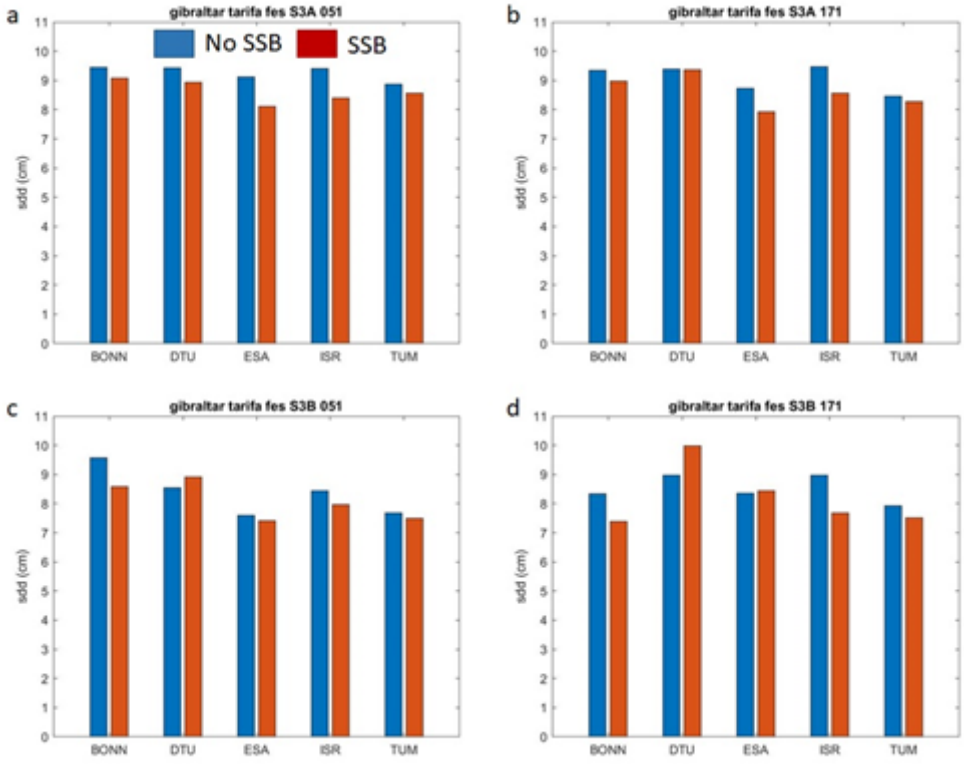


Figure 2.4.8. Same as in Figure 2.4.6 using Tarifa TG for S3A, tracks #051 (Fig. 2.4.8.a), #171 (Fig. 2.4.8.b), S3B, tracks #051 (Fig. 2.4.8.c) and #171 (Fig. 2.4.8.d).

The results show that SSB reduces the sdd in most of the tracks and retrackers. This is not the case of DTU retracker. The next analysis will be made applying the SSB to all the retrackers outputs.

2.4.1.3 Analysis of dry and wet tropospheric corrections from UPorto

The effect of using dry/wet tropospheric corrections from UPorto was analysed in this section. The along-track sdd in the [2 – 20] track segment was compared against the corrections available in the input files. The analysis was made using Huelva TG (Figure 2.4.9), Bonanza TG (Figure 2.4.10) and Tarifa TG (Figure 2.4.11) for all the available tracks from S3A and S3B. Quite similar sdds are observed in most of the cases with slight improvements in sdd (UPorto) in a few cases.

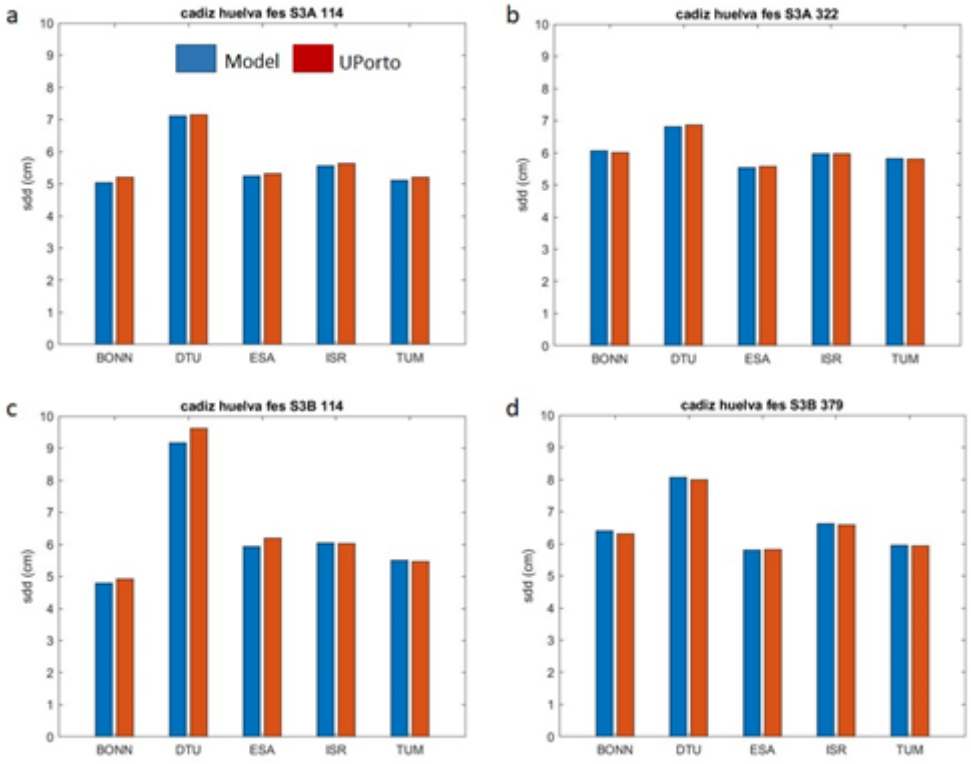


Figure 2.4.9. Along-track average of sdd in the [2 – 20] km track segment using dry/wet tropospheric corrections from input files (blue bars) and UPorto corrections (red bars). The analysis was made using Huelva TG station for S3A, tracks #114 (Fig. 2.4.9.a), #322 (Fig. 2.4.9.b), S3B, tracks #114 (Fig. 2.4.9.c) and #322 (Fig. 2.4.9.d).

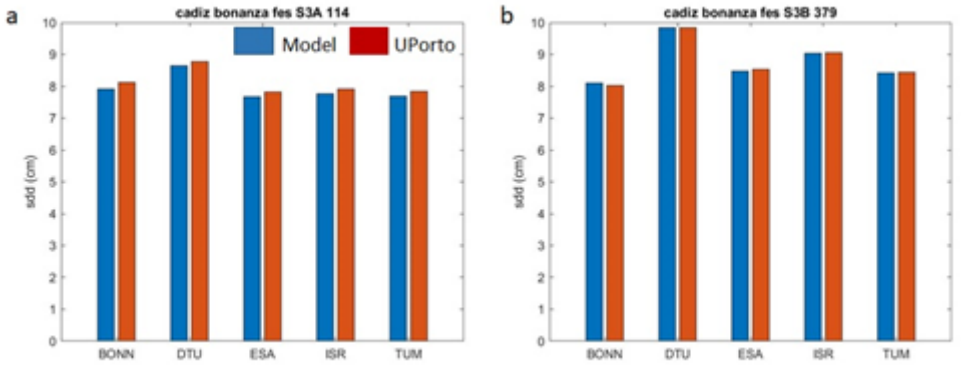


Figure 2.4.10. Same as in Fig. 2.4.9 using Bonanza TG station.

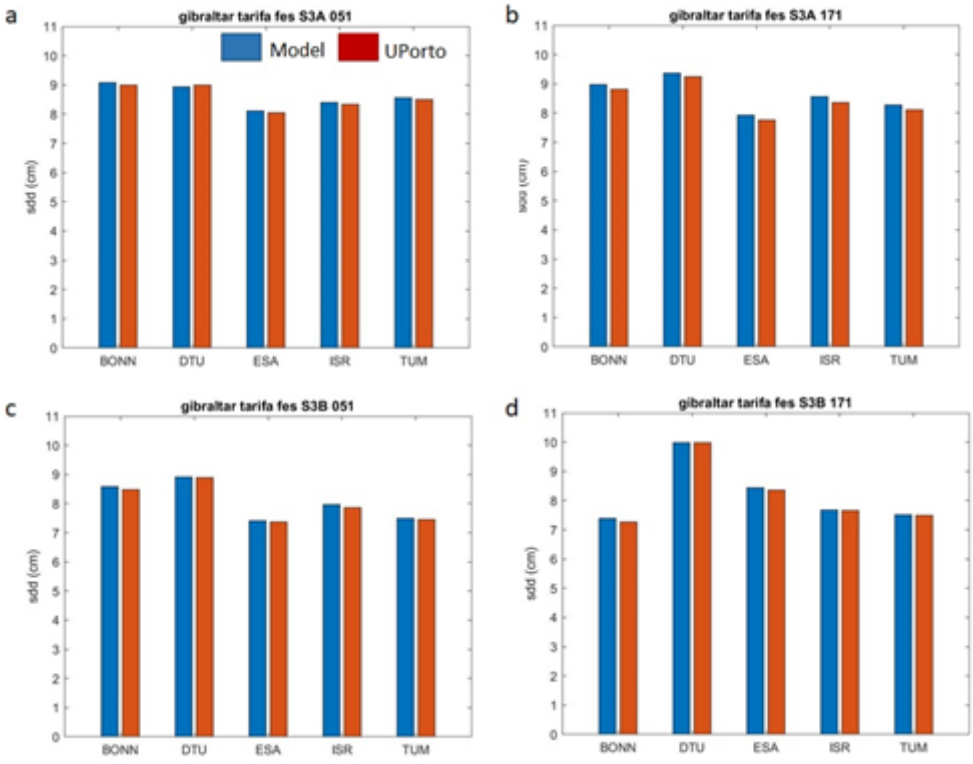


Figure 2.4.11. Same as in Fig. 2.4.9 using Tarifa TG station.

The results show that WTC/DTC from UPorto does not reduce the sdd in most of the tracks and retrackers. The next analysis will be made applying the UPorto corrections to all the retrackers outputs.

2.4.1.4 Summary of results

In this section, the comparison between retrackers is presented. The analysis was made using FES model (all the constituents), SSB correction and UPorto DTC/WTC corrections. Figure 2.4.12 shows the along-track %VD, sdd and r for S3A track #114 for all the retrackers using Huelva TG station. The %VD is quite high in all the cases, except for BONN in the whole track segment and for TUM in the [2 – 4] km track segment. The small percentage of valid data for BONN is due to a high number of NaNs for SWH along the track, which is affecting the computation of the SSB. This should be investigated further. The standard deviation of the differences is below 10 cm in the [2 – 20] km segment in all the cases. The Pearson's correlation coefficient is high in that segment for most of the tracks.

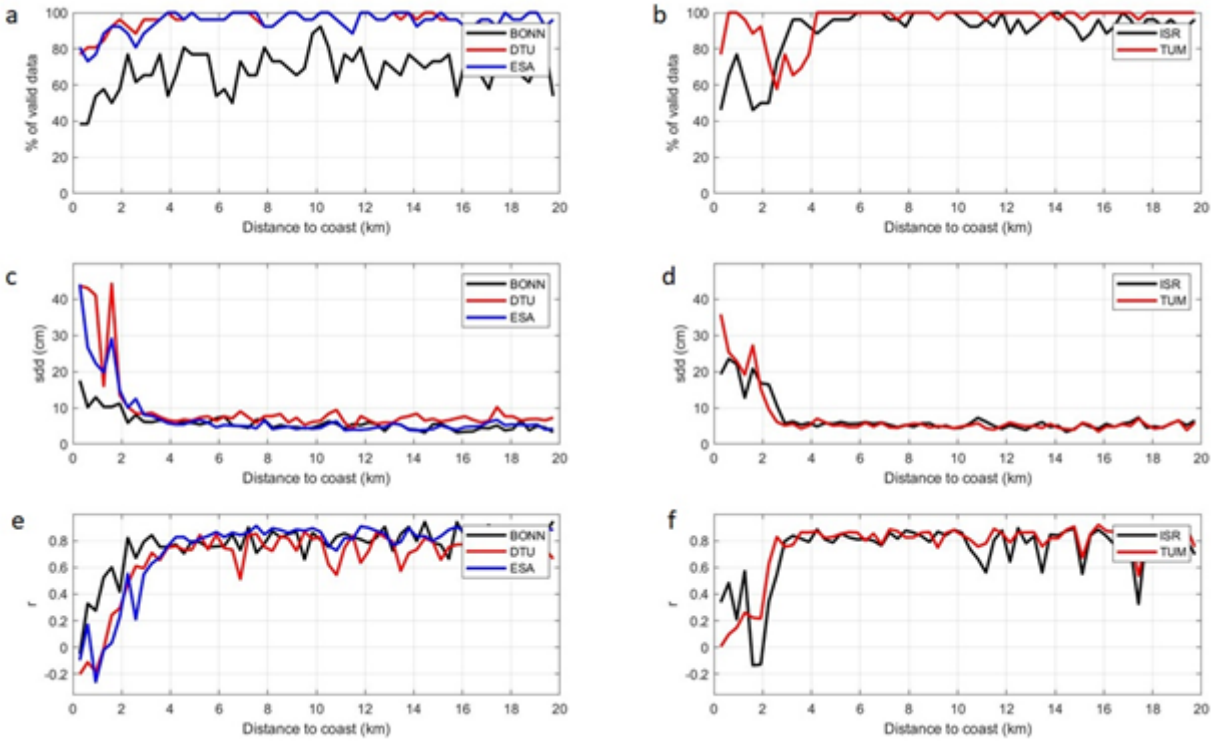


Figure 2.4.12. Along-track percentage of valid data (S3A: #114) for BONN, DTU and ESA retrackers (2.4.12.a), ISR and TUM (2.4.12.b). Same for sdd (2.4.12.c and 2.4.12.d) and r (2.4.12.e and 2.4.12.f). TG: Huelva.

Figure 2.4.13 gives the results for S3A track #322. Similar results are obtained, but in this case the good quality of SLA_Sat is observed even closer to the shore (almost one km to the coast). The results for S3B track #114 and #319 are given in Figure 2.4.14 and Figure. 2.4.15, respectively. The along-track behaviour of the statistics is similar to S3A with some exceptions. DTU shows low accuracy in most of track #114. Good quality data is mostly observed at even 1 km from the shore.

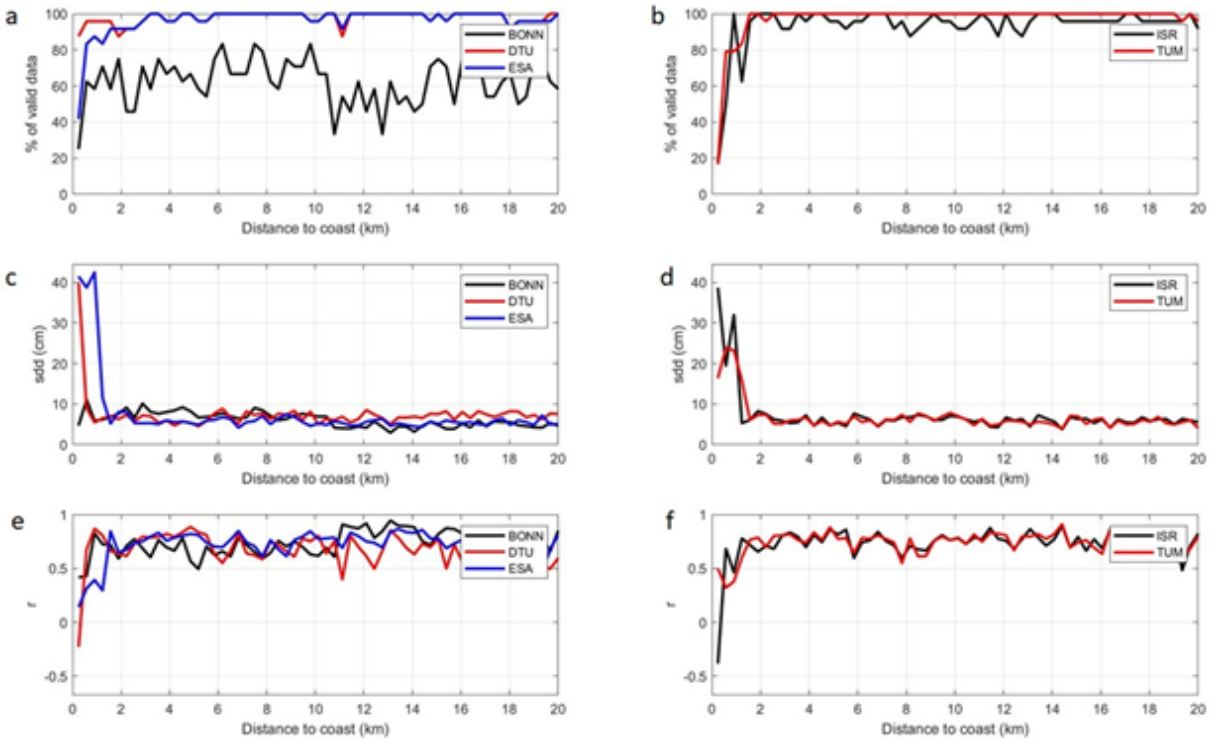


Figure 2.4.13. Same as Fig. 2.4.12 for S3A track #322.

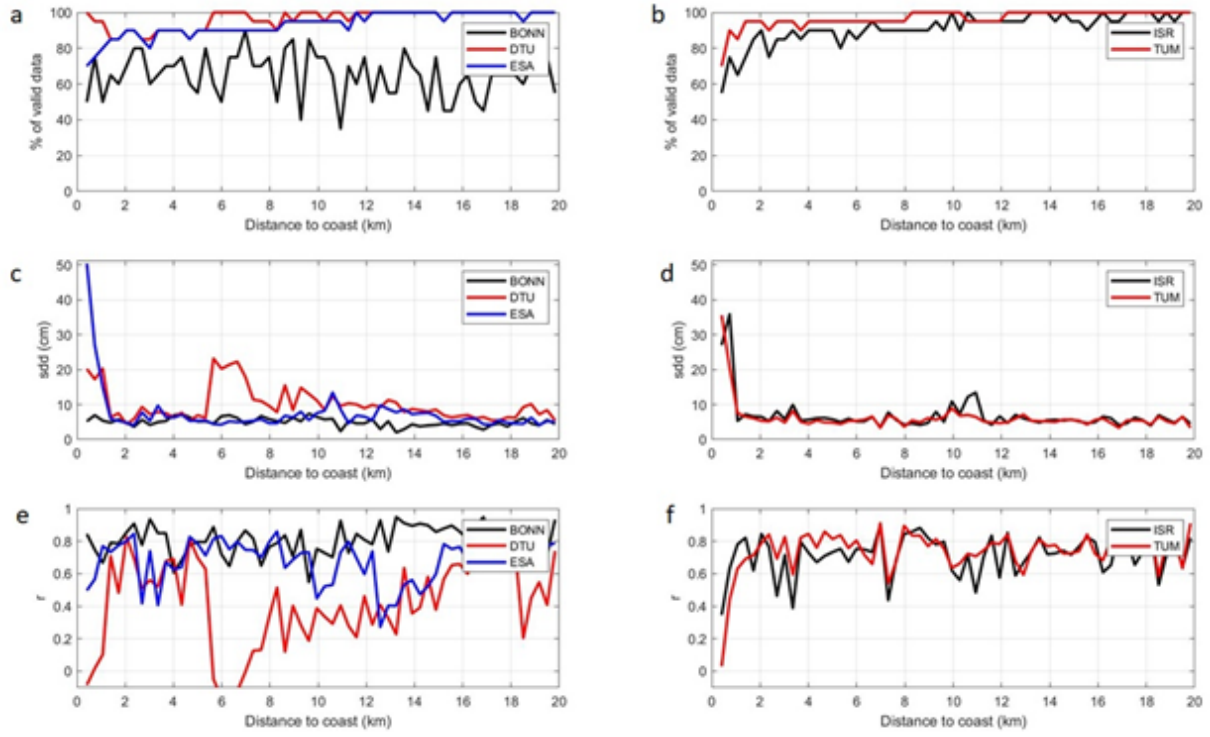


Figure 2.4.14. Same as Fig. 2.4.12 for S3B track #114.

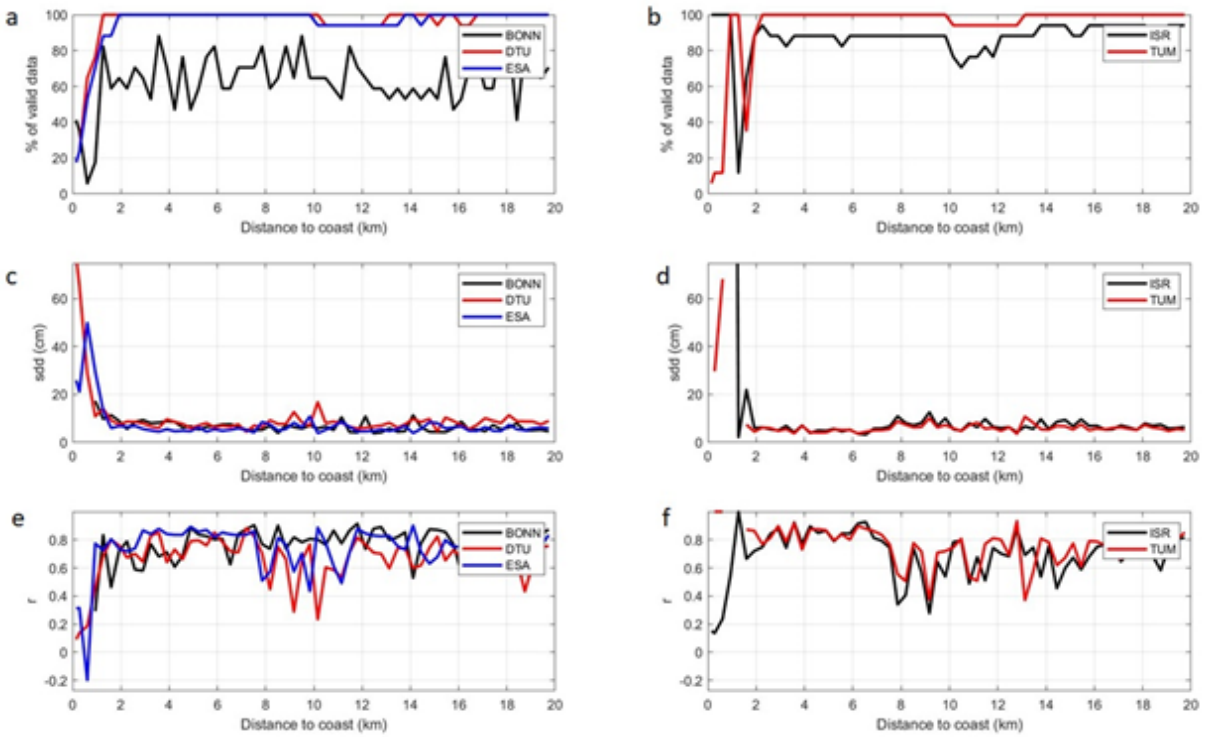


Figure 2.4.15. Same as Fig. 2.4.12 for S3B track #379.

Table 2.4.1 presents a summary of average sdd and r in the [2 – 20] track segment. BONN, ESA and TUM retrackers show the best performances in terms of sdd and r.

Table 2.4.1. Average of sdd and r for the tracks analysed using Huelva TG station.

	Sentinel-3A				Sentinel-3B			
	Track #114		Track #322		Track #114		Track #379	
	sdd	r	sdd	r	sdd	r	sdd	r
BONN	**5.2**	0.81	6.0	0.75	**4.9**	**0.82**	6.3	**0.78**
DTU	7.1	0.73	6.9	0.70	9.6	0.43	8.0	0.69
ESA	5.3	0.81	**5.6**	**0.77**	6.2	0.68	**5.8**	0.77
ISR	5.6	0.80	6.1	0.74	6.4	0.72	6.5	0.72
TUM	**5.2**	**0.82**	5.8	0.76	5.5	0.76	5.9	0.75

Figure 2.4.16 and Figure 2.4.17 give the comparison made with Bonanza TG station, for S3B track #114 and #379. The along-track sdd and r are a bit worse than in the previous tide gauge. This is clearly seen in Table 2.4.2. The same tracks give a lower accuracy than those obtained with Huelva TG station, maybe due to the location of the Bonanza TG in the mouth of the Guadalquivir estuary (Fig. 2.4.1).

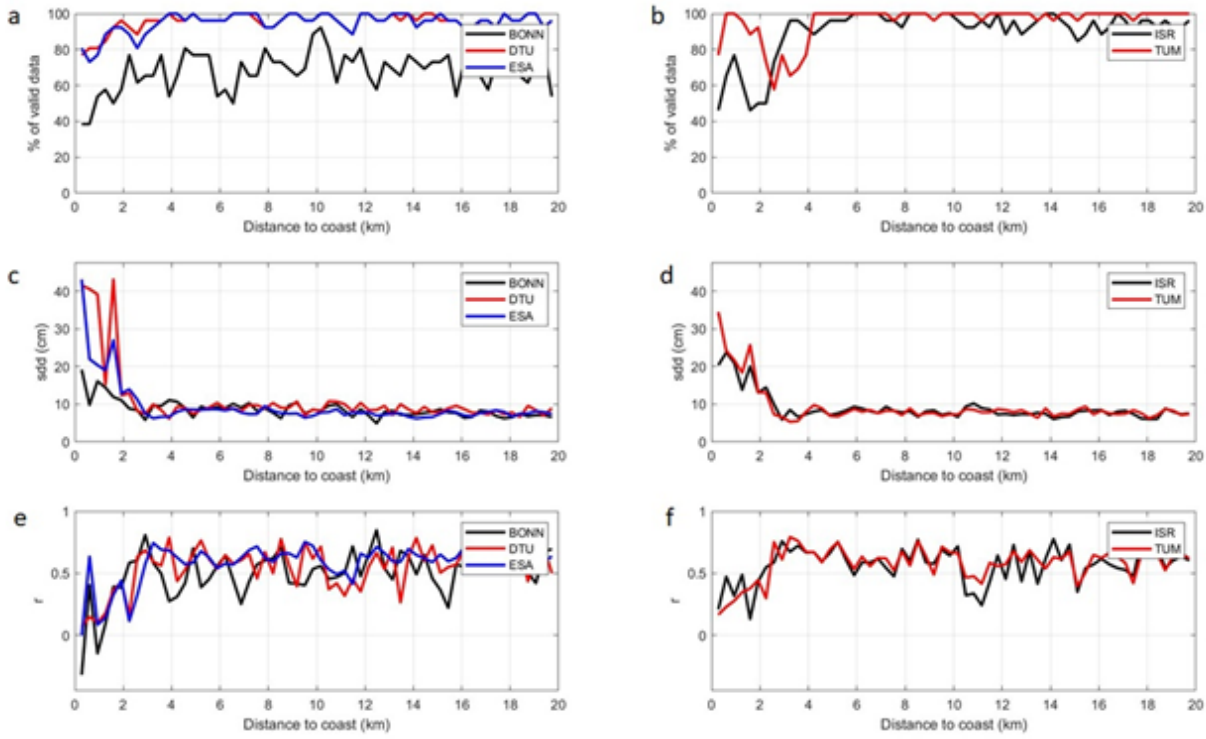


Figure 2.4.16. Along-track percentage of valid data (S3A: #114) for BONN, DTU and ESA retrackers (2.4.16.a), ISR and TUM (2.4.16.b). Same for sdd (2.4.16.c and 2.4.16.d) and r (2.4.16.e and 2.4.16.f). TG: Bonanza.

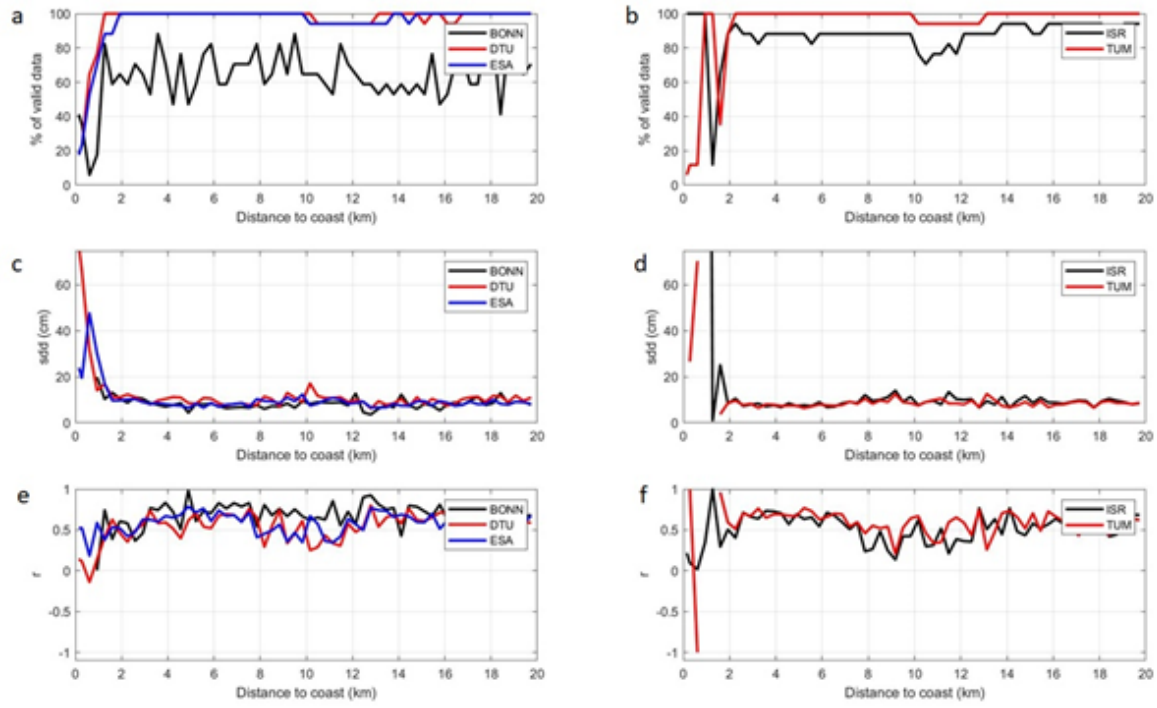


Figure 2.4.17. Same as Fig. 2.4.16 for S3B track #379.

Table 2.4.2 is a summary of the average sdd and r in the [2 – 20] track segment. BONN, ESA and TUM retrackers show the best performances in terms of sdd and r.

Table 2.4.2. Same as Table 2.4.1 for the comparison against Bonanza TG station.

	Sentinel-3A		Sentinel-3B	
	Track #114		Track #379	
	sdd	r	sdd	r
BONN	8.1	0.54	**8.0**	**0.68**
DTU	8.8	0.59	9.8	0.56
ESA	**7.8**	**0.61**	8.5	0.59
ISR	7.9	0.59	9.1	0.53
TUM	**7.8**	**0.61**	8.4	0.59

Finally, Figure 2.4.18 to 2.4.21 present the results obtained using Tarifa TG station in the Strait of Gibraltar. The length of the S3A tracks is 18 km (corresponding to the length of the tracks over the ocean in the strait). This is particularly interesting as there is one ocean-to-land and one land-to-ocean transition on each track (Fig. 2.4.1). In these tracks, the distance to the coast is estimated respect to the northern coast (i.e. Spain). Fig. 2.4.18 shows that sdd and r does not present a deterioration in the northern coast, which is observed in the southern transition.

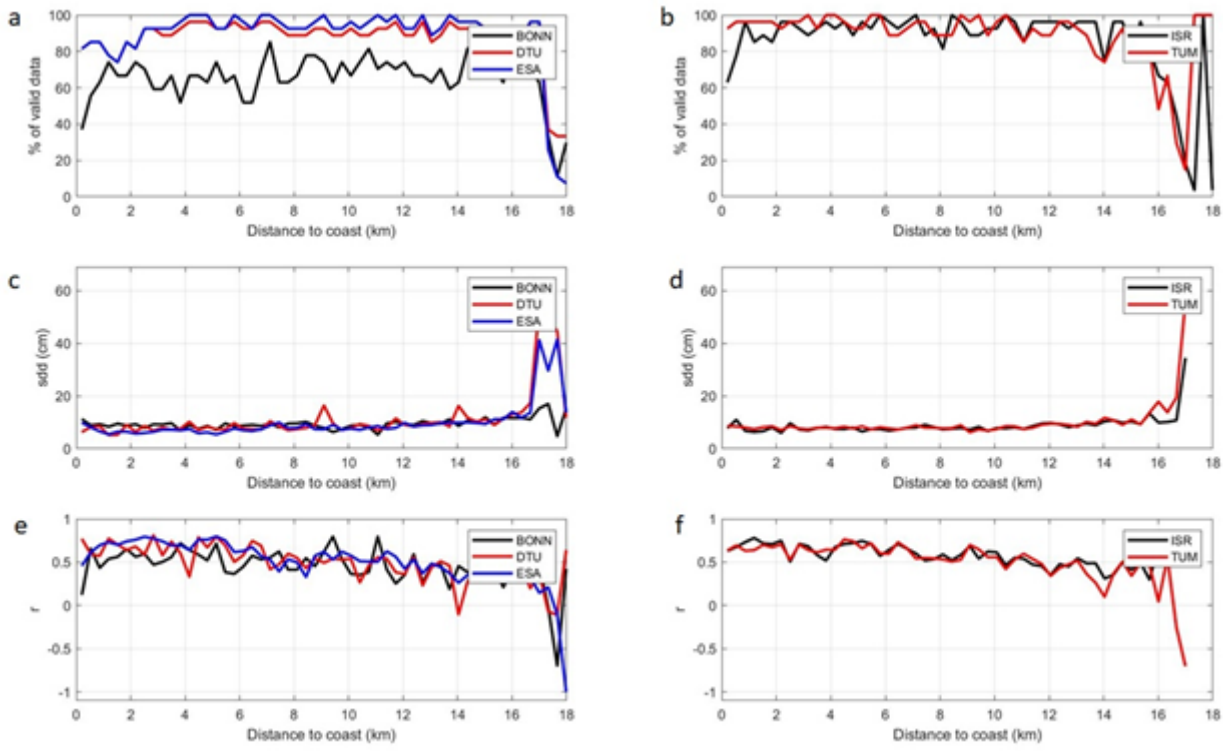


Figure 2.4.18. Along-track percentage of valid data (S3A: #051) for BONN, DTU and ESA retrackers (2.4.18.a), ISR and TUM (2.4.18.b). Same for sdd (2.4.18.c and 2.4.18.d) and r (2.4.18.e and 2.4.18.f). TG: Tarifa.

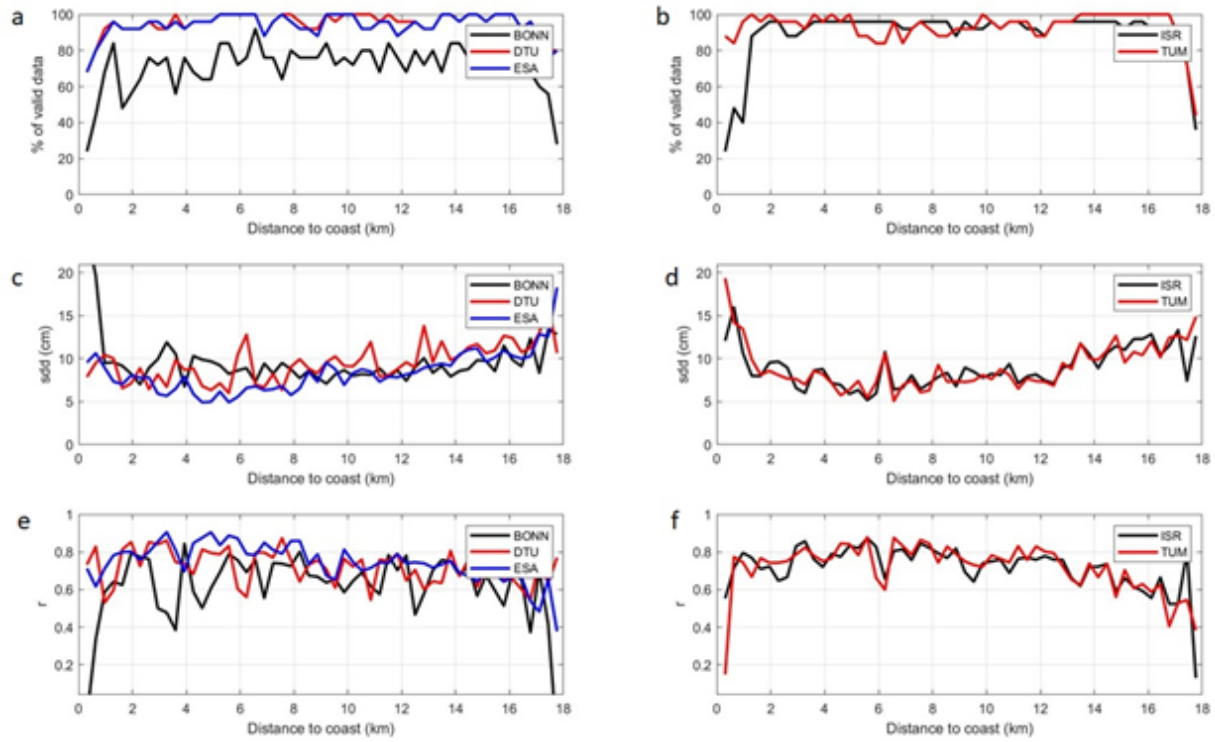


Figure 2.4.19. Same as Fig. 2.4.18 for S3A track #171.

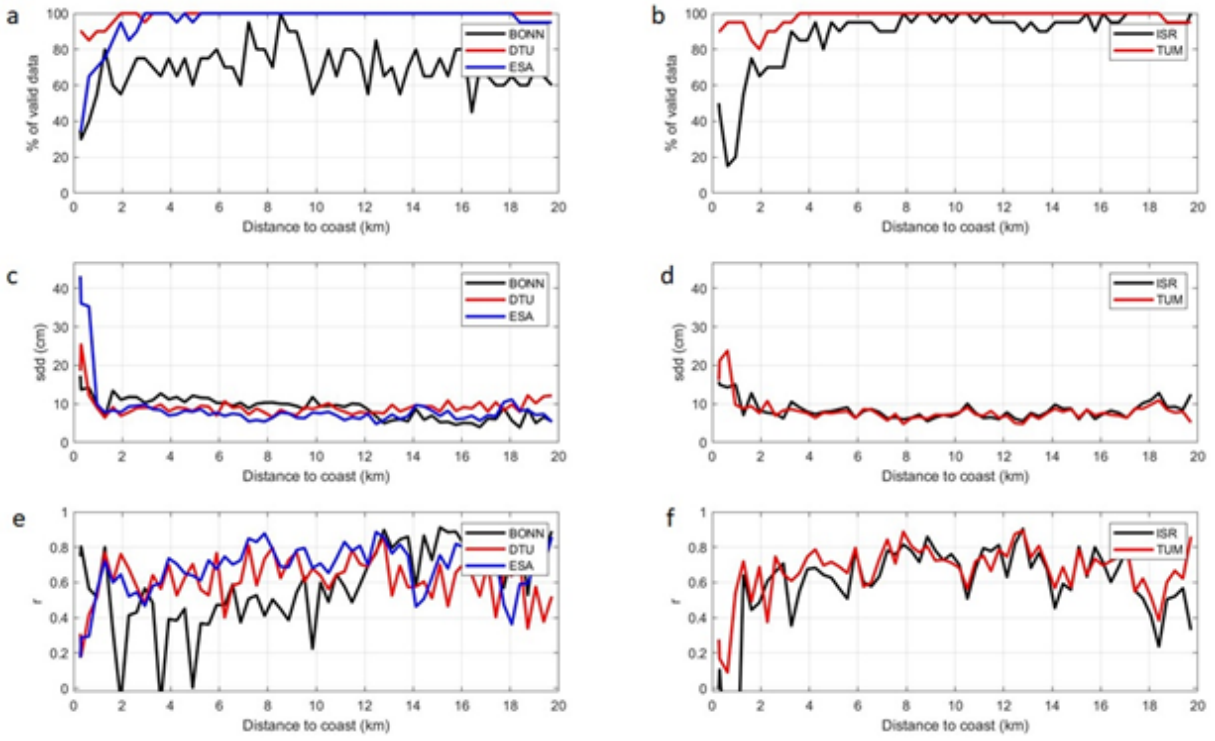


Figure 2.4.20. Same as Fig. 2.4.18 for S3B track #051.

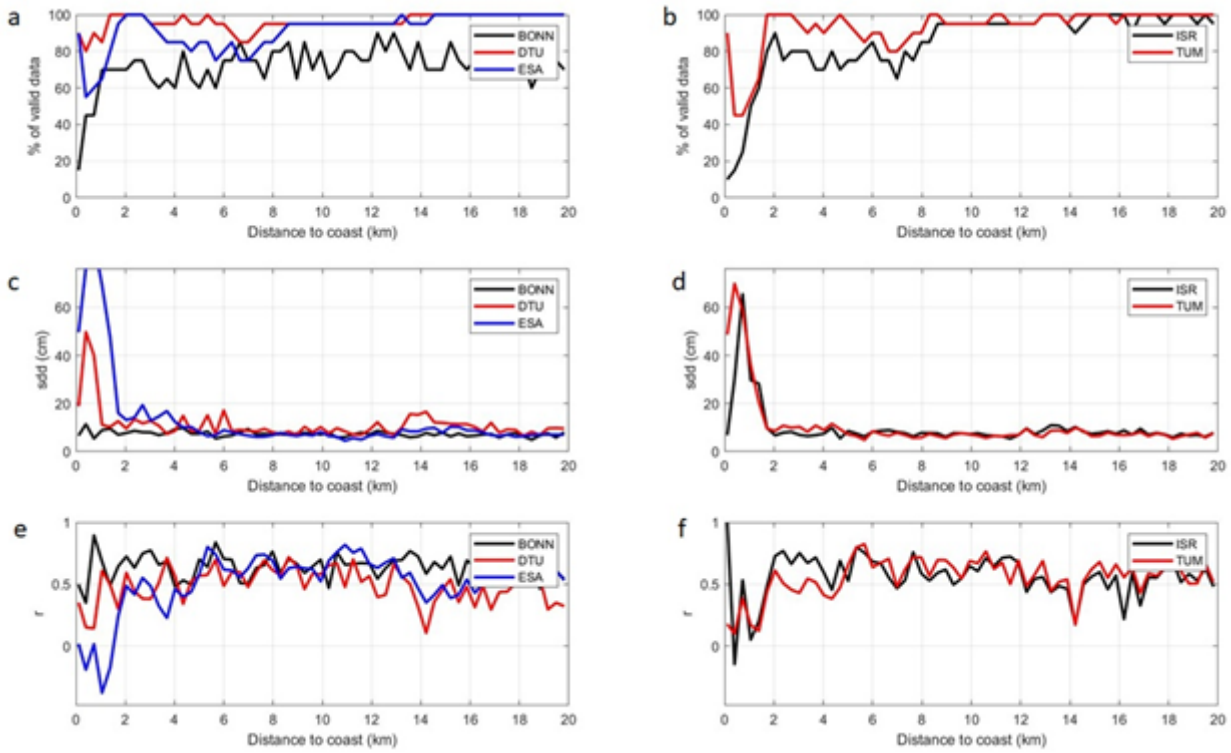


Figure 2.4.21. Same as Fig. 2.4.18 for S3B track #171.

Table 2.4.3 gives a summary of average sdd and r in the [2 – 16] track segment (S3A) and [2 – 20] km (S3B). ESA retracker gives the best results (sdd and r) in most of the tracks. This is not the case of S3B track #171 where BONN performs better.

Table 2.4.3. Same as Table 2.4.1 for the comparison against Tarifa TG station.

	Sentinel-3A				Sentinel-3B			
	Track #051		Track #171		Track #051		Track #171	
	sdd	r	sdd	r	sdd	r	sdd	r
BONN	9.0	0.48	8.8	0.66	8.5	0.60	**7.3**	**0.64**
DTU	9.0	0.51	9.2	0.73	8.9	0.62	10.0	0.51
ESA	**8.0**	**0.56**	**7.8**	**0.77**	**7.4**	**0.70**	8.3	0.58
ISR	8.3	**0.56**	8.4	0.74	7.9	0.65	7.7	0.58
TUM	8.5	0.54	8.1	0.76	7.5	**0.70**	7.5	0.59

2.4.2 Significant Wave Height (SWH): S3A and S3B

The analysis of SWH was made using the Gulf of Cadiz buoy. The track selection (S3A and S3B) was made considering a radius of 30 km from the buoy position. Figure 2.4.22 gives the track segments selected for the comparison: S3A track #322 and S3B tracks #144 / #379. Only two retrackers gave SWH as outputs: BONN and ISR. For comparison, we also used the wave heights from ESA retracker, available in the input files provided by IsardSat. We applied a thresholding retaining data in the range [0 – 15] m. Also, an outlier detection using the 3-sigma criterion.

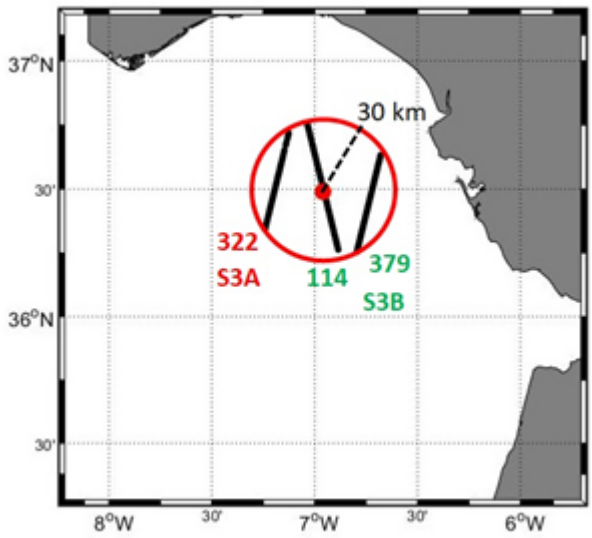
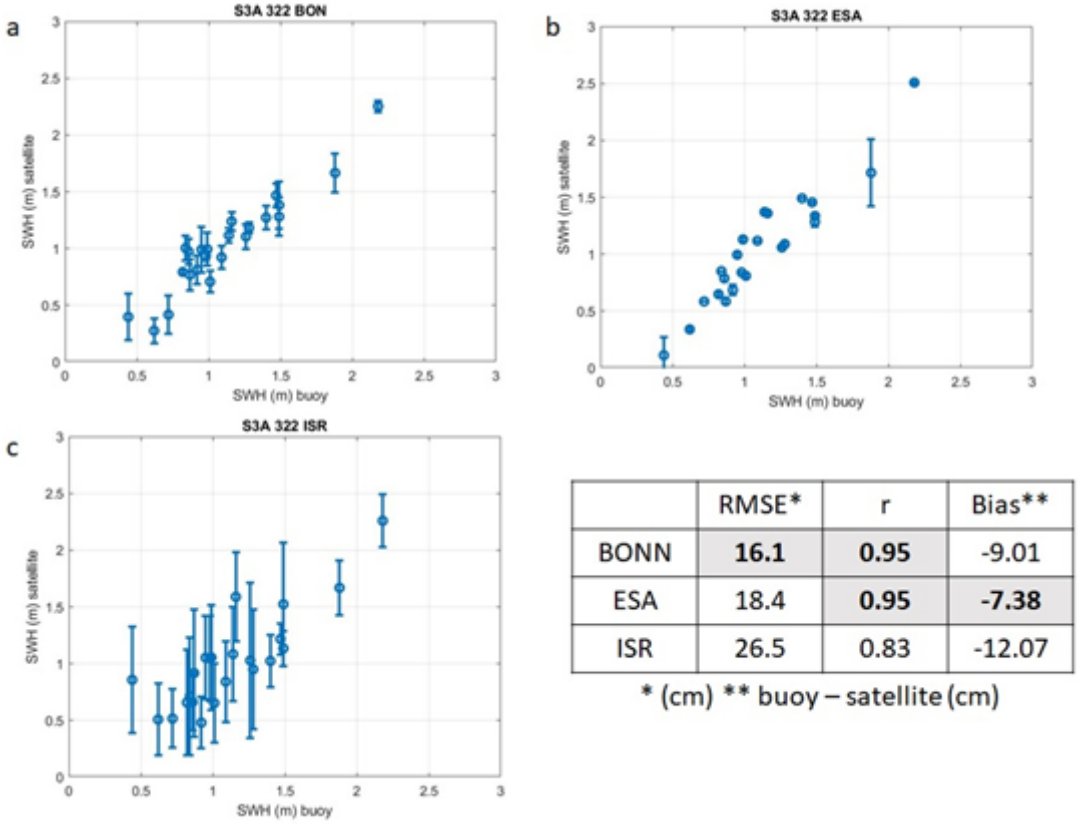


Figure 2.4.22. Segments of the tracks inside a radius of 30 km from the position of the Gulf of Cadiz buoy.

Time series of SWH from S3A / S3B were made by averaging the SWH values in the whole track segments inside the radius around the buoy. The standard deviation of these averages is extremely big in both cases (BONN and ISR retrackers). In order to reduce that uncertainty, we estimated the average SWH (satellite) using only the closest five along-track SWH measurements to the buoy position. The comparison between buoy / satellite pairs of SWH is shown in Fig. 2.4.23 for track S3A track #322 for BONN (Fig. 2.4.23.a), ESA (Fig. 2.4.23.b) and ISR (Fig. 2.4.23.c) retrackers. The standard deviation of the average SWH (satellite) is also given in the figures. Some statistics (bias, RMSE and r) are shown. BONN gives the best RMSE scores. The same r is observed for BONN and ESA. ESA is the less noisy dataset (standard deviation of the average SWH).



	RMSE*	r	Bias**
BONN	**16.1**	**0.95**	-9.01
ESA	18.4	**0.95**	**-7.38**
ISR	26.5	0.83	-12.07

* (cm) ** buoy – satellite (cm)

Figure 2.4.23. Scatter plot of SWH comparison (buoy vs. satellite) using S3A track #322 with the outputs of BONN (Fig. 2.4.23.a), ESA (Fig. 2.4.23.b) and ISR (Fig. 2.4.23.c) retrackers. A table with a summary of results is also given.

The results of the comparison made with S3B track #114 and #379 are shown in Table 2.4.4. BONN performs better (RMSE and r) than ESA and ISR for both tracks.

Table 2.4.4. Summary of statistics for S3B track #114 and #379.

Track #114

	RMSE*	r	Bias**
BONN	**18.4**	**0.95**	-5.74
ESA	23.1	0.87	**-3.67**
ISR	25.0	0.88	-9.07

Track #379

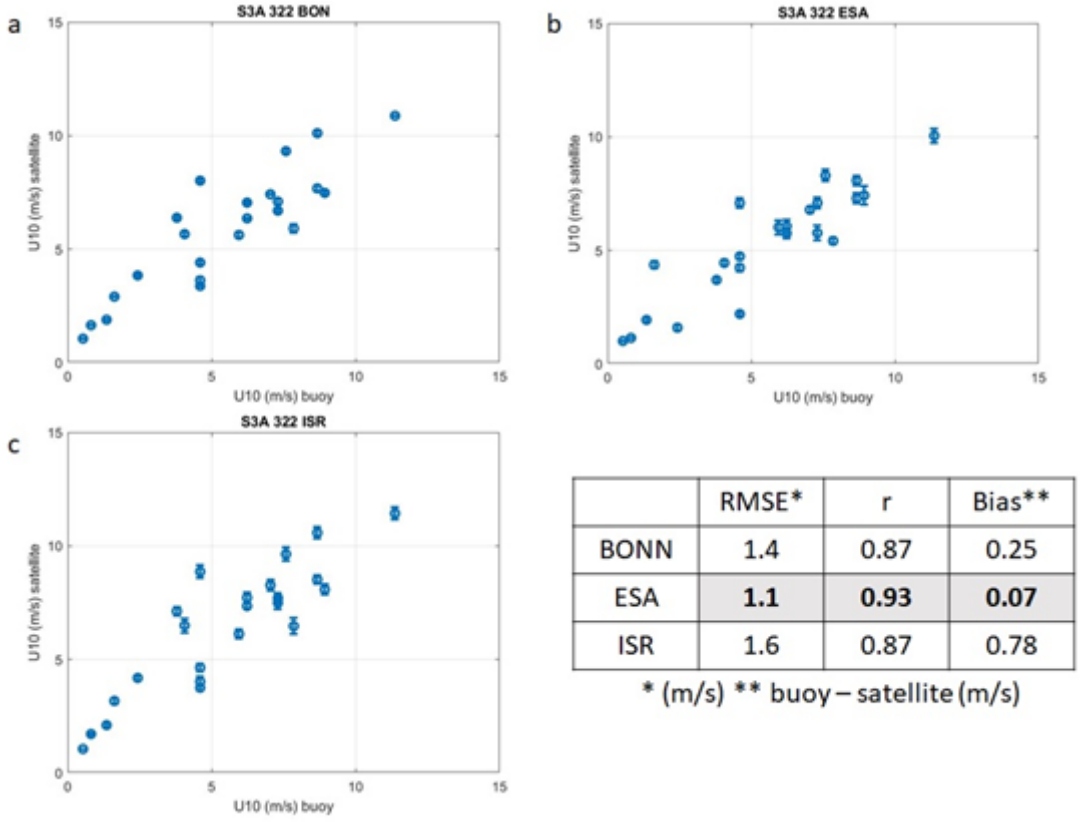
	RMSE*	r	Bias**
BONN	**21.7**	**0.97**	-12.10
ESA	34.7	0.88	9.75
ISR	43.8	0.80	**4.05**

* (cm) ** buoy – satellite (cm)

The number of along-track SWH measurements used for averaging is critical to assess the wave height obtained from the retrackers. BONN performs much better than ISR in terms of accuracy of SWH and correlation with in-situ data.

2.4.3 Wind speed (U10): S3A and S3B

The same procedure as in SWH was applied to U10 data analysis (using the same buoy). The output files of the retrackers do not give U10 as a field. Instead, they give sigma0. In order to compute U10 (following Abdalla`s algorithm) it is necessary to apply the atmospheric attenuation to sigma0. The atmospheric attenuation is a field available in the input files (IsardSat), and should be added to the uncorrected sigma0. Once applied, we estimated U10. Regarding data screening, an outlier detection using the 3-sigma criterion was used. Figure 2.4.24 shows the scatter plot for S3A track #322. ESA retracker gives the best scores.



	RMSE*	r	Bias**
BONN	1.4	0.87	0.25
ESA	**1.1**	**0.93**	**0.07**
ISR	1.6	0.87	0.78

* (m/s) ** buoy – satellite (m/s)

Figure 2.4.24. Same as Fig. 2.4.23 for U10.

Table 2.4.5 gives the scores of the statistical analysis made for S3B tracks #114 and #379. BONN retracker performs better for r and RMSE (track #379).

Table 2.4.5. Summary of statistics for S3B track #114 and #379.

Track #114

	RMSE*	r	Bias**
BONN	1.6	**0.91**	-0.93
ESA	**1.3**	0.89	**-0.07**
ISR	3.7	0.88	-3.00

Track #379

	RMSE*	r	Bias**
BONN	**2.0**	**0.82**	0.76
ESA	2.1	0.76	**-0.05**
ISR	3.2	0.84	-2.12

* (m/s) ** buoy – satellite (m/s)

2.4.4 Sea Level Anomaly (SLA): CS2

2.4.4.1 Analysis of SSB correction

The satellite-derived SLA was obtained along the satellite tracks surrounding the tide gauges (Huelva and Bonanza) during the analysed time period. The procedure to estimate the satellite-derived SLA was the same as in S3A / S3B. The tidal model analysed was FES (no information about the version was found) as no GOT model is available. The loading tide is not included in the tidal elevation (that is not the case of S3A / S3B), so it has to be included in the recipe to estimate SLA. Time series of CS2 data were built considering the along-track measurement distances to the coast. We extracted 21 time series at distances from 0 to 20 km to the coastline by taking, for each CS2 overpass, the average of the three SLA measurements at the closest 20-Hz positions to the selected distances. These time series were then used for comparison against in-situ data.

Figure 2.4.25 gives the results of the validation made using Huelva TG station for the 5 retrackers (no SSB correction was applied). The percentage of valid data is slightly smaller than in S3A / S3B datasets. ESA, ISR and TUM retrackers seem to keep more valid data along the track segments with an intense loss of data in the [0 – 2] km track segment. The standard deviation of the differences and r are bigger / smaller than those obtained in the previous analysis made with S3A / S3B. All the retrackers show extremely high sdd and low r in the [0 – 2] km track segment (due to the proximity to land) with the exception of BONN.

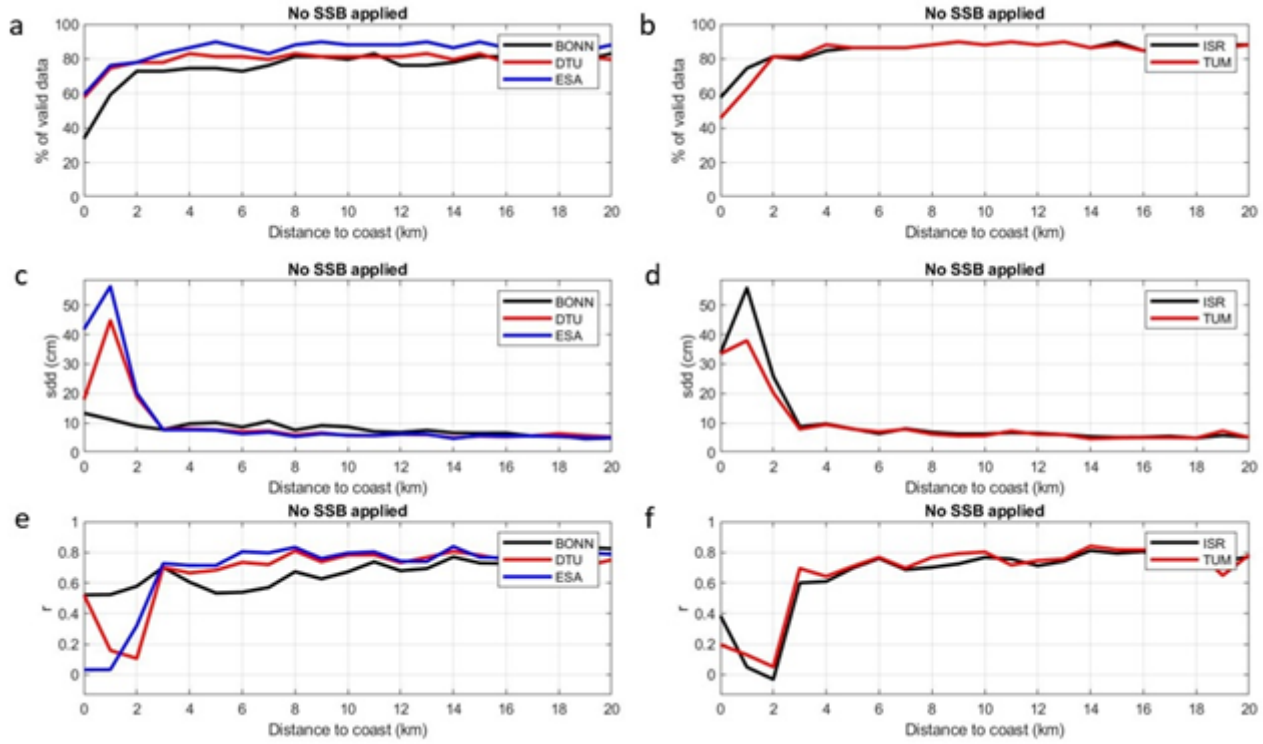


Figure 2.4.25. Along-track percentage of valid data for BONN, DTU and ESA retrackers (Fig. 2.4.25.a), ISR and TUM (Fig. 2.4.25.b) using Huelva TG station. Same for sdd (Fig. 2.4.25.c and Fig. 2.4.25.d) and r (Fig. 2.4.25.e and 2.4.25.f).

Figure 2.4.26 shows the results for the comparison with Bonanza TG station. The results are similar to Huelva TG.

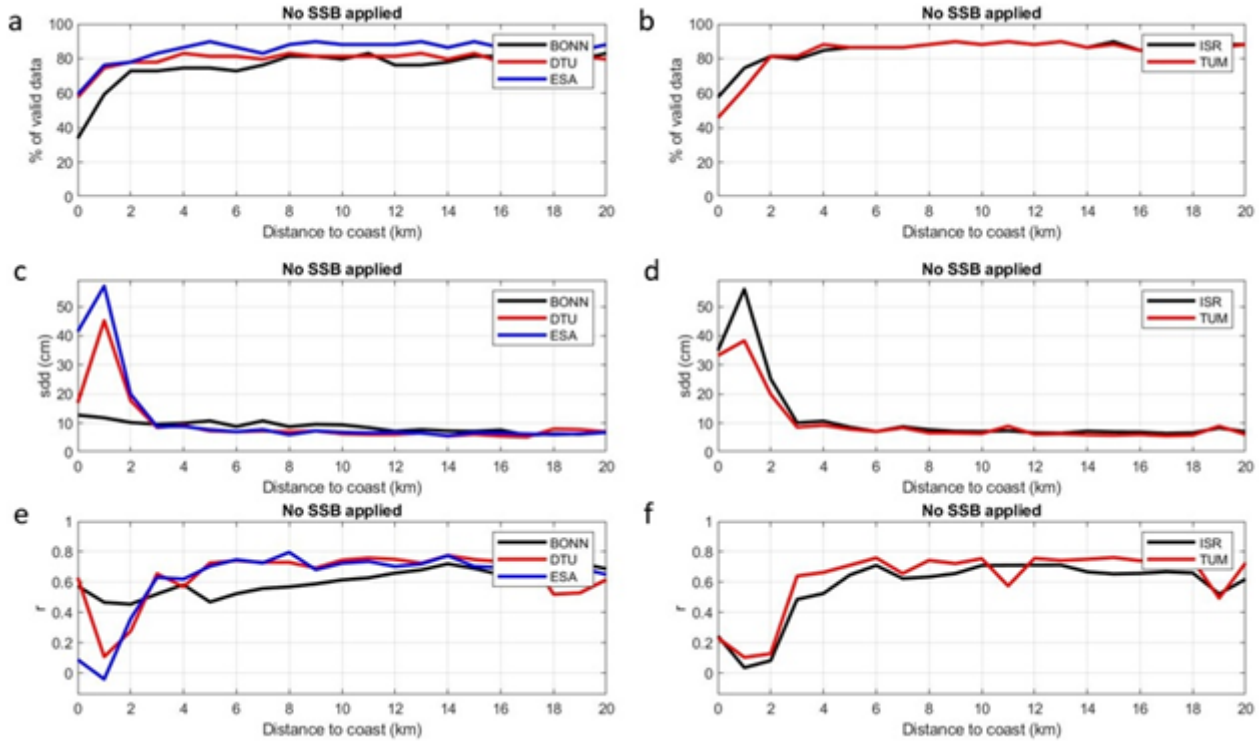


Figure 2.4.26. Same as Fig. 2.4.25 for Bonanza TG station.

Figure 2.4.27 presents the average of ssd for the [2 – 20] km track segment without / with the SSB correction applied for Huelva (Fig. 2.4.27.a) and Bonanza (Fig. 2.4.27.b) TG stations. The use of the SSB correction has a positive effect only for ISR and TUM retrackers (no effect for BONN). The SSB was obtained as:

- DTU, ESA: 5% of SWH (from ESA retracker available in the input files).
- BONN: 5% of SWH (output from the retracker).
- ISR: 5% of SWH (output from the retracker).
- TUM: its own solution for SSB (no SWH is available in this product).

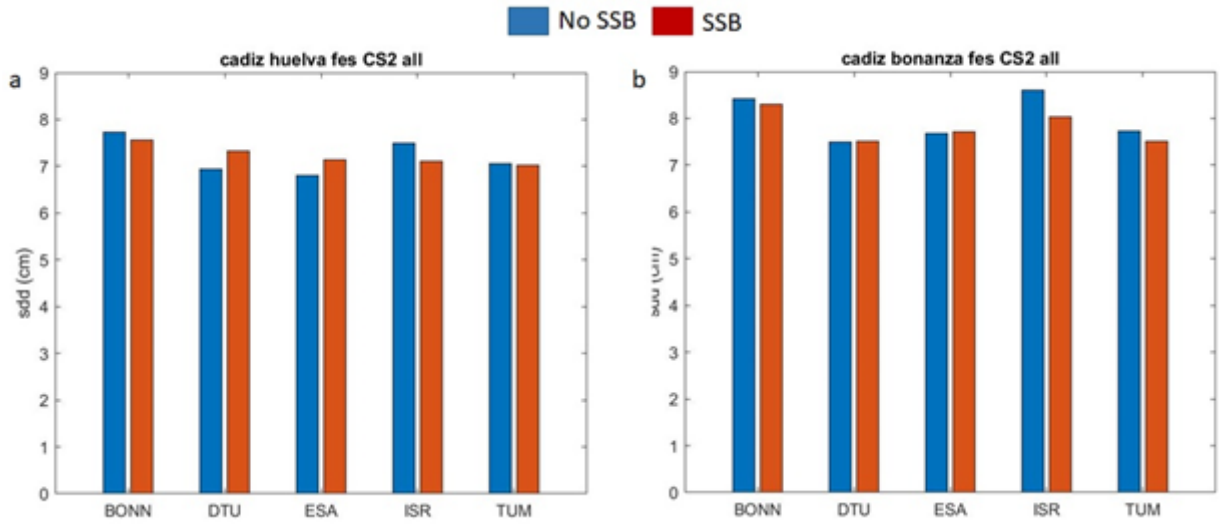


Fig. 2.4.27. Average of sdd in the [2 – 20] track segment for Huelva (Fig. 2.4.2.a) and Bonanza (Fig. 2.4.27.b) TG stations. Two datasets were compared: SLA without the SSB correction (blue bars) and with the SSB (red bars).

2.4.4.2 Analysis of dry and wet tropospheric corrections from UPorto

No differences were found in terms of rmse between the two options analysed. The summary of the average sdd for the 5 retrackers is shown in Figure 2.4.28. This analysis was made applying the SSB correction.

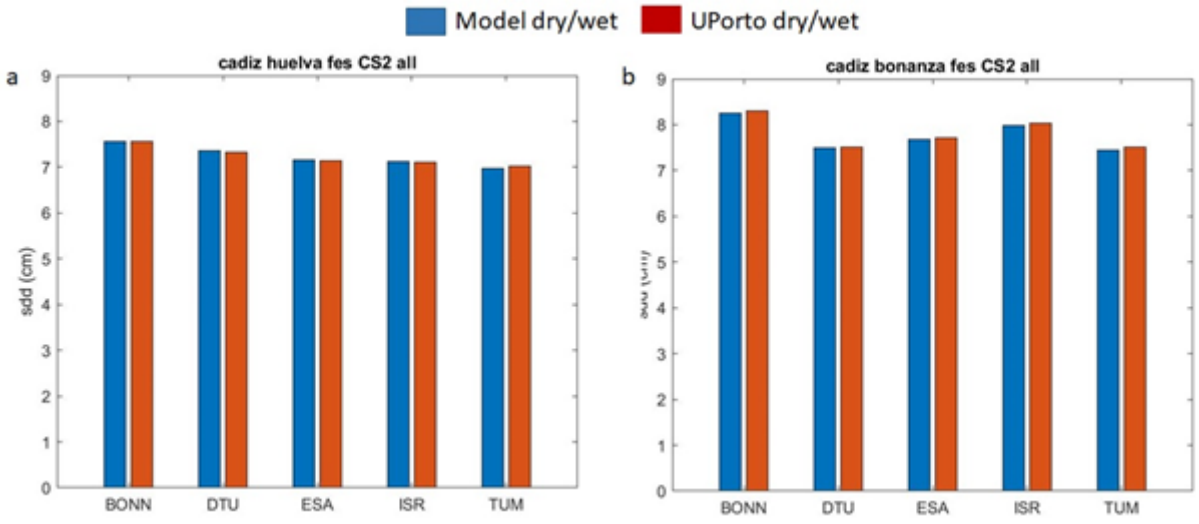


Figure 2.4.28. Along-track average of sdd in the [2 – 20] km track segment using dry/wet tropospheric corrections from input files (blue bars) and UPorto corrections (red bars). The analysis was made using Huelva (Fig. 2.4.28.a) and Bonanza (Fig. 2.4.28.b) TG stations.

2.4.4.3 Summary of results

Table 2.4.6 presents a summary of average sdd and r in the [3 – 20] track segment. ESA, ISR and TUM retrackers show the best performances in terms of sdd and r.

Table 2.4.6. Average of sdd and r for the tracks analysed using Huelva and Bonanza TG stations.

	Huelva		Bonanza	
	sdd	r	sdd	r
BONN	7.6	0.70	8.3	0.64
DTU	7.3	0.70	7.5	0.69
ESA	7.1	0.75	7.7	0.70
ISR	7.1	0.73	8.0	0.66
TUM	7.0	0.73	7.5	0.69

2.4.5 Significant Wave Height (SWH): CS2

The analysis of SWH was made using the Gulf of Cadiz buoy. The track selection was made considering a radius of 30 km from the buoy position. Figure 2.4.29 gives the position of the closest five measurements selected for the comparison of all the tracks available in CS2. Only two retrackers gave SWH as outputs: BONN and ISR.

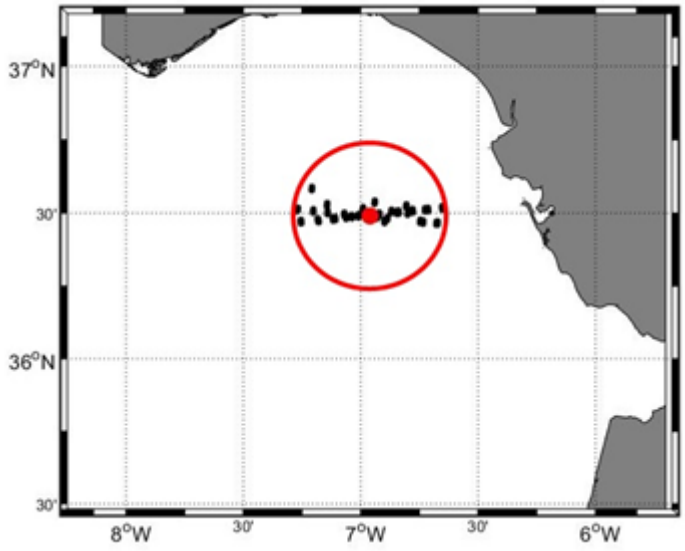
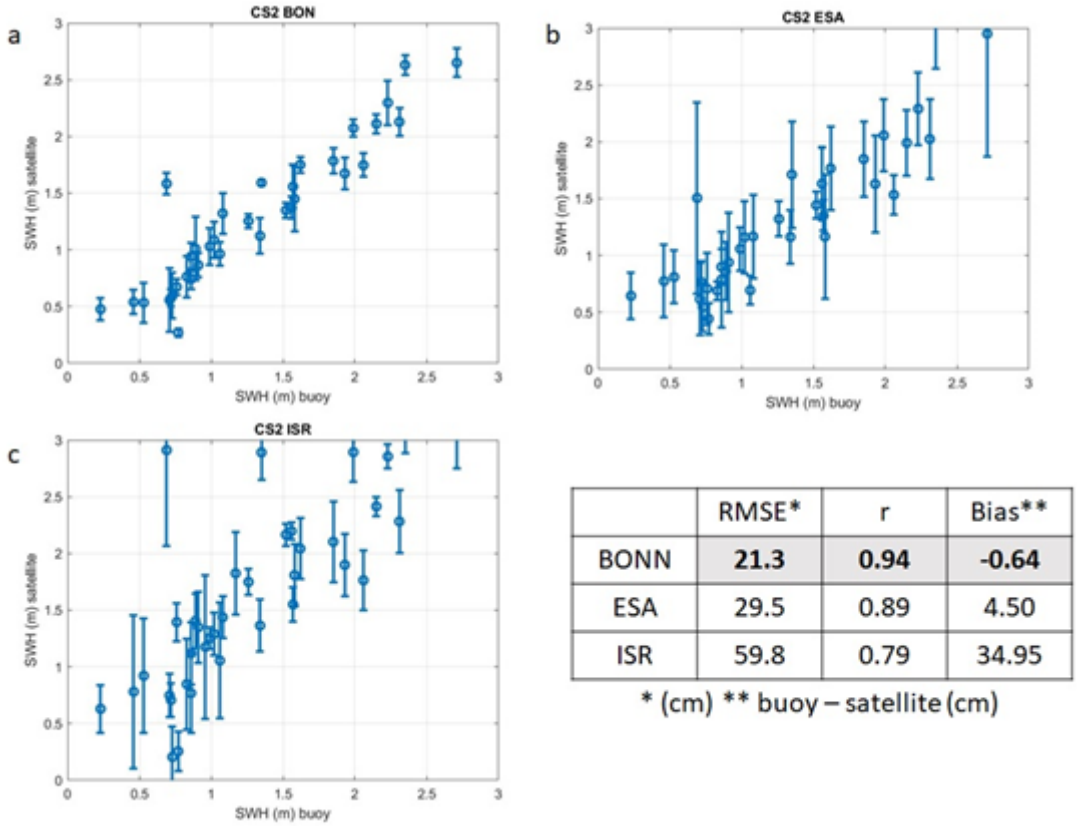


Figure 2.4.29. Closest measurements (5) of the tracks inside a radius of 30 km from the position of the Gulf of Cadiz buoy.

Time series of SWH from CS2 were made by averaging the closest (5) SWH values to the buoy for each track. The comparison between buoy / satellite pairs of SWH are shown in Fig. 2.4.30 for BONN (Fig. 2.4.30.a), ESA (Fig. 2.4.30.b) and ISR (Fig. 2.4.30.c) retrackers. The standard deviation of the average SWH (satellite) and some statistics (bias, RMSE and r) are also given. BONN retracker gives better scores than ESA and ISR as in S3A/S3B satellites.



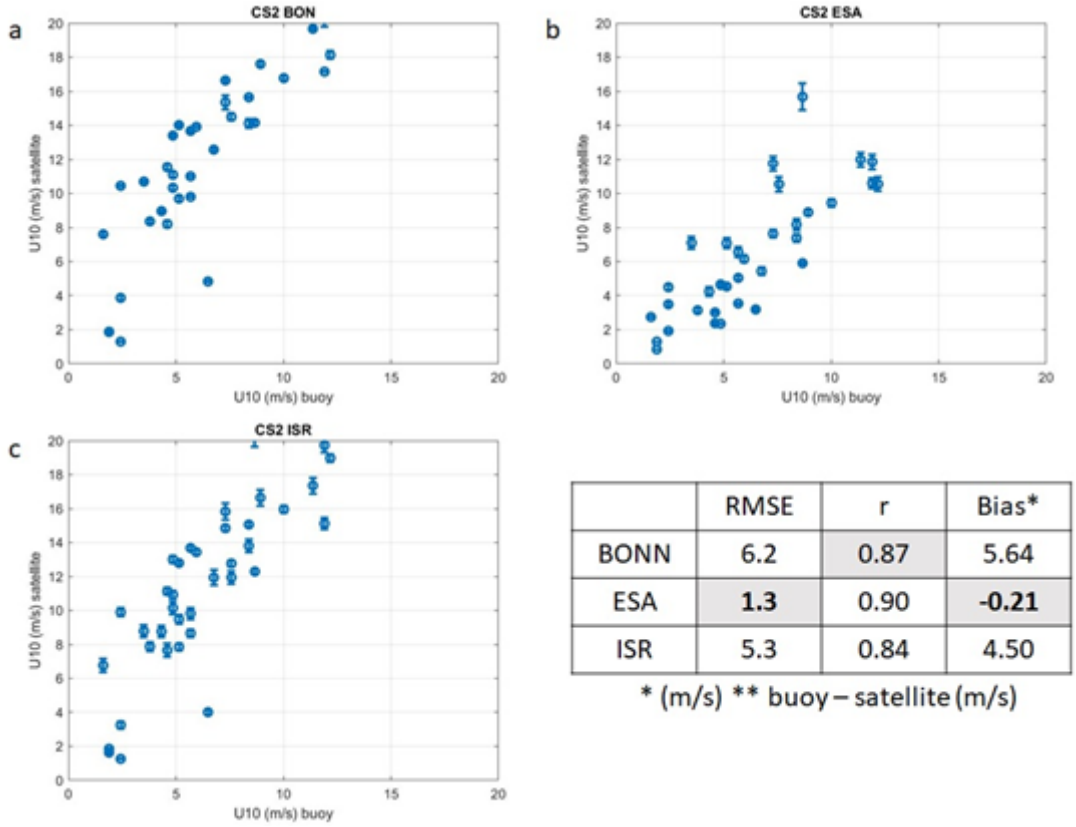
	RMSE*	r	Bias**
BONN	**21.3**	**0.94**	**-0.64**
ESA	29.5	0.89	4.50
ISR	59.8	0.79	34.95

* (cm) ** buoy – satellite (cm)

Figure 2.4.30. Scatter plot for CS2 (all the tracks inside the 30-km radius around the buoy are considered) using BONN (Fig. 2.4.30.a), ESA (Fig. 2.4.30.b) and ISR (Fig. 2.4.30.c) retrackers. The table gives the statistics.

2.4.6 Wind Speed (U10): CS2

The same procedure used in SWH was applied to U10 data analysis in order to obtain the time series. U10 was estimated by correcting the raw sigma0 by the atmospheric attenuation and applying Abdalla's algorithm. Regarding data screening, an outlier detection using the 3-sigma criterion was used. Figure 2.4.31 shows the scatter plot for the three retrackers. ESA retracker gives the best scores in terms of RMSE. BONN and ISR show a strong bias. This should be investigated further.



	RMSE	r	Bias*
BONN	6.2	0.87	5.64
ESA	**1.3**	0.90	**-0.21**
ISR	5.3	0.84	4.50

* (m/s) ** buoy – satellite (m/s)

Figure 2.4.31. Same as Fig. 2.4.30 for wind speed.

2.4.7 Summary Tables and Results

Below we summarise the key findings, according to the agreed metrics.

Sea Level Anomaly (FES model, UPorto DTC/WTC, SSB applied)

Table 2.4.7. Sentinel-3A SLA. Southwestern Spanish coast. Distance to coast: [2 – 20] km. Comparison against three tide gauges: Huelva, Bonanza, Tarifa.

S3A	Average STDD ± 1σ (cm)	Correlation
ESA	6.8 ± 1.8	0.70
TUM	7.0 ± 1.8	0.70
ISR	7.1 ± 1.9	0.69
Bonn	7.3 ± 2.1	0.65
DTU	8.1 ± 1.8	0.65

Table 2.4.8 Sentinel-3B SLA. Southwestern Spanish coast. Distance to coast: [2 – 20] km. Comparison against three tide gauges: Huelva, Bonanza, Tarifa.

S3B	Average STDD ± 1σ (cm)	Correlation
Bonn	7.0 ± 2.0	0.70
TUM	7.0 ± 1.7	0.68
ESA	7.3 ± 2.1	0.66
ISR	7.4 ± 2.0	0.64
DTU	9.3 ± 2.7	0.56

Table 2.4.9 CryoSat-2 SLA. Southwestern Spanish coast. Distance to coast: [2 – 20] km. Comparison against two tide gauges: Huelva, Bonanza.

CS2	Average STDD ± 1σ (cm)	Correlation
TUM	7.3 ± 2.9	0.71
ESA	7.4 ± 3.1	0.72
DTU	7.4 ± 2.7	0.70
ISR	7.6 ± 4.0	0.69
Bonn	7.9 ± 1.9	0.67

Significant Wave Height

Table 2.4.10 Sentinel-3A SWH. Gulf of Cadiz buoy.

S3A	RMSE (cm)	Correlation	Bias (buoy – satellite) (cm)
Bonn	16.1	0.95	-9.01
ESA	18.4	0.95	-7.38
ISR	26.5	0.83	-12.07

Table 2.4.11 Sentinel-3B SWH. Gulf of Cadiz buoy.

S3B	RMSE (cm)	Correlation	Bias (buoy – satellite) (cm)
Bonn	21.8	0.94	-9.09

ESA	28.3	0.89	0.57
ISR	33.8	0.85	-3.99

Table 2.4.12 CryoSat-2 SWH. Gulf of Cadiz buoy.

CS2	RMSE (cm)	Correlation	Bias (buoy – satellite) (cm)
Bonn	21.3	0.94	-0.64
ESA	29.5	0.89	4.50
ISR	59.8	0.79	34.95

Wind speed (U10)

Table 2.4.13 Sentinel-3A U10. Gulf of Cadiz buoy.

S3A	RMSE (m/s)	Correlation	Bias (buoy – satellite) (m/s)
ESA	1.1	0.93	0.07
Bonn	1.4	0.87	0.25
ISR	1.6	0.87	0.78

Table 2.4.14 Sentinel-3B U10. Gulf of Cadiz buoy.

S3B	RMSE (m/s)	Correlation	Bias (buoy – satellite) (m/s)
ESA	1.4	0.88	-0.23
Bonn	1.8	0.78	-0.31
ISR	3.5	0.78	-2.67

Table 2.4.15 CryoSat-2 U10. Gulf of Cadiz buoy.

CS2	RMSE (m/s)	Correlation	Bias (buoy – satellite) (m/s)
ESA	1.3	0.90	-0.21
ISR	5.3	0.84	4.50
Bonn	6.2	0.87	5.64

- The SLA validation for S3A and S3B is based on 3 tide gauges.
- For S3A, S3B and CS2 SLA, all retrieved STDD lay within 1sigma.
- For S3A SLA, the ESA data showed the “best” performance with the highest correlation and lowest STDD (6.8cm). The performance of TUM, ISR, and Bonn data was very similar (STDDs 7.0cm - 7.3cm). DTU showed the highest STDD (8.1cm).
- For S3B SLA, U Bonn and TUM showed the lowest STDD (7.0cm), followed by ESA (7.3cm), isardSAT (7.4cm) and DTU (9.3cm).
- For Cryosat-2 SLA, TUM showed the lowest STDD (7.3cm), followed by ESA and DTU (7.4cm), isardSAT (7.6cm) and U Bonn (7.9cm).
- For S3A, S3B and CS2 SWH (comparison against 1 buoy), U Bonn showed the lowest RMS, followed by the original ESA data and then isardSAT.
- For S3A and S3B U10 (comparison against 1 buoy), the original ESA data showed the lowest RMS, followed by U Bonn then isardSAT. There was some error in the calculation of U10 from Cryosat-2 data.

In conclusion, there was no clear pattern of one retracker performing better than the others in comparing Sea Level Anomaly data against tide gauge data, with a different order of results across the three satellite data sets (Sentinel 3A, Sentinel 3B and Cryosat-2).

For S3A and S3B SWH data, U Bonn data were found to give lower RMS values than ISR or ESA. Similarly, for Cryosat-2 SWH, U Bonn data showed a lower RMS than ISR

For S3A and S3B U10 data, the ESA data were found to give lower RMS values than UBonn or ISR, with ISR showing much higher RMS values for S3B.

2.5 Influence of land proximity and angle of approach (SatOC/SKYMAT)

2.5.1 Introduction

In this section we present the results of the analysis carried out by SKYMAT, looking at the influence of land proximity and angle of approach, on the uncorrected sea surface height (USSH).

We assess the performance of Sentinel 3 A & B CryoSat-2 from TUM, BONN, ISR, DTU and ESA (only for S3) by investigating the angle of approach to the coast line. We have analysed data from the German Bight / Southern Baltic Sea, and the California Regions of Interest

The angle of approach to the coast is computed by calculating the deviation angle between the direction of the satellite track and the direction of the gradient using the coast proximity parameter (Cipollini, 2011). Once the deviation angles associated with the USSH observations are computed, then they are binned at 30° intervals using less than 8 km distance to coast, hence a deviation angle value of 0-30° represents that the satellite track is normal to the coast, whereas 60-90° represents the satellite track that is parallel to the coast. Figure 2.5.1 illustrates how this is derived. Please note, that the coastal proximity parameter is a dimensionless measure of the effect of land over altimetric waveforms, where the values have a range between -1 and +1, such that a -1 refers to geographical locations that are unaffected by land (normally offshore) and 1 represents totally affected by land, hence close inshore. The coast proximity parameter is treated as an independent variable and has an advantage over the closest approach to the coast parameter of capturing differences in coastal morphology. In addition, we assess the deviation angle dependency in terms of data lost close to the coast, by computing the percentage of data lost as a function of deviation angle and distance to the coast.

The uncorrected sea surface height (USSH) is calculated from Orbit minus Range. No other corrections are applied. The noise is defined as successive differences of high frequency (20 Hz) USSH(n) observations along each of the tracks. For this analysis we provide this at 0.5km intervals.

In this document we present the results of the analysis, more complete details of the analysis procedure are given in the HYDROCOASTAL Product Validation Plan.

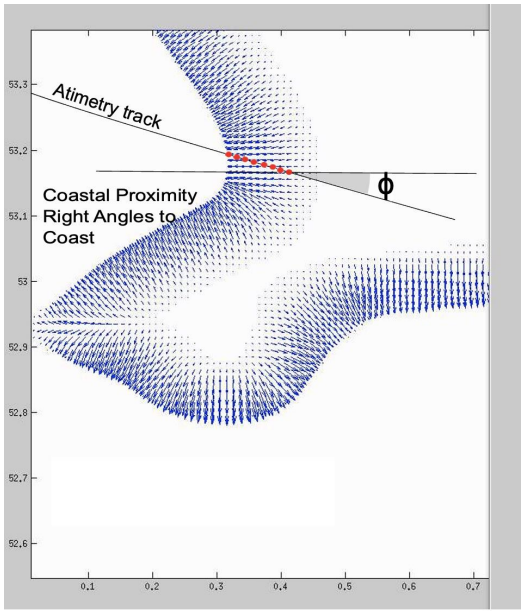


Figure 2.5.1. Illustration of how Angle of Approach (ϕ) is derived from the coastal proximity parameter (ref Cipollini, 2011)

2.5.2 Results: Angle of Approach and Coastal Proximity - Sentinel -3

The results of the analysis for the German Bight and Southern Baltic sea region, for the TUM, U Bonn, isardSAT, ESA and DTU retrackers are shown in Figure 2.5.2.

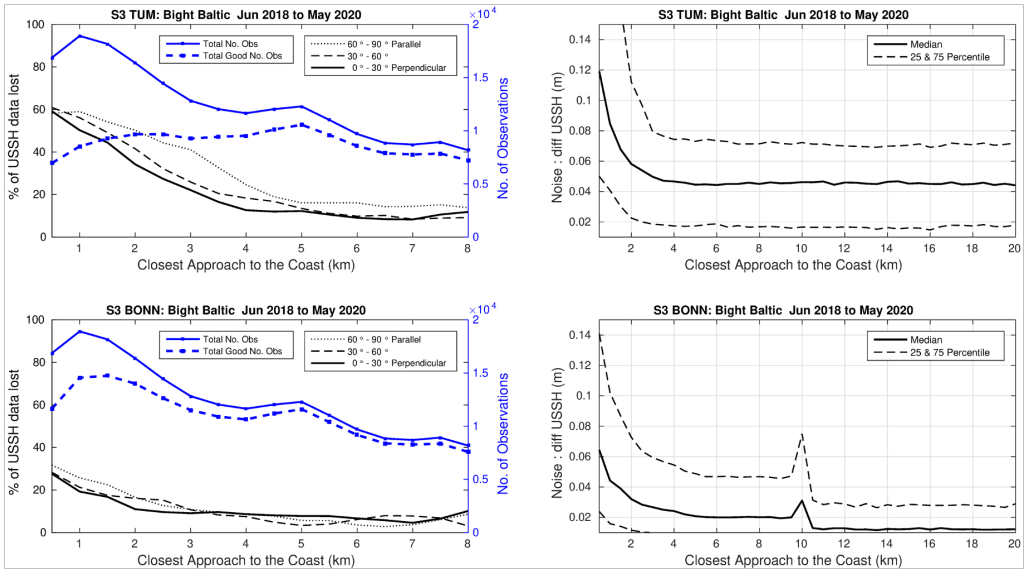
The TUM data show a relatively level noise value of just over 0.04m from 20 to 4km, rising to 0.06m at 2km, and then 0.12 at 1km. The percentage of data lost is ~10% for all angles of arrival from 8km to 5km, then rises to 40% lost for tracks with AoA 60°-90° at 3km, whereas it is ~20% for other AoA. Within 1km of the coast the loss of data for all AoA is 50% or more

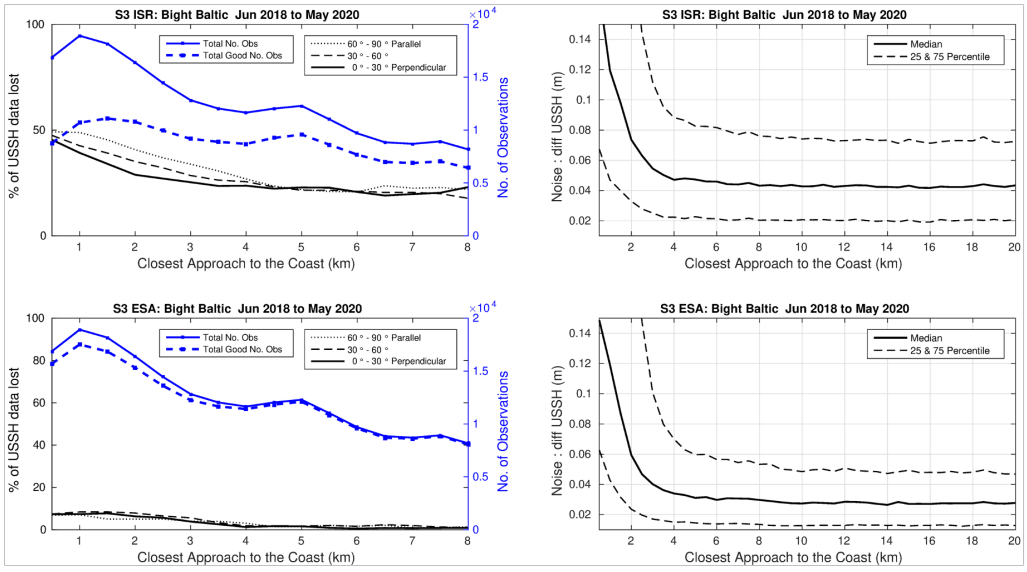
The UBonn data show a noise level of 0.01m at distances 20-11km from the coast, then there is a peak in noise level at 10km (0.03m), then 0.02m from 10km to 5km, before rising to 0.03m at 2km and 0.06m at 1km. The percentage of data lost is ~10% for all AoA to 3km, then rises to 20% for all AoA at 1km. There is no strong evidence for different behaviours for different AoA.

The isardSAT data show a relatively level noise value of just over 0.04m from 20 to 4 km, rising to 0.07m at 2km, and then 0.15 at 1km. The percentage of data lost is ~20% for all angles of arrival from 8km to 4km, then rises to 40% lost for tracks with AoA 60°-90° at 3km, 35% for AoA between 30°-60° and 30% for AoA between 0°-30°. Within 1km of the coast the loss of data for all AoA is ~50%. Note that only flags 1 and 2 were used to filter data. The use of flags 1-3 was recommended, but this was found to reduce the amount of data available severely.

The DTU data show a relatively level noise value of just over 0.06m from 20 to 4 km, rising to 0.07m at 2km, and then 0.11 at 1km. The percentage of data lost is very low, under 10% up to within 1km of the coast. There is no evidence of dependency on AoA.

Finally, the “standard” ESA data show a level noise value of 0.03m from 20 to 4 km, rising to 0.06m at 2km, and then 0.15 at 1km. Again, the percentage of data lost is very low, under 10% up to within 1km of the coast. There is no evidence of dependency on AoA.





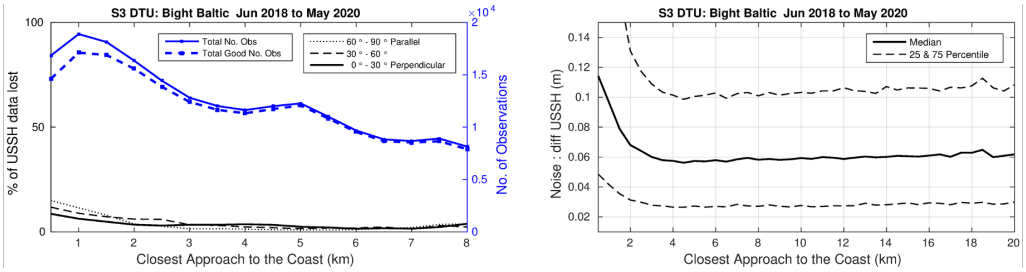


Figure 2.5.2. Results for Sentinel 3, top to bottom: TUM, U Bonn, ISR, ESA and DTU, for the German Bight and Southern Baltic. Left panels give percent of data lost (black) and number of observations (blue). Right panels give noise in uncorrected sea surface height.

The results of the analysis for the California coastal region, for the TUM, U Bonn, isardSAT, ESA and DTU retrackers are shown in Figure 2.5.3.

The TUM data show a relatively level noise value of 0.05m from 20 to 2km, rising to just under 0.1 at 1km. These data show a high dependency on AoA. The percentage of data lost is under 10% for all angles of arrival at 8km, but then rises to 30% lost for tracks with AoA 60°-90° from 5.5km to 3km and 50% at 2km. Data from tracks with AoA between 30°-60° have data loss of 25% at 3km and 40% at 2km. Data on tracks with AoA have a data loss of 15% at 3km, and 20% at 2km. Within 1km of the coast the loss of data for AoA > 30° is 60%, and 50% for tracks with AoA between 0° and 30°.

The UBonn data show a noise level of under 0.02m at distances 20-11km from the coast, then there is a peak in noise level at 10km (0.03m), then just over 0.02m from 10km to 5km, before rising to 0.03m at 2km and 0.05m at 1km. The percentage of data lost is 10% or under for AoA 0°-60° at all distances. It is higher for AoA 60°-90° within 7km of the coast, though under 20% until under 1km.

The isardSAT data show a relatively level noise value of just over 0.05m from 20 to 6 km, rising to 0.07m at 2km, and then 0.13 at 1km. The percentage of data lost is ~20% for all angles of arrival from 8km to 7km, then higher (20%) for tracks with AoA 60°-90° between 3.5km to 6km, and ~15% for AoA between 0°-60°. The percent data lost then rises for all AoA, becoming 40% for AoA 60°-90°, and 30% for other AoA at 1km from the coast

The DTU data show a slight decrease in noise approaching the coast, 0.09m at 20km and 0.07 at 4km. The USSH noise then increases to 0.08m at 2km, and 0.11 at 1km. The percentage of data lost for AoA 0°-30° remains at under 10% for all distances. It is higher for AoA of 30°-60° (10% at 4km, 20% at 1km), and 60°-90° (20% at 4km, and 25% at 1km)

Finally, the “standard” ESA data show a level noise value of just under 0.04m from 20 to 6 km, rising to 0.05m at 2km, and then 0.11 at 1km. The percentage of data lost is very low, under 10% up to within 1km of the coast. There is no evidence of dependency on AoA.

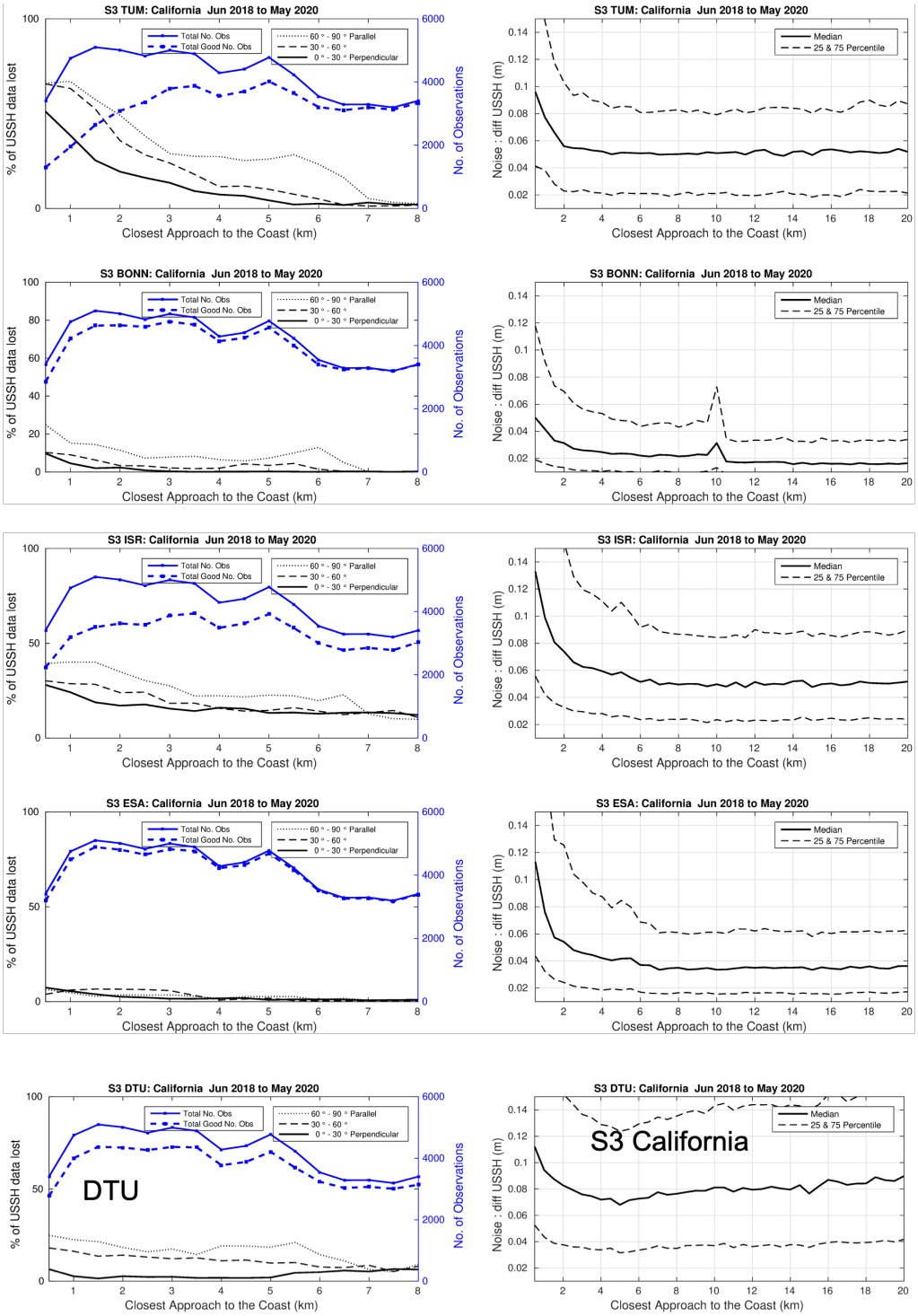


Figure 2.5.3. Results for Sentinel 3, top to bottom: TUM, U Bonn, ISR, ESA and DTU, for the California coast region of interest. Left panels give percent of data lost (black) and number of observations (blue). Right panels give noise in uncorrected sea surface height.

2.5.3 Results: Angle of Approach and Coastal Proximity - Cryosat-2

The results of the analysis of Cryosat-2 data for the German Bight and Southern Baltic sea region, for the TUM, U Bonn, isardSAT, ESA and DTU retrackers are shown in Figure 2.5.4.

The TUM data show a relatively level noise value of just over 0.04m from 20 to 4km, rising to 0.06m at 2km, and then 0.11 at 1km. The percentage of data lost is similar for all AoA, ~10% for all angles of arrival at 8km, rising to 30% at 3km, 50% at 2km, and 60% at 1km.

The UBonn data show a noise level of just under 0.02m at distances 20-11km from the coast, then there is a peak in noise level at 10km (0.03m), then just under 0.03m from 10km to 5km, before rising to 0.06m at 1km. The percentage of data lost is similar for all AoA, ~10% for all angles of arrival at 8km, rising to 20% at 2km, and 30% at 1km.

The isardSAT data show a relatively level noise value of just over 0.04m from 20 to 5 km, rising to 0.07m at 2km, and then 0.15 at 1km. The percentage of data lost is similar for all AoA, ~10% for all angles of arrival at 8km, rising to 30% at 3km, and 45% at 1km.

The DTU data show a relatively level noise value of just over 0.06m from 20 to 4 km, rising to 0.11 at 1km. The percentage of data lost is very low, under 10% up to within 1km of the coast. There is no evidence of dependency on AoA.

The results of the analysis for Cryosat-2 data for the California coastal region, for the TUM, U Bonn, isardSAT, ESA and DTU retrackers are shown in Figure 2.5.5.

The TUM data show a relatively level noise value of 0.05m from 20 to 2km, rising to 0.09m at 1km. The data show less dependence on AoA for this region than did the Sentinel-3 data. Data loss for all AoA is under 20% from 8km to 4km, rising to 40% at 2km and 60% at 1km.

The UBonn data show a noise level of just over 0.02m at distances 20-11km from the coast, then there is a peak in noise level at 10km (0.04m), then just over 0.03m from 10km to 5km, rising to 0.07m at 1km. The percentage of data lost is 20% or under for all AoA from 8km to 3km, increasing to ~30% at 1km.

The isardSAT data show a relatively level noise value of just over 0.05m from 20 to 5 km, rising to 0.13m at 1km. The percentage of data lost is ~20% for all angles of arrival from 8km to 2km, rising to 30% at 1km.

The DTU data show a slight decrease in noise approaching the coast, 0.09m at 20km and 0.08 at 2km. The USSH noise then increases to 0.11 at 1km. The percentage of data lost for all AoA and distances from the coast is 10%-20%.

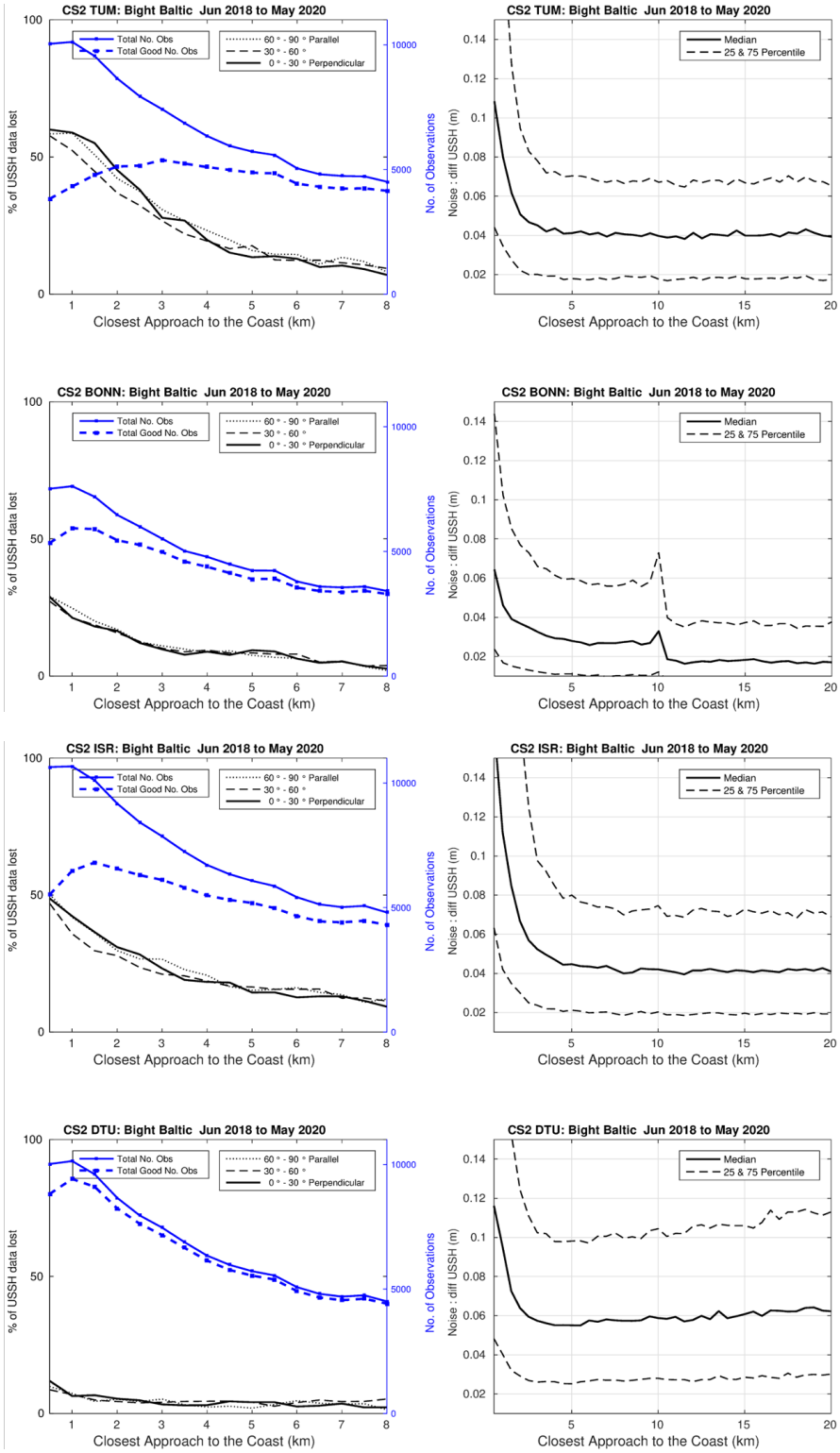


Figure 2.5.4. Results for Cryosat-2, top to bottom: TUM, U Bonn, ISR, and DTU, for the German Bight, Southern Baltic region of interest. Left panels give percent of data lost (black) and number of observations (blue). Right panels give noise in uncorrected sea surface height.

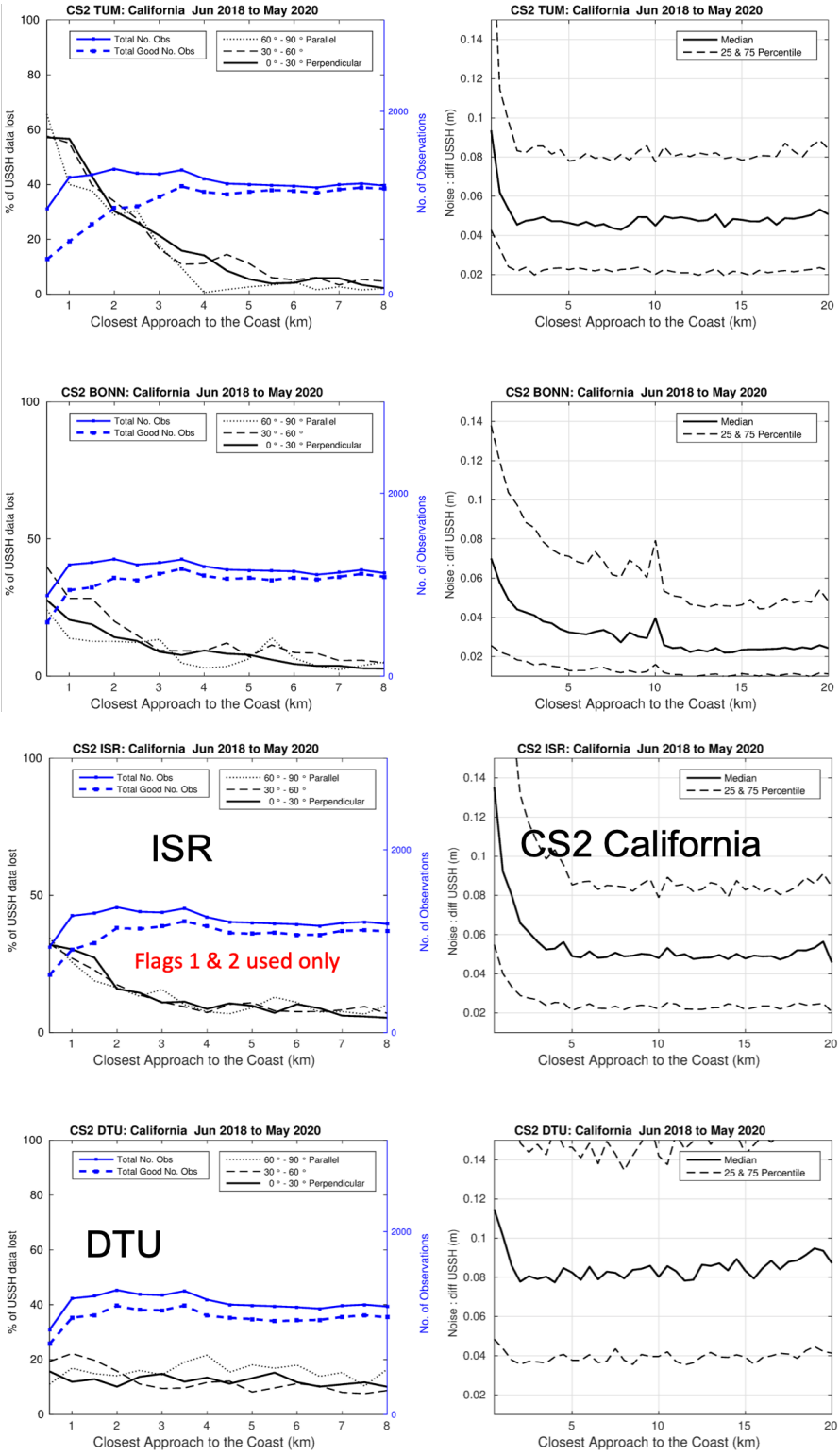


Figure 2.5.5. Results for Cryosat-2, top to bottom: TUM, U Bonn, ISR, and DTU, for the California coast region of interest. Left panels give percent of data lost (black) and number of observations (blue). Centre panels give noise in uncorrected sea surface height. Right panels have some information about flags.

2.5.4 Conclusions

General Conclusions

The analyses find largely consistent results for each re-tracker across the two regions considered (German Bight and Southern Baltic Sea, and the Californian Coast), in that more data are lost approaching the coast, and the noise in uncorrected sea surface height (USSH) increases towards the coast.

In general, for Sentinel-3 data, satellite tracks parallel to the coast appear to lose more data than those perpendicular to the coast. However, this difference is not observed so clearly in Cryosat-2 data.

Both Sentinel-3 & Cryosat-2 start to lose more data less than 5 km from the coast and rapidly lose data at less than 4 km.

Comparing Re-Trackers

In terms of the four new processors under evaluation for coastal data, the U. Bonn data show the lowest noise values, though there is a “spike” at 10 km from the coast. We understand this is due to a change in mode from ocean to coastal in the processing.

The TUM re-tracker came a close second, giving a consistent data loss approaching the coast as well as a steady noise level. The isardSAT data showed the next lowest noise levels. The DTU data showed the highest noise levels, but lower data loss. The ‘standard’ ESA product performed similarly to the isardSAT and TUM products.

Data Selection

Some analysis was done on options for data selection. The flagging for CS2 data appears to let through high noise values.

It was unclear if the good flags used from ISR are correct (i.e. Flags,1,2 &3) as when applied, they gave no loss of data towards the coast.

Alternatively, a threshold of 0.3m USSH noise gives a similar result of data loss towards the coast and median noise level.

2.5.5 Final Summary Tables

Below we give summary tables giving order of performances according to the agreed metrics.

Table 2.5.1. Sentinel 3A and 3B USSH (m) Noise v proximity [Bight –Baltic / California]

Order	Name	1 km	2 km	4 km	5-10 km	11-20 km
1	U. BONN	0.044 / 0.042	0.032 / 0.031	0.024 / 0.025	0.021 / 0.023	0.012 / 0.016
2	TUM	0.085 / 0.077	0.058/ 0.056	0.047 / 0.052	0.045 / 0.051	0.045 / 0.052
3	DTU	0.097 / 0.094	0.068 / 0.083	0.058 / 0.072	0.058 / 0.076	0.061 / 0.083
4	ESA	0.119 / 0.076	0.059 / 0.054	0.034 / 0.042	0.030 / 0.036	0.028 / 0.035
5	ISR	0.120 / 0.099	0.074 / 0.074	0.047 / 0.059	0.044 / 0.051	0.043 / 0.050

Table 2.5.2. Sentinel 3A and 3B % data loss v proximity, perpendicular to coast [Bight –Baltic / California]

Order	Name	1 km	2 km	4 km	8 km
1	DTU	6 / 3	3 / 3	4 / 2	4 / 6
2	ESA	7/5	6/3	1/2	1/1
3	U. BONN	19 / 5	11 / 2	9 / 0	10 / 0
4	ISR	39 / 24	29 / 18	24 / 16	23 / 12
5	TUM	50 / 38	34 / 19	13 / 7	12 / 2

Table 2.5.3. Cryosat-2 USSH (m) Noise v proximity [Bight –Baltic / California]

Order	Name	1 km	2 km	4 km	5-10 km	11-20 km
1	U. BONN	0.043 /0.051	0.035 /0.042	0.029 /0.036	0.027 /0.031	0.017 /0.024
2	TUM	0.065 /0.060	0.049 /0.046	0.042 /0.047	0.040 /0.046	0.040 /0.048
3	ISR	0.068 /0.066	0.053 /0.056	0.044 /0.050	0.041 /0.049	0.041 /0.049
4	DTU	0.066 /0.075	0.055 /0.066	0.053 /0.068	0.055 /0.074	0.059 /0.078

Table 2.5.4. Cryosat-2 % data loss v proximity, perpendicular to coast [Bight –Baltic / California]

Order	Name	1 km	2 km	4 km	8 km
1	U. BONN	25/24	19/14	10/11	3/4
2	DTU	27/32	13/15	5/17	4/11
3	ISR	33/31	26/17	11/9	3/1
4	TUM	47/41	35/22	15/12	3/1

2.6 Synthesis of Coastal Zone Validation Results. Conclusions and Recommendations.

2.6.1 Synthesis of Results - Sea Level Anomaly / Sea Surface Height

U Bonn, NOC and U Cadiz evaluated the Sea Surface Height provided by the different re-trackers by comparing the derived Sea Level Anomaly (SLA) against tide gauge data. Skymat evaluated the re-tracker performance by considering the along track noise in the uncorrected Sea Surface Height (USSH), defined as successive differences of high frequency (20 Hz) USSH(n) observations along each of the tracks. For this analysis it was provided at 0.5km intervals.

U Bonn

U Bonn evaluated SAR altimeter SLA data against in situ data from 9 tide gauges in the German Bight and Southern Baltic, with SSB applied, calculated from 5% of the re-tracker SWH, or using the re-tracker's own SSB (TUM).

The results for SLA for S3A and S3B are similar. For Sentinel 3A, the UBonn re-tracker shows the "best" performance with the highest correlation and lowest STDD (12 cm STDD) followed by DTU (15 cm STDD). For Sentinel 3B, DTU and U Bonn have the best (and same) STDD (13 cm).

For Cryosat-2, DTU is the best performing (25 cm) STDD followed by Bonn (27 cm).

NOC

NOC evaluated SAR altimeter SLA data against in situ data from 4 tide gauges on the California coast, with SSB applied, calculated from 5% of the re-tracker SWH, or using the re-tracker's own SSB (TUM).

For S3A SLA, TUM showed the "best" performance with the highest correlation and lowest STDD. The performance of ISR, DTU and the original ESA data was very similar with STDD's within the range of the Mean Absolute Differences. U Bonn showed the lowest correlation and highest STDD, but with a larger range of Mean Absolute Differences (no doubt due to the lower correlation with TG data from San Francisco and La Jolla).

For S3B SLA, U Bonn showed the "best" performance with the highest correlation and lowest STDD. The performance of DTU, the original ESA data, and isardSAT was similar with STDD's within the range of the Mean Absolute Differences. This time, in contrast to S3A, TUM showed the lowest correlation and highest STDD, but with a larger range of Mean Absolute Differences.

For Cryosat-2 SLA, with comparisons against only 1 tide gauge (so no MAD), the ESA data showed the highest correlation and lowest STDD, followed by DTU, ISR, and U Bonn. No results were retrieved for TUM.

U Cadiz

U Cadiz evaluated SAR altimeter SLA data against in situ data from 3 tide gauges on the south Spanish coast, with SSB applied, calculated from 5% of the re-tracker SWH, or using the re-tracker's own SSB (TUM).

There was no clear pattern of one retracker performing better than the others in comparing Sea Level Anomaly data against tide gauge data, with a different order of results across the three satellite data sets (Sentinel 3A, Sentinel 3B and Cryosat-2). The variation in retrieved STDD across re-trackers lay within 1 sigma.

Skymat

The Skymat analysis of along-track noise against proximity to the coast found that the UBonn data had lower values of noise in Uncorrected Sea Surface Height from 11-20km from the coast, to 1km from the coast. TUM showed the next lowest values of noise, followed by isardSAT and then DTU.

When looking at the percentage of data lost, DTU and U Bonn retrieved a higher percentage of valid data, followed by isardSAT and then TUM.

2.6.2 Synthesis of Results – Significant Wave Height and Ocean Surface Wind Speed (referenced to 10m)

Significant Wave Height

In comparisons of SWH against buoy data in the German Bight and Southern Baltic, by U Bonn, again different results were found for Sentinel 3A (8 buoys) and Sentinel 3B data (6 buoys). U Bonn showed the lowest RMS for S3A followed by ISR and the original ESA data. For S3B the smallest RMS is from ESA, followed by ISR and Bonn. For CryoSat-2 SWH (8 buoys), UBonn showed the lowest RMS followed by ISR and the original ESA data

The NOC evaluation of SWH against buoys in the California coastal region found that for S3A and S3B (comparison against 21 buoys), U Bonn showed the lowest RMS, followed by IsardSAT and the original ESA data. For Cryosat-2 SWH (comparison against 18 buoys), U Bonn showed a significantly lower RMS than ISR. SWH values were not available from the ESA data

Similar results were found from the SWH validation in Southern Spain (Cadiz buoy), with the UBonn re-tracker showing the lowest RMSE, followed by the original ESA data and then isardSAT.

10m Ocean Surface Wind Speed (U10)

In the UBonn evaluation of U10, wind speeds calculated from S3A and S3B data were validated against data from 11 buoys, and the original ESA data were found to show the lowest RMS, followed by U Bonn and then isardSAT.

In the NOC evaluation of U10 against buoys in the California coastal region for S3A and S3B data (comparison against 11 buoys), the original ESA data showed the lowest RMS, followed by U Bonn and isardSAT.

Again similar results were found from the U10 validation in Southern Spain (Cadiz buoy), with the original ESA data showing the lowest RMS, followed by U Bonn and isardSAT.

2.6.3 Conclusions and Recommendations

In conclusion, the evaluation of Sea Surface Height / Sea Level Anomaly was not able to identify one re-tracker that was consistently better performing than the others across the different regions of interest (German Bight / S Baltic, California, Southern Spain) and different satellites (S3A, S3B, and CS2). Similarly no re-trackers were found to consistently perform worse than the others.

Only the UBonn and the isardSAT re-tracker retrieve SWH and backscatter (from which wind speed is derived), and the analysis of SWH and wind speed data showed that the UBonn data showed a lower RMSE than the isardSAT data when compared to buoy data.

Finally, the Skymat analysis showed that the UBonn and DTU re-trackers retrieved the highest percentage of valid data as the coastline was approached.

The TUM and DTU re-trackers do not retrieve significant wave height or surface backscatter, from which the wind speed can be estimated, and these parameters are of interest in the coastal zone and required as part of a SAR altimeter coastal product. The analyses did not show that the TUM or DTU re-tracker provided significantly better performance in terms of estimation of sea surface height, so there is not a strong argument to implement different re-trackers for the retrieval of sea-surface height and sea-state parameters (SWH, U10).

Therefore, based on the superior performance of the U Bonn re-tracker in retrieving SWH and U10, and the fact that it also retained a higher percentage of valid data closer to the coast, we recommend that the UBonn re-tracker is implemented in the next phase of the project to generate the coastal zone data set. We note the issue with a discontinuity in performance at 10km from the coast, and recommend that U Bonn is asked to investigate an improved implementation to address this issue.

We also recommend that the coverage of the selected re-tracker for inland waters is extended to cover tidal estuaries close to the coast to provide continuity with inland water products.

It is important to note in assessing the results, that the processing of data for HYDROCOASTAL has generated a data set with a different datation to the original input Sentinel 3A, 3B and Cryosat-2 data. To match this datation, the ESA product L2 data have been interpolated to the new locations, which results in a certain amount of smoothing. This means that in the analyses the ESA data may demonstrate a lower along track noise and apparently lower variability, so the results of the analyses for the ESA data are not directly comparable to those of the data from the new HYDROCOASTAL re-trackers.

Nonetheless, none of the new re-trackers were found to perform significantly better than the original ESA product, therefore from this validation study, we cannot make a recommendation that another re-tracker should be implemented on the satellite ground segment processing chains.

3 Validation of the Test Datasets in different Inland Water Scenarios

3.1 Introduction

In this section, we present the results of the validation of the Test Dataset Geophysical parameters against other satellites and in situ data, in different Inland Water Scenarios.

These validation activities include the analysis of the influence of land proximity and ground-track orientation on SAR/SARin, analysis of retracking algorithms, analyses of the different algorithms proposed to produce the final datasets of water level (L3) and river discharge (L4), and validation results against gauging data.

A common methodology to define the validation metrics for the Inland Waters has been described in the PVP document.

3.2 Validation on the Rhine and Elbe rivers (U Bonn)

Each L3 Virtual Station (VS) in Rhine is validated with the closest in-situ station within a 30 km distance. For the validation, 15 minutes water level provided by BFG are used and interpolated in time to the L3 time steps. Outliers from the VS time series are eliminated with a 4-sigma criterion.

Statistical Metrics are:

- distribution of the bias-corrected RMSE between gauge and altimetry
- median of the above RMSE
- 95% CI for the median of RMSE (with bootstrap approach)
- NRMSE normalized RMSE
- NSE Nash-Sutcliffe coefficient
- R2 correlation
- MAD Median absolute deviation

Figure 3.2.1 shows the boxplot of RMSE(m) for the stations with data from all retrackers (10 stations), Default boxplot of matlab with 25th and 75th percentile (left) and with 5th and 95th percentile (right). We see differences between the two selections in the outlier identification.

Table 3.2.1 gives the confidence Interval (CI) at common stations and at all stations separately. Table 3.2.2 gives the same for the median. The number of stations used in the two cases are different because the ID of the time series is not the same as seen in Figure 3.2.2.

The final statistics in Tables 3.2.4 and 3.2.5 are very similar. For the common stations, TUM gives better NRMSE, NSE, R2, MAD (Table 3.2.4). For all the stations, DTU has the larger number of VSs and the best statistics in terms of RMSE, NRMSE, MAD (Table 3.2.5). Our recommendation is to select the DTU retracker for inland water.

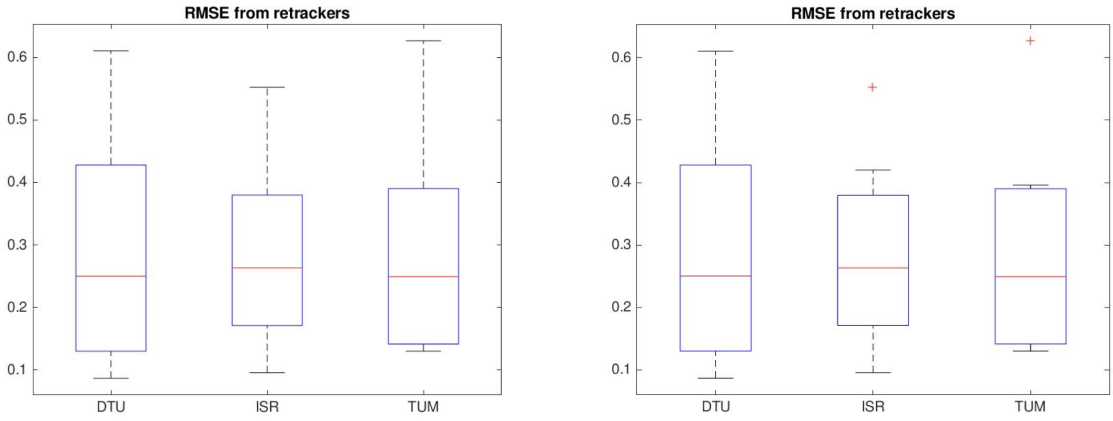


Figure 3.2.1. Boxplot of RMSE(m) for the stations with data from all retrackers (10 stations), Default boxplot of matlab with 25th and 75th percentile (left) and with 5th and 95th percentile (right)

Table 3.2.1. Confidence Interval (CI) at common stations and at all stations separately

	Common stations for all retrackers		All Stations	
	Lower CI [m] (0.025th q.)	Upper CI [m] (0.975th q.)	Lower CI [m] (0.025th q.)	Upper CI [m] (0.975th q.)
DTU	0.13	0.43	0.19	0.44
TUM	0.14	0.39	0.29	0.50
ISR	0.17	0.38	0.25	0.41

Table 3.2.2. Median of RMSE at common stations and at all stations separately

	Median of RMSE [m] (Common stations for all retrackers)	Median RMSE [m] (All Stations)
DTU	0.25	0.31
TUM	0.25	0.35
ISR	0.26	0.35

Table 3.2.3. Number of stations and records at common stations and at all stations separately

	Common stations for all retrackers		All Stations	
	# Stations	# Records	# Stations	# Records
DTU	10	243	21	474
TUM	10	244	18	384
ISR	10	241	17	384

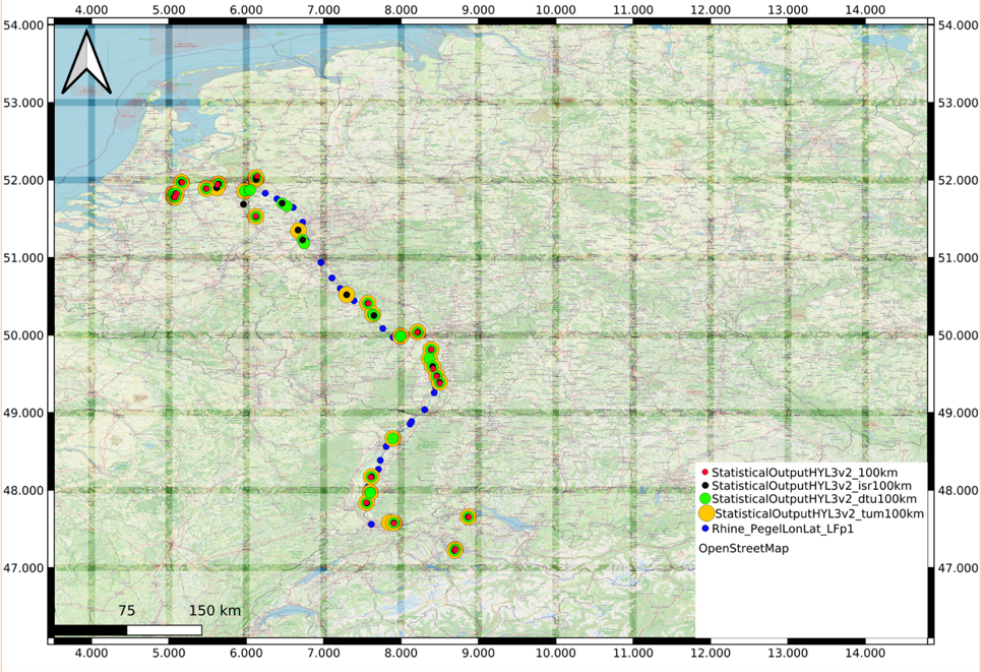


Figure 3.2.2. Location of VS. Common Virtual Stations (red), in-situ stations (blue), and VS for each retracker: DTU (green), TUM (orange) and ISR (black). We select 30 km, larger distances involve different hydrologic conditions.

Table 3.2.4. Statistics for median for common stations

MEDIAN	Common stations for all retrackers							
	RMSE[m]	NRMSE[%]	NSE	R2	MAD	95% CI,median lower	95% CI,median upper	n Stations
DTU	0.25	8.82	0.89	0.95	0.18	0.13	0.43	10
TUM	0.25	8.42	0.93	0.97	0.16	0.14	0.39	10
ISR	0.26	8.54	0.91	0.96	0.16	0.17	0.38	10

Table 3.2.5. Statistics for median for all stations

MEDIAN	All Stations							
	RMSE[m]	NRMSE[%]	NSE	R2	MAD	95% CI,median lower	95% CI,median upper	n Stations
DTU	0.30	11.11	0.88	0.95	0.17	0.19	0.44	21
TUM	0.35	15.25	0.81	0.94	0.23	0.29	0.50	18
ISR	0.35	13.11	0.9	0.96	0.18	0.25	0.41	17

3.3 Validation of Water Level Time Series (DGFI/TUM)

The HYDROCOASTAL Sentinel-3 Level 3 (L3) water level time series processed by DTU based on three different retrackers (DTU, TUM, and ISR) were compared to time series data from DGFI-TUM's "Database for Hydrological Time Series of Inland Waters" (DAHITI, www.dahiti.dgfi.tum.de) as well as to in-situ time series.

DAHITI (Schwatke et al., 2015) L3 data processing uses an extended outlier rejection and a Kalman filter approach. Currently, it provides altimetry-derived water level time series at more than 7000 virtual stations distributed globally. In DAHITI, all altimeter missions are retracked using an improved threshold retracker (ITR), i.e. SAR missions such as Sentinel-3 and Cryosat-2 are not handled differently from classical LRM missions. Moreover, DAHITI includes in-situ water level time series for validation purposes.

1091 L3 virtual stations (VS) had been provided for validation. No additional outlier filtering was applied to the L3 VS time series. For 1086 stations, corresponding DAHITI data exist (some were newly created). For 1031 VS in 8 river basins, the DAHITI data were considered being of sufficient quality and used for comparison. 246 VS in 6 river basins can be validated against in-situ data of 76 gauging stations. Only gauges connected to a VS according to the SWOT River Database (SWORD) are used, i.e. no comparison across dams or waterfalls were made. 240 of the 246 validated VS are within a distance of less than 100 km to the reference gauge. Table 3.3.1 lists the in-situ sources used, the number of gauges, and the distances to the validated VS.

Table 3.3.1. In-Situ data sources with number of gauges and minimum, median, and maximum distance to the validated virtual stations.

Source	**River**	**Number of gauges**	**Distance [km]**		
			Min	**Median**	**Max**
Bundesamt für Gewässerkunde (BfG)	Rhine	16	0.17	7.56	44.28
U.S. Geological Survey (USGS)	Mississippi	11	1.91	19.91	30.8
U.S. Army Corps of Engineers (USACE)	Mississippi	17	3.06	13.91	129.03
Agenzia Interregionale Fiume Po (AIPO)	Po	9	2.00	4.04	25.65
Republic Hydrometeorological Service of Serbia (HIDMET)	Danube	12	0.95	12.21	416.24
Zambezi River Authority (ZRA)	Zambezi	5	0.41	25.41	103.89
Hydro-Geochemistry of the Amazonian Basin (SO-HYBAM)	Amazon	6	5.62	128.7	202.95

3.3.1 Applied offsets

Prior to validation, each L3 time series was shifted by its offset (i.e., the median difference between the water level records for identical periods) to the corresponding DAHITI time series. Figure 3.3.1 shows the distribution of these offsets. The median offset was 0.16 m for DTU, 0.45 m for TUM, and 0.13 m for the ISR retracker. This also shows that the TUM retracker has a significantly different retracker bias than the other two retrackers.

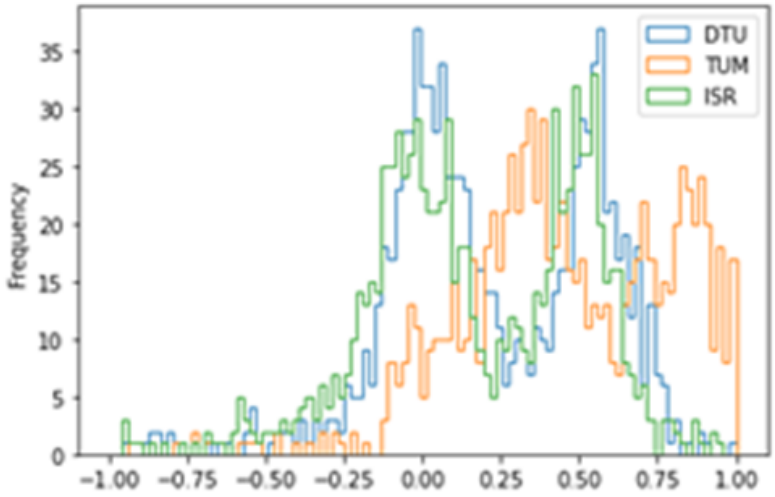


Figure 3.3.1. Distribution of estimated offsets of the L3 data with respect to DAHITI by used retracker.

3.3.2 Comparison to DAHITI

For the comparison of the L3 data with DAHITI, the root mean square deviation (RMSD), the RMSD normed by the amplitude between the 5th and 95th water level percentile (NRMSD), the Nash-Sutcliffe Efficiency (NSE), and the Pearson correlation coefficient (R2) were calculated. As not all retrackers provide data for all stations and we want to make a fair comparison as well as to check the completeness of the data, results are summarized for a reduced set of stations, for which all retrackers provide time series, and for all stations. Table 3.3.2 shows the results for the entire stations per retracker and Table 3.3.3 only for the stations with data from all retrackers. Of the three retrackers, the DTU approach shows the highest agreement with DAHITI and generates results for the largest number of stations.

Table 3.3.2. Median results of the comparison of the L3 data with DAHITI for all stations.

Median	RMSD [m]	NRMSD [%]	NSE	R2	n Stations
DTU	0.21	6.48	0.964	0.97	965
TUM	0.23	6.57	0.963	0.97	920
ISR	0.26	7.17	0.954	0.96	923

Table 3.3.3. Median results of the comparison of the L3 data with DAHITI river for stations with data from all retrackers.

Median	RMSD [m]	NRMSD [%]	NSE	R2	n Stations
DTU	0.18	5.69	0.970	0.98	845
TUM	0.22	6.02	0.968	0.97	845
ISR	0.24	6.57	0.961	0.97	845

Figure 3.3.2 shows the R2 against DAHITI for each retracker. While high correlations are reached for some rivers such as Amazon and Zambezi, large deviations can be observed at rivers with low annual signal (e.g., Yangtze, Danube, Rhine, Po) regardless of the used retracker. Some differences are caused by an incorrect VS placement at the edge of a water body or at a dam containing signals from the upstream and downstream part of the river (Figure 3.3.3 top). However, at most of the low correlating VS, the L3 time series is nearly constant not showing the signal contained in the DAHITI data (Figure 3.3.3 bottom). These deviations are either caused by incorrect retracking or the river model used in the L3 processor.

Table 3.3.4 lists results of the comparison with DAHITI per studied river, and Table 3.3.5 shows only the results for VS with data from all retrackers. Figure 3.3.4 shows the distribution of the RMSD per retracker for all stations and the stations with data from all retrackers. Tables 3.3.6 and 3.3.7 list the RMSD confidence interval and the “bootstrapped” median RMSD per retracker.

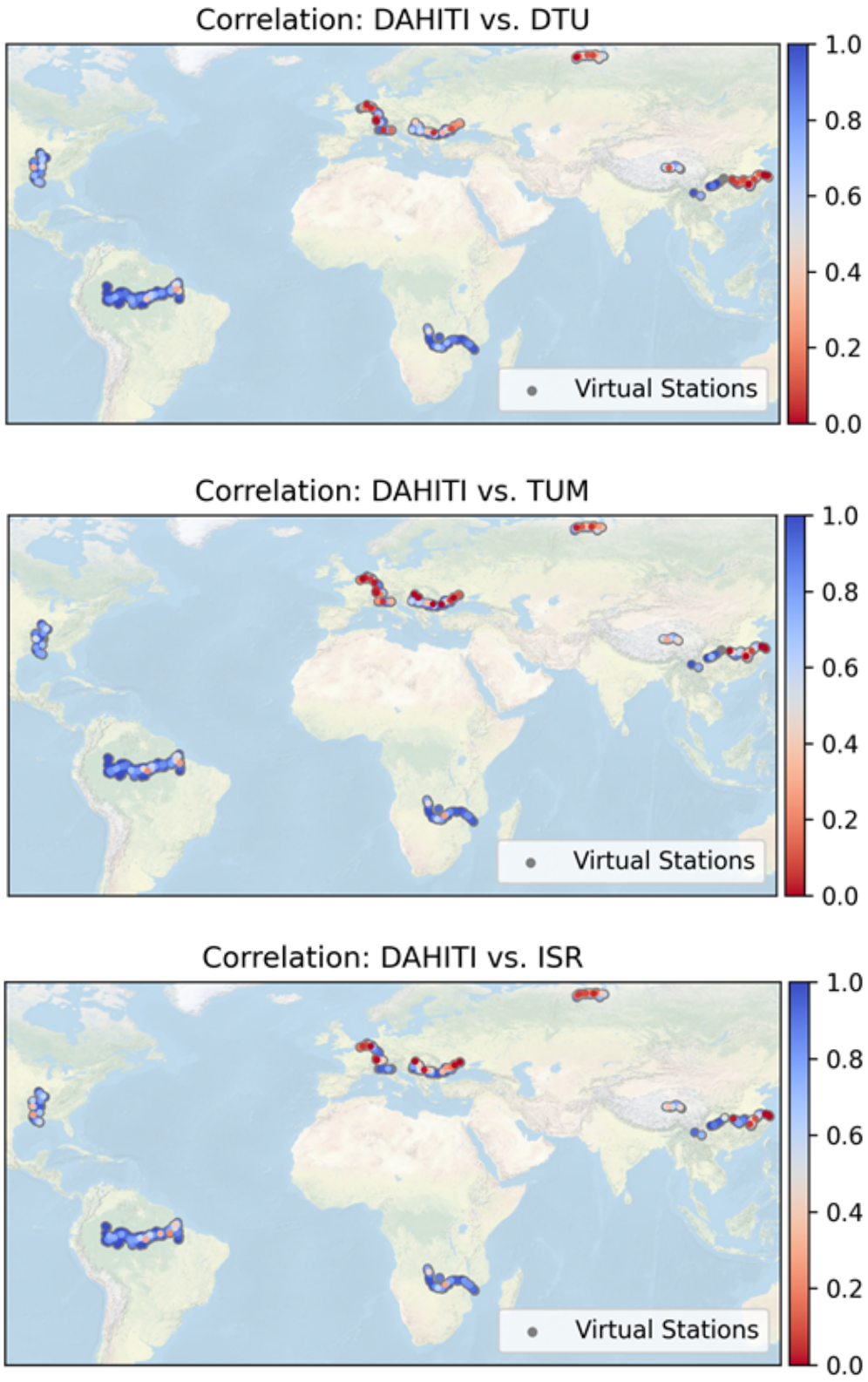


Figure 3.3.2. Correlation of the VS with three retrackers with DAHITI.

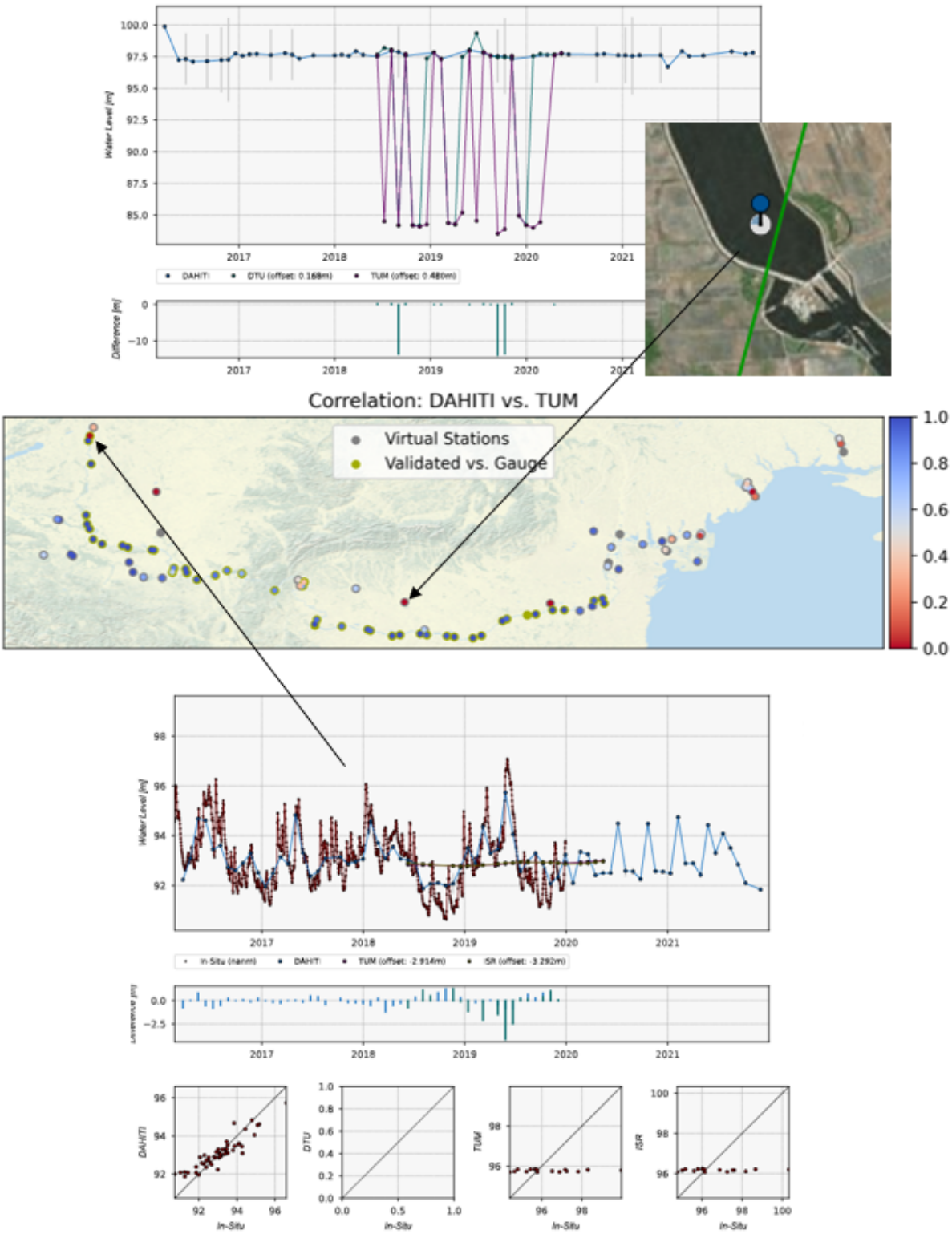


Figure 3.3.3. Example of Virtual Stations at the Danube River. Top: The time series contains data from up- and downstream a dam. Bottom: VS with constant L3 data not showing the signal contained in DAHITI.

Table 3.3.4. Median results of the comparison of the L3 data with DAHITI per studied river.

River	Retracker	RMSD [m]	NRMSD [%]	NSE	R2	n Stations
Ob	DTU	0.45	19.87	0.620	0.70	84
	TUM	0.46	20.90	0.583	0.72	78
	ISR	0.51	22.79	0.567	0.66	82
Rhine	DTU	0.29	21.37	0.596	0.80	40
	TUM	0.22	28.50	0.313	0.61	37
	ISR	0.29	27.48	0.373	0.61	36
Mississippi	DTU	0.33	5.73	0.976	0.98	76
	TUM	0.29	5.91	0.971	0.98	70
	ISR	0.37	6.70	0.960	0.97	72
Amazon	DTU	0.14	2.94	0.992	0.99	475
	TUM	0.17	3.30	0.991	0.99	460
	ISR	0.20	3.97	0.986	0.99	466
Yangtze	DTU	1.00	22.86	0.608	0.77	105
	TUM	0.79	19.74	0.699	0.79	92
	ISR	0.69	17.31	0.788	0.86	93
Danube	DTU	0.15	9.15	0.930	0.94	84
	TUM	0.20	10.05	0.921	0.93	83
	ISR	0.20	8.09	0.943	0.95	79
Po	DTU	0.44	14.05	0.835	0.92	12
	TUM	0.36	11.72	0.853	0.89	13
	ISR	0.23	8.49	0.942	0.97	10
Zambezi	DTU	0.10	3.41	0.990	0.99	89
	TUM	0.14	4.04	0.986	0.99	87
	ISR	0.17	4.99	0.979	0.98	85

Table 3.3.5. Median results of the comparison with DAHITI per river for stations with data from all retrackers.

River	Retracker	RMSD [m]	NRMSD [%]	NSE	R2	n Stations
Ob	DTU	0.47	19.91	0.617	0.71	67
	TUM	0.46	19.99	0.655	0.73	67
	ISR	0.47	20.63	0.643	0.71	67
Rhine	DTU	0.15	21.37	0.650	0.82	24
	TUM	0.19	26.99	0.475	0.69	24
	ISR	0.25	29.09	0.342	0.61	24
Mississippi	DTU	0.29	5.52	0.978	0.98	68
	TUM	0.31	6.23	0.969	0.98	68
	ISR	0.34	6.50	0.963	0.97	68
Amazon	DTU	0.14	2.93	0.993	0.99	442
	TUM	0.18	3.28	0.991	0.99	442
	ISR	0.20	3.94	0.987	0.99	442
Yangtze	DTU	0.87	21.24	0.670	0.82	77
	TUM	0.81	18.26	0.724	0.82	77
	ISR	0.70	17.31	0.788	0.86	77
Danube	DTU	0.13	8.35	0.934	0.94	75
	TUM	0.19	8.10	0.927	0.95	75
	ISR	0.20	7.97	0.946	0.96	75
Po	DTU	0.29	9.74	0.928	0.95	10
	TUM	0.29	10.80	0.912	0.94	10
	ISR	0.23	8.49	0.942	0.97	10
Zambezi	DTU	0.10	3.25	0.990	0.99	82
	TUM	0.13	3.80	0.988	0.99	82
	ISR	0.17	4.67	0.979	0.98	82

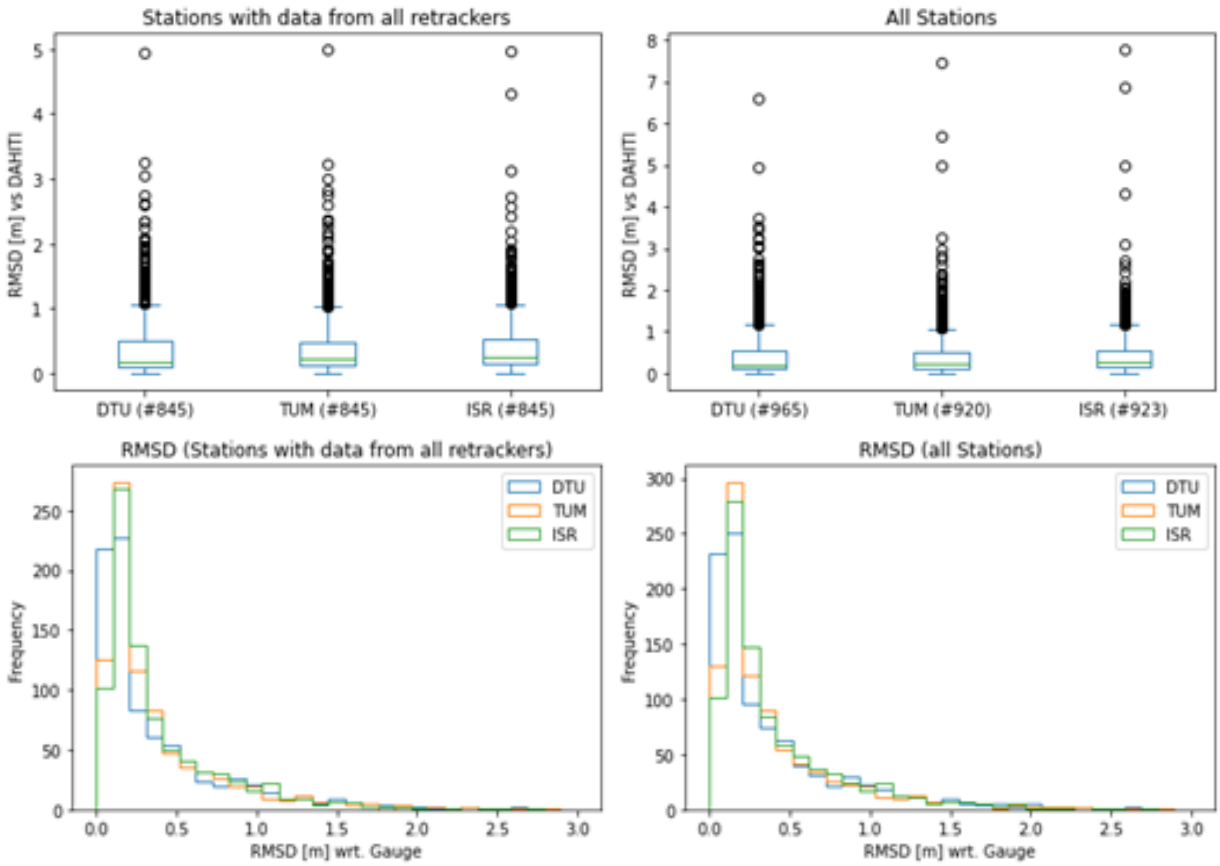


Figure 3.3.4. Distribution of RMSD between L3 data and DAHITI per retracker for the stations with data from all retrackers (left) and all stations (right).

Table 3.3.6. RMSD confidence interval (CI) per retracker.

	Stations with data from all retrackers		**All Stations**	
	Lower CI [m] (0.025th q.)	**Upper CI [m] (0.975th q.)**	**Lower CI [m] (0.025th q.)**	**Upper CI [m] (0.975th q.)**
DTU	0.05	1.67	0.05	2.05
TUM	0.06	1.65	0.06	1.72
ISR	0.07	1.56	0.07	1.67

Table 3.3.7. Median RMSD per retracker calculated using a bootstrap approach.

	Median RMSD [m] (Stations with data from all retrackers)	**Median RMSD [m] (All Stations)**
DTU	0.17 ± 0.04	0.21 ± 0.03
TUM	0.20 ± 0.03	0.23 ± 0.02
ISR	0.22 ± 0.03	0.26 ± 0.02

3.3.3 Validation against in-situ gauges

For the validation of the L3 data with in-situ data, the root mean square error (RMSE), the RMSE normed by the amplitude between the 5th and 95th water level percentile (NRMSE), the Nash-Sutcliffe Efficiency (NSE), and the Pearson correlation coefficient (R2) were calculated.

Table 3.3.8 shows the results for the entire stations per retracker and Table 3.3.9 only for the stations with data from all retrackers. There is no significant difference between the three retrackers.

Figure 3.3.5 shows the distribution of RMSE for the three retrackers and DAHITI. Table 3.3.10 lists the validation results per studied river and Table 3.3.11 shows only the results for VS with data from all retrackers. Tables 3.3.12 and 3.3.13 list the RMSE confidence interval and “bootstrapped” median RMSE per retracker.

Table 3.3.8. Median results of the comparison of the L3 data with DAHITI for all stations.

Median	**RMSE [m]**	**NRMSE [%]**	**NSE**	**R2**	**n Stations**
DTU	0.35	8.10	0.947	0.97	221
TUM	0.34	7.88	0.947	0.97	216
ISR	0.33	8.07	0.947	0.96	208

Table 3.3.9. Median results of the comparison with DAHITI river for stations with data from all retrackers.

Median	**RMSE [m]**	**NRMSE [%]**	**NSE**	**R2**	**n Stations**
DTU	0.33	6.93	0.958	0.98	198
TUM	0.34	7.35	0.959	0.97	198
ISR	0.33	7.56	0.953	0.97	198

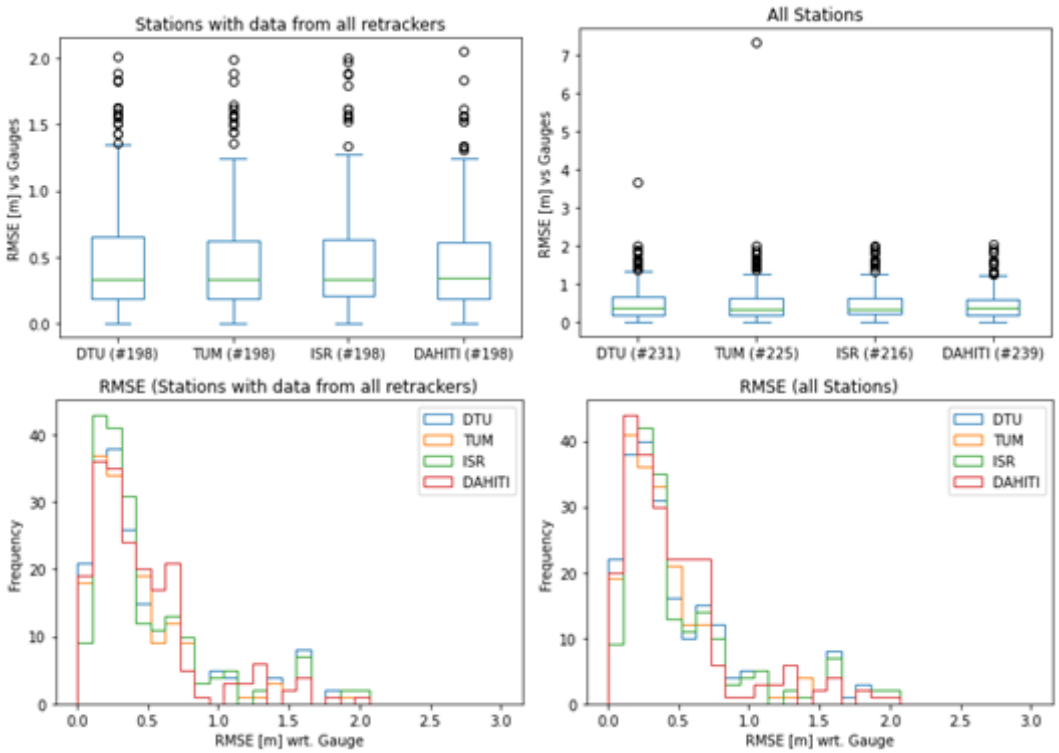


Figure 3.3.5. Distribution of RMSE between L3 and in-situ data per retracker and DAHITI for the stations with data from all retrackers (left) and all stations (right).

Table 3.3.10. Median validation results per studied river.

River	Retracker	RMSE [m]	NRMSE [%]	NSE	R2	n Stations
Rhine	DTU	0.40	16.54	0.695	0.85	28
	TUM	0.34	18.90	0.690	0.86	25
	ISR	0.31	12.82	0.840	0.91	23
Mississippi	DTU	0.29	4.70	0.981	0.99	61
	TUM	0.32	4.44	0.983	0.99	59
	ISR	0.31	4.83	0.981	0.99	58
Amazon	DTU	0.38	7.50	0.960	0.99	70
	TUM	0.39	5.39	0.978	0.99	68
	ISR	0.37	7.40	0.972	0.98	67
Danube	DTU	0.31	9.27	0.941	0.96	43

	TUM	0.35	9.94	0.925	0.94	44
	ISR	0.34	9.40	0.918	0.95	43
Po	DTU	0.32	15.93	0.810	0.93	12
	TUM	0.27	13.68	0.801	0.94	13
	ISR	0.26	7.12	0.956	0.96	10
Zambezi	DTU	0.66	13.40	0.796	0.96	7
	TUM	0.65	13.24	0.801	0.96	7
	ISR	0.70	13.53	0.793	0.96	7

Table 3.3.11. Median validation results per studied river for stations with data from all retrackers.

River	Retracker	RMSE [m]	NRMSE [%]	NSE	R2	n Stations
Rhine	DTU	0.27	13.29	0.870	0.88	17
	TUM	0.33	13.63	0.851	0.89	17
	ISR	0.26	18.79	0.574	0.81	17
Mississippi	DTU	0.28	4.70	0.981	0.99	57
	TUM	0.33	4.72	0.982	0.99	57
	ISR	0.31	4.80	0.982	0.99	57
Amazon	DTU	0.38	6.76	0.973	0.99	66
	TUM	0.40	5.39	0.978	0.99	66
	ISR	0.36	7.15	0.973	0.99	66
Danube	DTU	0.27	9.27	0.941	0.96	41
	TUM	0.31	9.79	0.926	0.94	41
	ISR	0.33	9.40	0.918	0.95	41
Po	DTU	0.30	11.66	0.892	0.95	10
	TUM	0.25	8.79	0.940	0.96	10
	ISR	0.26	7.12	0.956	0.96	10
Zambezi	DTU	0.66	13.40	0.796	0.96	7

	TUM	0.65	13.24	0.801	0.96	7
	ISR	0.70	13.53	0.793	0.96	7

Table 3.3.12. RMSE confidence interval (CI) per retracker.

	Stations with data from all retrackers		**All Stations**	
	Lower CI [m] (0.025^{th} q.)	**Upper CI [m] (0.975^{th} q.)**	**Lower CI [m] (0.025^{th} q.)**	**Upper CI [m] (0.975^{th} q.)**
DTU	0.04	1.63	0.04	1.75
TUM	0.03	1.62	0.03	1.64
ISR	0.08	1.63	0.08	1.77

Table 3.3.13. Median RMSE per retracker calculated using a bootstrap approach.

	Median RMSE [m] **Stations with data from all retrackers**	**Median RMSE [m]** **All Stations**
DTU	0.33 ± 0.04	0.35 ± 0.04
TUM	0.34 ± 0.04	0.34 ± 0.04
ISR	0.33 ± 0.03	0.33 ± 0.03

In order to analyze if the retracker performance depends on river width, a correlation study was performed. Figure 3.3.6 shows the RMSE by river width. As expected, smaller rivers show higher probability for larger RMSE. No significant difference in performance for the different retrackers could be found.

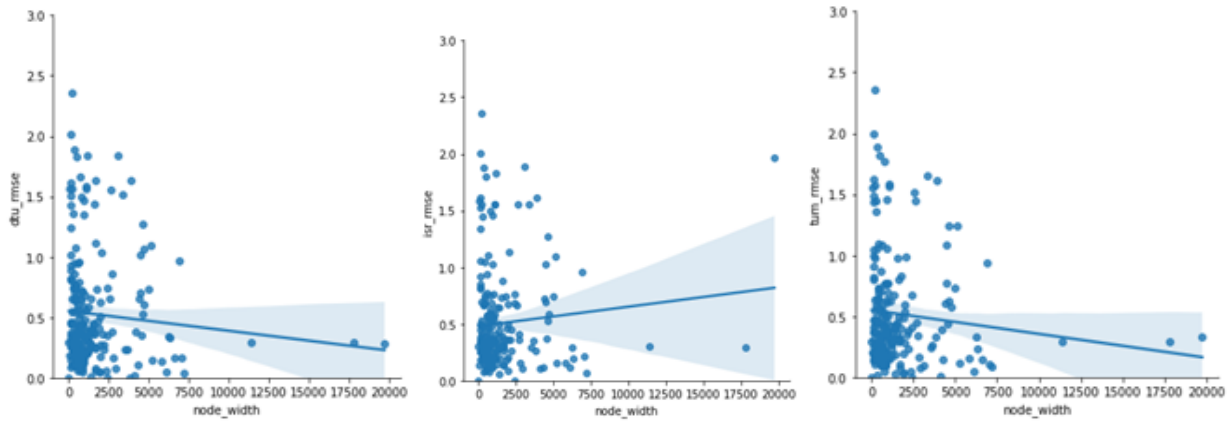


Figure 3.3.6. RMSE by river width per retracker (left: DTU, center: ISR, right: TUM)

3.3.4 Summary

The validation over eight globally distributed rivers shows that none of the three tested retrackers has a significantly higher accuracy than the others. The performance varies depending on the river flow system. However, the DTU retracker provides more valid time series overall than the other two retrackers. For some stations, the Hydrocoastal products do not improve on existing standard DAHITI products.

Unfortunately, only data from a single L3 algorithm (DTU) could be examined. However, the comparison with the DAHITI time series shows that the L3 Processor has a greater influence on the accuracy of the time series than retracking. Occasionally, the DTU L3 Processor shows an unrealistic behavior with no temporal variations at some stations. This may be due to the algorithm used as well as to different input data (compared to DAHITI).

3.4 Validation against in situ data over the Amazon Basin (AHL)

To be added in V2

3.5 Validation against in situ data for Amur, Yangtze and Zambezi (DTU)

Validation of the DTU L3 inland water product was carried out for the Amur, Yangtze and Zambezi Rivers. Table 3.5.1 provides a summary of the data used in the validation.

Table 3.5.1. Summary of datasets used in the validation

River	L3 filename	In-situ data provider
Amur	hydrocoastal_S3_amur_L3_v1_2.nc	Russia Hydrometeorological Service; provided by Elena Zakharova
Yangtze	hydrocoastal_S3_yangtze_L3.nc	Ministry of Water Resources of China; Classified data; validation performed by Liguang Jiang
Zambezi	hydrocoastal_S3_zambezi_L3_v1_2.nc	Zambezi River Authority

Because of limited temporal overlap between the DTU L3 datasets and the available in-situ records, validation was only possible at 4 stations in the Amur (Nikolaevsk, Komsomolsk, Khaborovsk and Jalinda), 6 stations in the Zambezi (Sesheke, Senanga, Nana's Farm, Matongo Platform, Lukulu, Chavuma) and 7 stations in the Yangtze (Jiujiang, Datong, Chenjiawa, Luoshan, Shigu, Gangtuo3, Zhimenda).

The vertical reference system of the in-situ data is unknown. Therefore, only relative heights were compared in the validation. The validation workflow consisted of the following steps:

1.) For each L3 data point, corresponding in-situ readings were extracted from a time period L3 timestamp ±1 day. If more than one in-situ data point was found within this time period, the in-situ data were averaged. This resulted in *nobs* data pairs L3/in-situ for each station and each retracker.

2.) The average in-situ and L3 water surface elevation (WSE) was calculated from the *nobs* data points for each station and retracker. Subsequently, relative WSE (rWSE) of L3 data ($rWSE_{L3}$) and in-situ data ($rWSE_{IS}$) were calculated as:

$$rWSE_{L3,i} = WSE_{L3,i} - \frac{1}{nobs}\sum_{i=1}^{nobs} WSE_{L3,i} \quad (3.6.1.)$$

$$rWSE_{IS,i} = WSE_{IS,i} - \frac{1}{nobs}\sum_{i=1}^{nobs} WSE_{IS,i} \quad (3.6.2.)$$

3.) Root mean squared error (RMSE) of relative WSE and weighted RMSE of relative WSE were calculated according to the following formulas:

$$RMSE = \sqrt{\frac{1}{nobs}\left(rWSE_{L3,i} - rWSE_{IS,i}\right)^2} \quad (3.6.3.)$$

$$wRMSE = \sqrt{\frac{1}{nobs}\left(\frac{rWSE_{L3,i} - rWSE_{IS,i}}{\sigma_{L3,i}}\right)^2} \quad (3.6.4.)$$

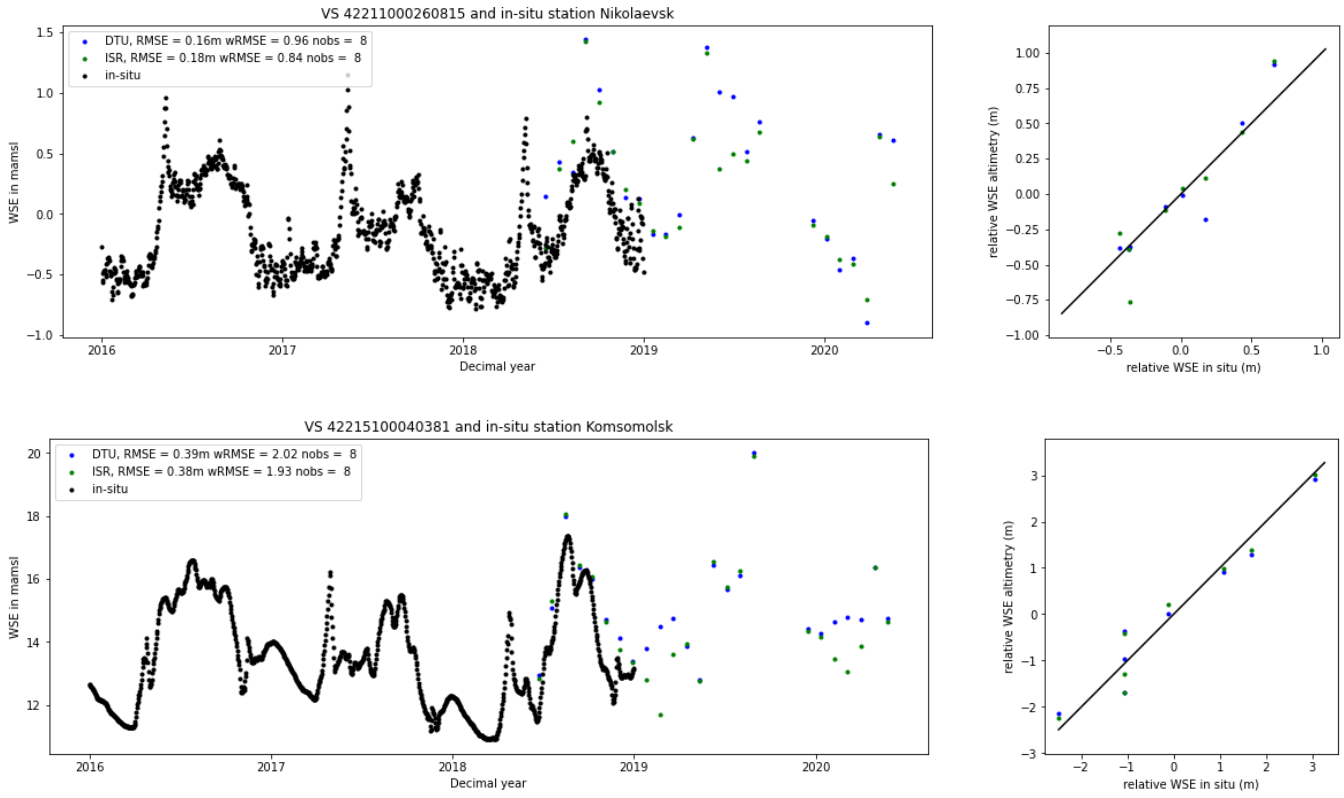
where $\sigma_{L3,i}$ is the standard deviation reported for this data point in the L3 product.

Table 3.5.2 provides an overview of the validation results for the Amur River.

Table 3.5.2. Overview of Amur validation results

In-situ Station	L3 VSID	nobs DTU	RMSE DTU (m)	wRMSE DTU (-)	nobs ISR	RMSE ISR (m)	wRMSE ISR (-)	nobs TUM	RMSE TUM (m)	wRMSE TUM (-)
Nikolaevsk	42211000260815	8	0.16	0.96	8	0.18	0.84	0		
Komsomolsk	42215100040381	8	0.39	2.02	8	0.38	1.93	0		
Khaborovsk	42218100010011	6	0.16	2.19	0			0		
Jalinda	42275000030091	0			8	1.20	1.62	5	0.37	0.48
All stations		**22**	**0.24**	**1.72**	**24**	**0.59**	**1.46**	**5**	**0.37**	**0.48**

Figure 3.5.1 shows performance for individual stations, including time series plots and scatter plots.



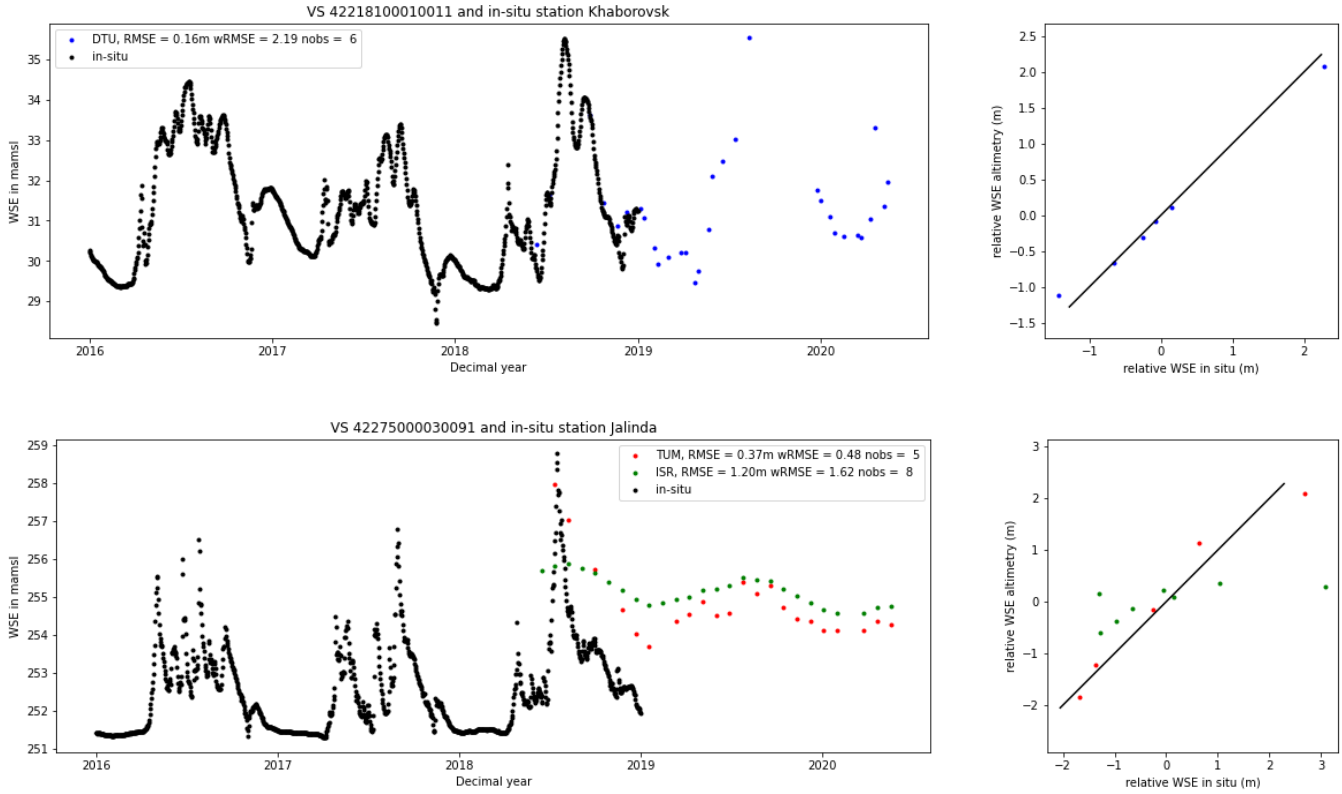


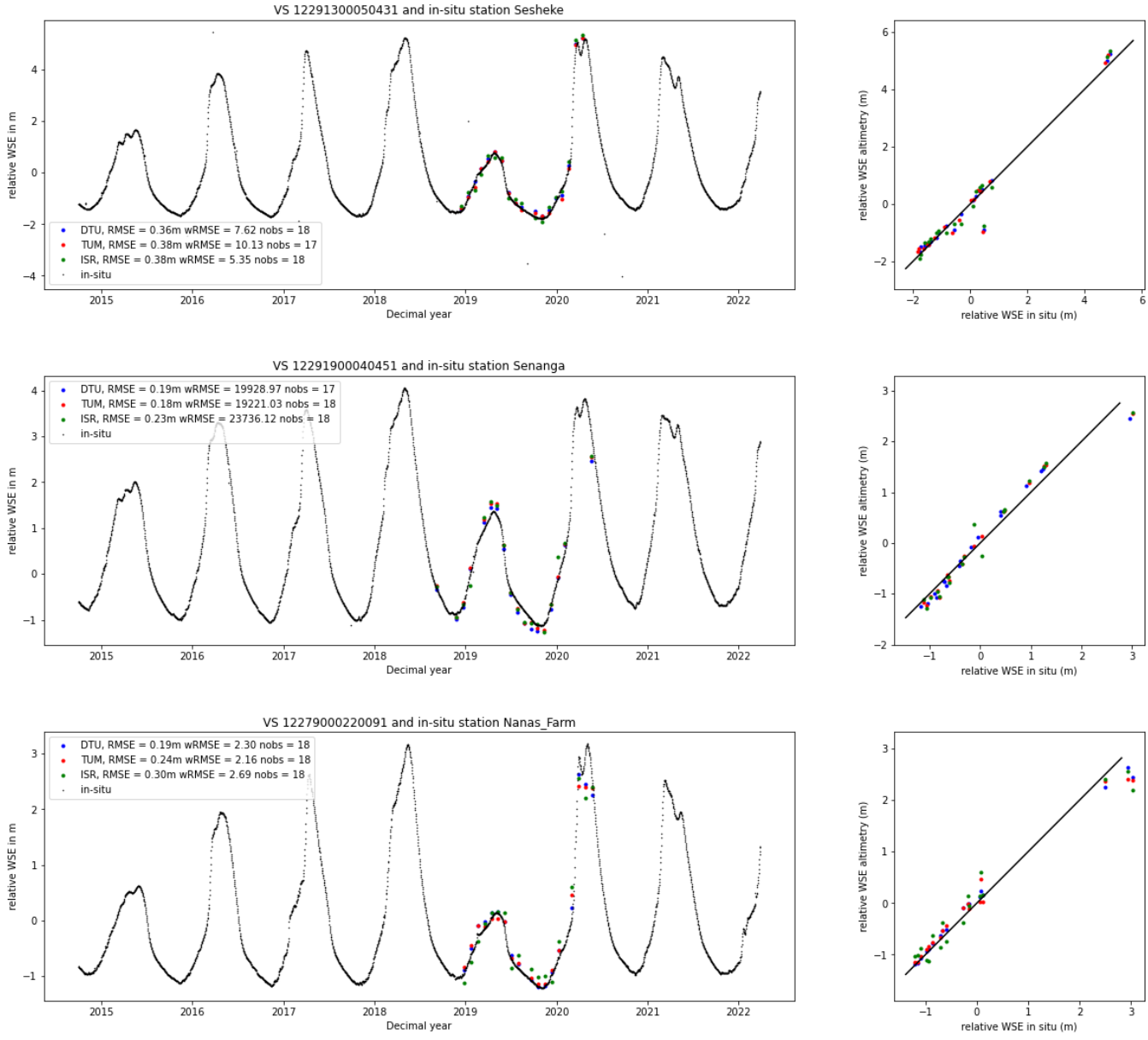
Figure 3.5.1. Validation plots for Amur River stations

Table 3.5.3 provides an overview of the validation results for the Zambezi River.

Table 3.5.3. Overview of Zambezi validation results

In-situ Station	**L3 VSID**	**nobs DTU**	**RMSE DTU (m)**	**wRMSE DTU (-)**	**nobs ISR**	**RMSE ISR (m)**	**wRMSE ISR (-)**	**nobs TUM**	**RMSE TUM (m)**	**wRMSE TUM (-)**
Sesheke	12291300050431	18	0.36	7.62	18	0.38	5.35	17	0.38	10.13
Senanga	12291900040451	17	0.19	19929.00	18	0.23	23736.00	18	0.18	19221.00
Nana's Farm	12279000220091	18	0.19	2.30	18	0.30	2.69	18	0.24	2.16
Matongo Platform	12293000040291	20	0.65	5.99	20	0.65	6.08	20	0.64	5.74
Lukulu	12293000110321	21	0.29	5.38	21	0.36	4.80	21	0.29	5.55
Chavuma	12297000160661	19	0.64	23.33	19	0.65	13.15	19	0.6	13.15
All stations		**113**	**0.39**	**3328.9**	**114**	**0.43**	**3961.4**	**113**	**0.39**	**3209.6**

Figure 3.5.2 shows performance for individual stations, including time series plots and scatter plots.



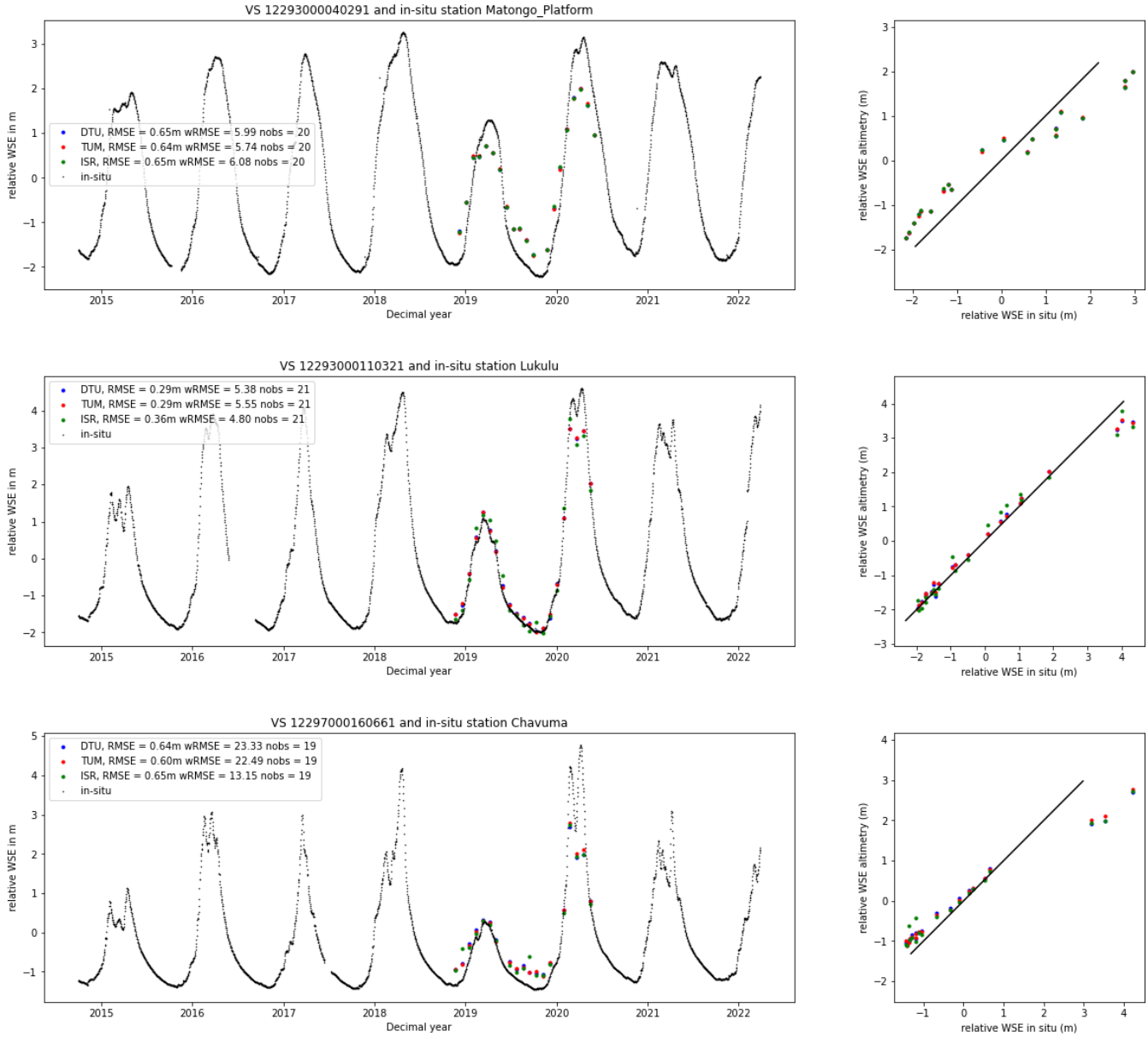


Figure 3.5.2. Validation plots for Zambezi River stations

Table 3.5.3 provides an overview of the validation results for the Yangtze River.

Table 3.5.3. Overview of Yangtze validation results

In-situ Station	L3 VSID	nobs DTU	RMSE DTU (m)	wRMSE DTU (-)	nobs ISR	RMSE ISR (m)	wRMSE ISR (-)	nobs TUM	RMSE TUM (m)	wRMSE TUM (-)
Jiujiang	43450100070201	23	2.09	27.94	23	1.56	5.33	22	2.72	5.46

Datong	43430700240511	8	2.05	7.67	0			0		
Chenjiawa	43491100200571	20	1.16	14.70	20	7.07	42.45	21	1.51	40.31
Luoshan	43470000190131	22	3.38	11.02	0			0		
Shigu	43499300110601	14	0.47	12.41	15	1.54	5.00	17	6.81	85.00
Gangtuo3	43499500340261	9	1.09	1.75	9	0.89	1.65	8	10.65	266.73
Zhimenda	43499500520121	12	0.35	0.82	12	1.58	44.90	10	0.88	6.43
All stations		**85**	**1.42**	**8.06**	**56**	**2.77**	**23.50**	**56**	**4.96**	**99.62**

Figure 3.5.3 shows performance for individual stations as time series plots.

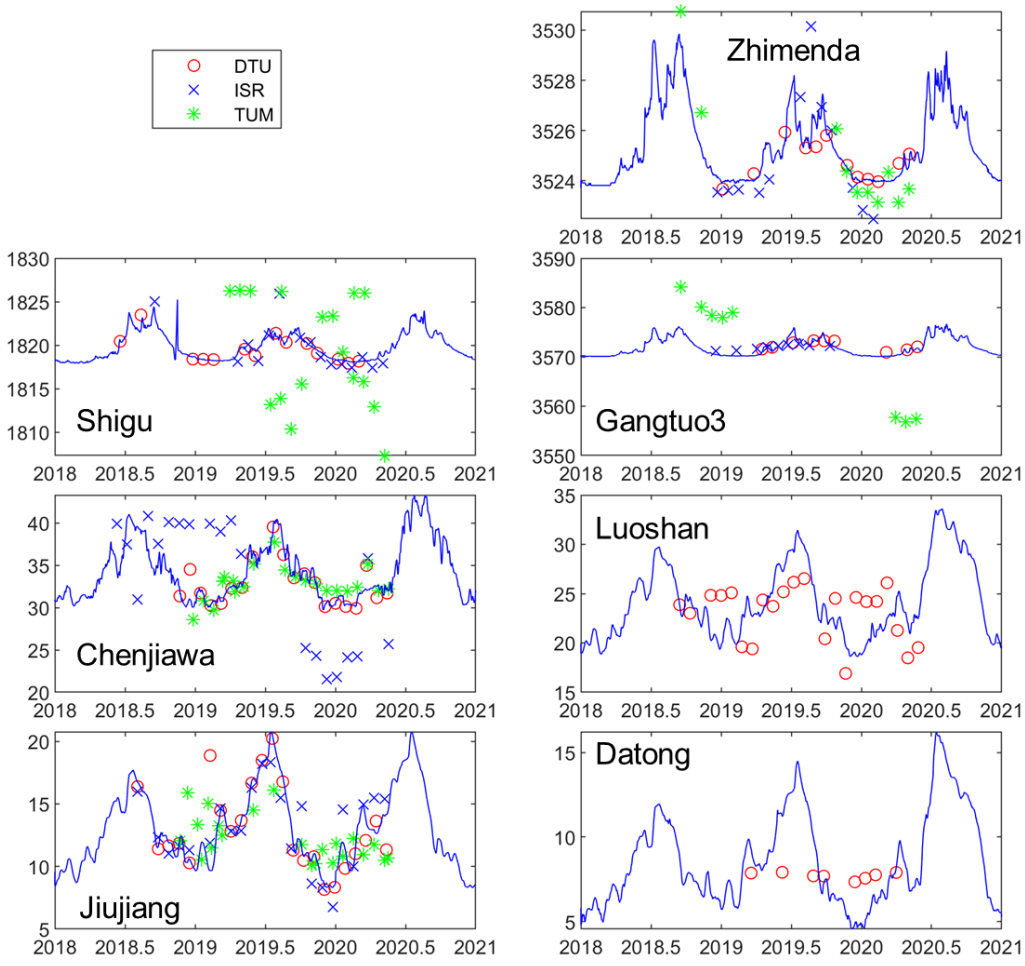


Figure 3.5.3. Validation plots for Yangtze River stations

We can summarize the findings of this validation exercise as follows:

- The DTU retracker performs best on the Amur, but the comparison is not entirely fair because the three retrackers cover different subsets of the Amur stations.
- All retrackers cover the same stations in the Zambezi. Performance is very similar, with DTU and TUM outperforming ISR by a small margin.
- The DTU retracker performs clearly better than both TUM and ISR on the Yangtze River.
- Uncertainty estimates provided in the L3 product are too optimistic. wRMSE values exceed 1 in almost all cases and are often higher than 5. This indicates that the provided uncertainty estimate is not a realistic estimate of the uncertainty of WSE estimates derived from satellite altimetry.

The overall conclusion is therefore that the DTU retracker performs best on the investigated river targets.

3.6 Validation against in situ data for Ob and Rhine Rivers (NUIM)

As NUIM was planned to participate only in the discharge retrieval activities, the goodness of the water level time series (WLTS) retrieved with different retrackers for an accurate discharge retrieving was the priority interest of L3 validation activities. WLTS performance was evaluated via comparison of the discharge retrieved from L3 WLTS using the rating curve method with in situ discharge.

Additionally, due to specifics related to the seasonal ice, the Hydrocoastal WLTS for the Ob River were compared with in situ observations of the water level at two gauging stations.

For the discharge retrieving using the rating curve method, the WLTS were split into calibration period and validation period. The length of the periods depends on the mission (S3A or S3B) used for WLTS construction at a given virtual station (VS). For several VS the CAL/VAL split was not possible due to missing L2/L3 retrievals.

3.6.1 The Ob River L4 Validation results

Table 3.6.1. Validation scores of the discharge retrievals using rating curve method for VS located on the Ob River ± 65 km from the gauging station Salekhard.

st_short_Id	period	retracker	Rcorr	NRMSE%	Bias, m3/s	Validation Nobs	NS
31231000020081	cal/val	isr	0.99	8	583	7	0.97
S3A		tum	0.99	7	492	7	0.98
		dtu	0.99	9	1114	7	0.96
31231000020081	cal/val	Isr	0.99	11	1248	8	0.96
S3A		tum	0.99	10	1007	8	0.97
		dtu	0.98	17	2117	8	0.91

31221000030031	na	lsr	correct				
S3B		tum	short period	+10% gaps			
		dtu	short period	+20% gaps			
31210300020551	cal/val	isr	0.98	19	2040	8	0.88
S3A		tum	0.90	35	2829	8	0.59
		dtu	0.99	9	217	8	0.97
31210300030141	cal	isr	0.96	21	351	13	0.92
S3B		tum	0.94	29	1380	13	0.84
		dtu	0.97	19	121	13	0.93
31210500010501	cal	isr	0.99	12	90	15	0.97
S3B		tum	0.99	8	30	15	0.99
		dtu	0.94	24	212	15	0.89
31210500050261	cal	isr	0.96	21	-6	12	0.92
S3A		tum	0.92	28	109	14	0.85
		dtu	0.99	13	97	14	0.97
31210500050471	cal	isr	0.99	11	6	15	0.98
S3B		tum	0.98	13	65	15	0.96
		dtu	0.97	16	62	15	0.95
31210700040161	na	isr	noisy	+ many outliers			
S3A		tum	noisy				
		dtu	noisy				

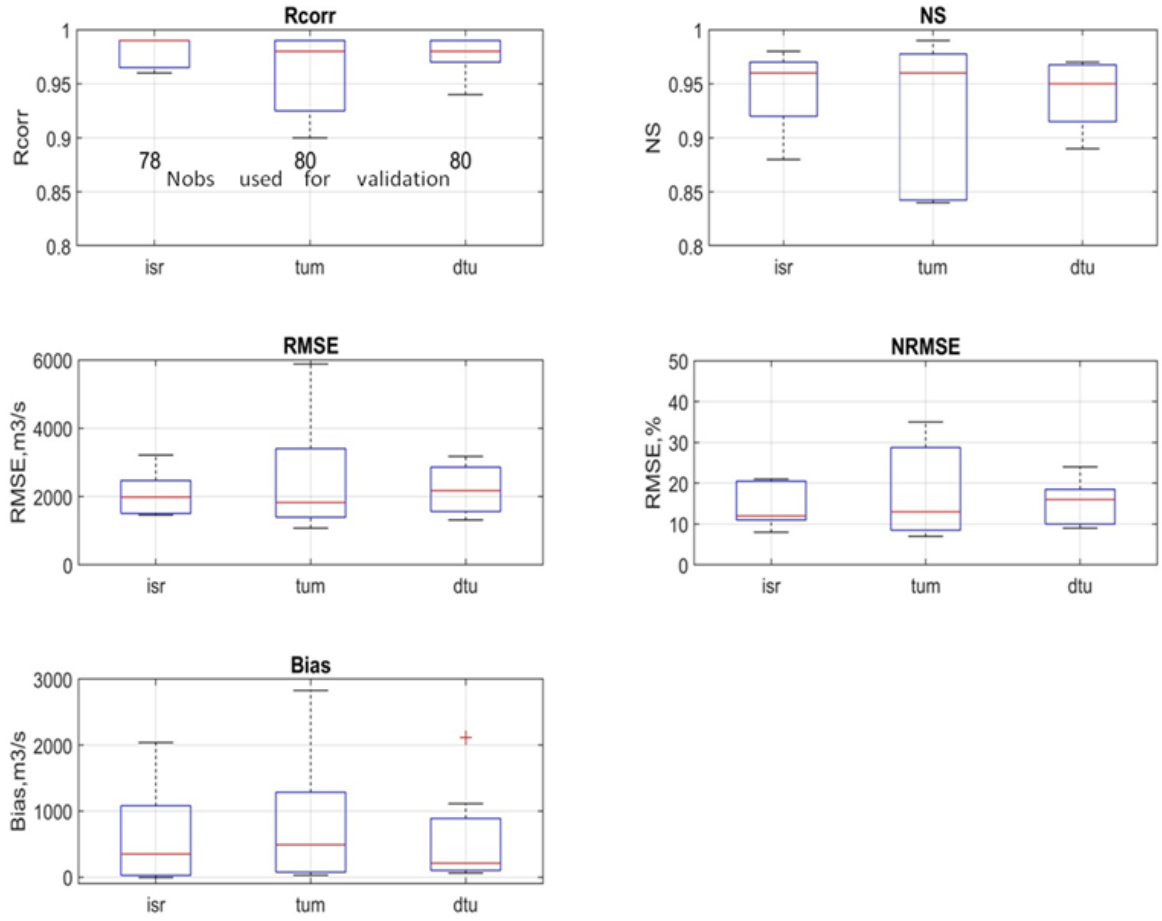


Figure 3.6.1. Summary boxes of the discharge retrieving statistics for the Ob River using only WLTS where all 3 retrackers are available.

3.6.2 The Ob River L3 validation results

During discharge retrieving, it was noted that the performance of the retrackers during the winter is different, that is why an additional comparison of the Hydrocoastal WLTS with the in situ water level at 2 gauging stations within Salekhard river reach was carried out.

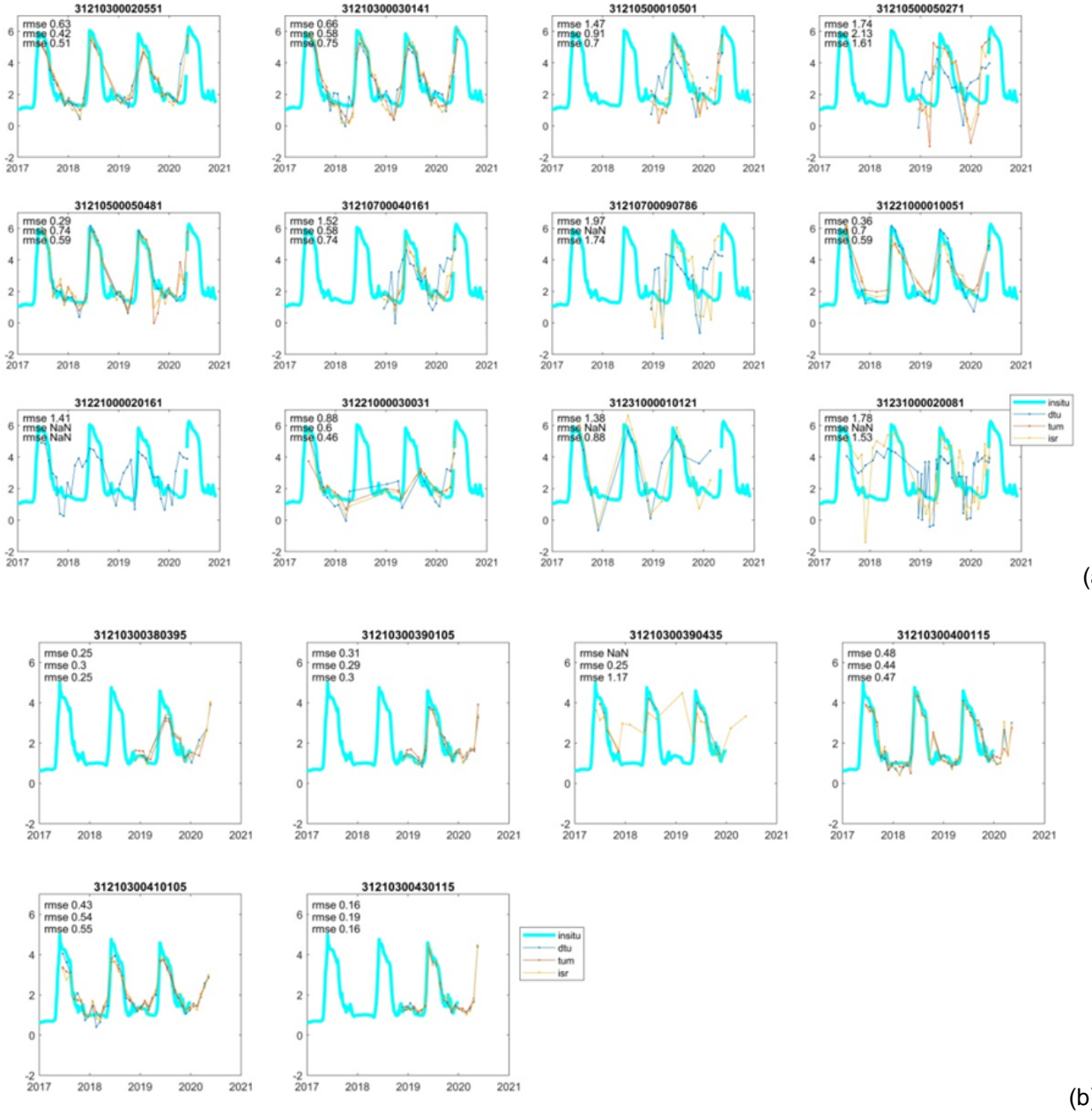


Figure 3.6.2. Comparison of the Hydrocoastal WLTS with in situ water level observations at Salekhard (a) and Aksarka (b) gauging stations.

Table 3.6.2. Validation scores for 18 WLTS for virtual stations located the Ob River within ± 65 km from the gauging stations Aksarka and Salekhard

GS Station/ VS_Id	retracker	R	NS	RMSE	NRMSE	Bias*	Nobs
Aksarka	dtu	0.97	-0.01	0.25	0.09	0.00	11
31210300380395	tum	0.97	-0.20	0.30	0.11	0.00	11
	isr	0.96	0.26	0.25	0.09	0.00	11
Aksarka	dtu	0.97	0.16	0.31	0.09	0.00	13
31210300390105	tum	0.98	0.27	0.29	0.09	0.00	13
	isr	0.97	0.17	0.30	0.09	0.00	13
Aksarka	dtu	na	na	na	na	na	0
31210300390435	tum	0.98	-0.34	0.25	0.07	0.00	14
	isr	0.44	-2.61	1.17	0.33	0.07	18
Aksarka	dtu	0.93	0.54	0.48	0.14	0.01	34
31210300400115	tum	0.94	0.53	0.44	0.13	0.01	34
	isr	0.93	0.53	0.47	0.14	0.01	34
Aksarka	dtu	0.95	0.41	0.43	0.13	0.01	34
31210300410105	tum	0.94	0.01	0.54	0.16	0.02	34
	isr	0.92	0.10	0.55	0.16	0.02	34
Aksarka	dtu	0.99	0.55	0.16	0.05	0.00	15
31210300430115	tum	0.99	0.53	0.19	0.05	0.00	15
	isr	0.99	0.56	0.16	0.05	0.00	15
Salekhard	dtu	0.92	0.81	0.63	0.13	0.02	32
31210300020551	tum	0.97	0.91	0.42	0.09	0.01	32
	isr	0.95	0.89	0.51	0.11	0.01	32
Salekhard	dtu	0.91	0.81	0.66	0.14	0.03	40
31210300030141	tum	0.94	0.89	0.58	0.12	0.02	40
	isr	0.92	0.84	0.75	0.15	0.03	40
Salekhard	dtu	0.50	-0.46	1.47	0.33	0.18	16
31210500010501	tum	0.86	0.75	0.91	0.20	0.06	18

	isr	0.89	0.78	0.70	0.16	0.04	20
Salekhard	dtu	0.12	-1.28	1.74	0.40	0.28	17
'50271'	tum	0.44	0.14	2.13	0.49	0.32	17
	isr	0.56	0.26	1.61	0.37	0.19	17
Salekhard	dtu	0.99	0.98	0.29	0.06	0.01	36
'50481'	tum	0.91	0.84	0.74	0.16	0.04	36
	isr	0.94	0.88	0.59	0.13	0.02	36
Salekhard	dtu	0.45	-0.25	1.52	0.34	0.18	20
31210700040161	tum	0.93	0.79	0.58	0.13	0.02	20
	isr	0.87	0.69	0.74	0.16	0.04	20
Salekhard	dtu	0.23	-0.46	1.97	0.44	0.35	20
'90786'	tum	na	na	na	na	na	0
	isr	0.57	0.28	1.74	0.39	0.23	20
Salekhard	dtu	0.99	0.97	0.36	0.08	0.01	22
'10051'	tum	0.95	0.73	0.70	0.15	0.04	21
	isr	0.96	0.84	0.59	0.13	0.03	22
Salekhard	dtu	0.58	-0.14	1.41	0.30	0.19	36
'20161'	tum	na	na	na	na	na	0
	isr	na	na	na	na	na	0
Salekhard	dtu	0.73	0.53	0.88	0.20	0.07	17
31221000030031	tum	0.80	0.30	0.60	0.14	0.03	17
	isr	0.91	0.82	0.46	0.11	0.02	17
Salekhard	dtu	0.73	0.52	1.38	0.32	0.16	13
'10121'	tum	na	na	na	na	na	0
	isr	0.96	0.86	0.88	0.21	0.06	13
Salekhard	dtu	0.33	-0.58	1.78	0.38	0.32	45
31231000020081	tum	na	na	na	na	na	0
	isr	0.65	0.36	1.53	0.32	0.20	46

na - WLTS are not available; * Relative bias and NRMSE are obtained by normalizing the absolute bias and RMSE on the gauging station seasonal WL magnitude obtained as WL difference between 5%-95% quantiles.

3.6.3 The Rhine River L4 validation results

On the Rhine River, the validation of the discharge retrievals was made for 4 virtual stations. Daily discharge observations on three gauging stations (Worms, Mainz and Kaub) were used. Additionally, a comparison with the discharge retrievals from WLTS provided by Bonn University and based on SAMOSA (Ocean), OCOG and 80Hz SAMOSA+ retrackers was performed.

Table 3.6.3. Validation scores of the discharge retrievals using rating curve method for VS located on the Rhine River

Gauging Station	**Retracker**	**RC NRMSE/R/NS**
Worms S3A 156	SAMOSA	19% /0.84/0.62 *7% /0.98/0.94**
	OCOG	22% /0.81/0.86 *6% /0.99/0.96**
	SAMOSA+	13% /0.93/0.83 *8% /0.97/0.93**
	HyCo dtu tum isr	21%/0.83/0.45 16%/0.91/0.68 --
	HyCo2 dtu tum isr	43% /0.64/<0 *15%/0.88/0.73** 12%/0.92/0.82 16%/0.88/0.70
Mainz S3A 156	SAMOSA	19%/ 0.84/0.62 *6% /0.98/0.96**
	OCOG	22% / 0.81/0.46 *6% /0.98/0.96**
	SAMOSA+	13% /0.93/0.83 *4% /0.99/0.98**
	HyCo dtu tum isr	5%/0.99/0.97 7%/0.99/0.94 5%/0.99/0.96
Kaub S3B 156	SAMOSA	12% /0.99/0.93

	OCOG	126% /0.22/ <0 *16% /0.98/0.90**
	SAMOSA+	14% /0.98/0.90
	HyCo dtu tum isr	14%/0.99/0.94 12%/0.99/0.95 --
Kaub S3B 179	SAMOSA	384% /0.19/ <0 *4% /0.99/0.99**
	OCOG	600% /0.44/ <0 *6% /0.97/0.94**
	SAMOSA+	4% / 0.99/0.99
	HyCo dtu tum isr	16%/0.99/0.89 -- 13%/0.97/0.94

* without outliers

3.6.4 Summary

- On the Ob River, the WLTS obtained from ISR retracker provide better estimation of the discharge. However, the main drawback of the ISR retracker is frequent failure during the ice melt period.
- During the ice melt the DTU retracker overestimates frequently the water level by 1-2 meters, however it performs better than other retrackers during the flood peak.
- Over the Rhine River, in "quasi-ideal" VS configuration (perpendicular track-river orientation, large channel, flat banks, narrow floodplain, absence of oxbow lakes and secondary channels) all retrackers perform equally good, comparable with the performance of enhanced 80Hz SAMOSA+ processing provided by ESA GPOD Service.
- In more complex VS configurations (high banks, narrow (~200 m) channel, parallel track-river orientation) WLTS obtained with DTU retracker are less accurate and have higher number of outliers.
- In spite of a higher number of gaps, ISR water level time series are more suitable for the discharge retrievals on both Ob and Rhine Rivers. However, using only ISR retracker a risk of having gaps during critical periods of ice melt and flood water level rise on Arctic rivers is sufficiently high. The TUM retracker demonstrated better performance during these periods.

3.7 Validation against in situ data for Po and Mississippi Rivers (CNR-IRPI)

The L3 product is evaluated against water level measurements acquired at the in situ stations and available from public platforms (https://www.agenziapo.it/content/monitoraggio-idrografico-0 for the Po River and https://waterdata.usgs.gov/nwis for the Mississippi River). Specifically, for the satellite validation analysis, hourly and sub-hourly in situ data have been used to be sure not to add further uncertainties in the comparison between the two time series. The following metrics have been used for the analysis based on the PVP for water level comparison: coefficient of correlation, R, Median Absolute Deviation, MAD, and the standard deviation of the error, std. All the metrics have been evaluated in terms of relative heights to avoid influence of difference in datum (often unknown for the in situ measurements) or the distance between the virtual station and the gauged station. The relative heights have been computed by removing the mean of both the temporal series when they are consistent in time.

Along the Po River (Figure 3.7.1) the three retrackers show a different number of tracks, with a highest number for the TUM and the lowest for the ISR. The comparison has been carried out by comparing each virtual station against the in situ stations located upstream and downstream. In terms of the performances, Tables 3.7.1, 3.7.2 and 3.7.3 report the results in terms of R, MAD and STD. At the end of each table the mean and the median of all the results are reported for all the tracks and for the tracks in common between the retrackers datasets. No sensitive differences are detectable from the analysis. In most of the cases the data over the virtual stations are superimposed to the in situ stations as shown in Figure 3.7.2. Only in a few cases the comparison is not good as represented in Figure 3.7.3 where DTU and TUM show constant values not considered reliable to represent the time series. Speculating on the results, isardSAT retracker seems to show performances higher than the other two, even if it has a smallest coverage in terms of tracks.

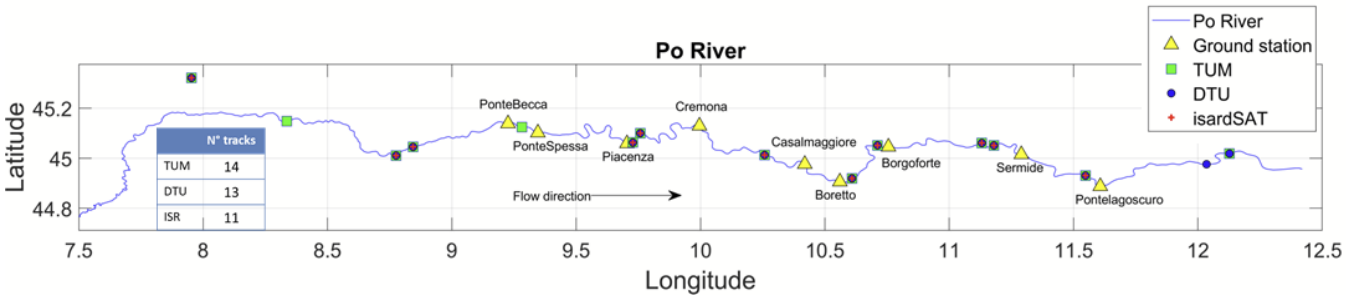


Figure 3.7.1. Po River: location of the in situ stations and the satellite tracks analyzed by the different retrackers (TUM, DTU and isardSAT). The table included in the figure shows the number of tracks for each retrackers along the main course.

Table 3.7.1. Performances in terms of coefficient of correlation between the different time series derived by the several retrackers and the in situ stations located upstream and downstream the virtual stations.

	DTU		ISR		TUM	
Satellite tracks	**Upstream**	**Downstream**	**Upstream**	**Downstream**	**Upstream**	**Downstream**
21406100040241	0.98	0.98	0.996	0.996	0.99	0.99
21406100070511	0.99	0.999	0.99	0.99	0.99	0.998

21406100080261	0.91	0.91	0.90	0.90	0.90	0.90
21406100100721	0.998	0.997	0.997	0.99	0.998	0.996
21406100290415	0.80	-	-	-	0.63	-
21406100310075	0.51	-	-	-	-	-
21406300030551	0.996	0.99	0.996	0.99	0.996	0.99
21406300070481	0.97	0.97	0.97	0.97	0.97	0.97
21406500030481	0.96	0.94	0.91	0.90	0.91	0.90
21406500040161	0.30	0.29	0.92	0.91	0.31	0.30
21406700040121	-	0.88	-	0.87	-	0.90
21406700040481	-	0.88	-	0.88	-	0.86
21406500070481	-	-	-	-	0.96	0.96
21406900070391	-	-	-	-	-	0.40
Mean	**0.84**	**0.88**	**0.96**	**0.94**	**0.87**	**0.85**
Median	**0.97**	**0.95**	**0.98**	**0.94**	**0.97**	**0.93**
ONLY COMMON TRACKS						
Mean	**0.89**	**0.88**	**0.96**	**0.94**	**0.89**	**0.88**
Median	**0.98**	**0.96**	**0.98**	**0.94**	**0.98**	**0.93**

Table 3.7.2. Performances in terms of Median Absolute Deviation between the different time series derived by the several retrackers and the in situ stations located upstream and downstream the virtual stations.

	DTU		**ISR**		**TUM**	
Satellite tracks	**Upstream**	**Downstream**	**Upstream**	**Downstream**	**Upstream**	**Downstream**
21406100040241	0.07	0.08	0.09	0.09	0.07	0.09
21406100070511	0.13	0.05	0.20	0.08	0.20	0.05
21406100080261	0.16	0.05	0.14	0.06	0.14	0.06

21406100100721	0.06	0.16	0.11	0.16	0.09	0.18
21406100290415	0.59	-	-	-	0.62	-
21406100310075	0.65	-	-	-	-	-
21406300030551	0.11	0.13	0.10	0.15	0.11	0.17
21406300070481	0.18	0.20	0.26	0.21	0.15	0.19
21406500030481	0.18	0.43	0.24	0.59	0.14	0.38
21406500040161	0.29	0.60	0.09	0.37	0.29	0.60
21406700040121	-	0.27	-	0.38	-	0.26
21406700040481	-	0.24	-	0.28	-	0.26
21406500070481	-	-	-	-	0.14	0.12
21406900070391	-	-	-	-	-	0.25
Mean	**0.24**	**0.22**	**0.15**	**0.24**	**0.19**	**0.22**
Median	**0.17**	**0.18**	**0.12**	**0.19**	**0.14**	**0.19**
ONLY COMMON TRACKS						
Mean	**0.46**	**0.51**	**0.36**	**0.42**	**0.45**	**0.51**
Median	**0.31**	**0.42**	**0.27**	**0.42**	**0.26**	**0.40**

Table 3.7.3. Performances in terms of standard deviation between the different time series derived by the several retrackers and the in situ stations located upstream and downstream the virtual stations.

	DTU		**ISR**		**TUM**	
Satellite tracks	**Upstream**	**Downstream**	**Upstream**	**Downstream**	**Upstream**	**Downstream**
21406100040241	0.36	0.32	0.21	0.15	0.25	0.18
21406100070511	0.30	0.08	0.36	0.18	0.35	0.08
21406100080261	0.49	0.46	0.51	0.49	0.51	0.49
21406100100721	0.14	0.16	0.18	0.22	0.14	0.18

21406100290415	1.00	-	-	-	1.10	-
21406100310075	1.33	-	-	-	-	-
21406300030551	0.21	0.25	0.22	0.25	0.22	0.26
21406300070481	0.30	0.33	0.30	0.33	0.28	0.32
21406500030481	0.24	0.69	0.33	0.75	0.32	0.68
21406500040161	1.49	2.11	0.61	1.04	1.50	2.12
21406700040121	-	0.56	-	0.49	-	0.45
21406700040481	-	0.54	-	0.54	-	0.58
21406500070481	-	-	-	-	0.37	0.30
21406900070391	-	-	-	-	-	2.16
Mean	**0.58**	**0.55**	**0.34**	**0.44**	**0.50**	**0.65**
Median	**0.33**	**0.39**	**0.31**	**0.41**	**0.33**	**0.38**
ONLY COMMON TRACKS						
Mean	**0.19**	**0.19**	**0.20**	**0.21**	**0.18**	**0.20**
Median	**0.16**	**0.16**	**0.13**	**0.21**	**0.15**	**0.17**

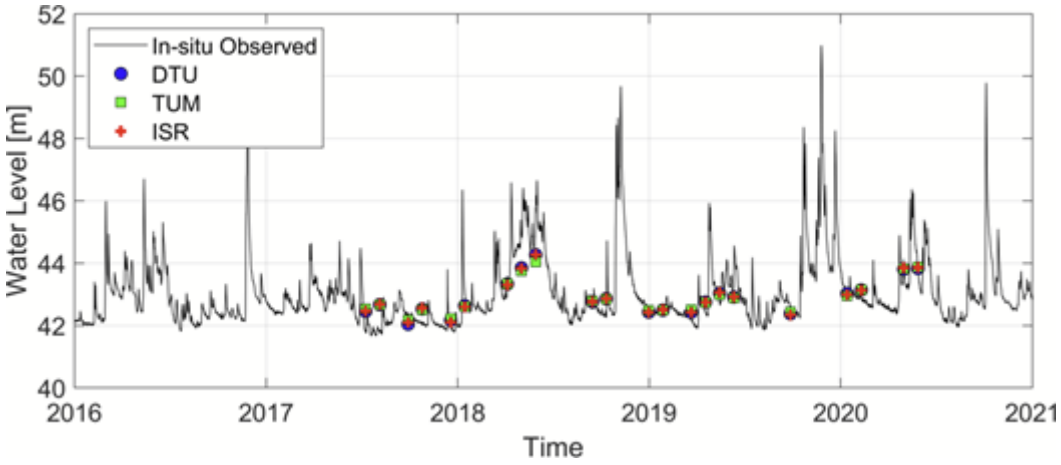


Figure 3.7.2. Comparison in terms of time series of the in-situ measurements acquired at Casalmaggiore station and the virtual tracks derived by the different retrackers for the track 21406300030551.

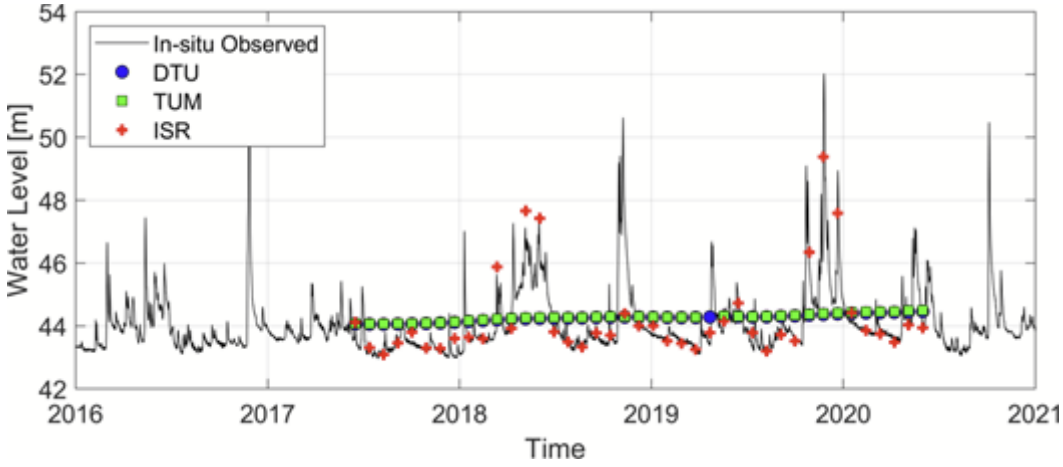


Figure 3.7.3. Comparison in terms of time series of the in-situ measurements acquired at Ponte Spessa station and the virtual tracks derived by the different retrackers for the track 21406500040161.

Following the same approach, the validation over the Mississippi has been carried out by comparing the several retrackers derived at each virtual station illustrated in Figure 3.7.4b against the in-situ stations upstream and downstream along the main course for which the sub-hourly water level data are available. The results of the analysis are reported in Tables 3.7.4, 3.7.5 and 3.7.6. As previously the different retrackers are very similar to each other and no appreciable differences are found. Based on the selected metrics ISR retracker seems the best even if it has a lower number of tracks.

The two analyses over the Po and the Mississippi River underlined not evident differences between the retrackers in the definition of reliable time series of water levels. Speculating on the results, even if it shows a lower number of available tracks, ISR seems to be more accurate, but two basins cannot be exhaustive to this conclusion and additional test areas should be analyzed to confirm it.

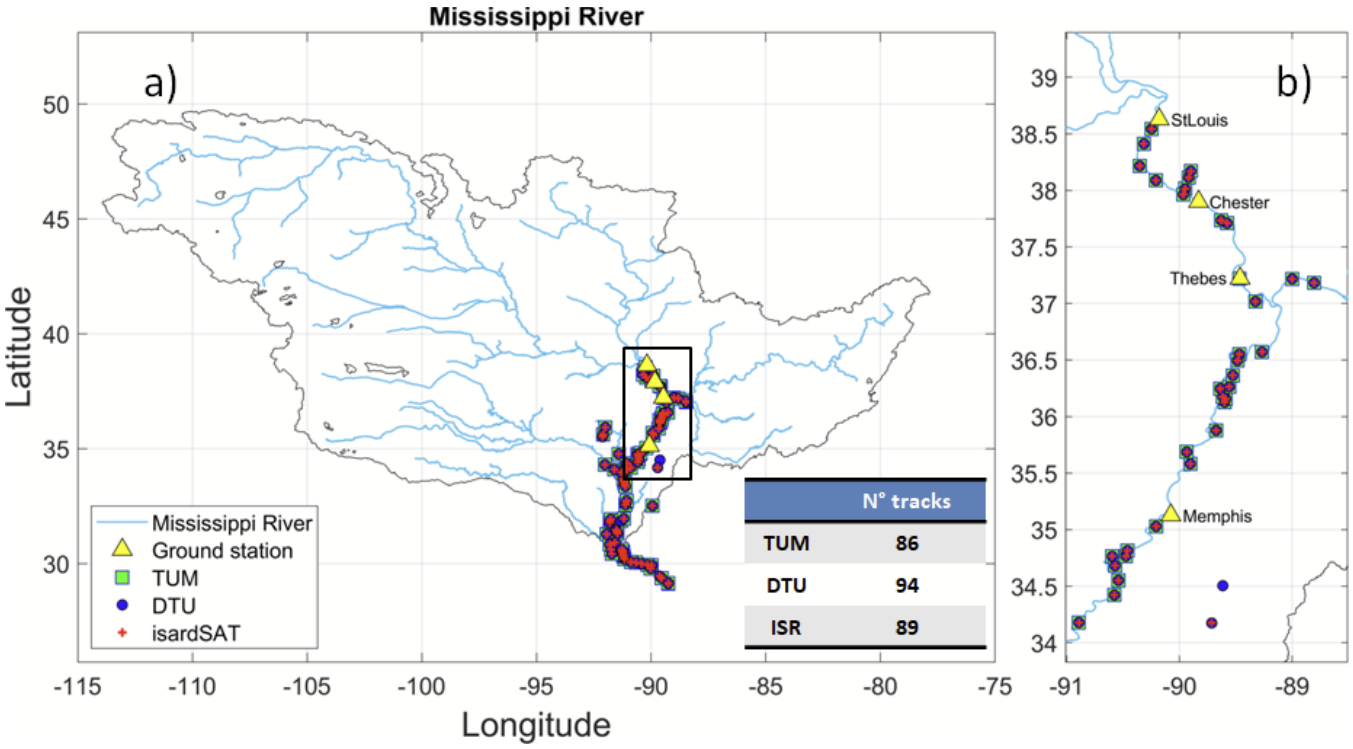


Figure 3.7.4. Mississippi basin: a) location of the in-situ stations and the satellite tracks analyzed by the different retrackers (TUM, DTU and isarSAT), b) in situ and virtual stations considered in the validation analysis. The table included in the figure shows the number of tracks for each retrackers along the main course.

Table 3.7.4. Performances in terms of coefficient of correlation between the different time series derived by the several retrackers and the in situ stations located upstream and downstream the virtual stations.

	DTU		ISR		TUM	
Satellite tracks	**Upstream**	**Downstream**	**Upstream**	**Downstream**	**Upstream**	**Downstream**
74253000030581	0.996	0.99	0.996	0.99	0.99	0.99
74253000080321	0.85	0.83	0.98	0.96	0.96	0.96
74253000100071	0.995	0.995	0.995	0.995	0.99	0.996
74254000040741	0.96	0.96	0.99	0.996	0.99	0.997
74255000010381	0.997	0.97	0.997	0.97	0.99	0.96
74255000020761	0.997	0.97	0.99	0.96	0.99	0.96
74255000030081	0.91	-	0.98	-	0.98	-
74255000070211	0.93	0.86	0.94	0.85	0.91	0.84

74257000020141	0.64	0.98	0.66	0.98	0.67	0.98
74257000030401	0.57	0.82	0.59	0.93	0.49	0.90
74257000070041	0.68	0.99	0.68	0.99	0.63	0.98
74259000020401	0.69	0.98	0.68	0.99	0.60	0.98
74259000030061	0.67	0.99	0.67	0.98	0.61	0.99
74259000030161	0.66	0.99	0.65	0.97	0.72	0.98
74259000040451	0.68	0.98	0.67	0.99	0.72	0.99
74259000050351	0.68	0.95	0.65	0.97	0.57	0.95
74259000070451	0.64	0.90	0.68	0.97	0.67	0.97
74259000100501	0.62	0.90	0.71	0.82	0.58	0.93
74259000110241	0.45	0.97	0.48	0.995	0.49	0.99
74259000140241	0.46	0.997	0.46	0.995	0.45	0.99
74270100100181	0.999	-	0.998	-	0.997	-
74270100130281	0.995	-	0.996	-	0.99	-
74270500020051	0.90	-	0.99	-	0.99	-
74270500020331	0.99	-	0.99	-	0.99	-
74270700060801	0.89	-	0.89	-	0.91	-
74270700220161	0.99	-	0.99	-	0.99	-
74270900010181	0.99	-	0.99	-	0.99	-
74270900030201	0.98	-	0.98	-	0.98	-
Mean	**0.81**	**0.95**	**0.83**	**0.96**	**0.82**	**0.96**
Median	**0.89**	**0.97**	**0.96**	**0.98**	**0.94**	**0.98**

Table 3.7.5. Performances in terms of Median Absolute Deviation between the different time series derived by the several retrackers and the in situ stations located upstream and downstream the virtual stations.

	DTU		ISR		TUM	
Satellite tracks	**Upstream**	**Downstream**	**Upstream**	**Downstream**	**Upstream**	**Downstream**
74253000030581	0.14	0.30	0.14	0.27	0.17	0.35
74253000080321	0.09	0.18	0.16	0.22	0.11	0.20
74253000100071	0.12	0.16	0.14	0.14	0.20	0.13
74254000040741	0.18	0.13	0.16	0.15	0.20	0.10
74255000010381	0.05	0.52	0.07	0.47	0.06	0.52
74255000020761	0.12	0.61	0.18	0.46	0.18	0.56
74255000030081	0.31	-	0.27	-	0.32	-
74255000070211	0.53	1.57	0.53	1.57	0.68	1.63
74257000020141	1.62	0.39	1.58	0.54	1.44	0.50
74257000030401	1.75	0.48	1.78	0.34	2.01	0.37
74257000070041	1.93	0.35	1.91	0.29	1.80	0.40
74259000020401	1.48	0.47	1.68	0.50	1.71	0.55
74259000030061	1.71	0.43	1.65	0.39	1.79	0.38
74259000030161	1.20	0.34	1.28	0.34	1.29	0.37
74259000040451	1.06	0.50	1.32	0.44	1.34	0.47
74259000050351	1.68	0.69	1.57	0.55	1.72	0.52
74259000070451	1.15	0.52	0.97	0.53	1.23	0.53
74259000100501	1.62	0.46	1.15	0.50	1.44	0.50
74259000110241	1.44	0.20	1.19	0.21	1.06	0.19
74259000140241	1.39	0.27	1.41	0.31	1.44	0.28
74270100100181	0.12	-	0.16	-	0.14	-

74270100130281	0.19	-	0.25	-	0.24	-
74270500020051	0.31	-	0.32	-	0.30	-
74270500020331	0.25	-	0.28	-	0.27	-
74270700060801	0.89	-	0.81	-	0.86	-
74270700220161	0.19	-	0.19	-	0.20	-
74270900010181	0.20	-	0.19	-	0.29	-
74270900030201	0.24	-	0.31	-	0.24	-
Mean	**0.78**	**0.45**	**0.77**	**0.43**	**0.81**	**0.45**
Median	**0.42**	**0.43**	**0.43**	**0.39**	**0.50**	**0.40**

Table 3.7.6. Performances in terms of Standard Deviation between the different time series derived by the several retrackers and the in situ stations located upstream and downstream the virtual stations.

	DTU		**ISR**		**TUM**	
Satellite tracks	**Upstream**	**Downstream**	**Upstream**	**Downstream**	**Upstream**	**Downstream**
74253000030581	0.283	0.38	0.285	0.38	0.32	0.38
74253000080321	1.31	1.40	0.50	0.62	0.55	0.59
74253000100071	0.285	0.244	0.278	0.238	0.30	0.227
74254000040741	0.77	0.73	0.35	0.227	0.37	0.207
74255000010381	0.207	0.69	0.201	0.70	0.41	0.68
74255000020761	0.184	0.79	0.27	0.75	0.29	0.80
74255000030081	0.79	-	0.37	-	0.43	-
74255000070211	0.68	1.68	0.66	1.70	0.75	1.74
74257000020141	1.90	0.49	1.88	0.55	1.89	0.56
74257000030401	2.26	1.69	2.22	1.09	2.46	1.22
74257000070041	1.97	0.52	1.98	0.50	1.98	0.59

74259000020401	1.79	0.74	1.80	0.65	1.89	0.71
74259000030061	1.88	0.57	1.89	0.61	1.94	0.57
74259000030161	1.83	0.65	1.81	0.84	1.72	0.81
74259000040451	1.72	0.75	1.78	0.66	1.71	0.71
74259000050351	1.74	1.00	1.73	0.95	1.86	1.02
74259000070451	1.77	1.33	1.67	0.92	1.71	0.91
74259000100501	1.94	1.26	1.98	1.70	1.90	1.09
74259000110241	2.29	0.65	2.24	0.295	2.24	0.35
74259000140241	2.25	0.284	2.25	0.307	2.28	0.31
74270100100181	0.233	-	0.248	-	0.274	-
74270100130281	0.325	-	0.345	-	0.40	-
74270500020051	1.18	-	0.43	-	0.52	-
74270500020331	0.35	-	0.45	-	0.38	-
74270700060801	1.26	-	1.26	-	1.28	-
74270700220161	0.41	-	0.42	-	0.43	-
74270900010181	0.39	-	0.39	-	0.39	-
74270900030201	0.61	-	0.60	-	0.64	-
Mean	**1.16**	**0.83**	**1.08**	**0.72**	**1.12**	**0.71**
Median	**1.22**	**0.73**	**0.63**	**0.65**	**0.70**	**0.68**

3.8 Validation against in situ data over the Canadian Lakes Reindeer and Wollaston (DTU)

L2 and L3 products are validated against in situ gauge data for Reindeer and Nonacho lakes. Figure 3.8.1 includes L2 data, L3 time series and in situ gauge data for Reindeer lake with S3A measurements. Data for each track is indicated by colour. Similar figures are produced for Nonacho lake and for S3B for both lakes. Focus is on the relative water variations to avoid bias issues between datums and the median error between the time series and the in situ data has been added to the in situ time series.

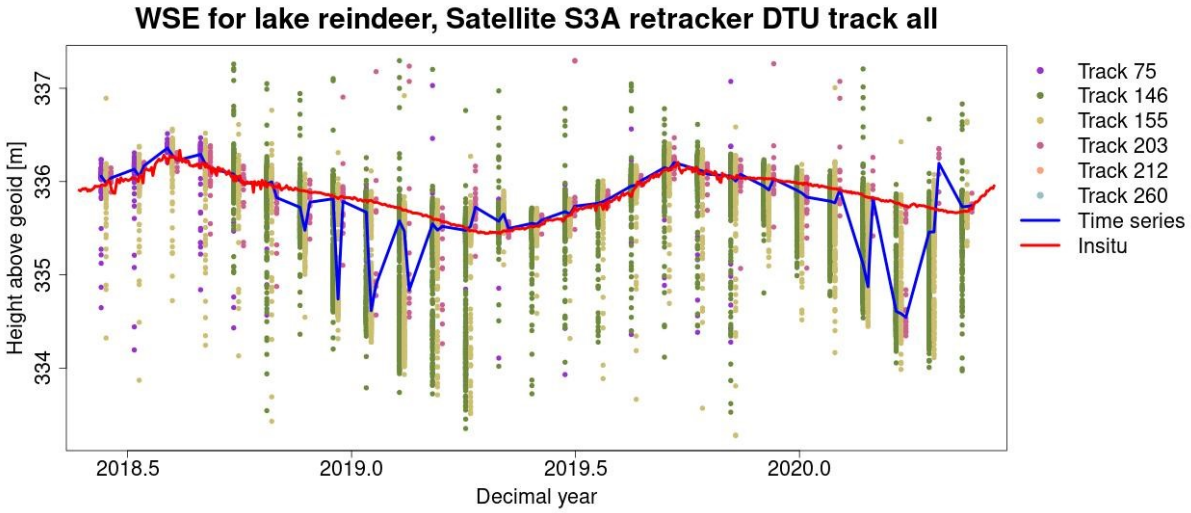


Figure 3.8.1. L2 and L3 products for Sentinel 3A for Reindeer lake. Track number is indicated by color. Blue line is the L3 time series product and red line is the in situ gauge water level.

3.8.1 Validation of L2 water level for Reindeer lake

Table 3.8.1 summarises the performance of the three retrackers across S3A and S3B for lake Reindeer. Each retracker has produced between 12000 and 14000 data points across the lake. The RMSE for Sentinel 3A lies between 1.14 m and 1.28 m, where TUM has the lowest RMSE. The percentage of points within 0.5 meters of the in situ water level is indicated by p_valid_05. The percentage lies within 69% and 79%, where ISR has the lowest percentage and DTU the highest. P_valid_02 indicates the percentage of points within 0.2 meters of the in situ water level. TUM has the smallest percentage of points within 0.2 meters (35%) while the DTU retracker is significantly better with 65%. The MAD, median absolute deviation, is significantly lower for the DTU retracker than for TUM and ISR.

Table 3.8.1. Statistics for L2 product for satellites Sentinel 3A and Sentinel 3B for Reindeer lake.

SAT	Retrack	n_valid	RMSE	p_valid_05	p_valid_02	minOC	maxOC	MAD
S3A	DTU	14044	1.28m	79.05%	64.84%	31%	100%	0.12m
S3A	TUM	12249	1.14m	75.35%	35.01%	31%	100%	0.27m
S3A	ISR	12027	1.22m	68.90%	42.34%	31%	100%	0.25m
S3B	DTU	9609	1.59m	76.54%	62.67%	31%	99%	0.13m
S3B	TUM	8699	1.24m	77.91%	42.01%	31%	99%	0.30m
S3B	ISR	8399	1.30m	68.04%	34.34%	31%	99%	0.30m

For satellite Sentinel 3B 8400-9600 measurements are available. TUM has the smallest RMSE of 1.24 m while DTU has the largest RMSE of 1.59 m. The percentage of points within 0.5 meters of the in situ data is highest for TUM (78%) followed by DTU (77%) while ISR has significantly lower percentage of 68%. DTU has the highest

percentage of points lying within 0.2 meters of in situ water level (63%), where TUM and ISR are 10% smaller. For S3B the MAD is also significantly smaller for DTU than for the two other retrackers.

Figure 3.8.2 shows histograms for the residuals between in situ water level and L2 water levels. Residuals extend beyond the span of the histogram. The DTU retracker has the largest part of the measurements within 20 cm of the in situ water levels. This indicates that the DTU retracker has a smaller fraction of outliers.

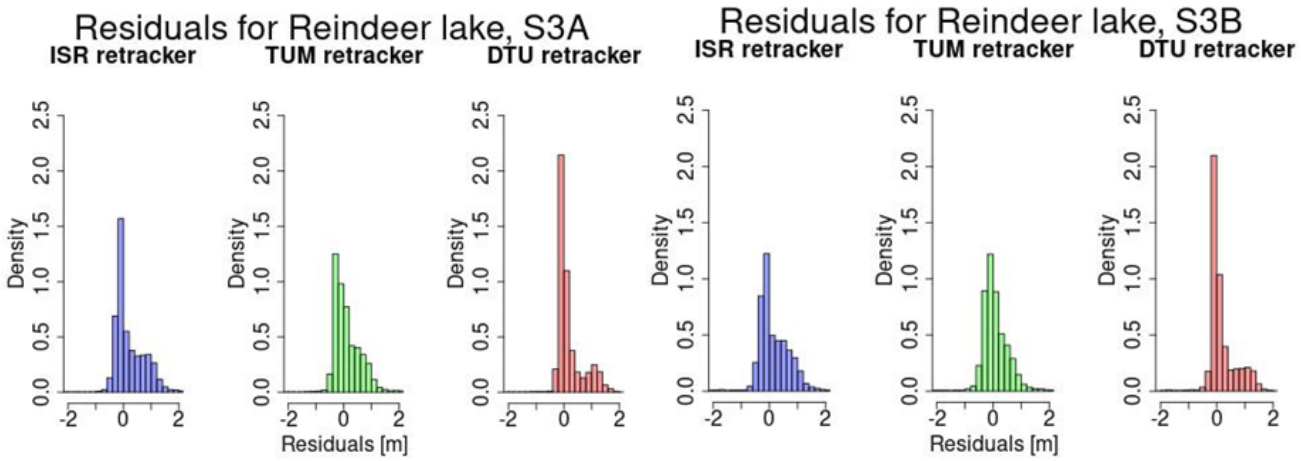


Figure 3.8.2. Histograms of residuals for Reindeer lake for satellite S3A and S3B.

3.8.2 Validation of L2 water level for Nonacho lake

Figure 3.8.3 illustrates the WSE and the L3 time series for the ISR retracker for Nonacho lake. Due to the size and shape of Nonacho lake compared to Rendeer lake, the L2 products contain less measurements. No reference datum is available for the in situ gauge data, and the median error between the time series and the in situ data has therefore been added to the in situ WSE for visualisation. The focus is therefore on the relative water variations.

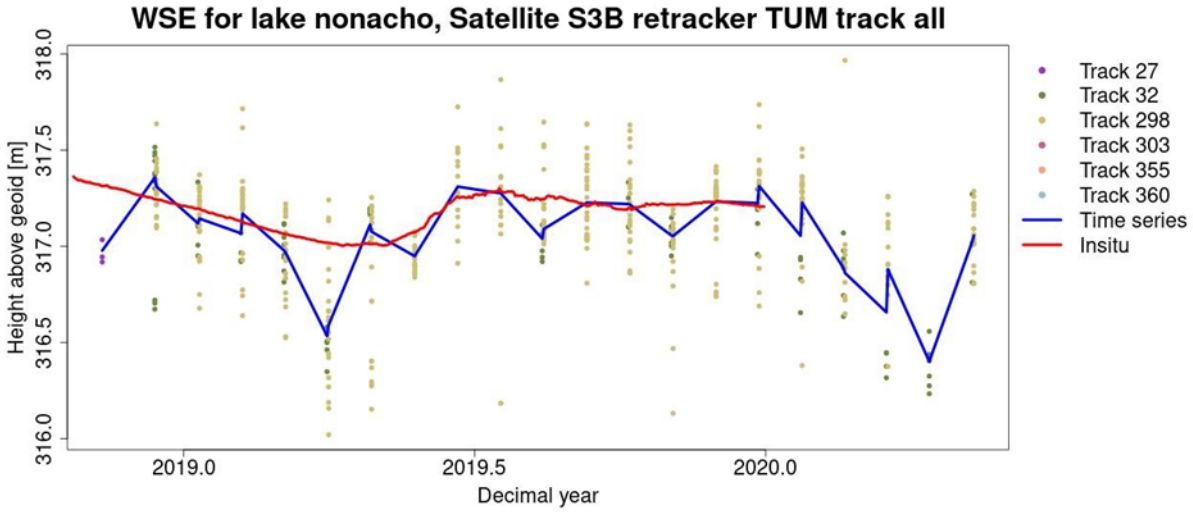


Figure 3.8.3. L2 and L3 products for Sentinel 3A for Reindeer lake. Track number is indicated by color. Blue line is the L3 time series product and red line is the in situ gauge water level.

Table 3.8.2 shows statistics for lake Nonacho. For Sentinel 3A the RMSE and MAD are lowest for DTU, and the percentage of points within 0.2 m and 0.5 m are largest for DTU. The RMSE for retracker ISR are significantly larger than for DTU and TUM. The MAD is 0.14 m and 0.15 m for ISR and TUM respectively.

Table 3.8.2. Statistics for L2 product for satellites Sentinel 3A and Sentinel 3B for Nonacho lake.

SAT	Retrack	n_valid	RMSE	p_valid_05	p_valid_02	minOC	maxOC	MAD
S3A	DTU	845	0.74m	89.36%	82.23%	33%	97%	0.06m
S3A	TUM	723	0.79m	85.77%	61.40%	33%	97%	0.15m
S3A	ISR	796	1.12m	82.36%	59.50%	33%	97%	0.14m
S3B	DTU	477	0.93m	88.22%	80.63%	32%	97%	0.06m
S3B	TUM	393	1.08m	88.73%	59.59%	32%	97%	0.16m
S3B	ISR	479	0.79m	85.38%	64.96%	31%	97%	0.12m

For Sentinel 3B the RMSe is largest for TUM (1.08 m), and smallest for ISR (0.79 m). Looking at the percentage of points within 0.2 m and 0.5 m, DTU performs better than the two other retrackers. The MAD is 0.06 m for the DTU retracker and is 0.12 m and 0.16 m for ISR and TUM respectively.

Figure 3.8.4 shows histograms of the residuals between the in situ water level and the retracked water levels. The DTU retracker provides the largest portion of the measurements close to the in situ values.

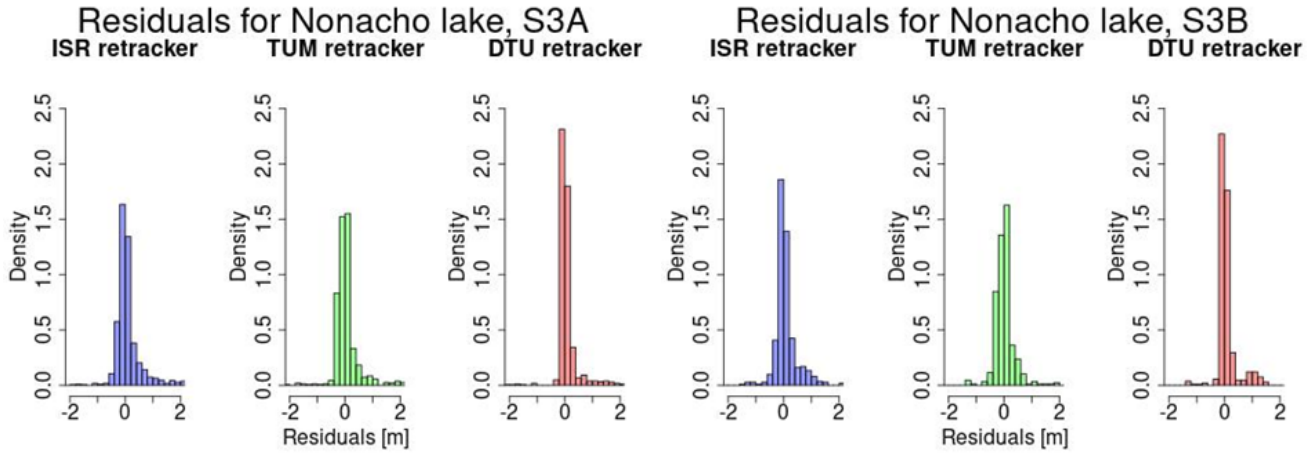


Figure 3.8.4. Histograms of residuals for Nonacho lake for satellite S3A and S3B.

3.8.3 Validation of L3 water level for Reindeer lake

The L3 product is a state-space model that estimates the water level each time a track passes over the lake. In Table 3.8.3, n_valid describes how many instances of L2 data are available, which is the number of instances where a L3 water level is estimated. For S3A the TUM retracker has 80 instances while DTU and ISR have 81, which indicates that one satellite crossing has not been produced by TUM. The TUM retracker has the smallest RMSE of 0.33 m and ISR has the largest RMSE of 0.44 m. The bias describes the median off-set between the L3 time series and the in situ. The bias is notably larger for the TUM retracker than for the others.

The correlation indicates how well the relative water level variations between L3 and in situ agrees. A high correlation is favorable as the time series thereby captures the seasonal variations. Looking at the percentage of points lying within 0.5 m and 0.2 m of the in situ, the DTU retracker is favorable, with p_valid_05 of 78% and p_valid_02 of 65%. The medSD is the median standard deviation of the L3 water level. This indicates how certain the time series are determined based on the L2 water level data. All medSD lies between 0.1 m and 0.2 m.

Table 3.8.3. Statistics for L3 product for satellites Sentinel 3A and Sentinel 3B for Reindeer lake.

SAT	Retrack	n_valid	RMSE	Bias	Corr	p_valid_05	p_valid_02	medSD	MAD
S3A	DTU	81	0.36m	0.21m	0.58	78.35%	65.03%	0.01m	0.05m
S3A	TUM	80	0.33m	0.39m	0.71	74.11%	38.28%	0.02m	0.15m
S3A	ISR	81	0.44m	0.16m	0.67	66.73%	46.99%	0.02m	0.11m
S3B	DTU	60	0.38m	0.24m	0.44	76.84%	62.58%	0.01m	0.08m
S3B	TUM	60	0.29m	0.42m	0.67	76.49%	43.48%	0.02m	0.13m
S3B	ISR	60	0.42m	0.20m	0.53	66.53%	39.31%	0.02m	0.18m

3.8.4 Validation of L3 water level for Nonacho lake

The RMSE is less than 0.1 m for the DTU retracker for both satellites, while it is above 0.1 m for the other two retrackers. The very largest bias for Nonacho lake occurs as no reference datum is available for the in situ data. The correlation is best for the DTU retracker, with values of 0.89 and 0.73, indicating that the seasonal variations are reflected in the time series. Inspecting the percentage of points within 0.5 m and 0.2 m of the in situ, DTU performs best while ISR performs poorest. The median standard deviation is slightly better for the DTU retracker where it is 0.02 m and 0.03 m, and 0.04m for the ISR and TUM retrackers.

Table 3.8.4. Statistics for L3 product for satellites Sentinel 3A and Sentinel 3B for Nonacho lake.

SAT	Retrack	n_valid	RMSE	Bias	Corr	p_valid_05	p_valid_02	medSD	MAD
S3A	DTU	53	0.06m	-331.33m	0.89	90.47%	82.66%	0.02m	0.03m
S3A	TUM	51	0.12m	-311.14m	0.71	87.09%	61.88%	0.04m	0.08m
S3A	ISR	54	0.17m	-311.38m	0.66	83.19%	59.29%	0.04m	0.10m

S3B	DTU	25	0.09m	-311.33m	0.73	88.01%	80.44%	0.03m	0.04m
S3B	TUM	26	0.16m	-311.17m	0.63	87.81%	64.11%	0.04m	0.09m
S3B	ISR	27	0.22m	-311.37m	0.59	85.13%	64.78%	0.04m	0.08m

3.8.5 Conclusion

The above analyses of L2 product for the retrackers TUM, ISR, and DTU find largely the same statistics for performance for Nonacho and Reindeer lakes. Generally, S3A contains more L2 data points, but this can be due to the overlap of the satellite tracks across the lakes. A difference in performance between the retrackers are clearly seen in the p_valid parameters. The DTU retracker detects a larger amount of the water levels within 0.2 m of the in situ than the two other retrackers. This is supported by the histograms visualising the distribution of residuals.

Validation of the L3 product shows that the seasonal changes are overall captured by the time series based on the correlation coefficient. All correlation coefficients are above 0.44 with the highest correlation coefficient of 0.89. The median standard deviation for the time series is largely the same, indicating that the time series are equally reliable. The percentage of points that lies within 0.5 m and 0.2 m of the in situ are largest for the DTU retracker. This is an effect of the performance of the L2 product.

Overall the DTU retracker performs better than the ISR and TUM retrackers. The DTU retracker is produced to perform well at inland waters, and the slightly better performance of the DTU retracker is therefore expected. The TUM and ISR retrackers are difficult to rate, as the performance varies between the lakes and the satellites.

3.9 Synthesis of Inland Water Validation Results. Conclusions and Recommendations.

3.9.1 Synthesis of Results

Note that these results refer only to Sentinel 3A and Sentinel 3B data. So far no Inland Water products from Cryosat-2 have been evaluated.

U Bonn, TUM, DTU Environment, NUIM, CNR-IRPI, and DTU Space evaluated the retracked S3A / S3B SAR altimeter data as L2 data (along-track product), L3 data (water level time series at specified locations), or L4 data (river discharge), over 11 different inland water regions. This summary considers the performance of the different re-tracking algorithms to Level 2.

L2 data from 3 re-trackers were under consideration: DTU, isardSAT and TUM.

U Bonn

U Bonn evaluated L2 SAR altimeter water level data against river gauges in the River Rhine, based on L3 time series generated for a number of virtual stations within 30km range of the in-situ river gauges. The performance of the three re-trackers was found to be very similar, with root mean square error ranging between 25 to 35 cm, and normalised rmse from 8.5% to 15%.

For the 10 stations where data from all re-trackers were available, the TUM re-tracker data give “better” NRMSE, NSE (Nutcliff coefficient), R2 (correlation), and MAD (mean absolute deviation). If all stations are considered, the DTU re-tracker provides data for the largest number of virtual stations, returns the largest number of data records, and the “best” statistics in terms of RMSE, NRMSE, and MAD.

U Bonn also used an in-house methodology to compute L3 time-series from the along-track retracked data and from standard Copernicus products. Differences were noticed, which arise both from the different methodologies but also from missing data. The first are minor, the second are not relevant for the validation, but needed for the U Bonn study case.

TU Munich (TUM)

TUM compared the L3 products, produced by DTU from the L2 data, using data from the DAHITI data base. Data for 8 river basins were considered, 1091 virtual stations were available in the HYDROCOASTAL data, 1086 corresponding stations were created with DAHITI data, and 246 of these could be validated against in-situ gauge data.

Little difference was found between the accuracy of data from the different re-trackers, with results (in terms of best performing re-tracker) varying between river basins and different types of river topography, with median rmse of 33-34 cm across all common stations. A “bootstrapped” median root mean square difference was also calculated and found that the DTU data gave the lowest median rmsd, followed by TUM and isardSAT. Again, it was found that DTU re-tracker provided data for the largest number of virtual stations, and returned the largest number of data records. It was suggested that the process to generate L3 data from L2 data could have a greater impact on data quality than the re-tracker used to generate L2 data.

In summary, across all basins and stations, the data derived from DTU were found to show the lowest rmsd, and provide most stations, so this re-tracker was recommended for global reprocessing.

DTU Env

DTU Env considered data from the Amur, Yangtze and Zambezi rivers. Because of limited data availability over the Amur (both satellite and in-situ), there are insufficient data to provide a good statistical basis for a comparison. Over the Yangtze River, the DTU data were found to be clearly better than both TUM and isardSAT data. For the Zambezi River, performance of the different re-trackers was very similar, with DTU and TUM outperforming isardSAT by a small margin.

NUIM

NUIM evaluated river discharge data (L4) for the River Ob and Rhine, and also river level time series data (L3) for the Ob. For this summary we focus on results from the River Ob, as the Rhine is fully covered by others (U Bonn and DAHITI).

It was found that the DTU re-tracker retrieved most data, and gave stable discharge values independent of local conditions, but also had most outliers and poorer performance in locations with complex geomorphology. The isardSAT re-tracker gave better accuracy and better results in complex relief, but had more data gaps and failed frequently during ice melt. Finally, the TUM re-tracker showed higher accuracy, better handling of relief, and best results during ice melt – but there were more gaps in the data and the re-tracker had problems where there were lakes close to the main river.

When recommending a re-tracker for operational tasks – the DTU re-tracker will provide the most complete data set but with lower accuracy than isardSAT or TUM. For research tasks, it should be noted that the TUM re-tracker shows higher accuracy but with higher dispersion in low accuracy tail, whereas the DTU and isardSAT re-tracker provide a “smoother” Time Series.

It was concluded that an ideal solution would be to use all three re-trackers to provide the most complete coverage under all conditions and situations.

CNR-IRPI

CNR-IRPI evaluated L3 data for the Po and Mississippi rivers. Based on the selected metrics ISR retracker seems the best even if it has a lower number of tracks. The results showed similar performances from the different re-trackers, with no appreciable differences between them. DTU again retrieved more records.

There were some strange results at some locations, with fixed water levels being returned.

DTU-Space

DTU-Space evaluated L2 for Nonacho and Reindeer lakes in North America. Both lakes are frozen for part of the year. Whilst there were insufficient data to support a recommendation for the best performing re-tracker, it was noted that none of the re-trackers gave realistic data when the lakes were ice covered.

3.9.2 Conclusions and Recommendations

Sentinel 3A and Sentinel 3B data

There is a consistent finding across the evaluations with larger data sets that the DTU re-tracker was on average found to be the best performing, in terms of accuracy of data retrieved (lowest NRMSE, RMSD) and the number of valid points retrieved.

However, there were some locations and types of river topography (complex geomorphology) in which the DTU retracker was not found to provide good estimates. In cases of complex relief the TUM and isardSAT re-trackers gave better results, and the TUM re-tracker provided better results than other re-trackers during ice melt. Therefore different re-trackers may be preferred for different types of inland water topographies and environments.

Therefore the main recommendation from this evaluation is that the DTU re-tracker is the preferred option for generating the global scale data set from Sentinel 3A and 3B data. If it is possible, other re-trackers could be used to provide coverage in specific types of topography and environment where the DTU re-tracker did not perform well (for instance in complex terrain and during ice melt).

An option for future development consideration could involve some form of classification scheme for inland waters, and the selection of different re-trackers most suitable for different situations. This could also involve the tuning of current re-trackers for these different classifications.

Cryosat-2 Data

Cryosat-2 data have not yet been evaluated (SAR or SARIN mode). There is extensive coverage of the Amazon basin in the SARIN mode, and it is anticipated that retracked SARIN mode data could be effective in regions to enable the selection of the desired target where there are a number of reflecting sites close to the satellite track. CryoSat-2 SARIN mode covers the Rhine river since 2017 and L3 time-series built at UBonn using the data processed by the SAMOSA+ retracker at GPOD/Earth Console are available for comparison.

3.10 Discharge retrievals from Hydrocoastal L3 product for the test sites (NUIM)

The aim of the study is an evaluation of the methods of the discharge retrieving from altimetry measurements. In the phase 1, the discharge retrieving was done using rating curve method, Bjerklie equation and Manning equation.

The water level time series obtained using DTU retracker was selected for current test, as this retracker was chosen for elaboration of the L3 product in the phase 2 of the project.

The rating curve method (RC) is the classical empirical method used by hydrological services for daily discharge estimation from in situ measurements of the water level. The rating curves are the relationships established between measurements of the water level at a gauging station (GS) and simultaneous instrumental measurements of the discharge effectuated several times a year. Similar rating curves can be established between altimetry derived water level at a virtual station (VS) and daily in situ discharge (Qinsitu) at a nearest gauging station.

Manning method is based on hydraulic equations involving such parameters as water depth, width, slope and river channel friction. Bjerklie method is based on semi-empirical equation obtained after simplification and calibration of the Manning equation.

Comparison of the methods' performance was effectuated on two big (Ob, Mississippi) and two medium size (Rhine, Po) rivers. On the Ob, Mississippi and Po rivers, the reaches up to 150 km length around a reference gauging station were selected. On the Rhine River, three short river reaches characterised by different geomorphologic conditions and located only in 2-28 km from corresponding GS were selected.

All available L3 time series (TS) were verified for their suitability for the L4 production, i.e. for goodness of the Qinsitu - Halti relationship, and only the L3 TS expressing an accuracy < 30% (NRMSE) with the RC method for calibration period were used for discharge retrieving by Bjerklie and Manning methods.

For the Bjerklie and Manning equations, the water width was estimated from high resolution Landsat 8 (30 m), Sentinel-2 (20 m) and Sentinel-1 (10 m and 40 m) images. For each river reach, a collection of 8-15 images acquired ±4 days around the Sentinel-3 river overflights were created and the river width near a "reference" VS were retrieved semi-automatically. The width was, then, related to the altimetric water height as described in Zakharova et al., 2020. The water slope was retrieved from the measurements of the water height at virtual stations located in less than 150 km each to other. For this, only VSs demonstrating a good relationship with the Qinsitu were selected indicating potential similarity of the fluvial morphology at these VS and at the GS. For the Rhine and Mississippi rivers VSs belonging to the same altimeter ground track were used. For the Ob and Po rivers, the orientation of the river reaches did not allow to find such a pair and VSs with 3- and 2-day over-flight differences were chosen.

A set of Bjerklie equations published in Bjerklie et al., (2003, 2005) was tested. The configuration from Bjerklie et al. (2003) provided the best results. Bjerklie's equation contains a coefficient linked to a channel hydraulic conductance. This term varies significantly from reach to reach. Using values suggested by Bjerklie, highly biased estimates of the discharge (> 150% error) were obtained at all test sites. An adjustment of the conductance during calibration period using Qinsitu allowed to improve remarkably the results. In the Manning equation, a characteristic of local channel conditions also exists (a channel friction term). This term also depends on local channel morphology and requires some calibration. Factors such as channel deposits, roughness, sandbanks, aquatic vegetation, channel shape and sinuosity, flow depth and Reynolds number, affect the value of effective channel friction. The effective friction may vary in time depending on variability of the above-mentioned conditions. A calibration of the friction term for different ranges of the water level may be beneficial for many river reaches. In L4 production routine an additional module was developed and for all target sites the test of advanced calibration of the friction was performed in order to evaluate a potential improvement of accuracy of discharge estimation using this modification.

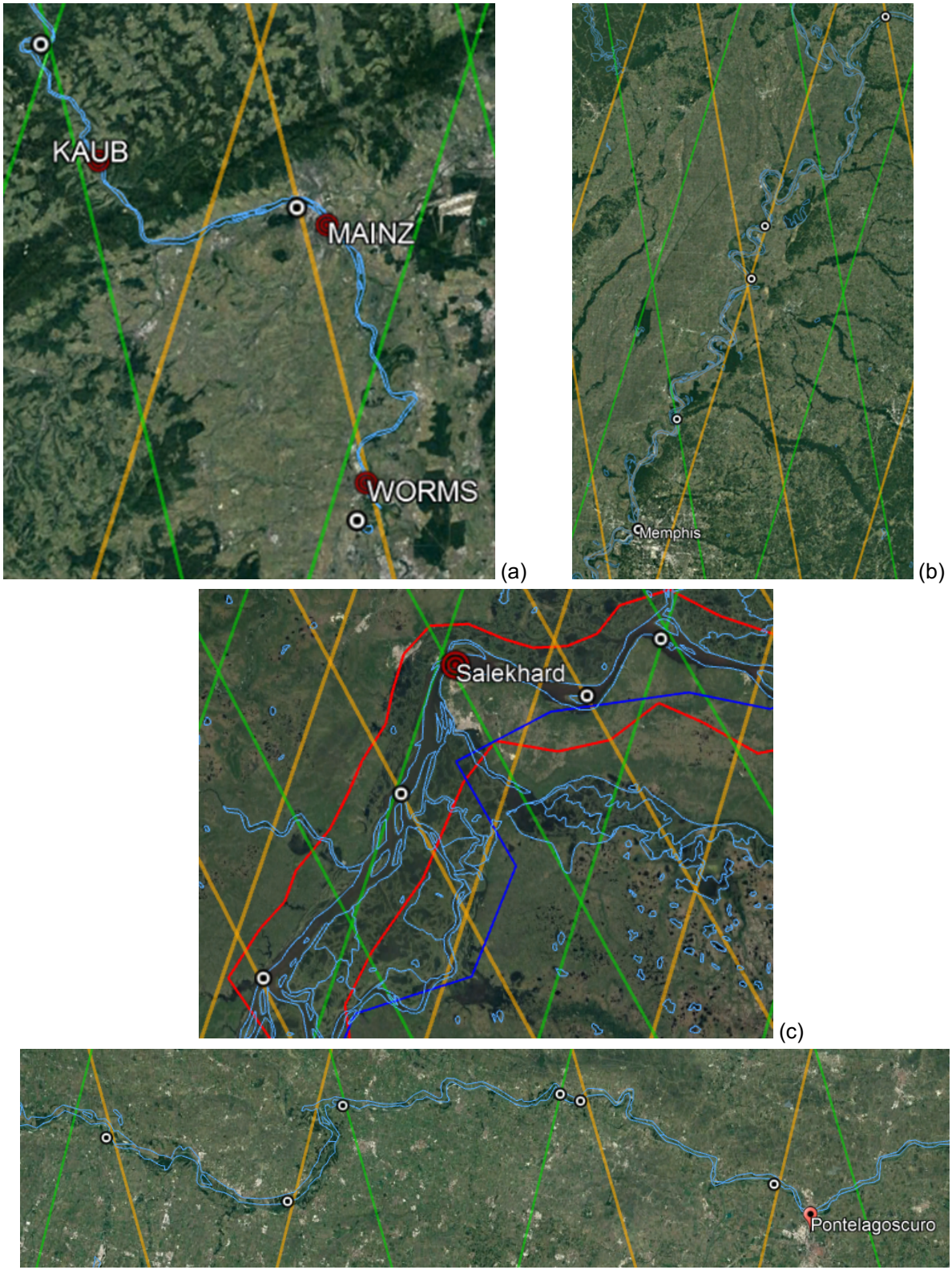


Figure 3.10.1. Location of gauging and virtual stations used for the L4 product retrieving on the Rhine (a), Mississippi (b), Ob (c) and Po (d) rivers.

A comparison of the accuracy of discharge estimation by different methods including the modification in the Manning method (Manning adv) is presented in Table 3.10.1. The number of points used for estimation of scores is referred to the validation period (if was possible). This number can differ for RC and Bjerklie/Manning methods as the second virtual station used for the slope evaluation can have gaps at time of observation at main ("reference") virtual station.

A configuration of virtual and gauging stations used for each test river is presented in Figure 3.10.1. For Sentinel-3A virtual stations, the length of the L3 time series allowed to divide the datasets on calibration and validation periods. For the Sentinel 3B TS, the available number of observations was less than 15. In this case, the provided scores correspond to the period of calibration. As the length of the TS is short, the relative statistics (NRMSE) were estimated by normalisation of absolute values on average in situ discharge calculated only for dates used for validation (and not on the mean annual discharge).

Table 3.10.1. Accuracy of the discharge estimation at test sites

River/ GS	RC	Bjerkie	Manning smpl	Manning adv
Ob, Salekhard 1	S3A-S3A	cal/val		
Rcorr	0.99	0.97	0.99	0.99
RMSE, m^3/s	1829	2630	1662	1868
NRMSE	12%	17%	11%	12%
Bias, m^3/s	-423	-673	-882	-850
NS	0.97	0.93	0.97	0.96
Nobs	16	14	14	14
Ob, Salekhard 2	S3A-S3B	cal/val	val=cal	val=cal
Rcorr	0.97	0.95	0.93	0.96
RMSE, m^3/s	2312	2892	3227	2510
NRMSE	17%	22%	24%	19%
Bias, m^3/s	-160	559	119	184
NS	0.93	0.90	0.87	0.92
Nobs	11	11	11	11
Rhine, Mainz	S3A-S3A	cal/val		
Rcorr	0.99	0.97	0.96	0.95
RMSE, m^3/s	93	221	181	154
NRMSE	6%	15%	13%	11%
Bias, m^3/s	-39	-135	-120	-42
NS	0.96	0.79	0.86	0.90
Nobs	17	16	16	16
Rhine, Worms	S3A-S3A	cal/val		
Rcorr	0.89	0.87	0.86	0.91
RMSE, m^3/s	259	282	255	272
NRMSE	20%	22%	20%	21%
Bias, m^3/s	-88	-158	76	-177
NS	0.69	0.63	0.70	0.65
Nobs	17	16	16	16

Rhine, Kaub	S3B-S3B	val=cal		
Rcorr	0.99	0.99	0.99	0.99
RMSE, m^3/s	214	311	292	95
NRMSE	13%	19%	18%	6%
Bias, m^3/s	-66	6	-3	-26
NS	0.94	0.87	0.89	0.99
Nobs	11	11	11	11
Mississippi, Memphis	S3A-S3A	cal/val		
Rcorr	0.98	0.97	0.97	0.95
RMSE, m^3/s	2015	2247	2489	3334
NRMSE	8%	8%	9%	13%
Bias, m^3/s	157	-408	-503	-831
NS	0.96	0.94	0.93	0.88
Nobs	15	15	15	15
Po, Pontelagoscuro	S3A-S3A	cal/val		
Rcorr	0.98	0.99	0.99	1.00
RMSE, m^3/s	266	293	180	157
NRMSE	18%	20%	12%	11%
Bias, m^3/s	16	-95	-15	21
NS	0.95	0.84	0.94	0.96
Nobs	16	10	10	10

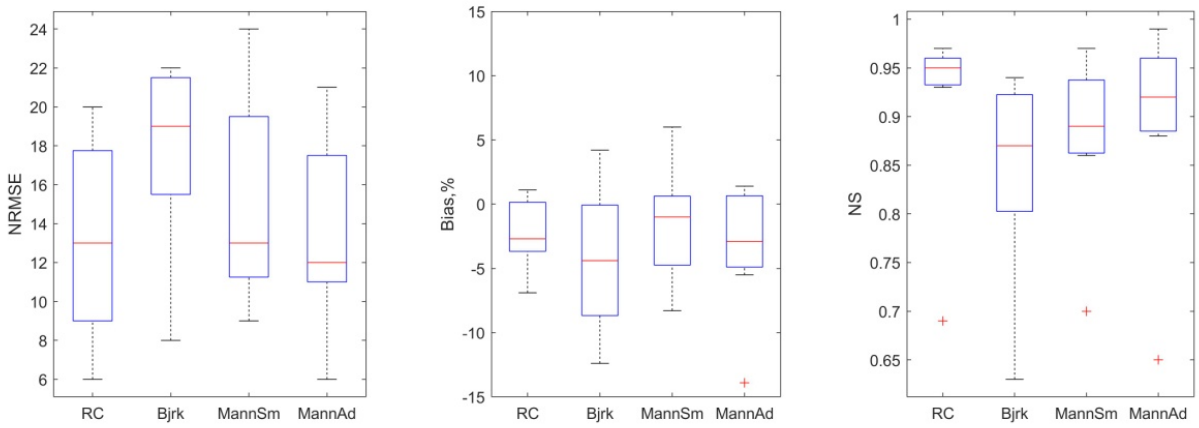


Figure 3.10.2. Comparison of the accuracy of the Hydrocoastal L4 product retrieved using different methods. Relative bias was obtained by division of absolute value on average Qinsitu calculated for dates from validation dataset.

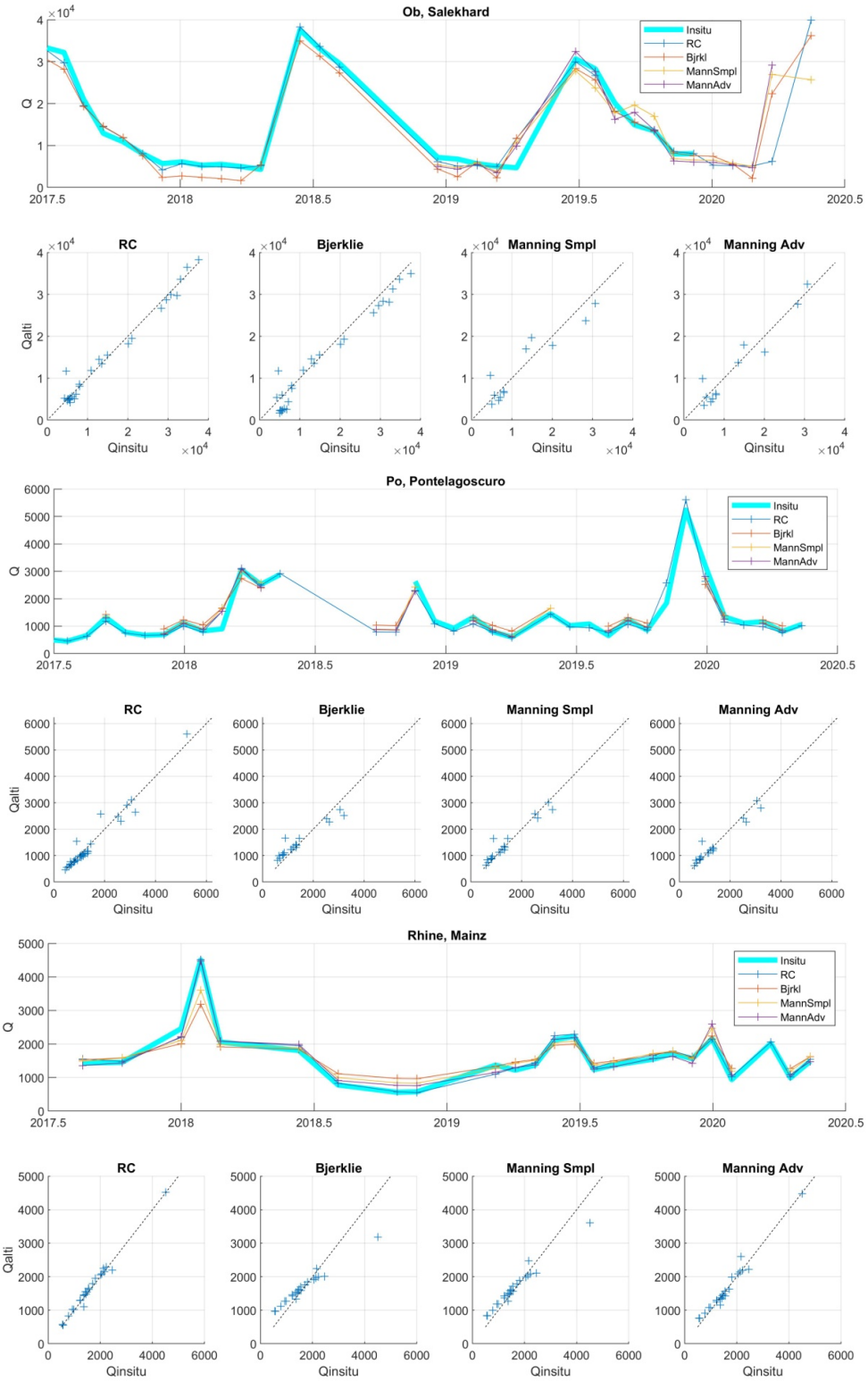


Figure 3.10.3. L4 time series and 1:1 scatterplots for discharge retrieved by different methods for the Ob River at Salekhard (a), Po River at Pontelagoscuro (b) and Rhine River at Mainz (c).

Summary

1. As expected, the rating curve method overperforms the Bjerklie and Manning methods.

2.The Manning method demonstrated performance comparable with the rating curve method and in several cases provided even more accurate discharge estimates. In half of the cases the advanced calibration of friction parameter allowed for up to 5-12 % (NRMSE) of accuracy improvement comparing to simple calibration routine.

3. The Bjerklie method is also able to provide meaningful discharge estimations. However, in all test sites in order to correct the high bias the calibration of the conductance term is necessary. After the calibration, the L4 product with uncertainties of order of 15-25% can be obtained. It is worth to note that the Bjerklie method represents very well middle-flow conditions, while it has a tendency to provide discharge underestimation during the high flow and overestimation during the low flow.

4. In spite of the higher accuracy and easy realisation, the rating curve method has an important drawback - a recalibration of the rating curves depending on mission and retracker. For this, for each mission the simultaneous in situ discharge data are required. With shrinking hydrological ground network, the availability of the simultaneous in situ Q observations became a problem for many remote areas of the World. Contrary to the rating curve, the Manning method allows an estimation of the parameters, which have a physical meaning (friction, channel shape, depth). For stable river channels, in absence of high erosion, these parameters can be used in the Manning formulation with any altimetry mission, independently on instrument and retracker. The parameters can be even distributed for quite a long geomorphologically homogeneous river reach giving a chance for application of the method with a succession of altimetry missions (ex. ENVISAT - Sentinel3).

Improvement of sampling frequency of Hydrocoastal L4 product.

The 27-days sampling frequency of the Sentinel-3 L4 product limits its application for climate monitoring and operational use. However, for many large and even middle size rivers the sampling frequency can be improved after combination of the discharge retrievals as it was done, for example, in Zakharova et al. (2019). A test of densification of the sampling frequency effectuated for the Mississippi and Po rivers showed that after the launch of Sentinel 3B, starting from 2019, the final L4 product can provide up to 7 Q retrievals in a month with 2-8 days interval and 18% of overall accuracy (expressed as NRMSE) for the Po River at Pontelagoscuro. For the Mississippi test site, the observing frequency can be increased up to 4-5 times a month and 4-9 days interval. The final Hydrocoastal L4 product at Memphis site will be of very high (9%) NRMSE accuracy (Figure 3.10.4).

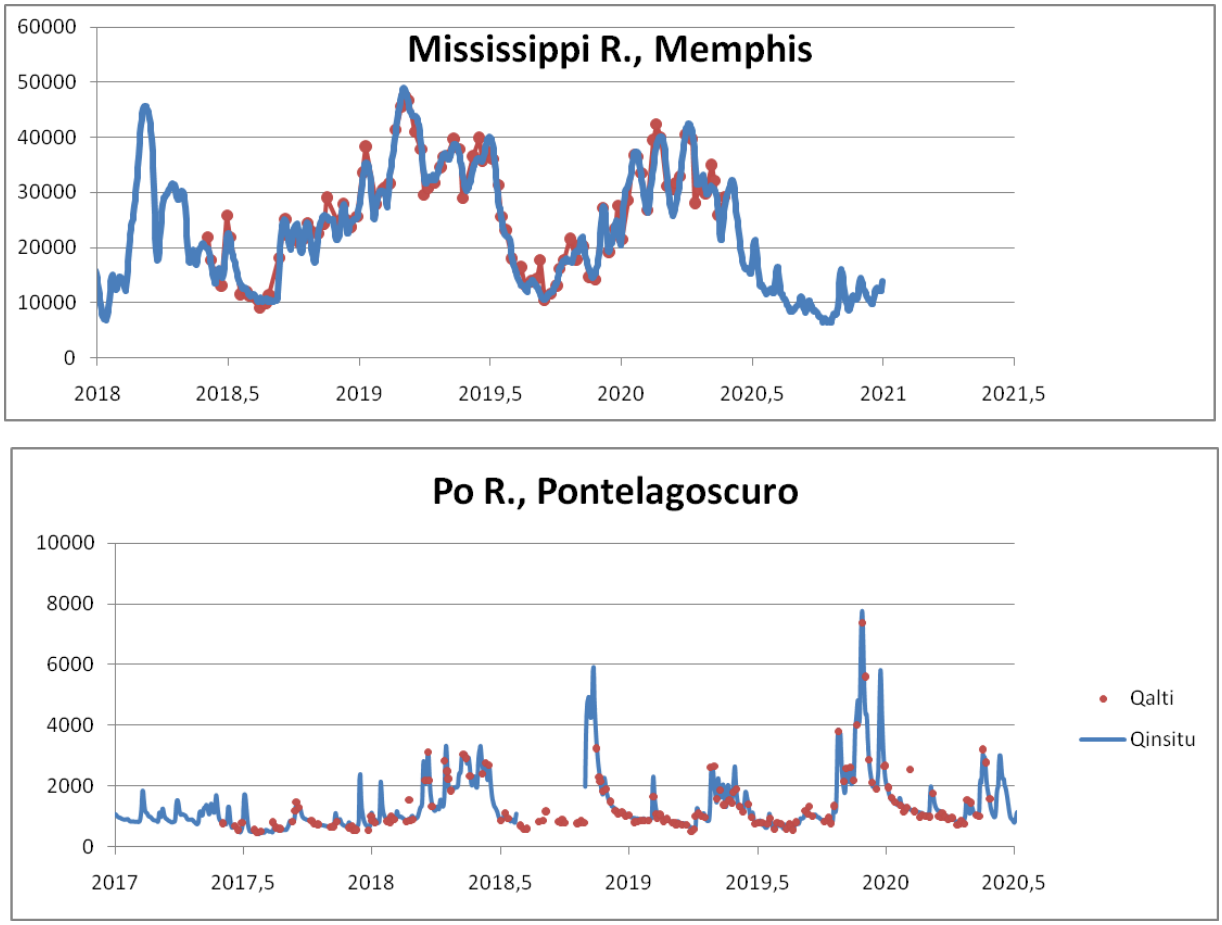


Figure 3.10.4. Hydrocoastal L4 product after discharge densification.

4 Validation of new DTC and WTC over CZ and IW regions (UPorto)

This section describes the assessment results obtained by UPorto for the Dry Tropospheric Correction (DTC) and the Wet Tropospheric Correction (WTC) over Coastal Zones (CZ) and Inland Water (IW) regions developed in WP2200 of the HYDROCOASTAL project.

This refers to the validation performed by UPorto in the four selected test areas (California, Caspian Sea, Danube River and Amazon River). In addition, specific behaviour of the corrections in other regions is being reported.

Additional independent validation of the corrections shall also be performed by other partners (in WP2500) in all test areas.

Subsection 4.1 provides background on the computation of these corrections while subsections 4.2 and 4.3 present the validation results for the WTC and DTC, respectively. Finally, sub-section 4.4 summarises the main conclusions on the validation of the WTC and DTC.

4.1 Computation of DTC and WTC over CZ and IW regions

4.1.1 Introduction

For the computation of the tropospheric corrections, the selected 17 project regions of interest (ROI) have been grouped into three different region categories: **Coastal**, **Lake** and **River** as follows:

Coastal	**Lake**	**River**
California	Caspian	Amazon
Bight-baltic	Nonacho	Amur
Cadiz	Reindeer	Danube
Greece		Mississippi
Gibraltar		Ob
		Po
		Rhine
		Yangtze
		Zambezi

Figure 4.1.1 illustrates the seventeen ROI and the corresponding S3A tracks over these regions.

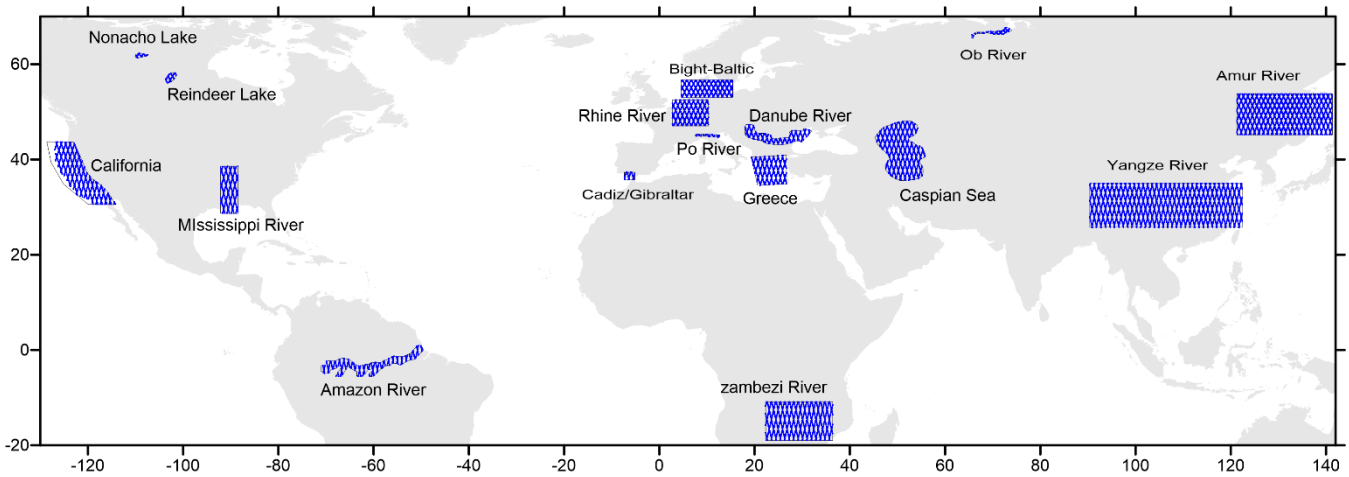


Figure 4.1.1. Location of the 17 project ROI and the corresponding S3A Land products data coverage.

Altimeter 20 Hz data from S3A, S3B and CS2 (SAR and SARIn) have been made available by IsardSAT in the project MEGA drive for all ROI. These data covered the following three periods:

- **Period 1** – From 01 Dec 2018 (MJD 58453) until 30 November 2019 (MJD 58817) for the Caspian Sea.
- **Period 2** - From 01 June 2017 (MJD 57905) until 31 May 2020 (MJD 59000) for the following ROI: Rhine, Ob and Po.
- **Period 3** – From 01 June 2018 (MJD 58270) until 31 May 2020 (MJD 59000) for all other ROI.

Considering the complex structure of the data in the MEGA drive, where data were split into 12 directories, each corresponding to a different 3-month time slot, each containing up to 17 ROI subdirectories, for use in this task data were first reorganized into a single directory for each (mission, ROI) pair.

For each region category (Coastal, Lake, River) the same methodology has been used in the computation of the two tropospheric corrections, which will be described in this section.

The procedure adopted in the computation of DTC and WTC over CZ and IW regions is implemented in two steps:

- Step 1: Compute DTC and WTC from the ECMWF ERA5 Numerical Weather Model (NWM) (Fernandes et al, 2021).
- Step 2: Compute WTC from the GNSS-derived Path Delay Plus (GPD+) (Fernandes and Lázaro, 2016, Lázaro et al., 2020).

Step 1: Computation of DTC and WTC from NWM

Step 1 inputs:

- Altimeter 20 Hz data from S3A, S3B and CS2 (SAR and SARIn) from the L2e reduced products made available in the project MEGA drive. The following fields have been used: time, lat, lon, WTC from the onboard Microwave Radiometer (MWR) and various flags (radiometer flags and surface type). In some

cases, data from the original satellite L2 products have also been used, in order to check the validity of the fields present in the L2e products.

- ERA5 model:
  - single layer fields at 0.25ºx0.25º spatial resolution and 3h intervals: Sea Level Pressure (SLP), Total Column Water Vapour (TCWV) and 2m-Temperature (2T).
  - Wet Path Delay (WPD) vertical profiles at 2ºx2º spatial resolution and 6h intervals, previously computed from 3D pressure level fields (temperature and specific humidity), for use in the modelling of the WPD vertical variability.
  - Model orography at 0.25ºx0.25º spatial resolution.
- Geoid model: EIGEN-6C4 (Förste et al., 2014). For use in this project, a file generated in the scope of the SHAPE project has been shared. It is a global grid at 0.05ºx0.05º spatial resolution, containing the EIGEN-6C4 geoid heights with respect to WGS84 and using the tide-free system. To be in line with EGM2008 global grids, a constant zero-degree term of -41 cm has been applied (Fernandes and Vieira, 2019).
- Mean Sea Surface (MSS): DTU21 at 1'x1' spatial resolution, available from ftp.space.dtu.dk/pub/DTU21/1_MIN.
- Water "occurrence" product based on Landsat imagery, from the Global Surface Water Explorer (GSW), https://global-surface-water.appspot.com/download (Pekel et al., 2016). Original resolution (1") has been resampled to 57.6".
- Digital Elevation Models (DEM):
  - Altimeter Corrected Elevations 2 (ACE2) (3") (Berry et al., 2008).
  - TanDEM-X (3") (Rizzoli et al., 2017), (Wessel et al., 2018).
- Distance from coast: global grids computed from a global netCDF grid (2'), with distances to the nearest GSHHG (Global Self-consistent, Hierarchical, High-resolution Geography Database) shoreline (Wessel and Smith WH,1996).
- Lake polygons: from HydroLakes https://hydrosheds.org/page/hydrolakes (Messager et al., 2016).
- River profiles:
  - From Karina Nielsen (DTU): Rhine, Amur, Po, Zambezi.
  - From the SWOT a priori River Database (SWORD) version v2, https://zenodo.org/record/5643392, (Altenau et al., 2021) for all River ROI.

Step 2: Computation of WTC from the GNSS-derived Path Delay Plus (GPD+) algorithm

Step 2 inputs:

- WTC from on-board MWR (only for S3) and corresponding MWR validity flag (previously computed in Step 1), specifying if a given MWR observation is valid or not.
- WTC from ERA5, from step 1, to be used as first guess in the Objective Analysis (OA) procedure implemented in GPD+.
- WTC from Global Navigation Satellite Systems (GNSS).
- WTC from scanning Imaging MWR (SI-MWR).
- ERA5 model:

  - WPD vertical profiles at 2ºx2º spatial resolution and 6h intervals, previously computed from 3D pressure level fields, for use in the modelling of the WPD vertical variability.
  - Model orography at 0.25ºx0.25º spatial resolution.
- ACE2 (3") (Berry et al., 2008).
- Altimeter data: time, lat, lon, MWR WTC, various flags (20 Hz).

4.1.2 Provided corrections and instructions of use

Corrections are computed for all points of the HYDROCOASTAL test data set files. Therefore, there is a one-to-one point correspondence between input and UPT correction files. The corrections are provided in netCDF files with the following fields

- **time_20_ku, lat_20_ku, lon_20_ku** – as in input files
- **upt_dry_tropo** – DTC_UPT, DTC from the ERA5 model, at surface height (h_surf), in metres.
- **gpd_wet_tropo** - WTC_GPD, WTC from the GPD+, at surface height (h_surf), in metres.
- **gpd_wet_tropo_flag** – Flag specifying the data sources used in the estimation of the GPD+ WTC. (0) – Valid on-board MWR value, eventually scaled; (1) – Estimate from on board MWR observations only; (2) – Estimate from SI-MWR observations only; (3) – Estimate from MWR and SI-MWR observations; (4) – Estimate from GNSS observations only; (5) – Estimate from MWR and GNSS observations; (6) – Estimate from SI-MWR and GNSS observations; (7) – Estimate from on-board MWR, SI-MWR and GNSS observations; (8) – No observations exist, estimate is from the NWM used as first guess (ERA5).
- **h_surf** – height at which the DTC and WTC have been computed, in meters.
- **geoid_EIGEN_6C4** - in meters.

Corrections are tuned for water applications. A buffer of width Δ, used in the definition of h_ref, has been created around each ROI.

- For coastal ROI, corrections are referred to mean sea level (h_ref=0) for all ocean points (surface_type=0) and for points up to distance from coast $D \leq \Delta$ (Δ=20km). For the remaining points, h_ref is the height of the adopted DEM (ACE2).
- For Lake ROI, corrections are referred to the mean lake level (h_ref=mean lake level) for all points inside the lake and for points up to distance from lake border $D \leq \Delta$. For the remaining points, h_ref is the height of the adopted DEM (TanDEM-X for Nonacho and Reindeer Lakes, ACE2 for the Caspian Sea). Values of Δ=50km and Δ=20km have been adopted for the Caspian Sea and for the Nonacho and Reindeer Lakes, respectively.
- For River ROI, corrections are referred to the height of the closest point in the river profile (h_ref=river profile height) for all points up to distance from river profile $D \leq \Delta$, Δ = max(2.0, river_width*1.5) km. For the remaining points, h_ref is the height of the adopted DEM (ACE2).

Instructions of use of the DTC and WTC UPT corrections:

- Both DTC and WTC are computed at the provided surface altitude, h_surf.

- To convert the DTC to another height (H), the following formula (4.1.1) can be used, with all variables in metres:

$$DTC(H) = DTC(\text{h\_surf}) + 0.0002563(\text{H} - \text{h\_surf}) \qquad (4.1.1)$$

- To convert the WTC to another height (H), the following formula (4.1.2) can be used, with all variables in metres and k=3000:

$$WTC(H) = WTC(\text{h\_surf})e^{\frac{\text{h\_surf}-H}{k}} \qquad (4.1.2)$$

When retracked and more accurate heights are available, it is recommended that these altitude conversions are applied, in particular for the DTC, for which the impact is more pronounced. However, it is important to note that this step can only be performed for small altitude differences, otherwise significant errors will be introduced, mainly in the DTC.

In order to quantify these errors, over the Yangtze ROI (which can have very high altitudes), if the DTC at zero level (available in the products) is reduced for the actual heights using the above linear equation, this step can introduce very significant errors. Figure 4.1.2 represents the differences between the DTC reduced from zero level to the height of each point, using the linear equation above, and the DTC_UPT originally computed at the actual height of each point by UPorto.

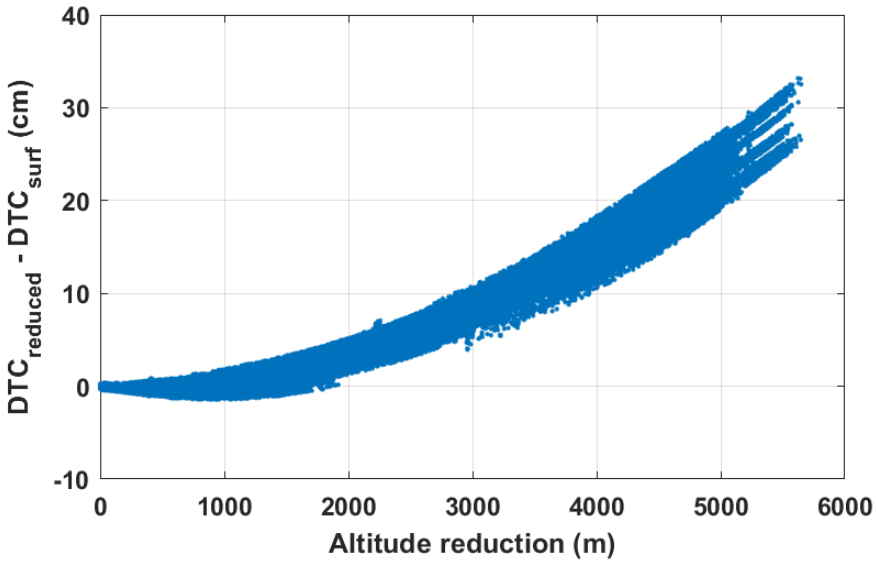


Figure 4.1.2. Differences between DTC reduced from zero level to the height of each S3A point, using a linear height reduction, and the DTC_UPT (DTCsurf in the Y axis label), originally computed at this height, for the Yangtze ROI.

Considering all S3A measurements over the Yangtze ROI, at heights around 1000 m, the differences between the DTC computed at the actual surface height and that reduced from zero level, using a linear reduction, have a mean and a standard deviation of 0.3 and 0.8 cm, respectively. If heights around 1500 m are considered, the same

statistics are 1.6 and 1.0 cm, respectively. These errors increase with the altitude reduction. For an altitude reduction from zero level to 4000 m, the corresponding DTC differences are larger than 10 cm (see Figure 4.1.2). These figures show the very significant impact of having corrections originally computed at the actual River or Lake height, as those computed by UPorto.

4.2 Validation of the WTC

In this task, well established methodologies for the assessment of WTC datasets (Fernandes and Lázaro, 2016, 2018, Lázaro et al., 2020) are adopted in the validation of the new WTC, namely:

a) Comparison with the MWR-derived WTC present in products (only for S3) – for coastal regions (California) and large lakes (Caspian Sea).

b) Comparison with the WTC from the ECMWF operational model – for all regions.

c) Comparison with GNSS-derived WTC – this provides information mainly about algorithm performance in the coastal regions and over IW regions with abundant number of GNSS stations.

For a better illustration of the results, data have been grouped into 27-day cycles (S3) or 27&29-day sub-cycles (CS2, using the RADS convention for sub-cycle numbering).

In the current S3 products, two types of model-derived WTC are provided, both from the ECMWF Op. model: one computed at zero level, **WTC_ECM$_{zero}$** (mod_wet_tropo_cor_zero_altitude field) and another at the altimeter measurement level, **WTC_ECM$_{meas}$** (mod_wet_tropo_cor_meas_altitude field).

For S3, over all regions, **WTC_GPD** is compared with the corresponding WTC provided in the S3 products at measurement level (**WTC_ECM$_{meas}$**) and with the MWR-derived WTC, **WTC_MWR** (rad_wet_tropo_cor field).

For CS2, in all cases, the **WTC_GPD** is compared with the only model-derived WTC present in the products, referred to the model orography (**WTC_ECM$_{oro}$**).

In the validation of the WTC and the DTC, statistics have been computed for “water” points. The criteria used to select these “water” points are:

- Coastal ROI: ocean points (surface_type=0) or points up to distance from coast $D \leq \Delta$ (Δ=10km) and GSW occurrence value larger than 10%.
- Lake ROI: points inside the lake and points up to distance from lake border $D \leq \Delta$ (Δ=10km) and GSW occurrence value larger than 10%.
- River ROI: points up to distance from river profile $D \leq \Delta$, Δ = max (2.0, river_width*1.5) km and GSW occurrence value larger than 10%.

In addition to these criteria, passes that were not inside the predefined ROI, i.e., passes that are in the ROI files but outside the ROI limits have been excluded. This excludes for example, a few passes over open ocean in the Southwestern part of the California ROI. Moreover, in the sequel of some data errors detected in ocean tracks far from the coast, for ocean points, analysis has been limited to points up to distances from coast of 150 km.

4.2.1 Caspian Sea

Figure 4.2.1 illustrates the ACE2 DEM over the Caspian Sea region, showing that the Sea is surrounded by mountains, with high relief in particular in the west and southern parts of the region. Figure 4.2.2 shows the location of S3A tracks for cycle 47 and the interleaved S3B tracks for cycle 28 (S3B cycle starts 10 days later than S3A cycle).

A major step in the estimation of the GPD+ WTC is the establishment of the MWR validity flag, specifying if a given WTC from the on-board MWR is considered valid or not. Figure 4.2.3 shows the MWR validity flag for S3A cycle 47. It can be seen that all land points and coastal points at distances from lake border less than 25 km are considered invalid (green points). Additional points are rejected based on statistical criteria (pink points).

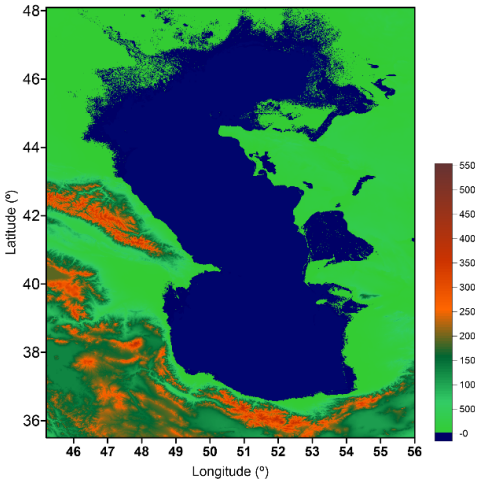


Figure 4.2.1. ACE2 DEM over the Caspian Sea region.

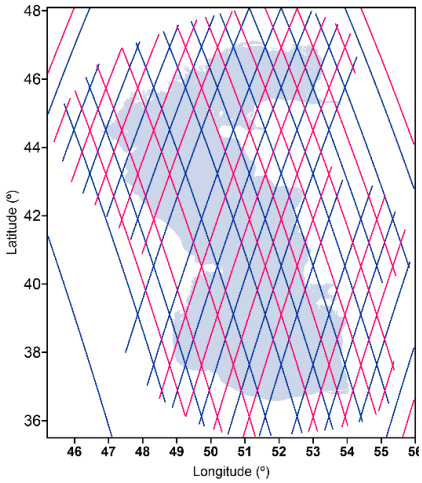


Figure 4.2.2. Location of S3A points for cycle 47 (blue) and S3B cycle 28 (red).

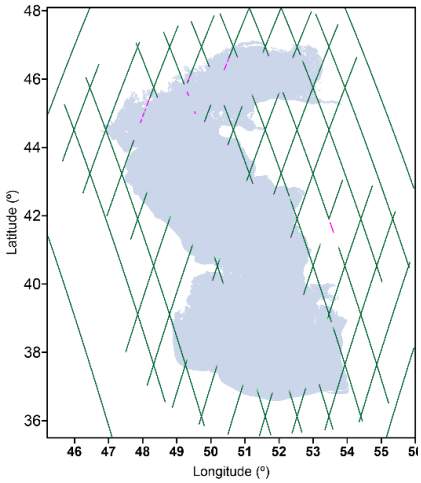


Figure 4.2.3. MWR validity flag for S3A cycle 47: green – land contamination; pink: statistical rejection criteria.

Figure 4.2.4 depicts the GPD source flag (gpd_wet_tropo_flag) for S3A cycle 47, providing information on the data sources used in the estimation of each GPD+ WTC. For each 20Hz S3A/B ground-track point the WTC_GPD is:

- The S3A/B MWR-derived WTC (eventually scaled after calibration) for all S3 points with valid MWR values - gpd_wet_tropo_flag = 0 (red points);
- A new estimate obtained from data combination of all available observations for all S3 points with invalid MWR - gpd_wet_tropo_flag with values from 1 to 7 (green points);
- The ERA5 model-derived WTC in the absence of observations - gpd_wet_tropo_flag = 8 (blue points).

For CS2, since no on-board MWR exists, all points will have an MWR validity flag ≠ 0, meaning that GPD+ algorithm will try to obtain a new WTC estimate, for all along-track points, from the available external observations. In the absence of those, the ERA5 WTC will be adopted.

For S3A, it can be observed that most points for which the WTC has been considered invalid and are located inside or close the lake borders will have a WTC_GPD estimated from a combined value of observations (green points in Figure 4.2.4). These observations can either be valid on-board MWR from the closest points (Figure 4.2.5), SI-MWR observations (Figure 4.2.6) or GNSS observations (Figure 4.2.7). For the Caspian Sea, there is a single

GNSS station in the vicinity of the Sea (southern part of the region). In the absence of GNSS observations, most land points will have a GPD WTC equal to the value from the ERA5 model (blue points in Figure 4.2.4). Finally, all points for which the on-board MWR WTC has been considered valid, the GPD+ WTC preserves this value, apart from possible calibration factors (red points in Figure 4.2.4).

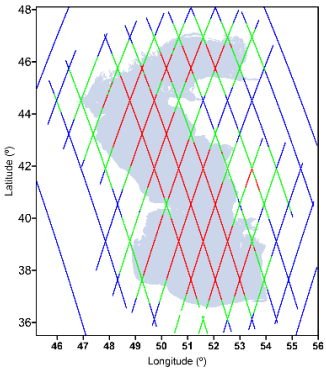


Figure 4.2.4. GPD Source flag for S3A cycle 47: 0 (red) – valid on-board MWR values; 1-7 (green) – estimated from observations; 8 (blue) – from ERA5 model.

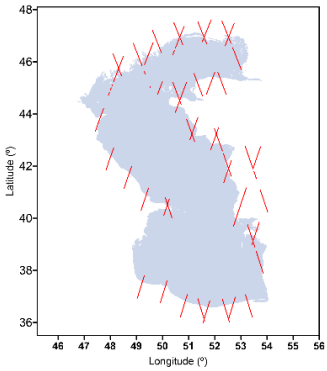


Figure 4.2.5. Points using on-board MWR observations.

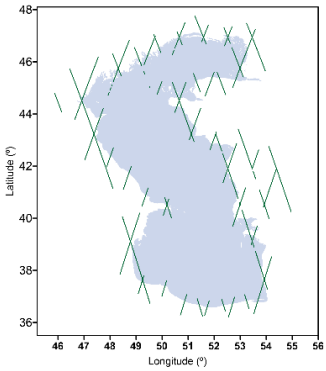


Figure 4.2.6. Points using observations from external imaging MWR.

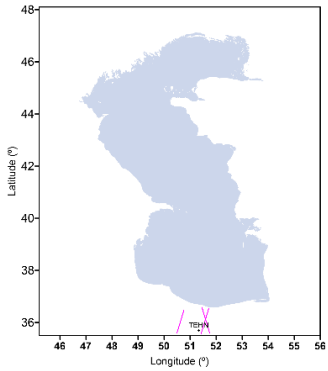


Figure 4.2.7. Points using GNSS observations (1 station).

Figure 4.2.8, Figure 4.2.9 and Figure 4.2.10 depict WTC_ECM$_{meas}$, WTC_MWR and WTC_GPD for S3A cycle 47 and S3B cycle 28, approximately 10-days apart, over a summer period. The major feature is the invalid WTC_MWR values around the Sea border and over land, which are not present in WTC_GPD.

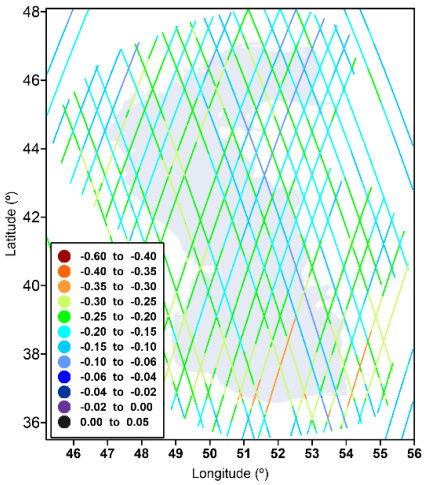


Figure 4.2.8. WTC_ECM$_{meas}$ over the Caspian Sea region, for S3A cycle 47 and S3B cycle 28.

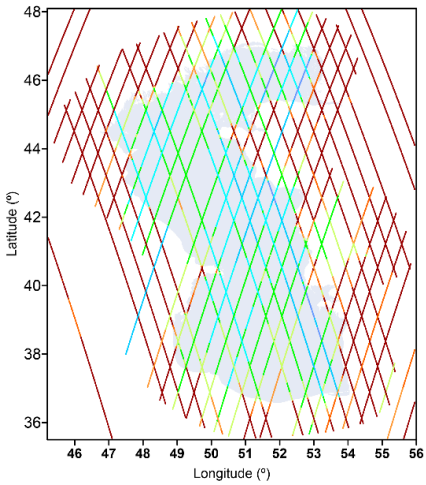


Figure 4.2.9. WTC_MWR over the Caspian Sea region, for S3A cycle 47 and S3B cycle 28.

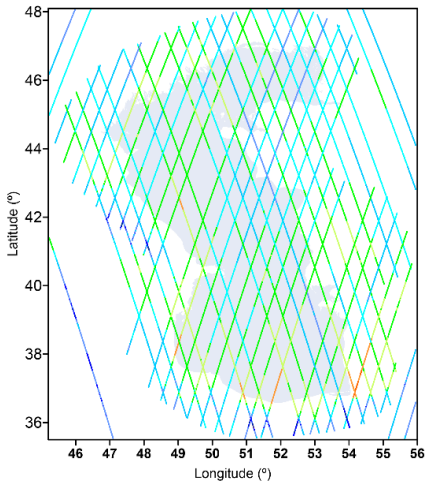


Figure 4.2.10. WTC_GPD over the Caspian Sea region, for S3A cycle 47 and S3B cycle 28.

After an overall assessment of the WTC over the Caspian Sea, a finer analysis is performed individually for each pass, in order to identify specific behaviours of the corrections. Figure 4.2.11 and Figure 4.2.12 illustrate examples of passes of S3A cycles 49 and 50. In these figures and in the most of the following pass examples, the WTC from the ERA5 model computed at surface height (WTC_UPT) is also shown for comparison with the model-derived WTC present in the HYDROCOASTAL products. In these examples, typically strong land contamination can be observed on WTC_MWR in the points close to the Sea borders. Moreover, these examples show that the model WTC and GPD can also be significantly different, with differences of several centimetres. For cycle 49 pass 367 (Figure 4.2.11), the RMS of the differences between WTC_ECM$_{meas}$ (blue) and WTC_GPD (black) over water points is 2.52 cm (see the differences between the blue and black lines). The corresponding value for cycle 50 pass 554 (Figure 4.2.12) is 2.46 cm.

For all examples of S3 or CS2 passes shown in this report, the very light blue and grey areas represent water and land, respectively, which can be confirmed in the spatial representation of each track, in the small bottom left panel in each figure.

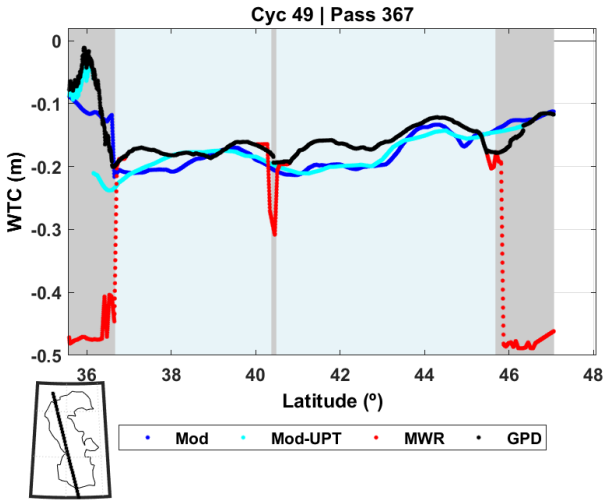


Figure 4.2.11. Various WTC for S3A cycle 49 pass 367: WTC_ECM$_{meas}$ (blue), WTC_UPT (cyan), WTC_MWR (red) and WTC_GPD (black).

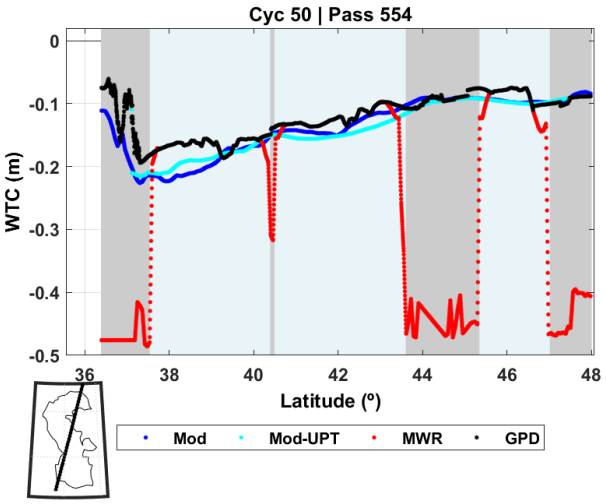


Figure 4.2.12. Various WTC for S3A cycle 50 pass 554: WTC_ECM$_{meas}$ (blue), WTC_UPT (cyan), WTC_MWR (red) and WTC_GPD (black).

Figure 4.2.13, Figure 4.2.14 and Figure 4.2.15 show the same WTC for S3A cycle 40 and S3B cycle 21, approximately 10 days apart, now over a winter period. In addition to the invalid WTC_MWR values around the Sea border, it can be observed that over this period most MWR values in the Northern part of the Sea are invalid (Figure 4.2.14).

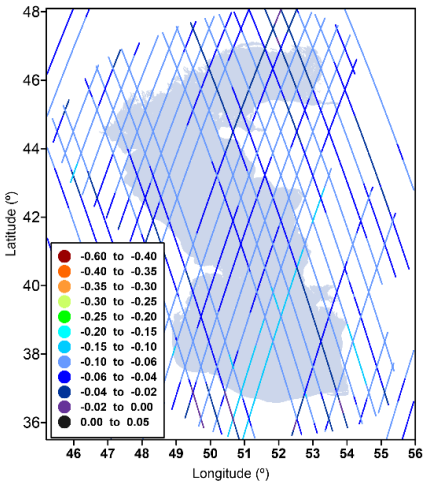


Figure 4.2.13. WTC_ECM$_{meas}$ over the Caspian Sea region, for S3A cycle 40 and S3B cycle 21.

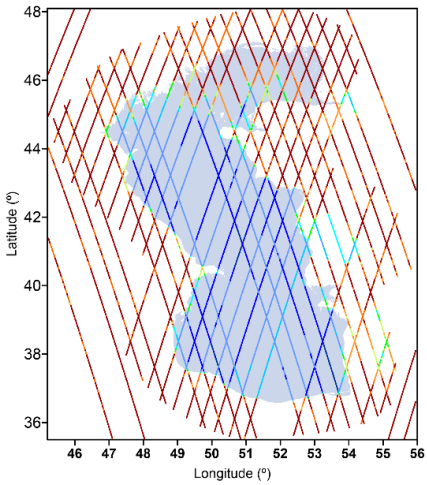


Figure 4.2.14. WTC_MWR over the Caspian Sea region, for S3A cycle 40 and S3B cycle 21.

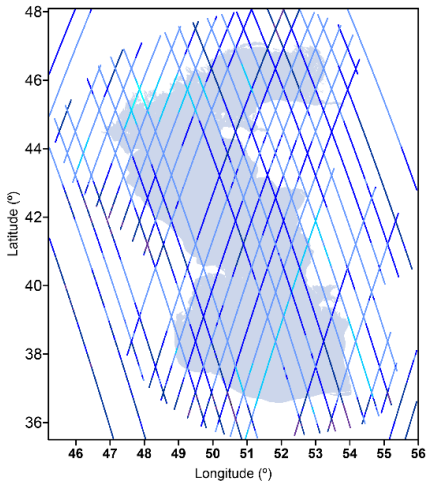


Figure 4.2.15. WTC_GPD over the Caspian Sea region, for S3A cycle 40 and S3B cycle 21.

Figure 4.2.16 and Figure 4.2.17 illustrate examples of passes for these S3A and S3B cycles, crossing the Northern part of the lake, where the strong contamination on the WTC_MWR is evidenced over this part of the sea. In all cases, WTC_GPD provides valid values over the whole region.

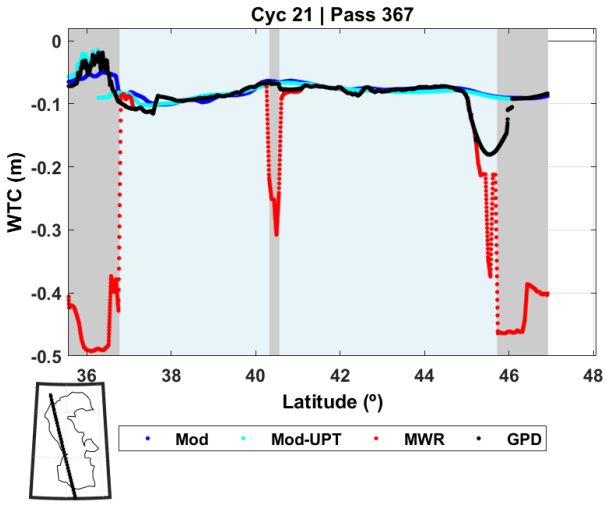


Figure 4.2.16. Various WTC for S3B cycle 21 pass 367: WTC_ECM$_{meas}$ (blue), WTC_UPT (cyan), WTC_MWR (red) and WTC_GPD (black).

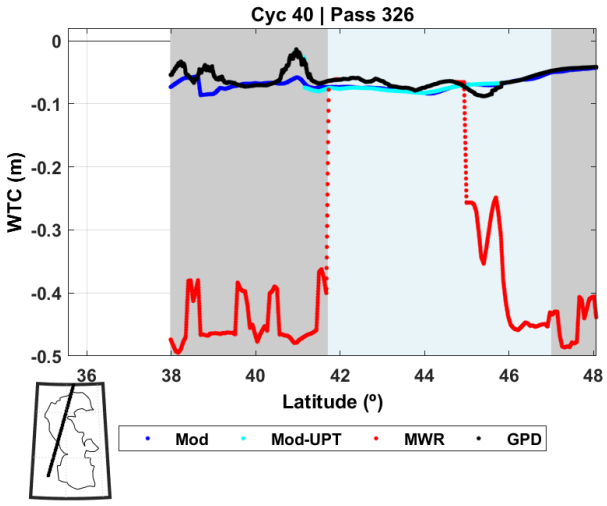


Figure 4.2.17. Various WTC for S3A cycle 40 pass 326: WTC_ECM$_{meas}$ (blue), WTC_UPT (cyan), WTC_MWR (red) and WTC_GPD (black).

Considering the whole S3A "water" points over the Caspian (347454 points) the statistics of the differences between WTC_GPD and WTC_MWR are: mean=3.01 cm, std=8.52 cm and max abs diff=43.81 cm. The corresponding values for the differences between WTC_GPD and WTC_ECM$_{meas}$ are: mean=0.33 cm, std =1.60 cm and max abs diff=9.99 cm. The large differences between WTC_GPD and WTC_MWR mainly reflect the differences in the points near the coast, since over the points where the MWR has been considered valid, the two corrections are equal.

Figure 4.2.18 and Figure 4.2.19 represent the WTC_ECM$_{oro}$ and the WTC_GPD respectively, over the Caspian region, for CS2 sub-cycle 118. In spite of the fact that no on-board MWR is present in CS2, it can be observed that the WTC_GPD can be significantly different from WTC_ECM$_{oro}$, since, as observed in Figure 4.2.6, there are various external SI-MWR observations over the Caspian, used in the estimation of WTC_GPD.

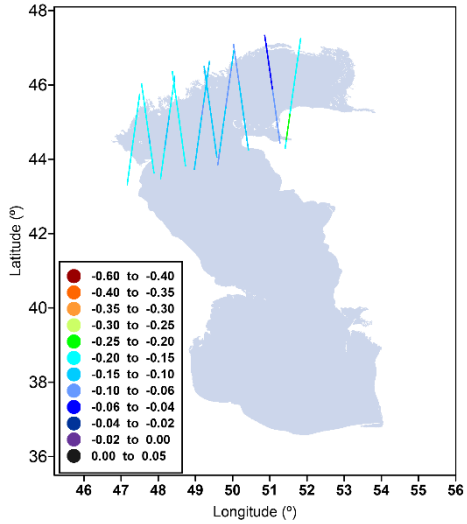


Figure 4.2.18. WTC_ECM$_{oro}$ over the Caspian region, for CS2 sub-cycle 118.

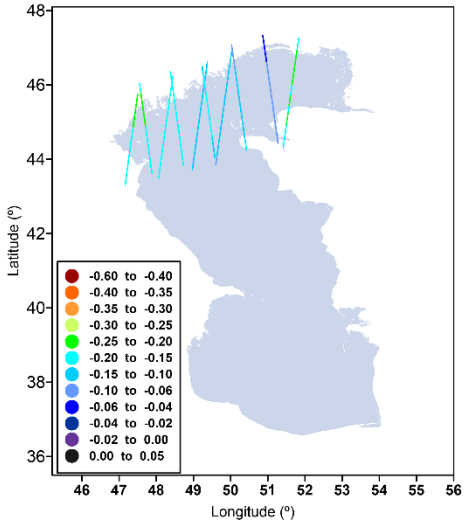


Figure 4.2.19. WTC_GPD over the Caspian region, for CS2 sub-cycle 118.

The pass examples given in Figure 4.2.20 and Figure 4.2.21 show that the differences between these two WTC (blue and black lines) can reach several centimetres. For the first case, the RMS of the differences between WTC_GPD and WTC_ECM$_{oro}$, over water points is 3.53 cm. For the second example this RMS is 2.84 cm.

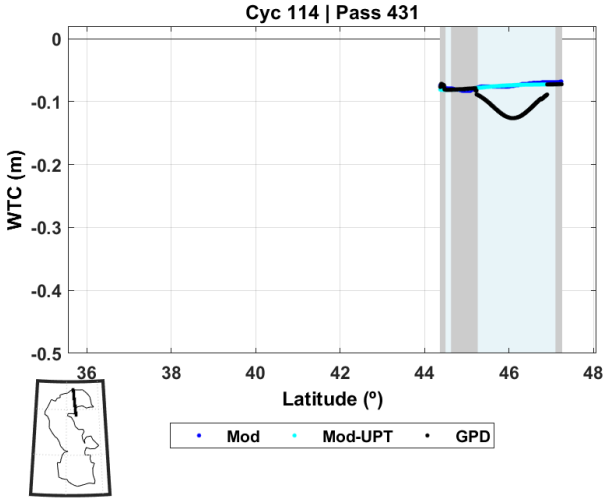


Figure 4.2.20. Various WTC for CS2 sub-cycle 114 pass 431: WTC_ECM$_{oro}$ (blue), WTC_UPT (cyan) and WTC_GPD (black).

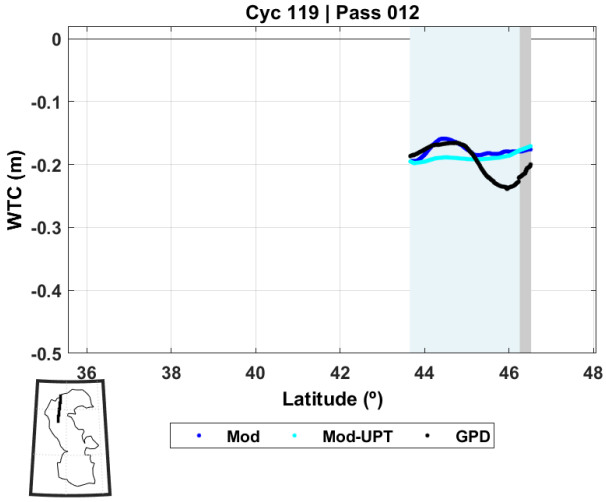


Figure 4.2.21. Various WTC for CS2 sub-cycle 119 pass 012: WTC_ECM$_{oro}$ (blue), WTC_UPT (cyan) and WTC_GPD (black).

Considering the whole CS2 “water” points dataset over the Caspian (186453 points), the statistics of the differences between GPD and model are: mean=-0.60 cm, std =1.17 cm and max abs diff=7.34 cm.

These statistics for S3 and CS2 over the Caspian Sea are a clear indication that using model-derived WTC in open ocean and coastal regions causes a significant degradation in the quality of the corresponding derived sea surface heights. Additionally, WTC retrieved from MWR in coastal regions is invalid and cannot be used.

4.2.2 California

Regarding the California region, Figure 4.2.22 illustrates the ACE2 DEM and Figure 4.2.23 represents the GPD source flag (gpd_wet_tropo_flag) for S3A cycle 47 over this region.

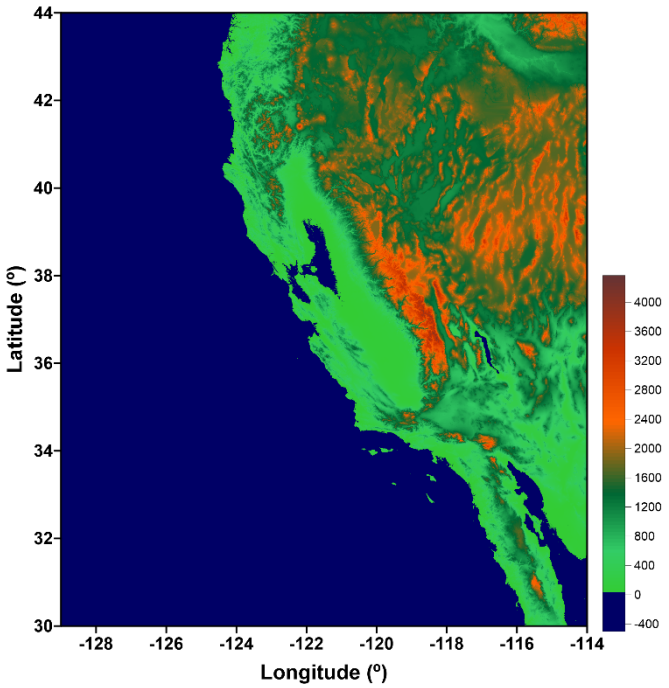


Figure 4.2.22. ACE2 DEM over the California region.

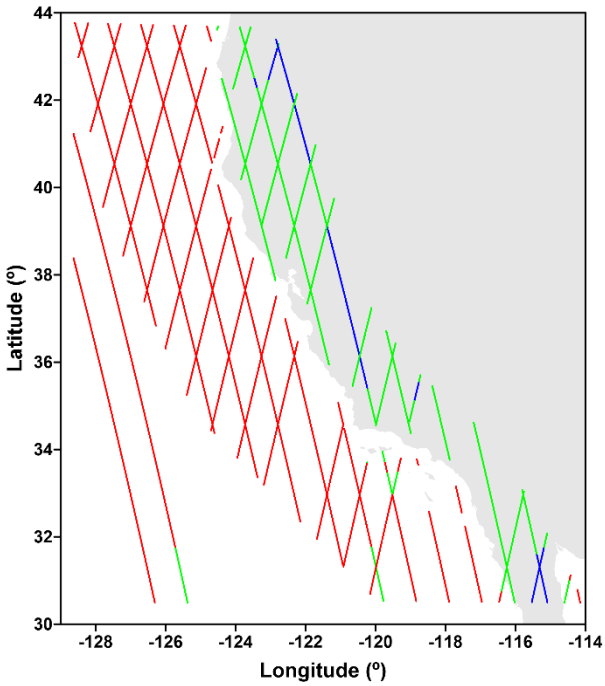


Figure 4.2.23. GPD Source flag for S3A cycle 47: 0 (red) – valid on-board MWR values; 1-7 (green) – estimated from observations; 8 (blue) – from ERA5 model.

Figures 4.2.24, 4.2.25 and 4.2.26 represent, for the same S3A cycle, the points for which observations exist to estimate the GPD+ WTC: valid on-board MWR from the closest points (Figure 4.2.24), SI-MWR observations (Figure 4.2.25) or GNSS observations (Figure 4.2.26). Contrary to the Caspian Sea, over California there is an abundant number of GNSS stations (68 for this cycle), providing observations to estimate WTC_GPD. In this region, virtually all coastal points are estimated from observations (green points in Figure 4.2.23) and only a few points have a WTC_GPD from a model value (blue points in Figure 4.2.23).

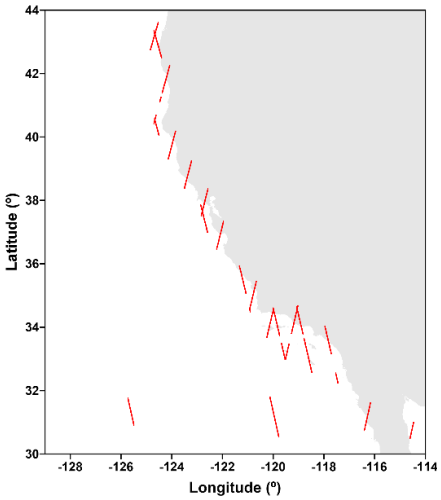


Figure 4.2.24. Points using on-board MWR observations.

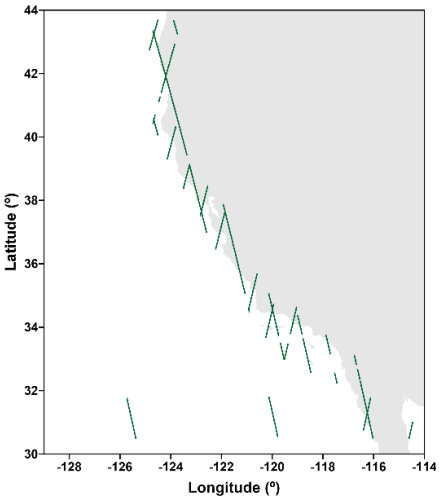


Figure 4.2.25. Points using observations from external imaging MWR.

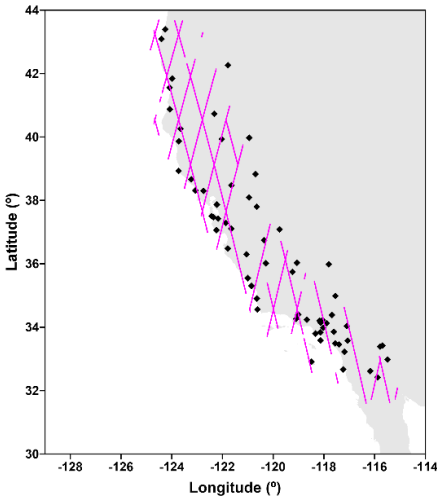


Figure 4.2.26. Points using GNSS observations (68 stations).

Figure 4.2.27, Figure 4.2.28 and Figure 4.2.29 represent WTC_ECM$_{meas}$, WTC_MWR and WTC_GPD for S3A cycle 47 and S3B cycle 28 over California. As for the Caspian, the major feature is the invalid WTC_MWR values around the coastline, which are not present in WTC_GPD.

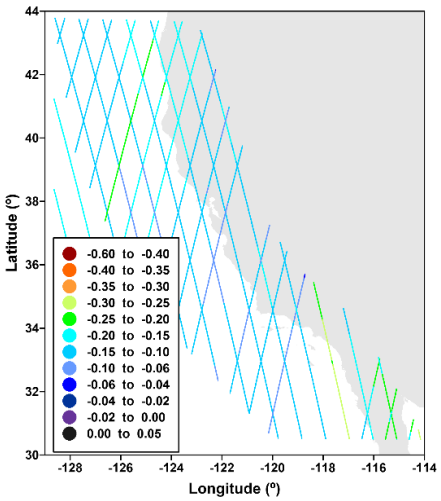


Figure 4.2.27. WTC_ECM$_{meas}$ over the Caspian Sea region, for S3A cycle 47 and S3B cycle 28.

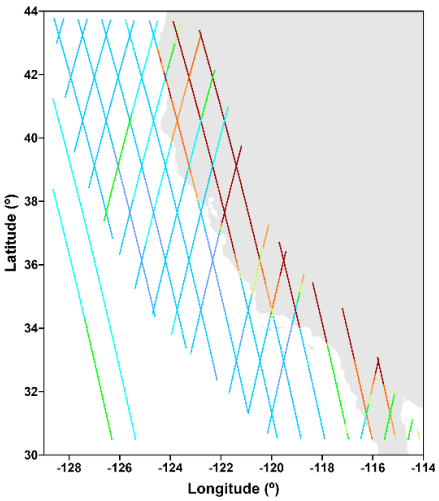


Figure 4.2.28. WTC_MWR over the Caspian Sea region, for S3A cycle 47 and S3B cycle 28.

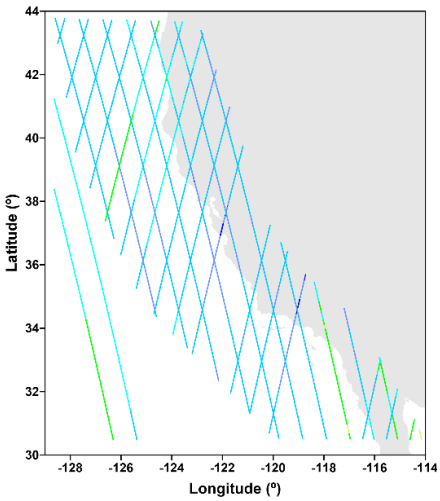


Figure 4.2.29. WTC_GPD over the Caspian Sea region, for S3A cycle 47 and S3B cycle 28.

Considering the whole S3A “water” points (532840) dataset over the California, the statistics of the differences between WTC_GPD and WTC_MWR are: mean=0.72 cm, RMS=4.06 cm and max abs diff=43.88 cm. The corresponding values for the differences between WTC_GPD and WTC_ECM$_{meas}$ are: mean= 0.53 cm, RMS =1.14 cm and max abs diff=8.45 cm. Considering the whole S3B “water” points (417973) dataset over this coastal region, the statistics of the differences between WTC_GPD and WTC_MWR are: mean=1.37 cm, RMS=5.13 cm and max

abs diff=43.84 cm. The corresponding values for the differences between WTC_GPD and WTC_ECM$_{meas}$ are: mean=0.36 cm, RMS =1.05 cm and max abs diff=5.84 cm.

Figure 4.2.30 and Figure 4.2.31 illustrate examples of the various WTC over two passes (S3A and S3B) in the California region, illustrating typical contamination in the MWR near the coast and differences in the model and MWR WTC, when compared with GPD. Over the first pass (S3A) the statistics of the differences between WTC_GPD and WTC_MWR are: RMS=1.07 cm and max abs diff=12.29 cm. The corresponding values for the differences between WTC_GPD and WTC_ECM$_{meas}$ are: RMS=2.45 cm and max abs diff=4.36 cm. For the second pass, the statistics of the differences between WTC_GPD and WTC_MWR are: RMS=5.35 cm and max abs diff=22.40 cm. The corresponding values for the differences between WTC_GPD and WTC_ECM$_{meas}$ are: RMS =2.15 cm and max abs diff=5.01 cm.

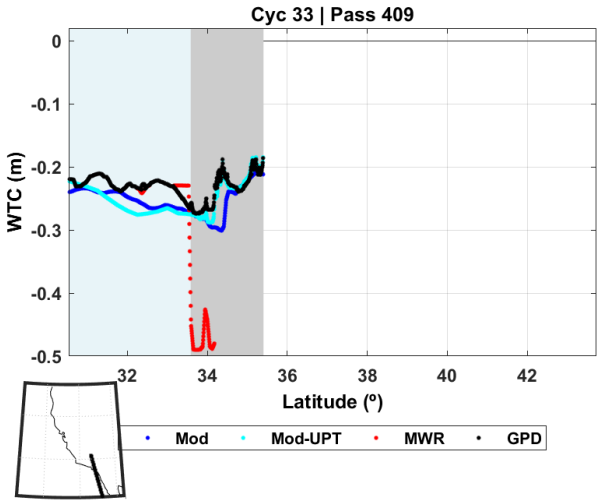


Figure 4.2.30. Various WTC for S3A cycle 33 pass 409: WTC_ECMmeas (blue), WTC_UPT (cyan), WTC_MWR (red) and WTC_GPD (black).

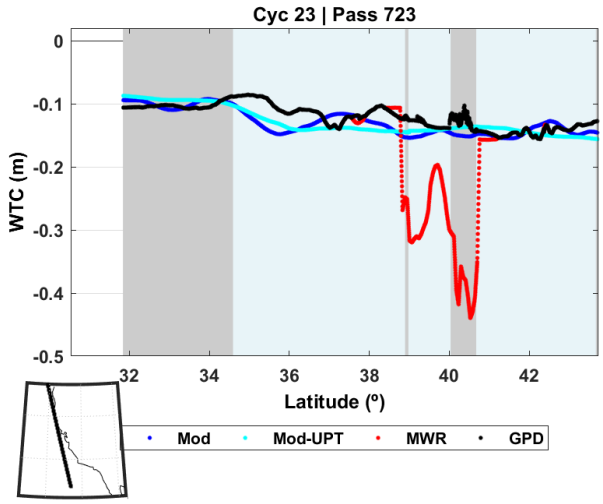


Figure 4.2.31. Various WTC for S3B cycle 23 pass 723: WTC_ECMmeas (blue), WTC_UPT (cyan), WTC_MWR (red) and WTC_GPD (black).

As for the Caspian Sea, it should be noted that the differences between WTC_GPD and WTC_MWR mainly reflect the differences in the points near the coast, since over open ocean the two corrections are equal.

Overall, these results show that the MWR observations cannot be used near the coast and, in these regions, should be replaced by valid values provided by GPD+. Although the overall statistics of the differences between WTC_GPD and WTC_ECM$_{meas}$ indicate small RMS values (smaller than 1.6 cm), there are passes where the RMS of the differences can reach several (up to 4) cm and maximum absolute differences up to almost 9 cm.

Figure 4.2.32 and Figure 4.2.33 show the WTC_ECM$_{oro}$ and the WTC_GPD respectively, over the California region, for CS2 sub-cycle 118. Again, although no on-board MWR exists in CS2, the WTC-GPD is different from WTC_ECM$_{oro}$. This is explained by the fact that, over this region, as observed in Figure 4.2.25 and Figure 4.2.26, there are abundant external SI-MWR and GNSS observations, used to estimate the WTC_GPD.

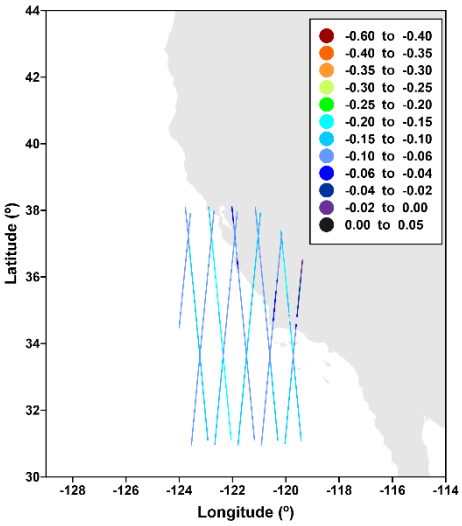


Figure 4.2.32. WTC_ECM$_{oro}$ over the Caspian region, for CS2 sub-cycle 118.

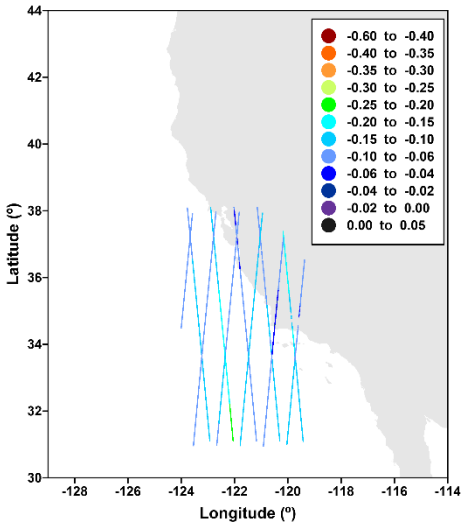


Figure 4.2.33. WTC_GPD over the Caspian region, for CS2 sub-cycle 118.

Figure 4.2.34 illustrates an example of the various WTC in the California region for CS2 sub-cycle 121, pass 677, where the WTC differences between WTC_ECM$_{oro}$ (blue) and the WTC_GPD (black) have an RMS of 2.15 cm, showing that the WTC_GPD can be significantly different from WTC_ECM$_{oro}$ over some passes. However, the overall differences between these two WTC are small. Considering the whole CS2 "water" points dataset over the California region (784605 points), the corresponding statistics of these differences are: mean=-0.13 cm, std=0.9 cm and max abs diff=5.76 cm.

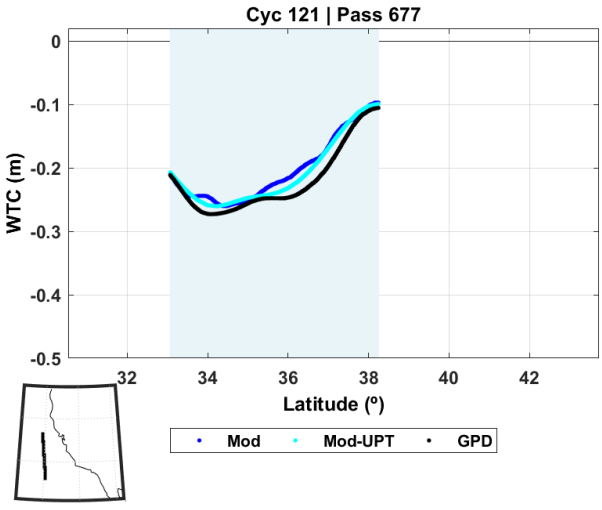


Figure 4.2.34. Various WTC for CS2 sub-cycle 121 pass 677: WTC_ECMoro (blue), WTC_UPT (cyan) and WTC_GPD (black).

Figure 4.2.35 shows the RMS of the differences between GNSS-derived WTC and WTC_MWR (red) and between GNSS and WTC_GPD (black) function of distance from coast. It can be observed that the land contamination present in WTC_MWR up to 20 km from the coast is not present in WTC-GPD.

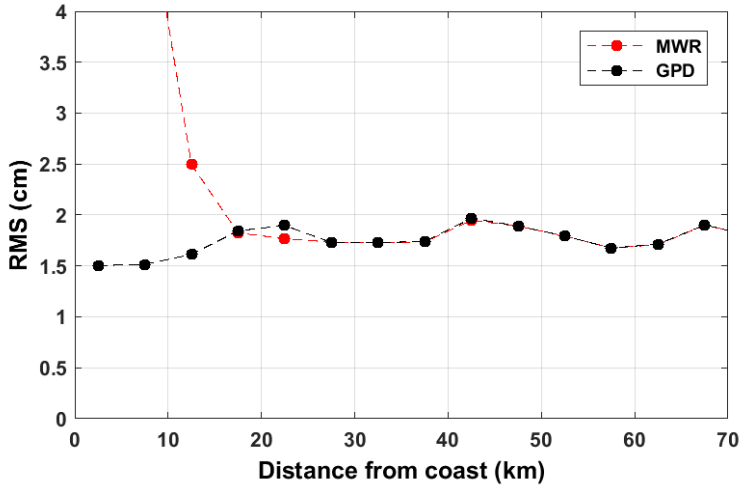


Figure 4.2.35. RMS of the differences between GNSS-derived WTC and WTC_MWR (red) and between GNSS and WTC_GPD (black) function of distance from coast.

Figure 4.2.35 shows an independent assessment of the coastal contamination in the WTC from MWR over California region, reinforcing that the MWR observations cannot be used near the coast and should be replaced by valid values provided by GPD+.

4.2.3 Danube

Figure 4.2.36 represents the ACE2 DEM over the Danube River and Figure 4.2.37 shows the Danube River profile in blue and the altitudes interpolated from ACE2 at the location of the river profile in red.

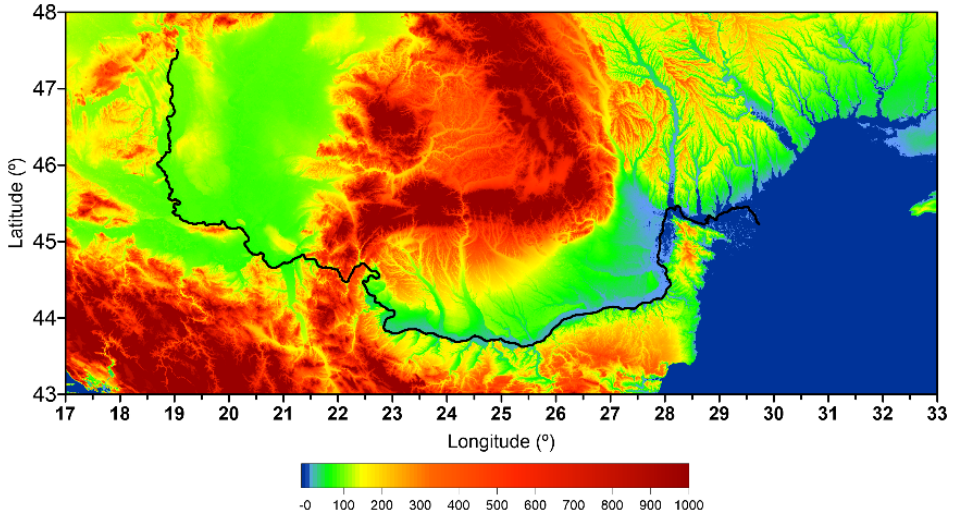


Figure 4.2.36. ACE2 DEM over the Danube River.

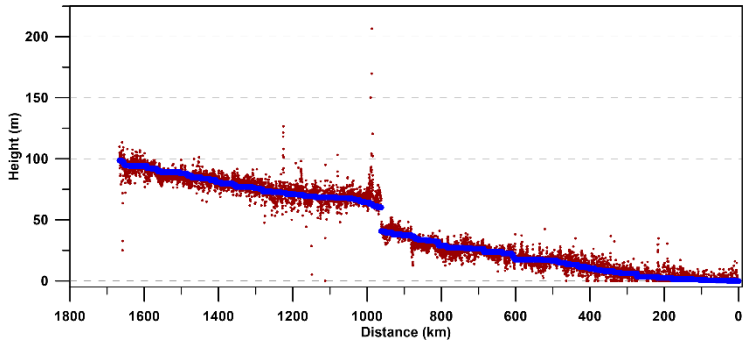


Figure 4.2.37. Danube River profile: Mean River profile (blue) and ACE2 DEM height (red). All values are above geoid, in metres.

Figure 4.2.38 represents the GPD source flag (gpd_wet_tropo_flag) for S3A cycle 47 over the Danube River. Figure 4.2.39 and Figure 4.2.40 represent, the points for which observations exist to estimate the GPD+ WTC over Danube from GNSS and SI-MWR observations, respectively, for the same S3A cycle. SI-MWR observations are mainly available in the sea region around the river mouth.

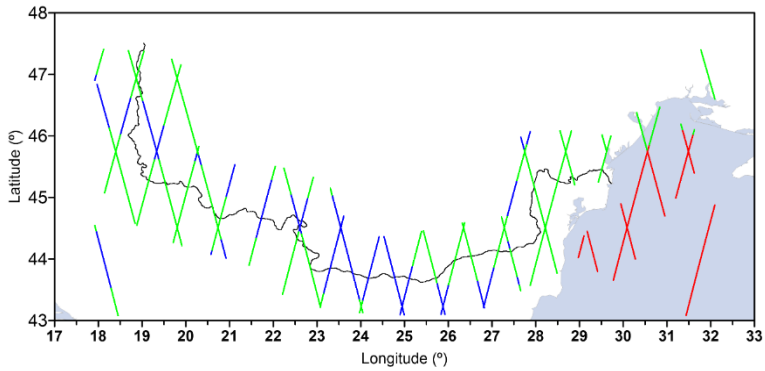


Figure 4.2.38. GPD Source flag for S3A cycle 47: 0 (red) – valid on-board MWR values; 1-7 (green) – estimated from observations; 8 (blue) – from ERA5 model.

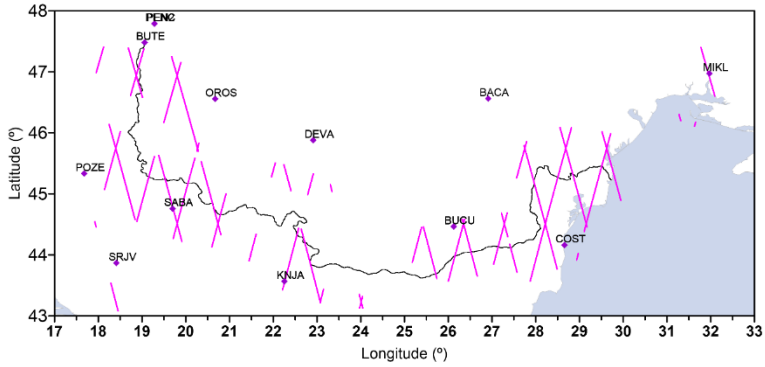


Figure 4.2.39. Points using GNSS observations (30 stations) on S3A cycle 47.

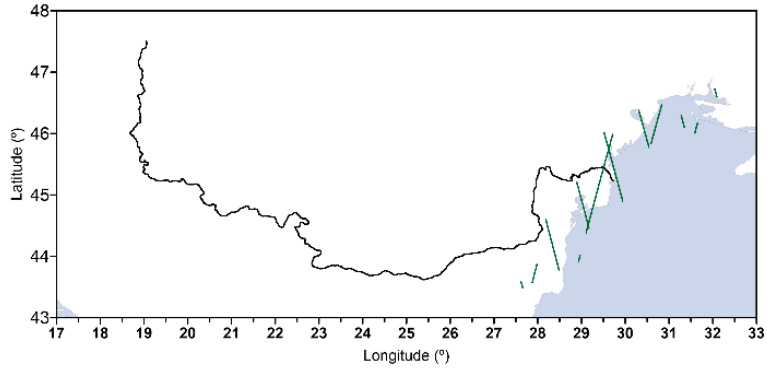


Figure 4.2.40. Points using SI-MWR observations on S3A cycle 47. Approximately the same points also use valid on-board MWR

Figure 4.2.41 and Figure 4.2.42 represent WTC_ECM$_{meas}$ and WTC_GPD, respectively, for S3A cycle 47 over Danube River. For narrow water bodies such as this river, WTC_MWR is always invalid and for this reason it is not represented and considered in the statistics.

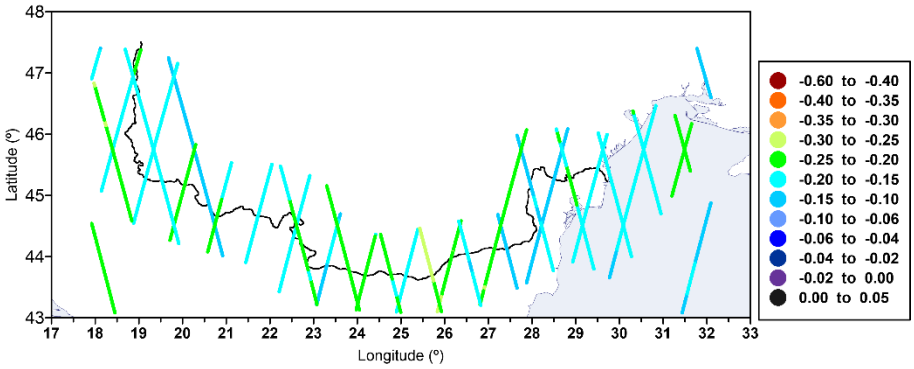


Figure 4.2.41. WTC_ECM$_{meas}$ over the Danube River, for S3A cycle 47.

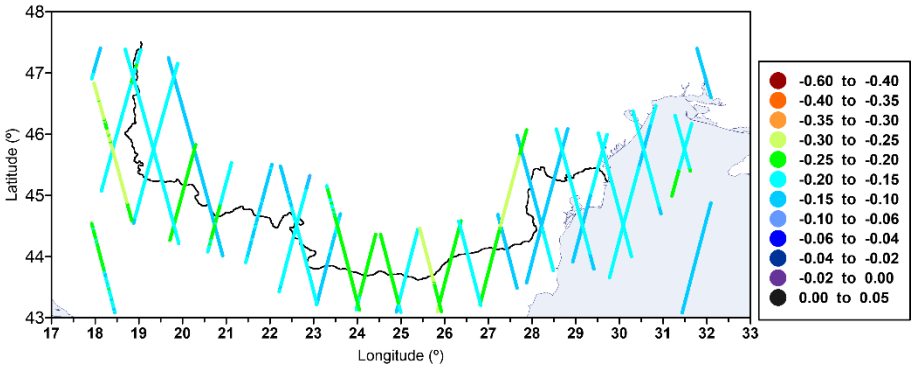


Figure 4.2.42. WTC_GPD over the Danube River, for S3A cycle 47.

Figure 4.2.43 and Figure 4.2.44 represent examples of various WTC for two S3A passes (cycle 46 pass 55 and cycle 58 pass 369) over the Danube River.

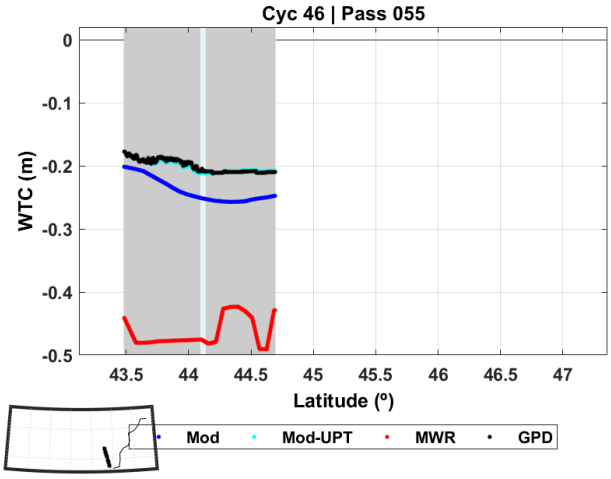


Figure 4.2.43. Various WTC for S3A cycle 46 pass 55: WTC_ECM$_{meas}$ (blue), WTC_UPT (cyan), WTC_MWR (red) and WTC_GPD (black).

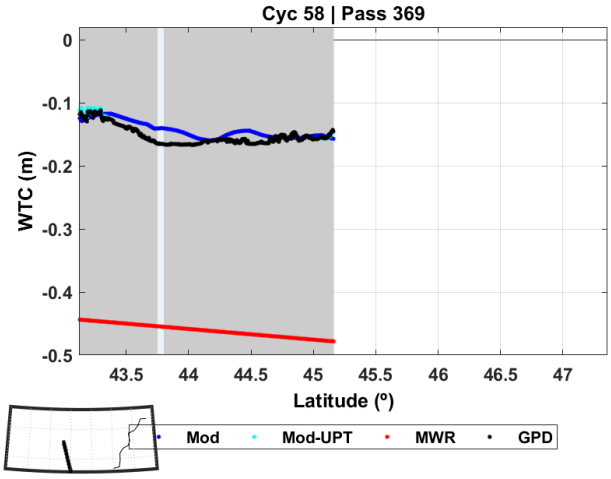


Figure 4.2.44. Various WTC for S3A cycle 58 pass 369: WTC_ECM$_{meas}$ (blue), WTC_UPT (cyan), WTC_MWR (red) and WTC_GPD (black).

For S3A cycle 46 pass 55, the statistics of the differences between WTC_GPD and WTC_ECM$_{meas}$ are: mean=4.41 cm, RMS=4.41 cm and max abs diff=4.69 cm. The same statistics for S3A cycle 58 pass 369 are: mean=-2.48 cm, RMS=2.48 cm and max abs diff=2.53 cm.

Considering the whole S3A “water” points (9752) dataset over the Danube River, the statistics of the differences between WTC_GPD and WTC_ECM$_{meas}$ are: mean=0.21 cm, RMS=0.78 cm and max abs diff=4.69 cm. Overall, the differences are small, however they can be very significant over some passes. Again, over this region, on one hand, the WTC from MWR is always invalid and, on the other hand, the model correction can have very significant errors over water.

Examples for CS2 over the Danube ROI are not shown in this report, since the project CS2 data are only available over sea (river mouth in the Black Sea).

4.2.4 Amazon

Figure 4.2.45 illustrates the ACE2 DEM over the Amazon River, while Figure 4.2.46 represents the Amazon River profile in blue and the altitudes interpolated from ACE2 at the location of the river profile in red.

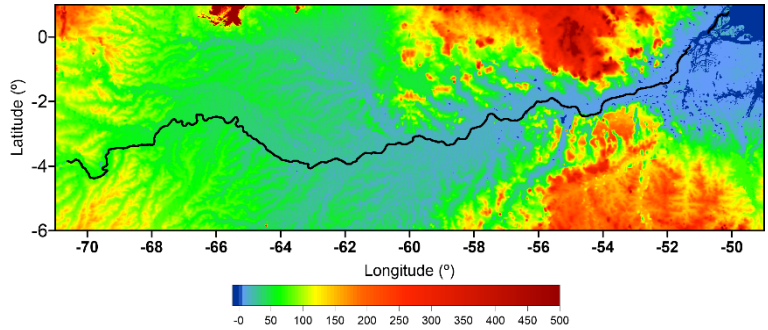


Figure 4.2.45. ACE2 DEM over the Amazon River.

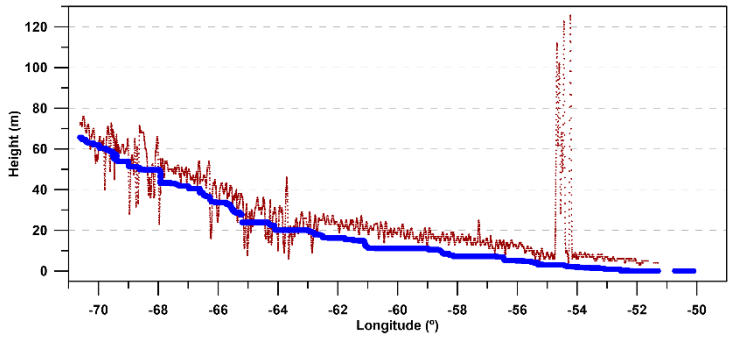


Figure 4.2.46. Amazon River profile: Mean River profile (blue) and ACE2 DEM height (red). All values are above geoid, in metres.

Figure 4.2.47 depicts the GPD source flag (gpd_wet_tropo_flag) for S3A cycle 48 over the Amazon River. The GPD+ WTC for this example is mainly from the ERA5 first-guess or from GNSS observations, the latter represented in Figure 4.2.48. As for the Danube River, WTC_MWR is always invalid and for this reason it is not represented and considered in the statistics.

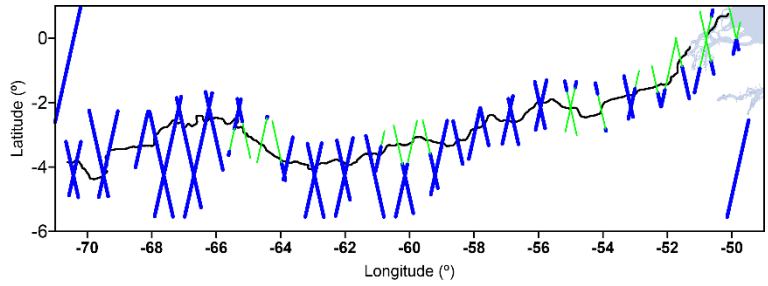


Figure 4.2.47. GPD Source flag for S3A cycle 48: 1-7 (green) – estimated from observations; 8 (blue) – from ERA5 model.

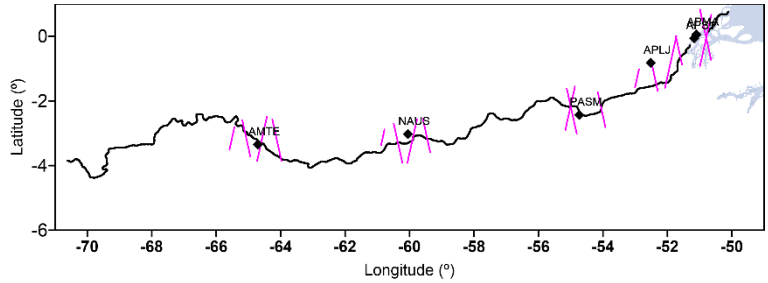


Figure 4.2.48. Points using GNSS observations (6 stations) on S3A cycle 48. Only two tracks in the river mouth use SI-MWR observations. No points use on-board MWR observations.

Figure 4.2.49 and Figure 4.2.50 illustrate the WTC_ECM$_{meas}$ and WTC_GPD, for S3A cycle 48 over the Amazon River, respectively.

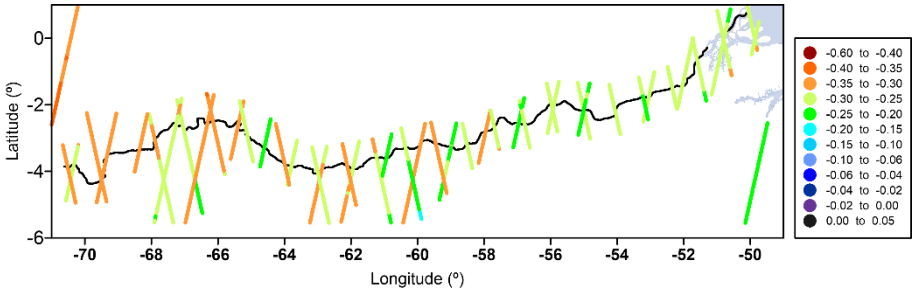


Figure 4.2.49. WTC_ECM$_{meas}$ over the Amazon River, for S3A cycle 48.

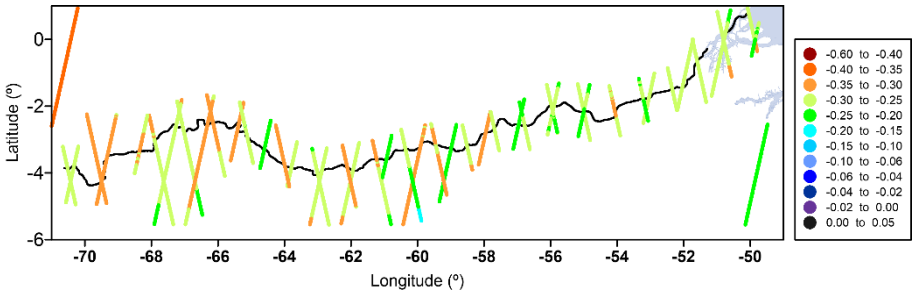


Figure 4.2.50. WTC_GPD over the Amazon River, for S3A cycle 48.

For a local analysis, Figure 4.2.51 and Figure 4.2.52 show examples of the various WTC over two S3A passes in the Amazon region (cycle 33 pass 491 and cycle 46 pass 106).

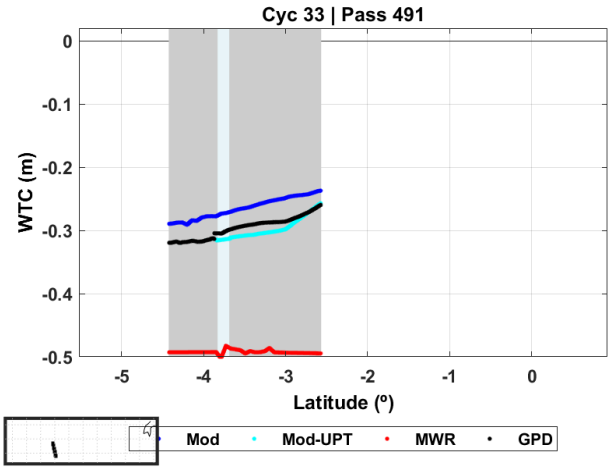


Figure 4.2.51. Various WTC for S3A cycle 33 pass 491: WTC_ECM$_{meas}$(blue) and WTC_GPD (black).

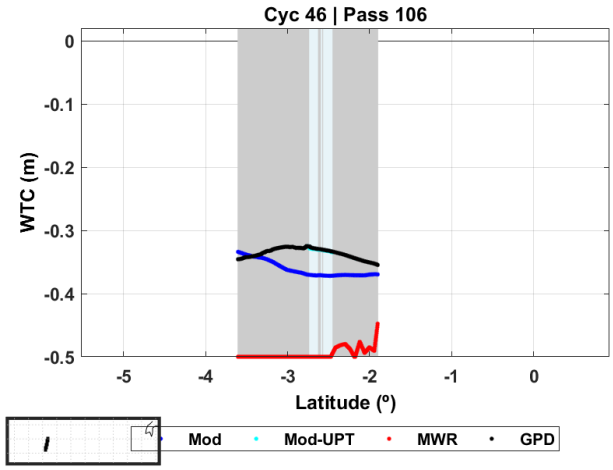


Figure 4.2.52. Various WTC for S3A cycle 46 pass 106 WTC_ECM$_{meas}$ (blue) and WTC_GPD (black).

For S3A cycle 33 pass 491, the statistics of the differences between WTC_GPD and WTC_ECM$_{meas}$ are: mean=-2.95 cm, RMS=2.95 cm and max abs diff=3.15 cm. The same statistics for S3A cycle 46 pass 106 are: mean=4.13 cm, RMS= 4.14 cm and max abs diff=4.49 cm.

Considering the whole S3A "water" points (55762) dataset over the Amazon River, the statistics of the differences between WTC_GPD and WTC_ECM$_{meas}$ are: mean=0.15 cm, RMS=1.10 cm and max abs diff=4.50 cm.

Figure 4.2.53 and Figure 4.2.54 illustrate various WTC for two examples of CS2 tracks over the Amazon River.

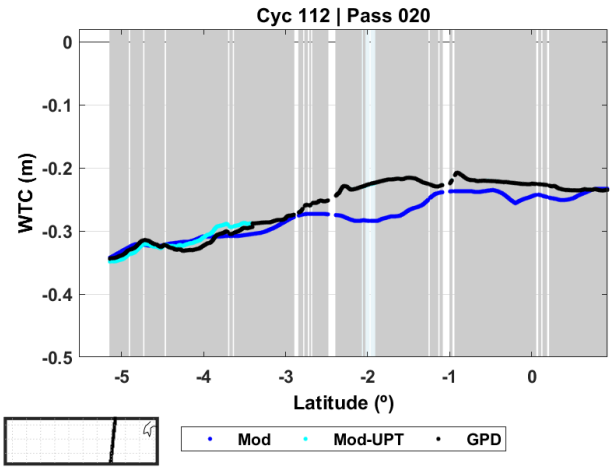


Figure 4.2.53. Various WTC for CS2 sub-cycle 112 pass 020: WTC_ECM$_{oro}$ (blue) and WTC_GPD (black).

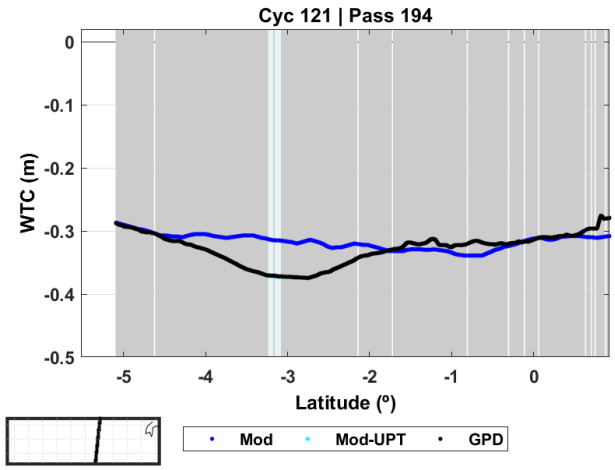


Figure 4.2.54. Various WTC for CS2 sub-cycle 121 pass 194 WTC_ECM$_{oro}$ (blue) and WTC_GPD (black).

Figure 4.2.51 and Figure 4.2.53 show regions where the GPD is not merely the ERA5 model first-guess, since it uses GNSS observations.

For CS2 sub-cycle 112 pass 20 (Figure 4.2.53), the statistics of the differences between WTC_GPD and WTC_ECM$_{oro}$ are: mean=5.82 cm, RMS= 5.82 cm and max abs diff=6.04 cm. The same statistics for CS2 sub-cycle 121 pass 194 (Figure 4.2.54) are: mean=-5.68 cm, RMS= 5.68 cm and max abs diff=5.75 cm.

Considering the whole CS2 "water" points (40193) dataset over the Amazon River, the statistics of the differences between WTC_GPD and WTC_ECM$_{oro}$ are: mean=0.07 cm, RMS=1.10 cm; max abs diff=6.04 cm.

For S3 and CS2 over the Amazon River, the results show that the overall differences between GPD and model are small (RMS values smaller than 1.1 cm), however over some specific regions these differences can be larger than 4 cm.

4.3 Validation of the DTC

In the validation of the DTC the following analyses are performed:

1. Comparison with DTC present in products.
2. Along-track analysis of DTC and water level profiles, inspecting unexpected behaviour of the correction, present in some current products.

In the current S3 products, two types of DTC are provided, both from the ECMWF Op. model: one computed at zero level, **DTC_Zero** (mod_dry_tropo_cor_zero_altitude field) and another at the altimeter measurement level, **DTC_meas** (mod_dry_tropo_cor_meas_altitude field).

For CS2, the **DTC_UPT** is compared with the only model-derived DTC present in the products, referred to the model orography (**DTC_oro**).

4.3.1 Caspian Sea

Figure 4.3.1 and Figure 4.3.2 illustrate, for S3A cycle 47, the DTC_UPT and DTC_meas, respectively, over the Caspian Sea. For this cycle, the statistics of the differences between these DTC, over "water" points are: mean=0.03 cm, std =0.11 cm; max abs diff=0.42 cm.

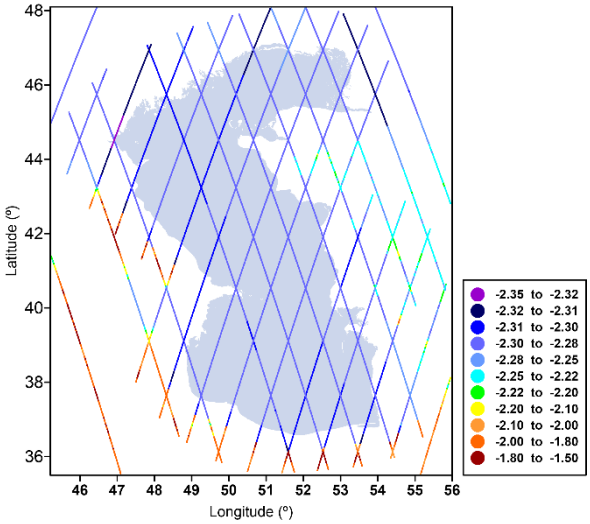


Figure 4.3.1. DTC_UPT over the Caspian Sea region, for S3A cycle 47.

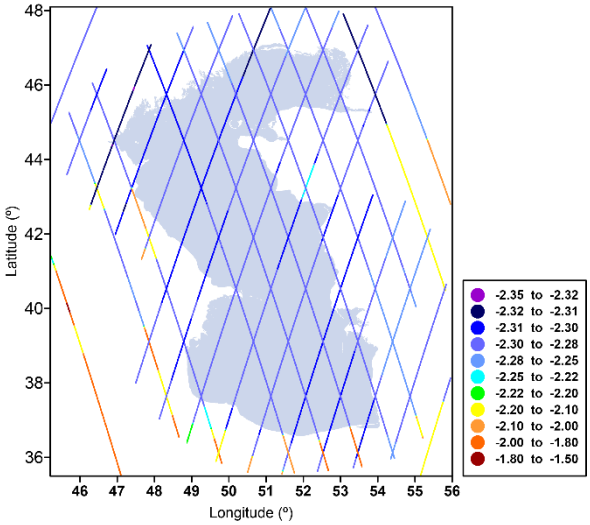


Figure 4.3.2. DTC_meas over the Caspian Sea region, for S3A cycle 47.

Figure 4.3.3 and Figure 4.3.4 represent examples of DTC for passes of S3A over the Caspian Sea, showing very small DTC differences between UPT and model over water (very light blue zones).

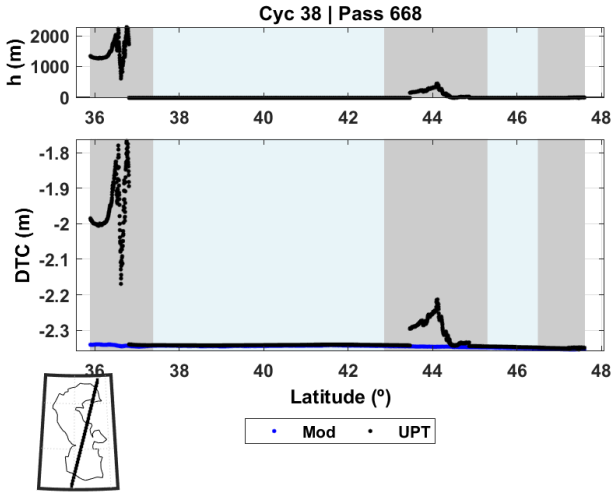


Figure 4.3.3. Various DTC for S3A cycle 38 pass 668: DTC_meas (blue) and DTC_UPT (black) over the Caspian Sea.

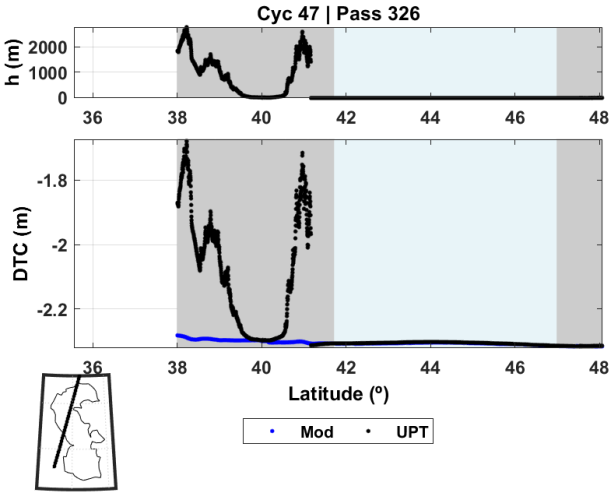


Figure 4.3.4. Various DTC for S3A cycle 47 pass 326: DTC_meas (blue) and DTC_UPT (black) over the Caspian Sea.

When the whole S3A “water” points dataset for the Caspian Sea is considered, the statistics of the differences between DTC_UPT and DTC_ECM$_{meas}$ are: mean=0.03 cm, std=0.33 cm and max abs diff=21.16 cm. The same statistics for S3B are 0.00, 0.18 and 19.01 cm, respectively.

Figure 4.3.5 and Figure 4.3.6 illustrate, for CS2 sub-cycle 118, the DTC_UPT and DTC_oro, respectively, over the Caspian Sea. The statistics of the differences between these DTC, over “water” points are: mean=-0.71 cm, std=0.07 cm; max abs diff=1.7 cm. When the whole CS2 “water” points dataset for the Caspian Sea is considered,

the corresponding statistics are: mean=-0.70 cm, std =0.16 cm; max abs diff=2.55 cm. The absolute mean difference of 0.7 cm is due to the altitude at which the corrections are computed. The DTC available in the CS2 products are computed at zero level, while the corrections computed at UPorto are available at the actual lake level (~28 m below zero level). Thus, over the Caspian Lake, the absolute DTC_UPT is systematically 0.7 cm larger than the absolute DTC present in the CS2 products.

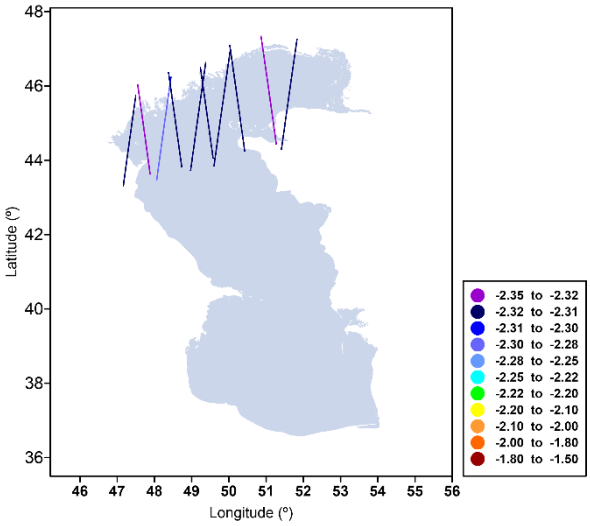


Figure 4.3.5. DTC_UPT over the California region, for CS2 sub-cycle 118.

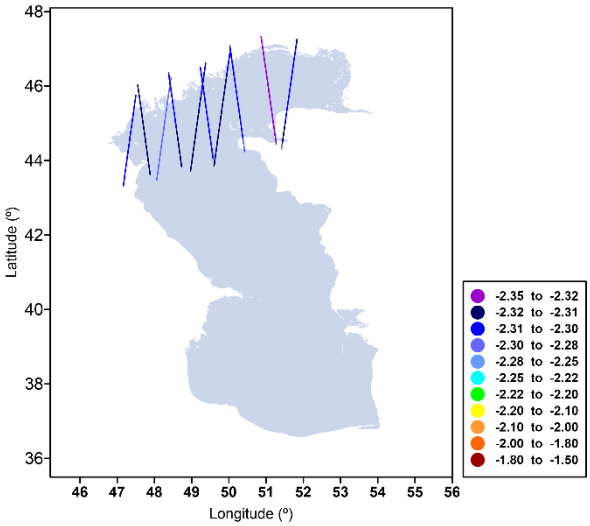


Figure 4.3.6. DTC_oro over the California region, for CS2 sub-cycle 118.

Figure 4.3.7 and Figure 4.3.8 show examples of passes of CS2 for the DTC over the Caspian, where the systematic difference is observed.

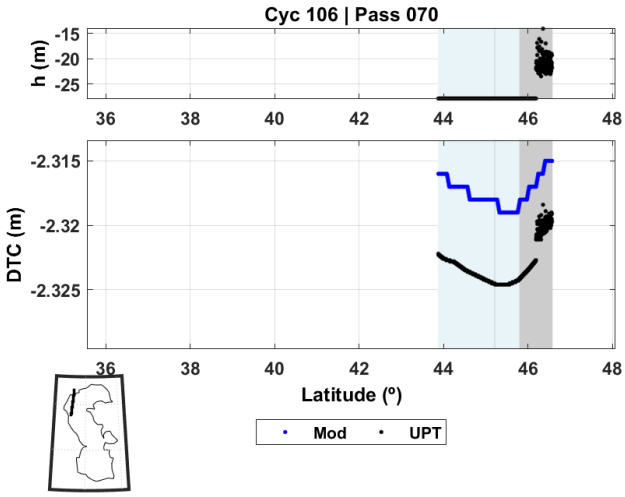


Figure 4.3.7. Various DTC for CS2 sub-cycle 106 pass 70: DTC_oro (blue) and DTC_UPT (black) over the Caspian Sea.

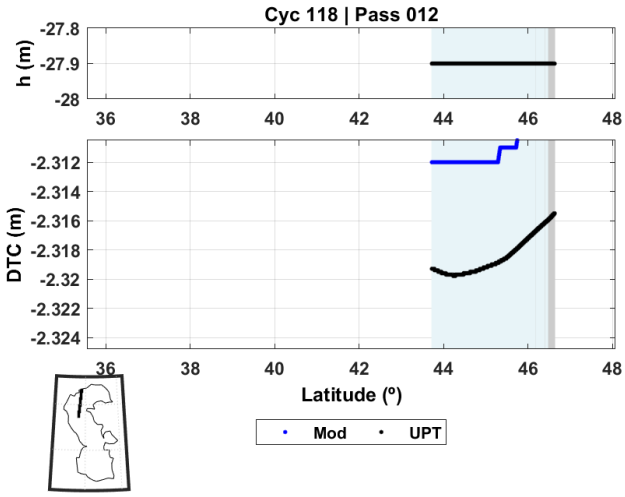


Figure 4.3.8. Various DTC for CS2 sub-cycle 118 pass 12: DTC_oro (blue) and DTC_UPT (black) over the Caspian Sea.

It can be concluded that, over the Caspian, considering “water” points only, DTC_UPT and DTC_meas present in S3 products are very close with sub-centimetric differences. However, DTC_oro present in CS2 products is provided at zero instead of the mean sea level of the Caspian Sea (approximately -28 m), which causes a systematic bias of 7 mm (DTC_oro less negative than DTC_UPT by 7 mm).

When “water” and land points surrounding the Sea are considered, the UPT and S3 corrections are provided at different levels. In the UPT corrections a buffer of width Δ=50 km has been created around the sea and, within this buffer the corrections are also given at the mean sea level of the Caspian Sea (-28 m). According to the product information, in S3 DTC_meas should be given at the altimeter measurement level. However, in some examples as those given in Figure 4.3.3 and Figure 4.3.4, this does not seem to be the case, as the corrections seem also to be provided at the mean lake level.

4.3.2 California

Figure 4.3.9 and Figure 4.3.10 illustrate, for S3A cycle 47, the DTC_UPT and DTC_meas respectively over the California region. For this cycle, the statistics of the differences between these DTC, over “water” points are: mean=0.03 cm, std =0.51 cm; max abs diff=1.68 cm.

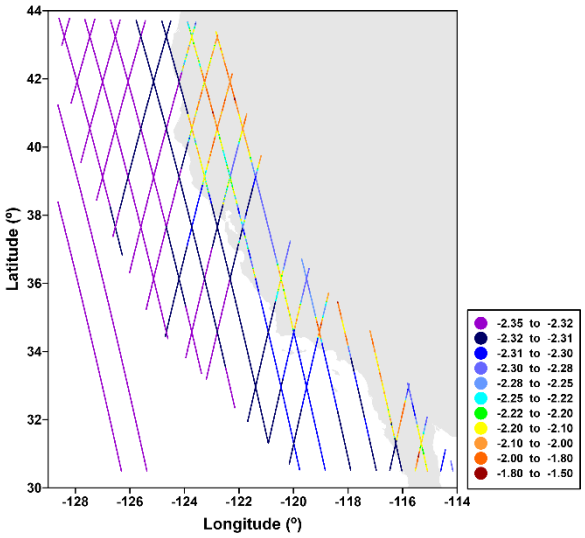


Figure 4.3.9. DTC_UPT over the California region, for S3A cycle 47.

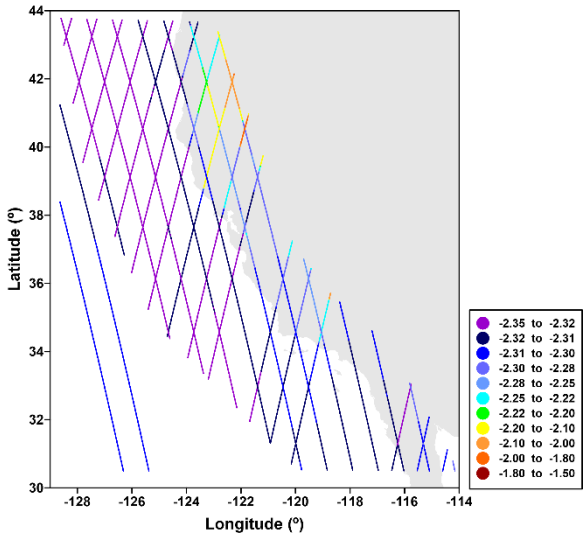


Figure 4.3.10. DTC_meas over the California region, for S3A cycle 47.

Figure 4.3.11 and Figure 4.3.12 show examples of passes of S3A, where the DTC differences over water are very small. Overall, over the oceanic and coastal “water” regions the corrections are very close, at the sub-decimetre level.

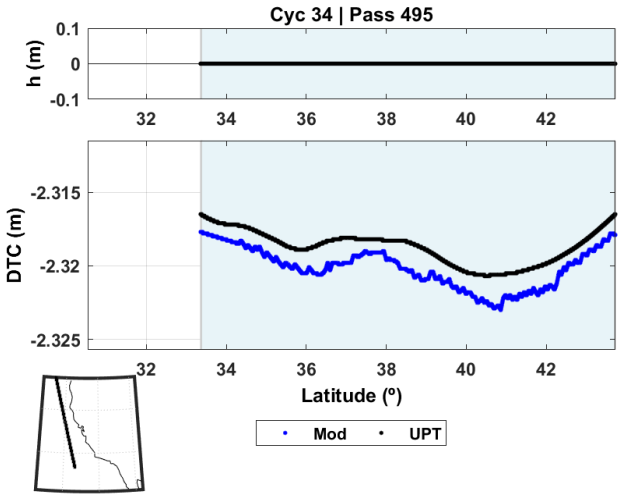


Figure 4.3.11. Various DTC for S3A cycle 34 pass 495: DTC_meas (blue) and DTC_UPT (black) over California.

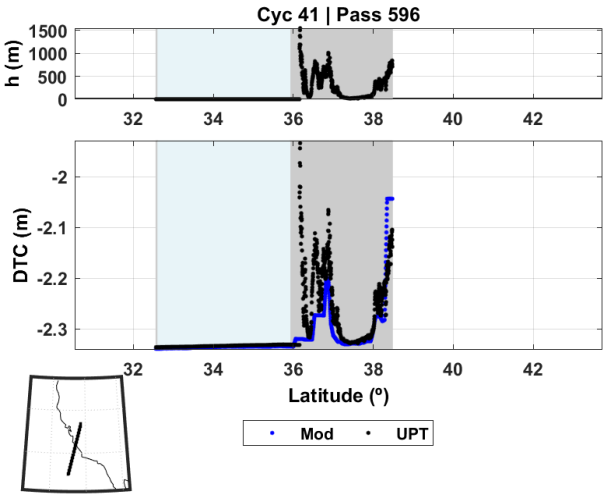


Figure 4.3.12. Various DTC for S3A cycle 41 pass 596: DTC_meas (blue) and DTC_UPT (black) over California.

Considering the whole S3A "water" points dataset over the California, the statistics of the differences between DTC_UPT and DTC_ECM$_{meas}$ are: mean=0.24 cm, RMS=0.27 cm. The same statistics for S3B are 0.25 cm and 0.28 cm, respectively. When only "ocean" points are considered (surface_type=0), the same statistics are obtained.

Figure 4.3.13 and Figure 4.3.14 illustrate, for CS2 sub-cycle 112, the DTC_UPT and DTC_oro respectively over the California region.

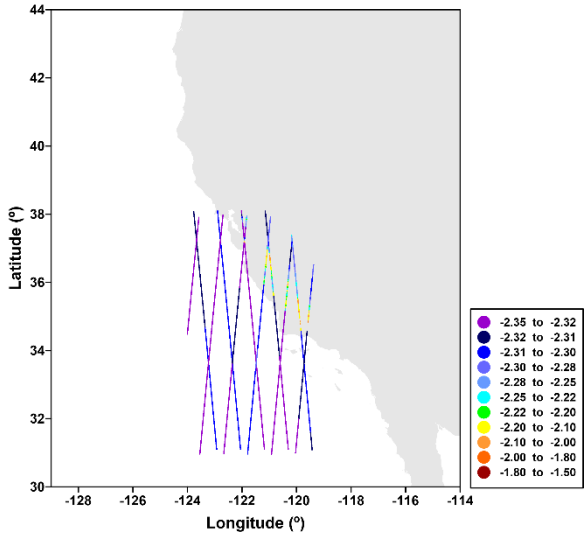


Figure 4.3.13. DTC_UPT over the California region, for CS2 sub-cycle 112.

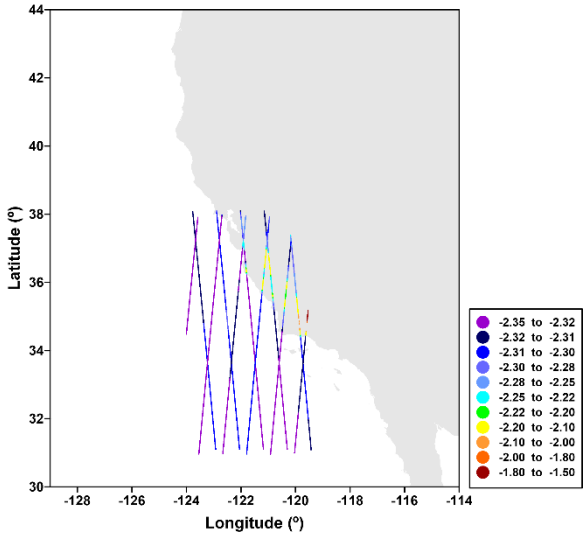


Figure 4.3.14. DTC_oro over the California region, for CS2 sub-cycle 112.

When the whole CS2 "water" points dataset for the California region is considered, the corresponding statistics of the DTC differences between UPT and model are: mean: -0.17 cm, std=1.76 cm and max abs diff=32.96 cm.

Figure 4.3.15 and Figure 4.3.16 Illustrate examples of DTC for passes of CS2 over the coast of California.

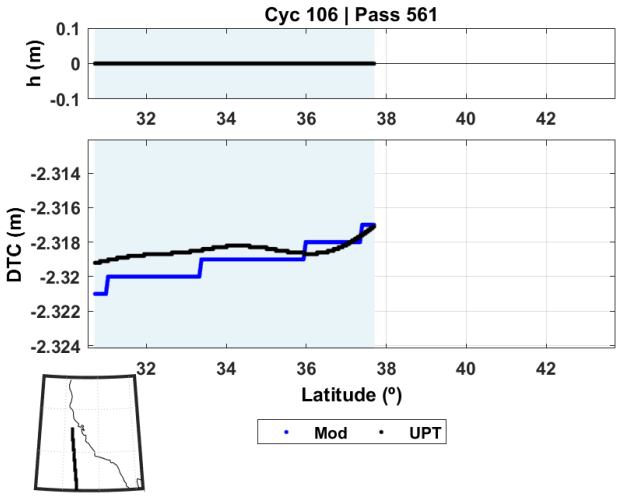


Figure 4.3.15. Various DTC for CS2 sub-cycle 106 pass 561: DTC_oro (blue) and DTC_UPT (black) over California.

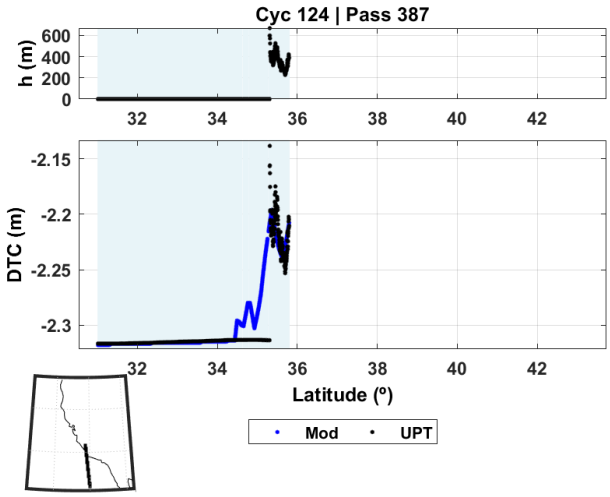


Figure 4.3.16. Various DTC for CS2 sub-cycle 124 pass 387: DTC_oro (blue) and DTC_UPT (black) over California.

For CS2 over the California region, it can be concluded that, over clean “ocean” points (surface_type=0), DTC_UPT and DTC_oro present in CS2 products are very close, at sub-centimetre level (Figure 4.3.15).

4.3.3 Danube

Figure 4.3.17 and Figure 4.3.18 show the DTC_UPT and DTC_meas, respectively, over the Danube River, for S3A cycle 47.

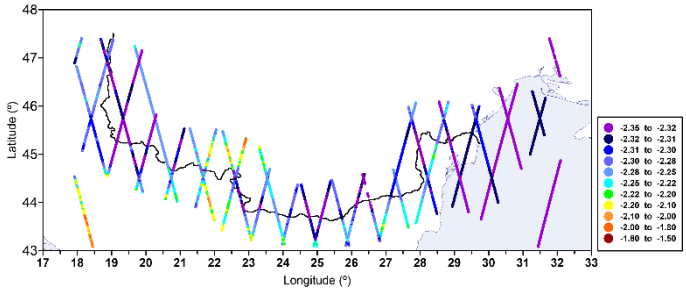


Figure 4.3.17. DTC_UPT over the Danube River, for S3A cycle 47.

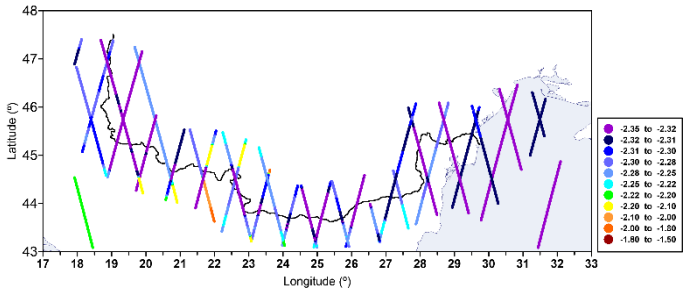


Figure 4.3.18. DTC_meas over the Danube River, for S3A cycle 47.

Figure 4.3.19 and Figure 4.3.20 represent examples of DTC for passes of S3A over the Danube River, showing also very small DTC differences over water surfaces (very light blue and narrow bars).

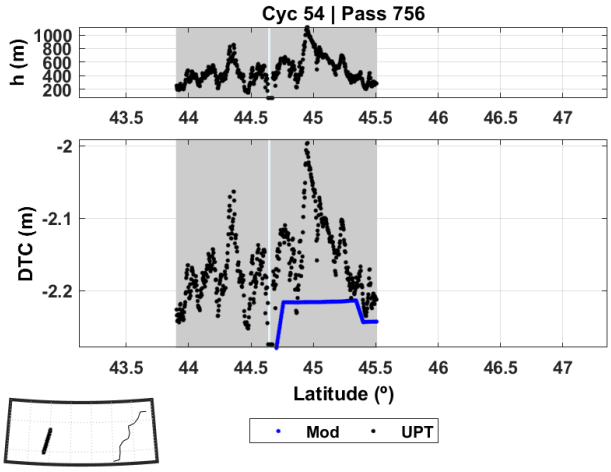


Figure 4.3.19. Various DTC for S3A cycle 54 pass 756: DTC_meas (blue) and DTC_UPT (black) over the Danube River.

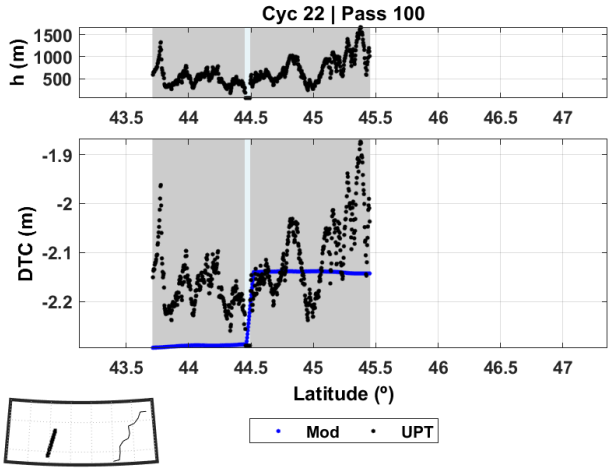


Figure 4.3.20. Various DTC for S3B cycle 22 pass 100: DTC_meas (blue) and DTC_UPT (black) over the Danube River.

Considering the whole S3A "water" points dataset over the Danube River, the statistics of the differences between DTC_UPT and DTC_ECM$_{meas}$ are: mean=0.02 cm, RMS=0.20 cm. The same statistics for S3B are -0.06 cm and 0.30 cm, respectively. For S3 over the Danube region, it also can be concluded that DTC_UPT and DTC_meas present in S3 products are very close, at sub-centimetre level.

4.3.4 Amazon

Figure 4.3.21 and Figure 4.3.22 illustrate the DTC_UPT and DTC_meas, respectively, over the Amazon River, for S3A cycle 48.

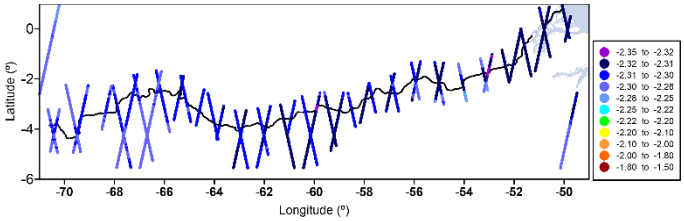


Figure 4.3.21. DTC_UPT over the Amazon River, for S3A cycle 48.

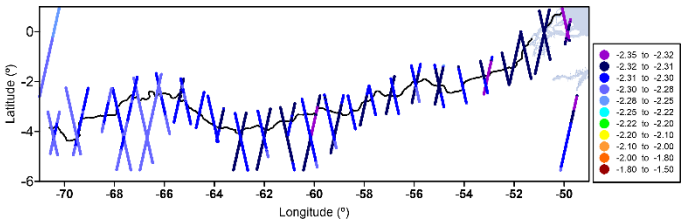


Figure 4.3.22. DTC_meas over the Amazon River, for S3A cycle 48.

Figure 4.3.23 and Figure 4.3.24 represent examples of DTC for passes of S3A crossing the Amazon River, showing also small DTC differences. The example of Figure 4.3.23 shows a difference of almost 1 cm over water points.

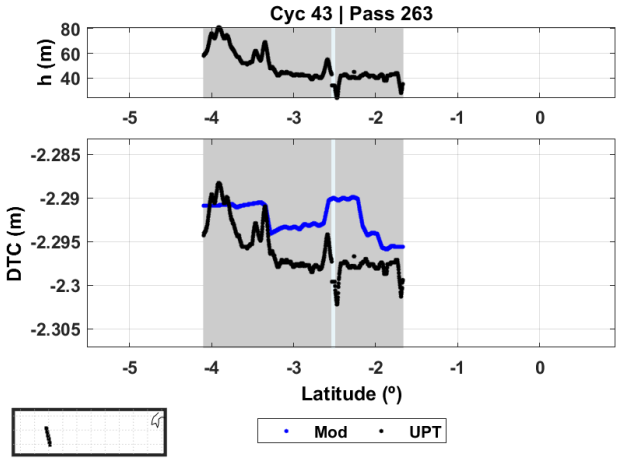


Figure 4.3.23. Various DTC for S3A cycle 43 pass 263: DTC_meas (blue) and DTC_UPT (black) over the Amazon River.

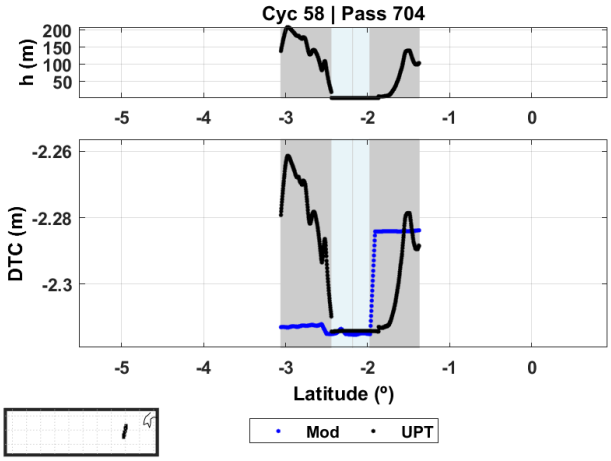


Figure 4.3.24. Various DTC for S3A cycle 58 pass 704: DTC_meas (blue) and DTC_UPT (black) over the Amazon River.

Considering the whole S3A "water" points dataset over the Amazon River, the statistics of the differences between DTC_UPT and DTC_ECM_{meas} are: mean=-0.11 cm, RMS=0.25 cm. The same statistics for S3B are -0.02 cm and 0.24 cm, respectively. These results show again differences at sub-centimetre level.

Figure 4.3.25 and Figure 4.3.26 represent examples of DTC for CS2 passes crossing the Amazon River, where the differences can reach 2 cm over water, however the overall statistics are very small.

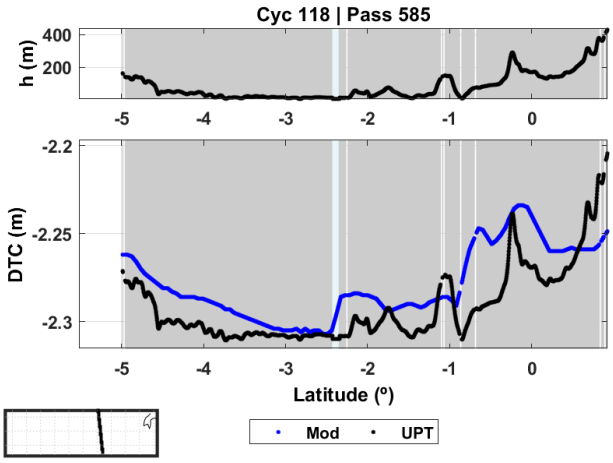


Figure 4.3.25. Various DTC for CS2 sub-cycle 118 pass 585 DTC_oro (blue) and DTC_UPT (black) over the Amazon River.

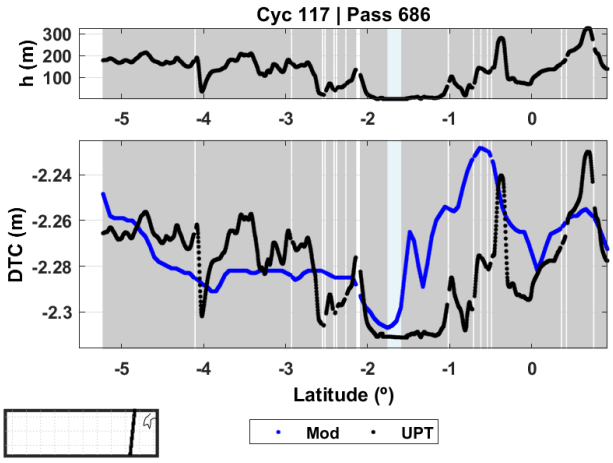


Figure 4.3.26. Various DTC for CS2 sub-cycle 117 pass 686: DTC_oro (blue) and DTC_UPT (black) over the Amazon River.

Considering the whole CS2 “water” points dataset over the Amazon River, the statistics of the differences between DTC_UPT and DTC_ECM_{oro} are: mean=-0.29 cm and RMS=0.40 cm, showing also differences at sub-centimetre level.

4.3.5 Other regions

The sub-sections above describe the validation performed by UPorto in the four selected test areas (Caspian Sea, California, Danube River and Amazon River). In addition, this sub-section reports specific behaviour of the dry tropospheric corrections for the S3 in the Yangtze River and for the CS2 in Amur River. Overall, the four regions reported above show DTC differences at sub-centimetre level. These two additional regions of interest (Yangtze and Amur Rivers) can have regions with very large differences between DTC_UPT and those present in the satellite products.

4.3.5.1 Yangtze River

Figure 4.3.27 and Figure 4.3.28 represent examples of DTC for S3A passes over the Yangtze River, showing very large DTC differences. These two examples show regions where the DTC differences between UPT and the corrections present in the S3 products can reach 40 cm (Figure 4.3.27) and 20 cm (Figure 4.3.28) over water surfaces (very light blue and narrow bars).

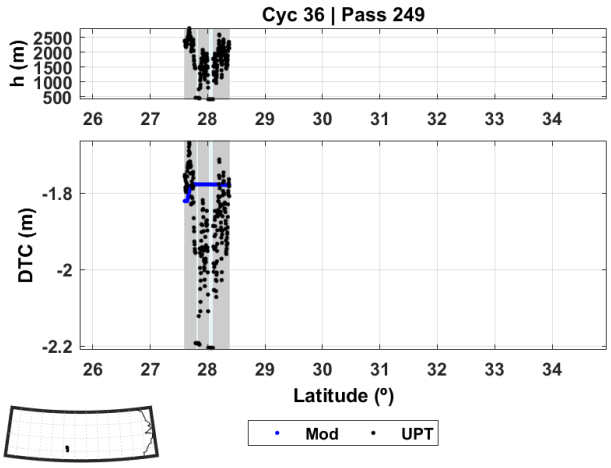


Figure 4.3.27. Various DTC for S3A cycle 36 pass 249: DTC_meas (blue) and DTC_UPT (black) over the Yangtze River.

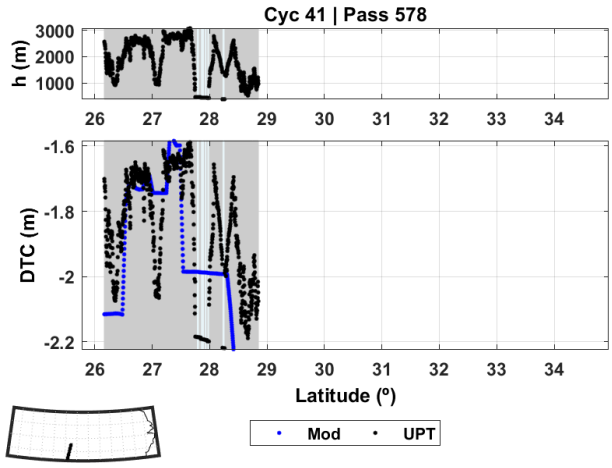


Figure 4.3.28. Various DTC for S3A cycle 41 pass 578: DTC_meas (blue) and DTC_UPT (black) over the Yangtze River.

Considering the whole S3A "water" points over the Yangtze the statistics of the differences between DTC_UPT and DTC_ECM_{meas} are: mean=-1.02 cm; RMS=5.02 cm and max abs diff=45.58 cm, showing very significant DTC differences with a direct impact in the water level measurement from altimetry.

4.3.5.2 Amur River

Figure 4.3.29 shows an example for a CS2 track over the Amur River, where the differences between the DTC_UPT and those present in the CS2 products are larger than 3 cm over the water points (represented by light blue bars and where the altitude in the top plot is flat).

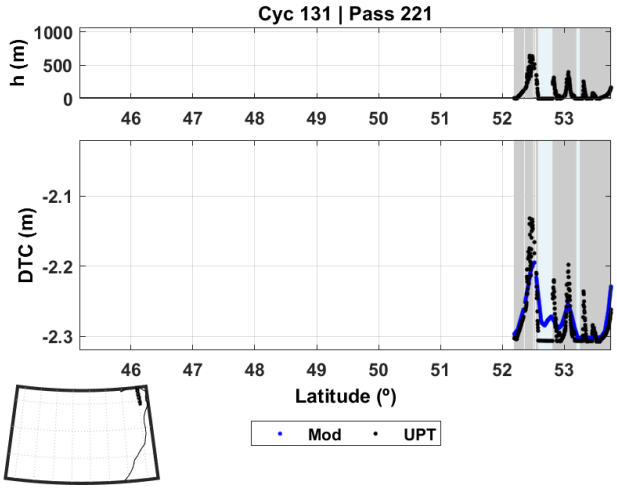


Figure 4.3.29. Various DTC for CS2 sub-cycle 131 pass 221: DTC_oro (blue) and DTC_UPT (black) over the Amur River.

Considering the whole CS2 “water” points over the Amur River, the statistics of the differences between DTC_UPT and DTC_ECM$_{oro}$ are: mean=-1.37 cm; RMS=2.11 cm and max abs diff=7.96 cm.

4.4 Main conclusions on the validation of the WTC and DTC

4.4.1 Main conclusions for the WTC

To summarize the main results of this report concerning the validation of the WTC, Table 4.4.1 shows the overall statistics of the WTC differences for the several analysed ROI.

Table 4.4.1. Overall statistics of the various WTC differences for the four regions of interest.

ROI	Satellite	WTC difference	NP	Mean (cm)	Standard deviation (cm)	RMS (cm)	Maximum absolute difference (cm)
Caspian	S3A	GPD-ECM	347454	0.33	1.60	1.63	9.99
		GPD-MWR		3.01	8.52	9.01	43.81
	CS2	GPD-ECM	186453	-0.60	1.17	1.31	7.34
California	S3A	GPD-ECM	532840	0.53	1.01	1.14	8.45
		GPD-MWR		0.72	3.99	4.06	43.88
	CS2	GPD-ECM	784605	-0.13	0.89	0.90	5.76
Danube	S3A	GPD-ECM	9752	0.21	0.75	0.78	4.69
Amazon	S3A	GPD-ECM	55762	0.15	1.09	1.10	4.50
	CS2	GPD-ECM	40193	0.07	1.10	1.10	6.04

Concerning the WTC from MWR, over rivers, the WTC differences between GPD and MWR are not analysed, since the WTC retrieved from MWR over narrow rivers is always invalid. For the Caspian and California regions, the RMS of the WTC differences between GPD and MWR is about 9 and 4 cm, respectively. For coastal regions and large lakes, these very significant differences mainly reflect the differences in the points near the coast, since over the points where the MWR has been considered valid, the two corrections are equal. Overall, these results show that the MWR observations cannot be used near the coast, due to the land contamination. The independent comparison with GNSS performed in California reinforces this conclusion.

Regarding the differences between GPD and model (WTC_ECM$_{meas}$ for S3 and WTC_ECM$_{oro}$ for CS2), although the overall statistics of the differences indicate small RMS values in the range of 0.8-1.6 cm, there are passes where the RMS can reach 4 cm and maximum absolute differences can reach almost 10 cm (S3A over Caspian). Examples of S3 and CS2 passes where these differences are systematically of various centimetres were shown. These values are a clear indication that the adoption of model WTC present in the S3 and CS2 products may lead to a significant increase in the altimeter derived sea/water surface heights errors over these regions.

Thus, in these regions, neither the MWR nor the model should be adopted. The adopted WTC should be an improved and continuous WTC such as GPD+. Moreover, studies carried out by Ablain et al. (2022) indicate that the WTC accounts for ~40% of the current errors in the SSH, therefore, improvements in the WTC are still required.

4.4.2 Main conclusions for the DTC

Regarding the validation of the DTC, in order to summarize the main results of this report, Table 4.4.2 shows the overall statistics of the DTC differences for the several analysed ROI.

Table 4.4.2. Overall statistics of the DTC differences for the selected regions of interest.

ROI	Satellite	DTC difference	NP	Mean (cm)	Standard deviation (cm)	RMS (cm)	Maximum absolute difference (cm)
Caspian	S3A	UPT-ECM	347454	0.03	0.33	0.33	21.16
	CS2		186453	-0.70	0.10	0.71	2.25
California	S3A		532840	0.24	0.10	0.27	2.53
	CS2		784605	-0.17	1.76	1.77	32.96
Danube	S3A		9752	0.02	0.20	0.20	4.50
Amazon	S3A		55762	-0.11	0.22	0.25	1.42
	CS2		40193	-0.29	0.28	0.40	2.50
Yangtze	S3A		20422	-1.02	4.92	5.02	45.58
Amur	CS2		3268	-1.37	1.60	2.11	7.96

Overall, it can be concluded that over clean water points DTC_UPT and DTC_meas (or DTC_oro for CS2) are very close, at sub-centimetre level. In particular, this is the case of coastal regions. However, for high and narrow inland water bodies, the differences between these two DTC can be of the order of the decimetre level.

Over Yangtze River, the overall statistics of the differences between DTC_UPT and DTC_ECM_{meas} (S3A) indicate an RMS value of 5 cm and in some passes these differences can be larger than 40 cm (cf. also Figure 4.1.2, for the potential errors caused by linear height reduction of the DTC for very large altitudes) For CS2, over regions such as the Amur River the DTC may have also very significant errors (of various centimetres).

In the case of the S3 products these very significant errors still present in the DTC may be due to the use of “wrong” altimeter-derived water heights, caused e.g. by poor retracker performance, mainly in narrow regions of the water body. A step shape of the S3 DTC observed in some regions was not well understood. In the case of CS2, the largest errors found may be due to “wrong” model orography.

Therefore, the adoption of model DTC over these regions may lead to a significant increase in the altimeter derived water surface heights errors, in particular over regions at high altitudes. Over these regions the DTC should be correctly referred to the actual surface height, as performed in the computation of the DTC by UPorto.

5 References

Ablain et al., (2022) Benefit of a second calibration phase to estimate the relative global and regional mean sea level drifts between Jason-3 and Sentinel-6A. 2020 Ocean Surface Topography Science Team Meeting (virtual).

Aldarias, A., Gómez-Enri, J., Laiz, I., Tejedor, B., Vignudelli, S., Cipollini, P. Validation of Sentinel-3A SRAL Coastal Sea Level Data at High Posting Rate: 80 Hz. IEEE Transactions on Geoscience and Remote Sensing, 58, 6, 3809-3821. doi: 10.1109/TGRS.2019.2957649. 2020.

Altenau et al., (2021) The Surface Water and Ocean Topography (SWOT) Mission River Database (SWORD): A Global River Network for Satellite Data Products. Water Resources Research. doi:10.1029/2021WR030054

Andersen, O. B., Nielsen, K., Knudsen, P., Hughes, C. W., Bingham, R., Fenoglio-Marc, L., Gravelle, M., Kern, M., & Polo, S. P. (2018). Improving the coastal mean dynamic topography by geodetic combination of tide gauge and satellite altimetry. Marine Geodesy, 41(6), 517–545. https://doi.org/10.1080/01490419.2018.1530320.

Berghuijs W. R., Sivapalan M., Woods R. A., Savenije H. G. (2014). “Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales”. Water Resources Research, vol. 50, num 7, p. 5638-5661, doi:10.1002/2014WR015692.

Berry, P. A. M., Smith, R. G., & Benveniste, J. (2008). ACE2: The New Global Digital Elevation Model. In S. P. Mertikas (Ed.), Gravity, Geoid and Earth Observation (Vol. 135, pp. 231–237). Chania, Greece: Springer. doi: 10.1007/978-3-642-10634-7_30

Bjerklie, D.M., Lawrence Dingman, S., Vorosmarty, Charles J., Bolster, Carl H., Congalton, Russell G., 2003. Evaluating the potential for measuring river discharge from space. Journal of Hydrology 278 (1–4), 17–38.

Bjerklie, D. M., S. L. Dingman, and C. H. Bolster (2005), Comparison of constitutive flow resistance equations based on the Manning and Chezy equations applied to natural rivers, Water Resour. Res., 41, W11502, doi:10.1029/2004WR003776.

Bouffard, J., Naeije, M., Banks, C.J., Calafat, F.M., Cipollini, P., Snaith, H.M., Webb, E., Hall, A., Mannan, R., Féménias, P. and Parrinello, T., 2018. CryoSat ocean product quality status and future evolution. *Advances in Space Research*, *62*(6), pp.1549-1563.

Bryden, H. L. Kinder, T. H. Steady two-layer exchange through the Strait of Gibraltar, Deep-Sea Research., doi: 10.1016/S0198-0149(12)80020-3. 1991.

Busker T., de Roo A., Gelati E., Schwatke C., Adamovic M., Bisselink B., Pekel J.-F., Cottam A.: A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry. Hydrology and Earth System Sciences, 23(2), 669-690, 10.5194/hess-23-669-2019, 2019

Calafat, F. M., P. Cipollini, J. Bouffard, H. Snaith, P. Féménias. Evaluation of new CryoSat-2 products over the ocean, Remote Sens. Environ., 191, 131-144, 2017.

Cipollini, P. (2011). A new parameter to facilitate screening of coastal altimetry data and corrections. Presented at the 5th Coastal Altimetry Workshop, San Diego, USA available from http://www.coastalt.eu/sites/default/files/sandiegoworkshop11/poster/P08_Cipollini_Castal_Proximity.pdf

Criado-Aldeanueva, F., García-Lafuente, J., Navarro, G., Ruiz, J. Seasonal and interannual variability of the surface circulation in the eastern Gulf of Cadiz (SW Iberia). Journal of Geophysical Research: Oceans, 114(C1). doi: 10.1029/2008JC005069. 2009.

Dinardo S., Fenoglio-Marc L., Buchhaupt C., Becker M., Scharro R., Fernandez J. Benveniste J. (2018). CryoSat-2 performance along the German coasts, AdSR special Issue CryoSat-2, https://doi.org/10.1016/j.asr.2017.12.018

Dinardo S., Fenoglio L., M. Becker; R. Scharroo; M. J. Fernandes; J. Staneva; S. Grayek; J. Benveniste (2020), A RIP-based SAR Retracker and its application in North East Atlantic with Sentinel-3, Advances in Space Research (2020), doi: https://doi.org/10.1016/j.asr.2020.06.004

Dinardo, S., 2020. Techniques and applications for Satellite SAR Altimetry over water, land and ice. PhD Dissertation Thesis. 56, Darmstadt, Germany, Technische Universität Darmstadt, ISBN 978-3-935631-45-7. https://doi.org/10.25534/tuprints-00011343.

Fabry P., Bercher N., Roca M., Martinez B., Fernandes J., Lázaro C., Gustafsson D., Arheimer B., Ambrózio A, Restano M, Benveniste J. (2016). "A step towards the characterization of SAR Mode Altimetry Data over Inland Waters – SHAPE Project". In “New era of altimetry, new challenges”, Ocean Surface Topography Science Team meeting (OSTST), 31 Oct – 4 Nov 2016, La Rochelle, France

Fabry, P. and Bercher, N. (2015). “Characterization of SAR Mode Altimetry over Inland Water”. In Proceedings of the Sentinel-3 for Science Workshop, 2-6 June, Venice, Italy

Fenoglio-Marc, L., Dinardo, S., Scharroo, R., Roland, A., Dutour, M., Lucas, B., Becker, M., Benveniste, J., Weiss, R. (2015): The German Bight: a validation of CryoSat-2 altimeter data in SAR mode, Adv. Space Res., doi: 10.1016/j.asr.2015.02.014

Fenoglio-Marc, L., Dinardo, S., Buchhaupt, C., Scharroo, R., Becker, M., and Benveniste, J. (2019). Calibrating the SAR Sea Surface Heights of CryoSat-2 and Sentinel-3 along the German coasts. In Proceedings of International Association of Geodesy Symposia

Fenoglio L., S. Dinardo, B. Uebbing, C. Buchhaupt, M. Gärtner, J. Staneva, M. Becker, A. Klose, J. Kusche, M. Becker. Investigating improved coastal Sea Level Change from Delay Doppler Altimetry in the North-Eastern Atlantic, Adv. Space Res., under review

Fernandes, M. J., Lázaro, C. (2016). GPD+ Wet Tropospheric Corrections for CryoSat-2 and GFO Altimetry Missions. Remote Sensing, 8(10), 851. doi:10.3390/rs8100851

Fernandes, M. J., Lázaro, C. (2018). Independent assessment of Sentinel-3A wet tropospheric correction over the open and coastal ocean. (2018) Remote Sensing, 10(3), 484. doi:10.3390/rs10030484

Fernandes M. J, Vieira T., (2019). Analysis of the tropospheric corrections present in the Sentinel-3A products, Technical Note 3 (TN-3), UPORTO_ESA_SHAPE_TN03_2019_003, V1.1.

Fernandes, M. J., Lazaro, C., Vieira, T. (2021) On the role of the troposphere in satellite altimetry, Remote Sensing of Environment 252 (2021) 112149, https://authors.elsevier.com/sd/article/S0034-4257(20)30522-8.

Förste, Christoph; Bruinsma, Sean.L.; Abrikosov, Oleg; Lemoine, Jean-Michel; Marty, Jean Charles; Flechtner, Frank; Balmino, G.; Barthelmes, F.; Biancale, R. (2014). EIGEN-6C4 The latest combined global gravity field model including GOCE data up to degree and order 2190 of GFZ Potsdam and GRGS Toulouse. GFZ Data Services. doi: 10.5880/icgem.2015.1

Fukumori, I., Menemenlis, D. Lee, T. A near-uniform basin-wide sea level fluctuation of the Mediterranean Sea. Journal of Physical Oceanography. doi: 10.1175/JPO3016.1. 2007.

García-Lafuente, J., Delgado, J., Criado-Aldeanueva, F., Bruno, M., del Río, J., Vargas, J. M. Water mass circulation on the continental shelf of the Gulf of Cadiz. Deep Sea Research Part II: Topical Studies in Oceanography, 53(11-13), 1182-1197. doi: 10.1016/j.dsr2.2006.04.011. 2006.

Garel, E., Laiz, I., Drago, T., Relvas, P. Characterisation of coastal counter-currents on the inner shelf of the Gulf of Cadiz. Journal of Marine Systems, 155, 19-34. doi: 10.1016/j.jmarsys.2015.11.001. 2016.

Gómez-Enri, J., Aboitiz, A., Tejedor, B., Villares, P. Seasonal and interannual variability in the Gulf of Cadiz: Validation of gridded altimeter products. Estuarine, Coastal and Shelf Science 96, 114-121. doi: 10.1016/j.ecss.2011.10.013. 2012.

Gómez-Enri, J., Escudier, R., Pascual, A., Mañanes, R. Heavy Guadalquivir River discharge detection with satellite altimetry: The case of the Eastern continental shelf of the Gulf of Cadiz (Iberian Peninsula). Advances in Space Research. DOI: 10.1016/j.asr.2014.12.039. 2015.

Gómez-Enri, J., P. Cipollini, M. Passaro, S. Vignudelli, B. Tejedor, J. Coca. Coastal altimetry products in the Strait of Gibraltar. IEEE Transactions on Geoscience and Remote Sensing. 54, 5455-5466. doi: 10.1109/TGRS.2016.2565472. 2016.

Gómez-Enri, J., Vignudelli, S., Cipollini, P., Coca, J., González, C.J. Validation of CryoSat-2 SIRAL sea level data in the eastern continental shelf of the Gulf of Cadiz (Spain). Advances in Space Research. doi: 10.1016/j.asr.2017.10.042. 2018.

Gómez-Enri, J., González, C.J. Passaro, M., Vignudelli, S. Álvarez, O., Cipollini, P., Mañanes, R., Bruno, M., López-Carmona, M.P., Izquierdo, A. Wind-induced cross-strait sea level variability in the Strait of Gibraltar from coastal altimetry and in-situ measurements. Remote Sensing of Environment, 221, 596-608. doi: 10.1016/j.rse.2018.11.042. 2019.

Gupta, Hoshin & Kling, Harald & Yilmaz, Koray & Martinez, Guillermo. (2009). Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling. Journal of Hydrology. 377. 80-91. 10.1016/j.jhydrol.2009.08.003.

Jiang, L., Nielsen, K., Andersen, O. B., Bauer-Gottwein, P. CryoSat-2 radar altimetry for monitoring freshwater resources of China. Remote Sensing of Environment, 200, 125-139. doi: 10.1016/j.rse.2017.08.015. 2017

Jiang, L., Madsen, H., Bauer-Gottwein, P. Simultaneous calibration of multiple hydrodynamic model parameters using satellite altimetry observations of water surface elevation in the Songhua River. Remote Sensing of Environment, 225, 229-247. doi: 10.1016/j.rse.2019.03.014. 2019

Jiang, L., Nielsen, K., Dinardo, S., Andersen, O. B., Bauer-Gottwein, P. Evaluation of Sentinel-3 SRAL SAR altimetry over Chinese rivers. Remote Sensing of Environment, 237, 111546. doi: 10.1016/j.rse.2019.111546. 2020

Kittel, C. M. M., Jiang, L., Tøttrup, C., Bauer-Gottwein P. Sentinel-3 radar altimetry for river monitoring – a catchment-scale evaluation of satellite water surface elevation from Sentinel-3A and Sentinel-3B. Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2020-165, 2020. Preprint under review for HESS

Kosuth P., Blitzkow D., Cochonneau G. (2006). “Establishment of an altimetric reference network over the Amazon basin using satellite radar altimetry (Topex/Poseidon)”, in the proceedings of the "15 years of progress in radar altimetry" Symposium, Venice, Italy.

Kouraev A.V., Zakharova E.A., Samain O., Mognard-Campbell N., Cazenave A. “Ob’ river discharge from TOPEX/Poseidon satellite altimetry data”. Remote Sensing of Environment, 93, 2004, pp. 238-245

Lacombe, H., Richez, C. The regime in the Strait of Gibraltar. In Hydrodynamics of Semi-Enclosed Seas, Jacques C.J. Nihoul (ed.), ISBN: 978-0-444-42077-0, 13-73. 1982.

Laiz, I., Gómez-Enri, J., Tejedor, B., Aboitiz, A., Villares, P. Seasonal sea level variations in the gulf of Cadiz continental shelf from in-situ measurements and satellite altimetry. Continental Shelf Research 53, 77-88. doi: 10.1016/j.csr.2012.12.008. 2013.

Lázaro, C., Fernandes, M. J., Vieira, T., and Vieira, E. (2020). A coastally improved global dataset of wet tropospheric corrections for satellite altimetry, Earth Syst. Sci. Data, 12, 3205–3228, https://doi.org/10.5194/essd-12-3205-2020.

Menemenlis, D. Fukumori, I. Lee, T. Atlantic to Mediterranean Sea Level Difference Driven by Winds near Gibraltar Strait. Journal of Physical Oceanography. doi: 10.1175/JPO3015.1. 2007.

Messager, M.L., Lehner, B., Grill, G., Nedeva, I., Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. Nature Communications, 7: 13603. https://doi.org/10.1038/ncomms13603

Nash J.E., Sutcliffe J.V. (1970). “River flow forecasting through conceptual models part I — A discussion of principles”, Journal of Hydrology, Volume 10, Issue 3, p. 282-290, ISSN 0022-1694, doi:10.1016/0022-1694(70)90255-6.

Passaro, M., P. Cipollini, S. Vignudelli, G. Quartly, and H. Snaith, (2014) “ALES: A multi-mission subwaveform retracker for coastal and open ocean altimetry", Remote Sensing of the Environment, vol. 145, pp. 173-189, 2014.

Passaro, M., Dinardo, S., Quartly, G.D., Snaith, H.N., Benveniste, J., Cipollini, P., Lucas, B. Cross-calibrating ALES Envisat and CryoSat-2 Delay-Doppler: a coastal altimetry study in the Indonesian Seas. Adv. Space Res. 58, 289–303. doi: 10.1016/j.asr.2016.04.011. 2016.

Passaro M., Rose S.K., Andersen O.B., Boergens E., Calafat F.M., Dettmering D., Benveniste J.: ALES+: Adapting a homogenous ocean retracker for satellite altimetry to sea ice leads, coastal and inland waters. Remote Sensing of Environment, 211, 456-471, 10.1016/j.rse.2018.02.074, 2018.

Pawlowicz, R., Beardsley, B., & Lentz, S. Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. Computers & Geosciences, 28(8), 929-937. 2002.

Pekel, JF., Cottam, A., Gorelick, N. et al. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422 (2016). doi:10.1038/nature20584

Peliz, A., Dubert, J., Marchesiello, P., Teles-Machado, A. Surface circulation in the Gulf of Cadiz: Model and mean flow structure. Journal of Geophysical Research-Oceans 112. doi: 10.1029/2007JC004159. 2007

Relvas, P., Barton, E. D. Mesoscale patterns in the Cape Sao Vicente (Iberian peninsula) upwelling region. Journal of Geophysical Research: Oceans, 107(C10), 28-1. doi: 10.1029/2000JC000456. 2002.

Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Borla Tridon, D., Bräutigam, B., Bachmann, M., Schulze, D., Fritz, T., Huber, M., Wessel, B., Krieger, G., Zink, M., and Moreira, A. (2017): Generation and performance assessment of the global TanDEM-X digital elevation model. ISPRS Journal of Photogrammetry and Remote Sensing, Vol 132, pp. 119-139.

Schwatke C., Dettmering D., Bosch W., Seitz F.: DAHITI – an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. Hydrology and Earth System Sciences 19(10): 4345-4364, 10.5194/hess-19-4345-2015, 2015

Schwatke C., Scherer D., Dettmering D.: Automated Extraction of Consistent Time-Variable Water Surfaces of Lakes and Reservoirs Based on Landsat and Sentinel-2. Remote Sensing, 11(9), 1010, 10.3390/rs11091010, 2019

Schwatke C., Dettmering D., Seitz F.: Volume Variations of Small Inland Water Bodies from a Combination of Satellite Altimetry and Optical Imagery. Remote Sensing, 12(10), 1606, 10.3390/rs12101606, 2020

Stevenson, R.E. Huelva Front and Malaga, Spain, Eddy chain as defined by satellite and oceanographic data. Deutsche Hydrographische Zeitschrift 30 (2), 51–53. doi: 10.1007/BF02226082. 1977.

Wessel P, Smith WHF (1996) A global, self-consistent, hierarchical, high-resolution shoreline database. J Geophys Res Solid Earth 101(B4):8741–8743. doi:10.1029/96JB00104

Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., Roth, A. (2018): Acuracy Assessment of the Global TanDEM-X Digital Elevation Model with GPS Data. ISPRS Journal of Photogrammetry and Remote Sensing. Vol. 139, pp. 171-182.

Zakharova EA, IN Krylenko, AV Kouraev, Use of non-polar orbiting satellite radar altimeters of the Jason series for estimation of river input to the Arctic Ocean, Journal of Hydrology, 2019, 568, 322-333.

Zakharova EA., Nielsen K., Kamenev G., Kouraev A., River discharge estimation from radar altimetry: Assessment of satellite performance, river scales and methods, Journal of Hydrology, 2020, 583, 124561.

List of Acronyms

ACE2 Altimeter Corrected Elevations (vers. 2)

AD Applicable Documents

AGC Automatic Gain Control

AH Alti-Hydro

AHP Alti-Hydro Product(s)

AI Action Item

AIM Action Item Management (tool)

AltiKa Altimeter in Ka band and bi-frequency radiometer instrument

AMSR-E Advanced Microwave Scanning Radiometer-Earth Observing System

ANA Agência Nacional de Águas (National Water Agency, Brazil)

AoA Angle of arrival

API Application Programming Interface

AR Acceptance Review

ASAP As Soon As Possible

ASCII American Standard Code for Information Interchange

ATBD Algorithm Technical Basis Document

ATK ALONG-TRACK S.A.S.

AVISO Archivage, Validation et Interprétation des données des Satellites Océanographiques

BfG German Federal Institute of Hydrology

BKG German Federal Agency for Cartography and Geodesy

BSH German Federal Maritime and Hydrographic Agency

BIPR Background Intellectual Property Right

CASH Contribution de l'Altimetrie Spatiale à l'Hydrologie (Contribution of Space Altimetry to Hydrology)

CCN Contract Change Notice

CFI Customer Furnished Item

CLASS NOAA/Comprehensive Large Array-Data Stewardship System

CoG Centre of Gravity

CNES Centre Nationales des Etudes Spatiales

CPP CryoSat-2 Processing Prototype (CNES)

CryoSat-2 Altimetry satellite for the measurement of the polar ice caps and the ice thickness

CRISTAL Copernicus polaR Ice and Snow Topography ALtimeter

CRUCIAL CRyosat-2 sUCcess over Inland wAter and Land

CSV Coma Separated Values

CTOH Centre de Topographie des Océans et de l'Hydrosphère (Centre of Topography of the Oceans and the Hydrosphere)

DAC Dynamic Atmospheric Correction

DAHITI Database for Hydrological Time Series of Inland Waters

DAO Data Access Object

DARD Data Access Requirement Document

DDM Delay-Doppler Map

DDP Delay-Doppler Processor

DEM Digital Elevation Model

DGC Doppler Ground Cell

DPM Detailed Processing Model

DPP Data Procurement Plan

DTC Dry Tropospheric Correction

DTU Danmarks Tekniske Universitet (Technical University of Denmark)

DVT Data Validation Table

ECMWF European Centre for Medium-Range Weather Forecasts

ECSS European Cooperation for Space Standardisation

EGM Earth Gravitational Model

ENVISAT ENVIronment SATellite

EO Earth Observation

EOEP Earth Observation Enveloppe Programme

EOLi Earth Observation Link

EOLi-SA EOLi-Stand Alone

EPN EUREF Permanent Network

ERA Iterim ECMWF ReAnalysis

ESA European Space Agency

EUREF IAG Reference Frame Sub-Commission for Europe

FBR Full Bit Rate

FFT Fast Fourier Transform

FR Final Review

FTP File Transfer Protocol

FCUP (from portuguese) *"Faculdade de Ciências da Universidade"*, Science faculty of the University of Porto

GDAL Geospatial Data Abstraction Library

GDR, [I-,S-] Geophysical Data Record, [Interim-, Scientific-]

GFZ Deutsche GeoForschungsZentrum (German Research Centre for Geosciences)

GIM Global Ionospheric Maps

GLOSS Global Sea Level Observing System

GNSS Global Navigation Satellite System

GOCE Gravity field and steady-state Ocean Circulation Explorer

GPD GNSS-derived Path Delay

G-POD Grid Processing on Demand

GPT2 Global Pressure and Temperature model (vers. 2)

GPP Ground Processing Processor

GPS Global Positioning System

GRACE Gravity Recovery And Climate Experiment

GRDC Global Runoff Data Centre

GRGS Groupe de Recherche de Géodésie Spatiale (Space Geodesy Research Group)

GRLM Global Reservoir and Lake Monitor

GSHHS Global Self-consistent, Hierarchical, High-resolution Shorelines

GTN-L Global Terrestrial Network - Lakes

HDF-EOS Hierarchical Data Format - Earth Observing System

HGT A SRTM file format

HWS High Water Stage

HYCOS Hycos Hydraulics & Control Systems

HYPE Hydrological Predictions for the Environment model

IAG International Association of Geodesy

IDAN Intensity-Driven Adaptive-Neighbourhood

IE Individual Echoes

IGS International GNSS (Global Navigation Satellite Systems) Service

IM Internal Meeting (e.g. not with the client)

IODD Input Output Data Document

IPF Integrated Processing Facility

ISD isardSAT

ITRF International Terrestrial Reference Frame

IRF Impulse Response Function

Jason-1 Altimetry satellite, T/P follow-on

Jason-2 Altimetry satellite, also knwon as the « Ocean Surface Topography Mission » (OSTM), Jason-1 follow-on

Jason-3 Altimetry satellite, Jason-2 follow-on

Jason-CS Jason Continuity of Service

KGE Kling-Gupta efficiency

KML Keyhole Markup Language

KO Kick Off

L1A Level-1A

L1B Level-1B

L1B-S, L1BS Level-1B-S (aka, Stack data)

L2 Level-2

L3 Level-3

L4 Level-4

LAGEOS Laser Geodynamics Satellite

LEGOS (french acr.) Laboratoire d'Études en Géophysique et Océanographie Spatiale (Laboratory for Studies in Geophysics and Spatial Oceanography)

LOTUS Preparing Land and Ocean Take Up from Sentinel-3

LPS Living Planet Symposium

LRM Low Resolution Mode

LSE Least Square Estimator

LWL Lake Water Level

LWS Low Water Stage

MARS Meteorological Archival and Retrieval System

MDL Minimum Description Length

MMSE Minimum Mean Square Error

MNDWI Modification of Normalised Difference Water Index

MoM Minutes of Meeting

MPC Mission Performance Centre

MRC Mekong River Commission

MTR Mid Term Review

MSS Mean Square Slope

MSS Mean Sea Surface

MWR Microwave Radiometer

NAVATT Navigation and Attitude

NDBC US National Data Buoy Center

NDVI Normalised Difference Vegetation Index

NDWI Normalised Difference Water Index

netCDF Network Common Data Form

NOAA National Oceanic and Atmospheric Administration
NR New Requirement (w.r.t. the SoW)
NRT Near Real-Time
NSE Nash-Sutcliffe model efficiency coefficient
NWM Numerical Weather Model
OCOG Offset Centre of Gravity
OPC One per Crossing
OSTM Ocean Surface Topography Mission (also known as Jason-2), is also the name of the satellites series T/P, Jason-1, Jason-2 and Jason-3
OVS Orbit State Vector
PDF Probability Density Function
PEACHI Prototype for Expertise on AltiKa for Coastal, Hydrology and Ice
PEPS Sentinel Product Exploitation Platform (CNES)
PISTACH (french acr.) Prototype Innovant de Système de Traitement pour les Applications Cotières et l'Hydrologie
PLRM Pseudo Low Rate Mode
PMP Project Management Plan
POCCD Processing Options Configuration Control Document
PR Progress Report
PRF Pulse Repetition Frequency
PSD Product Specification Document
PSMSL Permanent Service for Mean Seal Level
PTR Point Target Response
PVP Product Validation Plan
PVR Product Validation Report
PVS Pseudo Virtual Station(s)
RADS Radar Altimeter Database System
RB Requirements Baseline (document)
RCMC Range Cell Migration Curve
RCS Radar Cross Section
RD Reference Document
RDSAR Reduced SAR (also known as Pseudo-LRM)
RF Random Forest
RGB Red, Green, Blue
RID Review Item Discrepancy
RIP Range Integrated Power (of the MLD) sometimes referred as Angular Power Response (APR)
RMS Root Mean Square
rmse root mean square error
ROI (geographical) Region(s) Of Interest
RP Report Period (a month that is being reported into a Progress Report)
RSS Remote Sensing Systems
RWD River Water Discharge
RWL River Water Level
SAMOSA SAR Altimetry MOde Studies and Applications
SARAL In Indian "simple", in english "SAtellite for ARgos and AltiKa.
SARIn SAR Interferometric (CryoSat-2/SIRAL mode)
SAR Synthetic Aperture Radar
SARvatore SAR Versatile Altimetric Toolkit for Ocean Research & Exploitation
SCOOP SAR Altimetry Coastal & Open Ocean Performance
SDP Software Development Plan
SEOM Scientific Exploitation of Operational Missions
SHAPE Sentinel-3 Hydrologic Altimetry PrototypE
SI-MWR Scanning Imaging MWR
SLA Sea Level Anomaly
SLR Sampling Loss Rate
SME Small and Medium-sized Enterprise
SMHI Swedish Meteorological and Hydrological Institute
SNAP SeNtinel Application Platform
SOA State Of the Art
SONEL Système d'Obserbvations du Niveau des Eaux Littorales
SOW Statement Of Work
SPR Software Problem Reporting
SPS Sentinel-3 Surface Topography Mission System Performance Simulator
SRAL SAR Radar Altimeter
SRTM Shuttle Radar Topography Mission
SSB Sea State Bias
SSH Sea Surface Height
SSMI/IS Special Sensor Microwave Imager (SSM/I) Sounder
SSO Single Sign-On
Stack Matrix of stacked Doppler beams
STD Standard Deviation
STDD Standard Deviation of Differences
STM Sentinel-3 Surface Topography Mission

SUM Software User Manual
SWBD SRTM Water Body Data
SWH Significant Wave Height
TAI Temps Atomique International (International Atomic Time)
TBC To Be Confirmed
TBD To Be Done
TCWV Total Column Water Vapour
TDS Test Data Set
TMI Tropical Rainfall Measuring Mission (TRMM) Microwave Imager
TN Technical Note
T/P Topex/Poseidon (altimetry satellite)
TR Technical Risk
UNESCO United Nations Educational, Scientific and Cultural Organization
URL Uniform Resource Locator
USGS United States Geological Survey
USO Ultra Stable Oscillator
USSH Uncorrected Sea Surface Height
UTC Coordinated Universal Time
UWM Updated Water Mask
VS Virtual Station(s)
VH Vertical-Horizontal polarisation
VV Vertical-Vertical polarisation
WBS Work Breakdown Structure
WFR Water Fraction Ratio
WFRWF Water Fraction Ratio - Water content in Footprint
WMO World Meteorological Organization
WP Work Package(s)
w.r.t. with respect to
WTC Wet Tropospheric Correction
XML eXtensible Markup Language
ZP Zero Padding