sea state cci

Product Validation and Algorithm Selection Report (PVASR)

version 1.1, 25 November 2019

Contents

Author	Approved	Signature	Date
Marcello Passaro, Jean Bidlot, Francesco Nencioli, Graham D. Quartly, Björn Tings	Fabrice Ardhuin, Ellis Ash		25/11/2019
ESA Acceptance	Digitally signed by Craig Donlon Date: 2020.02.19 15:58:19 +01'00'		

Issue	Date	Comments
1.0	11/02/2019	First Version
1.1	25/11/2019	Updates following ESA review

List of Acronyms

ADP	Algorithm Development Plan
ATBD	Algorithm Theoretical Basis Document
cci	Climate Change Initiative
DD	Delay-Doppler
DtC	Distance to Coast
E3UB	End-to-End ECV Uncertainty Budget
ECV	Essential Climate Variable
FFT	Fast Fourier Transform
GDR	Geophysical Data Record
GTS	Global Telecommunication System
L4	Level 4
LRM	Low Rate Measurement
LUT	Look-Up Table
MAD	Median Absolute Deviation
MLE	Maximum Likelihood Estimator
NRCS	Normalized Radar Cross-Section
OSTST	Ocean Surface Topography Science Team
PHCP	Percentage of High Correlation
PLRM	Pseudo Low Rate Measurement
PTR	Point Target Response
PVASR	Product Validation and Algorithm Selection Report
RA	Radar Altimeters
RR	Round Robin
R.m.s.	Root mean square
RMSE	Root mean square error
S3A	Sentinel-3A
S3B	Sentinel-3B
SAR	Synthetic Aperture Radar
SI	Scatter Index
SSH	Sea Surface Height
S.D.	Standard Deviation
SWH	Significant Wave Height
WHALES	Wave Height Adaptive Leading Edge Subwaveform (retracker)
w.r.t	with respect to
WV	Wave (mode for SAR)

1. Introduction

This document presents the Product Validation and Algorithm Selection Report (PVASR) for the Sea State Climate Change Initiative (SS_cci), deliverable 2.1 of the project.

The SS_cci project is part of the ESA Climate Change Initiative, and aims to identify, produce and validate a sea state essential climate variable (ECV). Requirements for sea state have been expressed formally by GCOS only for a single variable: the significant wave height (SWH). SWH will therefore be the key parameter on which the evaluation of the algorithms will be based.

In order to identify the best performing algorithm or combination of algorithms, the SS_cci project is holding an open Round Robin (RR): an algorithm intercomparison exercise following the protocol reported in this document, which summarises the agreement found during the first months of the project within the Consortium. By maximising the number of users participating in the Round Robin exercise, ESA expects to identify the best algorithms for a future operational system. The chosen algorithm(s) will then be implemented in an end-to-end system to generate the SS_cci data records.

The SS_cci project aims at providing SWH estimations from two different classes of sensors: Satellite Radar Altimeters (RA) and Synthetic Aperture Radars (SAR). Given the very different nature of these measurements, two different Round Robin exercises are planned. Routines and techniques to evaluate algorithms for RA and SAR will be programmed by different teams and are described completely separated from each other in this document.

Concerning RA, the participation to the Round Robin was publicly advertised during the Ocean Surface Topography Science Team (OSTST) Meeting and the “25 Years of Progress in Satellite Altimetry” Conference in September 2018. Personal invitations were sent by email to the following scientists, authors and co-authors of recent publications focused on the retrieval of SWH from Satellite Altimetry: F.Birol (LEGOS), C. Buchhaupt (TU Darmstadt), L. Fenoglio-Marc (University of Bonn), S. Dinardo (EUMETSAT), W.H. Smith (NOAA), F. Peng and X. Deng (University of Newcastle), R. Roscher (University of Bonn), Xi-Feng Wang (ESST, Kyushu University), N. Kurtz (NASA), D. Sandwell (NOAA). Preliminary expressions of interest in participating were obtained from S. Dinardo (participation to Altimetry DD Round Robin), D. Sandwell (participation to Altimetry LRM Round Robin) and F. Peng (University of Newcastle).

Internal participation to the Altimetry Round Robin is guaranteed by TUM (participating to both LRM and DD Round Robin), CLS (participating to both) and IsardSAT (participating to DD Round Robin).

Internal participation to the SAR Round Robin is guaranteed by ODL, Ifremer and DLR. A personal invitation was sent by email to Xiaoming Li (Chinese Academy of Sciences).

1.1 Overview of document

This document is organised into the following sections:

Section 1: this section introduces the main purpose of the document and sets the general rules of the Round Robin experiment

Section 2: this section describes the input and output dataset for Satellite Altimetry and Synthetic Aperture Radar and the external dataset required to produce the statistics

Section 3: this section describes the procedures on which the Round Robin is based

Section 4: this section presents the results of the Round Robin for each participant. It is left empty in this first version

Section 5: this section describes the decision process followed to select the winning algorithms

Section 6: this section summarises the main findings

1.2 General Principles of the Round Robin for Altimetry

In order to make the Round Robin as transparent as possible, a set of preliminary rules and requirements for participation are defined.

1) The Round Robin exercise is a transparent process. The programming language of the Round Robin is MATLAB. TUM and PML will share the MATLAB code used for the Round Robin.

2) The criteria of the Round Robin must be quantitative.

3) The rules of the Round Robin have been approved collegially by the Consortium.

4) TUM, as leader of the Altimetry-Algorithm Development group, and the Science Leader of the SS CCI Project have the final word in clarifying disputes on methodology.

5) Proposed changes to the Round Robin methodology after the start of the exercise (KO+9) will not be considered.

6) The Round Robin assesses the quality of the Ku-band significant wave height at 20-Hz. The providers are invited to avoid bad practice such as: forcing to absolute zero the SWH, using external data to force retrievals at NaN. The participants shall provide a 1-0 flag to assess bad retrievals and shall describe the criteria used for it. If the algorithm allows, the authors shall provide the estimations of sigma0. Finally, the participants must suggest the best strategy to average from 20-Hz to 1-Hz data, since the final product will be distributed as 1-Hz measurements. The Round Robin will involve both internal consistency checks (outliers, along-track variability…) and validation with external data (buoys and models), as

described in this document.

1.3 Definitions

Here we report the definition of terminology used in the following sections of the document.

Closest point: “closest point” is defined as the median SWH of the 51 high-frequency (“51 20-Hz”) closest points to the buoy, including NaNs. Unrealistic estimations, i.e. outside the interval [-0.25 m, 25 m], are excluded.

Distance to Coast (DtC): “Distance to coast” is the distance of each 20-Hz point from the nearest coast, computed using the “Distance to Nearest Coastline: 0.01-Degree Grid: Ocean” available from http://pacioos.org. In this dataset, “Distances were computed with GMT using its intermediate-resolution coastline and then gridded globally at a spatial resolution of 0.04 degrees. Bilinear interpolation was then applied to increase the spatial resolution to 0.01 degrees.”

Outliers: outliers are considered points for which SWH=NaN, which lie outside [-0.25m, 25m] and/or which are more than 3*MAD away from the median of the closest 20 points.

Noise: noise is defined as the standard deviation of the 20-Hz SWH within a 1-Hz distance.

2. Description of data

2.1 Altimeter (LRM and DD)

2.2.1 Input Data

The Round Robin evaluation will focus on a subset of Jason-3 and Sentinel-3A mission data.

There following table provides the list of input products collected and made available to round-robin participants, for both missions:

Mission	Product	Version
Jason-3	SGDR	version D
Sentinel-3A	SR_1_SRA_BS_	Operational NTC (TBC: unless 2018 reprocessing can be provided by Eumetsat)
	SR_1_SRA_A_	Operational NTC (TBC: unless 2018 reprocessing can be provided by Eumetsat)

For the comparisons with models and in-situ data, participants will need to process 2 full years of data along a small number of selected tracks passing close to the in-situ data. The analysis will cover the period July 2016-June 2018 to cover a balanced mix of winter/summer conditions.

The input data from Jason-3 and Sentinel-3A are being collected by Ifremer repository only, to avoid later discrepancies if different sources are used.

The following tracks for Jason-3 have been selected:

Pass	Area	Potentially collocatable in-situ observations
026	Gulf of Mexico	42039, 42003, 42097, 42057
039	South Pacific - Canadian East Coast - Iceland	32012 44141, 44139, 41046, (42058), (42065), (TFGRV), (TFKGR)

045	New Zealand, Hawaii, Canadian West Coast	(NZban), 51209, 51002, 51100, 51000, 46004, 46208, 46183, (45150), (45141)
063	North Atlantic to Norwegian Sea	41040, (41066), FAWV3, LF3N, LF5I, FL7I, LF3N
065	US East Coast Offshore - Caribbean Sea	44174, (44024), 44008, 44004, 41001, 41002, 42057
070	English Channel - Celtic Sea	62105, 62107, 62023, 62094 (the Porthleven buoy could be added)
094	North Sea to Iceland	62129, 62130, 62134, LF4C, LF5C, 62102, 62161, 64046, FAWV3, FSWV2, TFBLK
113	West Ireland, Norwegian Sea	62095, 62105, 62106, 62048, (64046), 62302, 63118, 64041, 63115, 62169, 63117 63101
120	North Sea	EURO, 62142, 63058, 62144, 62289, 62142, 62148, 62127, 62131, 62165, 62293, 64045
127	Korea	21229, 22105, 22106, 22190, (22104), (22188)
137	Canary Islands to Finland (Portuguese and Finnish data will need to be acquired)	(13131), FARO, 62192, 62025, 62001, (62074), 62288, 62044, 62286, 25077, FIMR3
163	North Sea - Irish Sea - Brazil	62146, 62122, 62119, 62094, 31229, 31231, 31375
172	Western Europe	62095, 62081, 62029, 62163, (62001), 62025, 61417
206	US West coast offshore	46076, 46085, 46184, 46004, (46036), (46006), 46022, 46212, 46255, 46219, (46047)
243	US East Coast - Iceland	41010, 41025, 41062, 44087, 44066, 44097, 44090, 44020, 44018, 44005, (TFBLK), TFKGR
244	Norwegian Sea - Arabian Sea	LF3N, LF5I, LF7I, 23456, 23451, 23494, 23453

The following tracks for Sentinel-3A have been selected:

Pass	Area	Potentially collocatable in-situ observations
011	South Pacific to Gulf of Alaska	46006, 46085
074	Norwegian Sea - Spain	LFB2, (62046), 6201, 62X55, 62X54, 62163, 62083, (62084)
084	Antarctica to Western Canada	46006, 46036, (46208), 46183
099	Hawaii to Bering Sea	51001, 51003, 51101, 46073, (46035)
115	Portugal to Iceland	62085, 62191, 62084, 62082, 62029, 62081, 62095, TFSRT, TFSTD
142	Gulf of Alaska to Hawaii	51206, 51100, 51000, 56246, 46085
181	Southern California offshore	46047, 46219, 46255, 46249, (46053), (46087)
207	Caribbean Sea - Gulf of Mexico	42056, 42003, (42039), 41012
226	West of Canada to Antarctica	46206, (46002), 46059
250	US East coast offshore to Caribbean Sea	44017, 44066, (44087), (41001), 41002
285	Mediterranean to Norwegian Sea	61002, (61431), 64046, FAWV1, FAWV3, FAWV4
319	Caribbean to Canada	41300, 41044, (44150), 44024
355	Hawaii -Bering Sea	51000, 51100, 51004
371	West Scotland to the Mediterranean Sea	62048, 62047, 62091, (62301), Pembrokeshire buoy, 62078, 61281
393	India	23091, 23093, 23455
463	Caribbean to Florida	42058, 41010, 41004
474	Norwegian Sea - South Atlantic	(FAWV2), (FAWV3), 64045, 62105, 62095

478	US East Coast Offshore - South Pacific	44032, 44005, 44018, (44020), 41001, (41002), 32012
491	US East Coast	44032, 44005, 44018, (44020), 41001, (41002), 32012
493	South Pacific	32012
508	Gulf of Mexico - Lake Superior	42055, 42002, (45004)
513	Mediterranean Sea -North Sea	61021, 62142, 63058, 62144, 62148, 63059, 63111, 62302, 63118, 64041
514	New Zealand to Alaska	46075, NzBan
530	Portugal to Norwegian Sea	62200, FARO, 62025, 62001, 62293, 62162,62118, 62155, 62168, 62161, 62130, 62128, LF8F, 63057, LF4H, LF6T, LF5B, LF5E, LF5A, LF7I, LF5I, LF3N, LF5T
592	Labrador Sea to South eastern Pacific	45138, (44005), (44037), 44008, 41001
646	Iceland - Brazil	TFSRT, TFGRS, 31375, 31053
656	Gulf of Alaska to Hawaii	51209, (51003), 51001, 51101, (46066), 46078, 46264, 46080, 46076
695	US West Coast Offshore	46083, 46004, 46036, (46005), 46002, 46059
703	Korea	22189, 22106, 22190, 22105
741	North Sea - Mediterranean Sea	52134, 62116, 62104, 62122, 62119, 62120, 62154, 62123, 61001

2.2.2 Output Data

The output data will have to be given in NetCDF format, with individual files for each cycle and for each track within it. The following fields are necessary to participate in the Round Robin: time, latitude, longitude, Significant Wave Height, Quality Flag. All fields should be provided at a 20-Hz posting rate. Estimations of Backscatter coefficient and Range are encouraged, but not mandatory. The time record is to be exactly as provided on the original waveform product, as this may be used in collocating the coverage of different algorithms.

The algorithms that take part in the Round Robin will need to be fully described in the Algorithm Theoretical Basis Document (ATBD). External participants will also be requested to provide an analogous documentation.

2.2 Synthetic Aperture Radar

2.2.1 Input Data

The Round Robin evaluation will focus on a subset of Sentinel-1 Wave Mode data. For the comparisons with models and in-situ data, participants will need to process 1 full years of data. The analysis will cover the period of January 2019 to December 2019 to provide a balanced mix of winter/summer conditions. The Round Robin is planned for spring 2020.

The following list contains 60 relative orbits numbers with regular coverage of NDBC buoys within ~50km collocation distance:

Orbits Pacific:

	relative orbit number	**type**	**Sentinel-1 A/B**		**collocated NDBC buoy, S1A/S1B**	**wv1 wv2**	**collocation, km (A/B)**
1	001	descending	-	B	46070 B	2	57
2	005	ascending	A	B	46005 A/B	1	52/9
3	006	ascending	-	B	46066 B	2	50
4	008	ascending	-	B	51209 B	2	29
5	014	descending	A	B	51004 A 46085 B	1 1	34 47
6	020	ascending	A	B	46085 A 46006 B	2 2	20 28
7	028	descending	A	B	46005 A/B	1	21/23
8	029	descending	-	B	46078 B 46066 B	1 2	22 49
9	030	descending	-	B	46071 B	2	4
10	035	ascending	A	B	46078 A/B	1	26/41
11	043	descending	A	B	46184 A/B	1	36/29
12	049	ascending	A	B	46004 A	2	6/47
13	050	ascending	-	B	46075 B	2	36

14	051	ascending	A	B	46075 A/B	2	3/3
15	058	descending	-	B	46001 B	1	20
16	059	descending	A	B	46073 B 51209 A	1 1	47 52
17	064	ascending	-	B	46001 B 46080 B	2 2	27 57
18	066	ascending	A	B	52201 A/B	1	54/54
19	067	descending	A	-	52211 A	2	38
20	078	ascending	A	B	46005 A/B 46036 A/B 46208 B 43010 B 46002 A/B	2 1 2 1 2	13/14 0/52 30 52 35/39
21	087	descending	A	-	46246 A	1	45
22	088	descending		B	46072 B	2	44
23	093	ascending	A	B	46001 A 46246 B	1 1	28 30
24	101	descending	A	B	46004 A/B	2	39/39
25	102	descending	-	B	46006 B	1	20
26	103	descending	A	B	46071 A/B	1	37/37
27	110	ascending	A	B	51209 A/B	1	5/46
28	112	descending	-	B	41040 B	2	57
29	116	descending	-	B	46085 B	2	9
30	122	ascending	A	B	46184 A/B 46006 B 46085 B	2 1 1	38/23 56 19
31	131	descending	-	B	51208 B 46078 B	1 2	35 25
32	132	descending	-	B	52201 B	2	16
33	137	ascending	-	B	46078 B	1	57

34	138	ascending	-	B	46035 B	1	40
35	139	ascending	A	B	52201 A/B	2	17/24
	140	ascending	-	B	52211 B	2	57
36	147	descending	-	B	46070 B	1	46
37	151	ascending	-	B	46036 B 46205 B	2 2	40 51
38	153	ascending	-	B	46070 B	2	18
39	160	descending	-	B	46001 B	1	33
40	161	descending	A	B	46072 B 51209 A/B	1 2	37 50/25
41	174	descending	A	-	46036 A	1	21

Orbits Atlantic:

	relative orbit number	**type**	**Sentinel-1 A/B**		**collocated NDBC buoy, S1A/S1B**	**wv1 wv2**	**collocation, km (A/B)**
42	002	ascending	-	B	44139 B	1	40
43	018	ascending	-	B	44011 B 44150 B	1 2	35 43
44	025	descending	A	B	41049 A 41052 B 41049 B 41056 B	2 2 2 2	27 35 57 20
45	031	ascending	A	B	41041 A/B	1	39/39
46	033	ascending	-	B	44095 B 41001 B 41047 B 44099 B 41025 B	2 1 1 2 2	32 39 32 33 54
47	040	descending	-	B	41047 B	1	36
48	047	ascending	-	B	44137 B	1	27

49	054	descending	-	B	41044 B	1	30
50	062	ascending	-	B	44008 B 41048 B 42085 B	2 1 2	49 29 23 (near land)
51	068	descending	-	B	41041 B	2	44
52	076	ascending	-	B	44139 B	2	40
53	091	ascending	A	B	44011 A/B 41049 A/B	2 2	41/- 42/21
54	104	ascending	A	B	41041 A/B	2	47/39
55	113	descending	-	B	41002 B 41001 B 44066 B 44005 B	2 1 2 1	50 28 35 45
56	120	ascending	-	B	44137 B	2	22
57	135	ascending	-	B	41046 B 41047 B 41001 B 41115 B	2 1 1 2	0 46 50 10 (near land)
58	148	ascending	A	-	3100053 A	2	40 *EMONDNET
59	156	descending	A	B	41044 A/B	2	30/30
60	164	ascending	-	B	41052 B	1	45

As reference data two separate datasets are used:
1. In-situ measurements (ground truth) from buoys
2. Model measurements (ground truth) from MFWAM distributed by CMEMS

For some orbits buoy collocation from both Sentinel‑1 platforms A and B are possible. This means from the list of 60 relative orbit numbers 87 WV products are collocatable with buoys. For buoy validation about 1.100 vignettes for each wv1 and wv2 are expected, resulting in ~2.200 data samples in the buoy validation dataset. For model validation only vignettes within -60°<latitude<60° are considered to avoid ice coverage. For all 87 WV products about 100.000 vignettes for each wv1 and wv2 are expected, resulting in ~200.000 data samples in the model validation dataset. Each reference data sample must be temporally interpolated to the image acquisition time between the two neighboring measurements not further apart than six hours.

2.2.2 Output Data

The output data will have to be given in NetCDF format, with individual files for each orbit. The following fields are necessary to participate in the Round Robin: time, latitude, longitude, Significant Wave Height, Quality Flag. Estimations of Significant Wave Length and Significant Wave Direction are encouraged, but not mandatory. The time record is to be exactly as provided on the original waveform product, as this may be used in collocating the coverage of different algorithms. The algorithms that take part in the Round Robin will need to be fully described in the Algorithm Theoretical Basis Document (ATBD). External participants will also be requested to provide an analogous documentation.

2.3 Round Robin datasets

2.3.1 Altimeter (LRM and DD)

In-situ dataset

As part if its wave forecast verification activities within the Joint Commission for Oceanography and Marine Meteorology, ECMWF gathers wave observations. Most of the data are obtained from the data received via the Global Telecommunication System (GTS), where atmospheric and oceanographic data are disseminated to weather forecasting institutions. A few data sets are also supplied directly to ECMWF. Most data are from a wide range of moored buoys, except for data from the North Sea, the Norwegian Sea and the Gulf for Mexico. A basic quality control procedure is performed following Bidlot et al. (2002) and the resulting hourly time series have been available to the project.

Wave Model dataset

A long global wave model hindcast has been produced using the latest version of ECMWF wave model (CY46R1, ECMWF 2019). The model spatial resolution is 0.125x0.125 degrees and the output is hourly, making it ideal to be collocated with altimeter data passes. The necessary hourly wind forcing and sea ice cover information come from ECMWF latest reanalysis (ERA5). The hindcast covers the period from 1979 to present. The data is preferred to using directly ERA5 output instead, because, ERA5 wave model data are on a coarser resolution (0.36 degree), it is based on a older version of the wave model code and most of all, ERA5 has used altimeter data from (ERS-1, ERS-2, ENVISAT, Jason-2, Jason-2), rendering it not independent. The latest hindcast has been found to be even better than ERA5 (Jean Bidlot personal communication). Over the course of 2019, it will become a product made available alongside ERA5 data.

2.3.2 Synthetic Aperture Radar

In-situ measurements (ground truth) from buoys

The 50 km collocation distance defines the distance from the nearest border or corner of a 20km×20km S1-WV vignette to the buoy location, this means ~65 km distance from vignette center. Due to fact that the positions of the buoys are slightly changing in the time (will be checked in 2020), and also because S1 vignettes can be slightly shifted, an actual distance of about 55 km was considered.

For in-situ measurements 54 buoys providing wave heights were selected:
- 49 buoys from NOAA buoy stations network NDBC (National Data Buoy Center)
- 8 buoys from ECCC (Environment and Climate Change Canada)
- 1 buoy from EMODnet

The stationary NOAA buoys typically provide hourly measurements and the ECCC buoys store the data more unsystematically. For validation the data of each buoy are temporally interpolated. The data with a measurement time gap over 6h were excluded.

The following table lists all buoys with their corresponding geographic coordinate:

Name	Lat	Lon	Data source	Group
43010	10.051	-125.032	NDBC	Group-1 Alaska, Canada
46001	56.304	-147.920	NDBC	
46002	42.612	-130.537	NDBC	
46004	50.930	-136.100	ECCC	
46005	46.140	-131.070	NDBC	
46006	40.782	-137.397	NDBC	
46035	57.026	-177.738	NDBC	
46036	48.350	-133.940	ECCC	
46066	52.785	-155.047	NDBC	
46070	55.082	175.153	NDBC	
46071	51.125	179.012	NDBC	
46072	51.672	-172.088	NDBC	
46073	55.031	-172.001	NDBC	
46075	53.983	-162.041	NDBC	
46078	55.556	-152.582	NDBC	
46080	57.947	-150.042	NDBC	
46085	55.868	-142.492	NDBC	
46184	53.910	-138.850	ECCC	
46205	54.190	-134.320	ECCC	
46208	52.520	-132.690	ECCC	

46246	50.033	-145.200	NDBC	
51209	-14.264	170.493	NDBC	Group-2 South-West Pacific
52201	7.083	171.392	NDBC	
52211	15.268	145.662	NDBC	
51001	24.453	-162.000	NDBC	Group-3 Pacific Hawaii
51002	17.037	-157.696	NDBC	
51004	17.604	-152.364	NDBC	
51101	24.361	-162.075	NDBC	
51205	21.018	-156.425	NDBC	
51202	17.037	-157.696	NDBC	
51208	22.285	-159.574	NDBC	
41001	34.502	-72.522	NDBC	Group-3 North Atlantic, Sargasso Sea
41002	31.892	-74.930	NDBC	
41025	35.006	-75.402	NDBC	
41040	14.554	-53.045	NDBC	
41041	14.441	-46.033	NDBC	
41043	21.124	-64.830	NDBC	
41044	21.582	-58.630	NDBC	
41046	23.822	-68.384	NDBC	
41047	27.514	-71.494	NDBC	
41048	31.838	-69.585	NDBC	
41049	27.490	-62.938	NDBC	
44005	43.201	-69.128	NDBC	
44008	40.504	-69.248	NDBC	
44011	41.070	-66.588	NDBC	
44014	36.606	-74.840	NDBC	
44066	39.618	-72.644	NDBC	

44095	35.750	-75.330	NDBC	
44099	36.914	-75.720	NDBC	
44137	42.260	-62.000	ECCC	
44139	44.240	-57.100	ECCC	
44150	42.500	-64.020	ECCC	
41052	18.249	-64.763	NDBC	Group-4 North Atlantic, Caribbean Sea
41056	18.261	-65.464	NDBC	
41115	18.376	-67.280	NDBC	
42085	17.869	-66.532	NDBC	
3100053	-23.478	-43.984	EMONDNET	p

Model measurements (ground truth) from MFWAM distributed by CMEMS

For model "ground truth" MFWAM model data from CMEMS (COPERNICUS) with spatial resolution of 1/12 degree will be used. The results are stored in 3h time intervals and will be temporally interpolated.

http://marine.copernicus.eu/services-portfolio/access-to-products/?option=com_csw&view=details&product_id=GLOBAL_ANALYSIS_FORECAST_WAV_001_027

3. Round Robin procedures

3.1 Altimeter (LRM and DD)

Here we report the statistics chosen to evaluate the Round Robin. The evaluation of the Round Robin will be based on Tables reporting the key statistics for each of the participants. The statistics (unless differently specified) will have to be reported separately for each of the following categories:

GEOGRAPHICAL CATEGORIES

- Full Dataset: all data without any distinction
- Coastal Data: all data in which DtC < 20 Km, < 10 Km, < 5 Km
- Open Ocean Data: all data in which DtC > 20 Km

SEA STATE CATEGORIES

- Low Sea States: all data in which 0m<SWH<1m
- Average Sea States: all data in which 1.5m<SWH<2.5m
- High Sea States: all data in which SWH>6m
- Very High Sea States: all data in which SWH>12 m

3.1.1 Outlier Analysis

For the definition of Outliers please see the previous sections. The percentage of Outliers will be defined for each of the listed categories.

3.1.2 High Frequency Noise Analysis

The noise will be computed as the standard deviation of the 20-Hz SWH estimations within a 1-Hz block. For each category, the statistics will be obtained by computing the median of all the noise estimations.

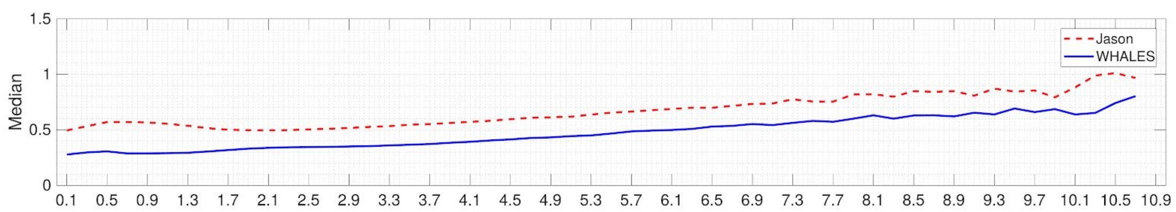

Figure 1: Visual example of noise analysis for varying SWH performed on a test area with Jason-3 data and an algorithm under development for the SS_cci project

3.1.3 Comparison with in-situ Data

Data from wave buoys at the time of each satellite passage will be compared to the satellite observations at the closest point, as defined in section 1.3. Statistics will be separated according to the geographical and sea state categories defined in section 3.1. The following statistics will be provided :

- Percentage of Cycles for High Correlation (PHCP). is a statistic designed to take into account both the correlation (Pearson correlation coefficient) between time series and the number of observations available. The test is performed in an iterative way: First of all, for the selected location, the correlation of the buoy time series with the entire set of altimetry retrievals is checked; if the correlation coefficient is lower than 0.9, then the cycle with the maximum discrepancy (quantified as the maximum of the absolute value of the difference) between the buoy value and altimeter retrieval is excluded. This exclusion is iterated until the correlation rises above 0.9, at which point the percentage of cycles left provides the measure of the general quality of the retracked altimetry values.
- Standard Deviation between SWH from altimetry at closest point and buoy.
- Slope of the linear fit (regression line of SWH altimetry -vs- SWH buoy scatter plot)
- Median bias between SWH from altimetry at closest point and buoy

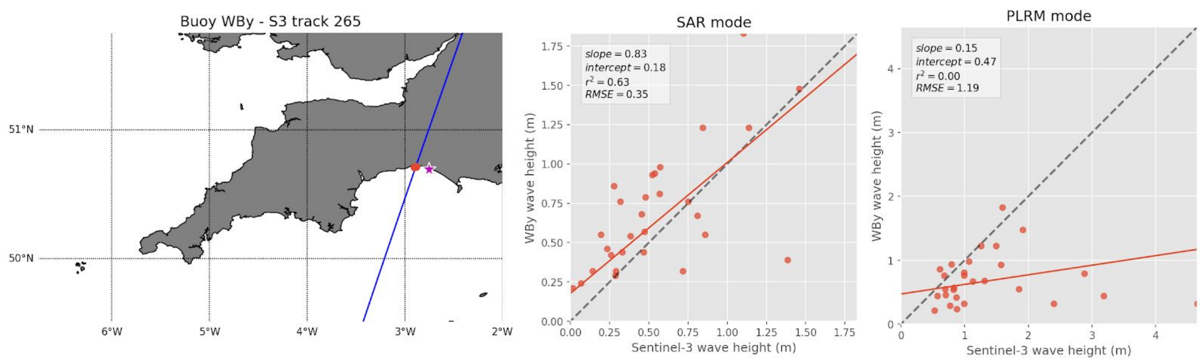

Figure 2 : Example of correlation plot between time series of buoy observations at West Bay, UK (magenta star on left map) and Sentinel-3A measurements closest to the buoy position along track 265 (red dot). Center and right panels show linear correlation curve and correlation parameters for measurements collected in SAR and PLRM mode, respectively.

3.1.4 Comparison with Model Data

Model grid points and altimetry will be coupled by considering the median of the SWH 20-Hz measurements from altimetry within the grid point. The Statistics will be divided according to the categories previously defined. The following statistics will be provided:
- Correlation
- Standard Deviation of the difference between SWH from altimetry and SWH from model
- Slope of the linear fit (regression line of SWH from altimetry -vs- SWH from model scatter plot)
- Median bias between SWH from altimetry and SWH from model.

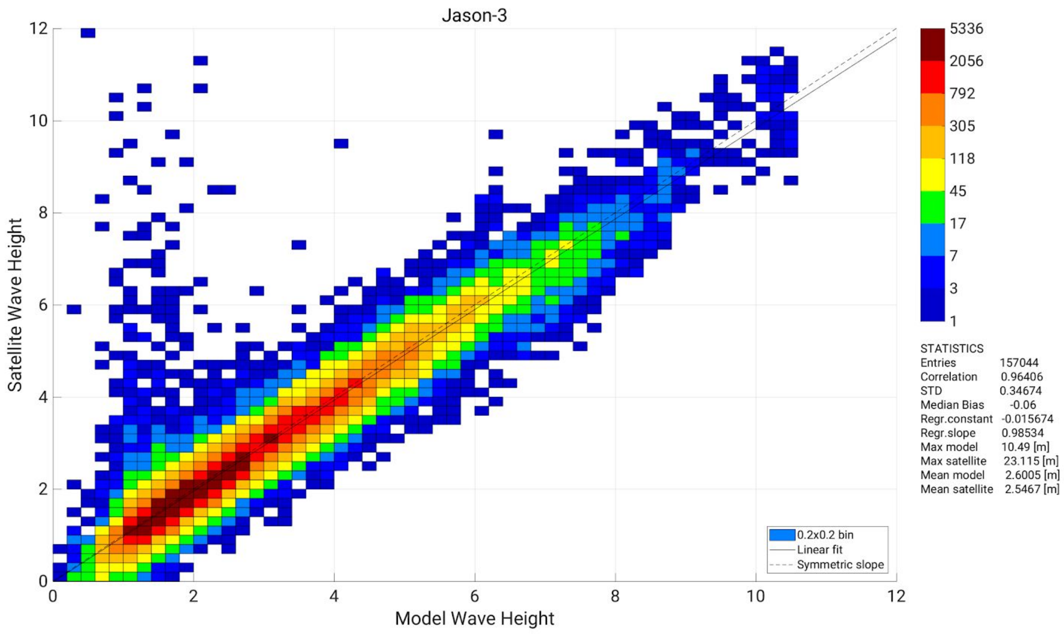

Figure 3: A density plot in a test area illustrating key statistics in the comparison between Jason-3 data and the Météo-France global wave model (MFWAM) available from the Copernicus Marine Service.

3.1.5 Representation of Scales of Variability

Spectra will be readily applied to SWH in the same way they are used for Sea Surface Height (SSH) to compare the performance of different retrackers. We will use FFT applied to the 20 Hz SWH data, using segments of at least 1024 points (~330 km), with Hamming weighting applied. Provided the retracker algorithms treat each waveform independently, and there is no along-track smoothing applied, the spectra at very short scales should be white noise, with a plateau level consistent with the value for "high-frequency noise" (section 3.1.2). At large scales, the spectra should be dominated by the real geographical variations in SWH, and providing all estimators are unbiased their spectra will converge.

The main assessment considered in this section is the nature of the spectra in the ranges 25-50 km and 50-100km, as these are the scales requested by GCOS, and are also the scales of models that might assimilate altimeter SWH data. Spectra will be assessed at these scales for all retrackers, with the results partitioned according to ocean basin and also by median wave height, in order to ascertain whether there are problems associated with differently shaped waveforms (i.e. leading edge slope).

3.2 Synthetic Aperture Radar

The statistics chosen to evaluate the SAR Round Robin are similar to the Altimetry Round Robin. As the amount of data samples with sea state above 12m is expected low, in contrast to the Altimeter Round Robin high and very high sea state is comprised into one category. Further, a category for rough sea state has been added. This means the statistics for the SAR Round Robin are categorized as follows:

GEOGRAPHICAL CATEGORIES

- Full Dataset: all data without any distinction, as the amount of ground-truth data samples compared to altimetry is less due to non continuous acquisitions.

SEA STATE CATEGORIES

- Low Sea States: all data in which 0m<SWH<1.5m
- Average Sea States: all data in which 1.5m<SWH<3m
- Rough Sea State: all data in which 3m<SWH<6m
- High and Very High Sea States: all data in which SWH>6m

- Wind sea : all data in which 0s<wave period<12s
- Swell: all data in which 12s<wave period<20s
- fore-runners: all data in which wave period>20s

3.2.1 Outlier Analysis

For the definition of Outliers please see the previous sections 1.3. The percentage of Outliers will be defined for each of the listed categories.

3.2.3 Comparison with in-situ Data

The collocation of in-situ buoy data with SAR measurements is only valid, when the distance between the respective buoys and and the SAR subscene are at most 100km apart from each other. The in-situ buoy data and SAR will be coupled by comparing the median of the SWH and integral spectral partition parameters measurements from SAR with the buoy's SWH. The Statistics will be divided according to the categories previously defined. The following statistics will be provided:

- Root Mean Square Error (RMSE)
- Scatter Index (SI)
- Median bias between SWH and integrated spectral parameters from SAR and from buoy data.

3.2.4 Comparison with Model Data

The same MFWAM model distributed by CMEMS used for altimetry dataset, will be used for SAR comparison.
Model grid points and SAR will be coupled by considering the median of the SWH and integral spectral partition parameter measurements from SAR within the grid point. The Statistics will be divided according to the categories previously defined. The following statistics will be provided:

- Root Mean Square Error (RMSE)
- Scatter Index (SI)
- Median bias between SWH and integrated spectral parameters from SAR and from model.

3.2.5 Optional comparison with CFOSAT altimeter data

Since CFOSAT is in a dust-down sun synchronous orbit, such as Sentinel1A/B, good time and space collocation exist (within about 20 minutes and 25km) between some SAR wave mode vignettes and CFOSAT nadir beam that has shown to have excellent quality, as compared to Jason 3 or Altika.
CFOSAT swh data from 2019 will be used as an additional optional dataset for SAR Round Robin validation, in case the statistics based on model and buoy data are not expressive enough. The collocating orbits with Sentinel-1 WV vignettes will be retrieved only then.

4. Round Robin results

4.1 Altimeter (LRM and DD)

This section is intentionally left blank and will be updated in the next version.

4.2 Synthetic Aperture Radar

This section is intentionally left blank and will be updated in the next version.

5. Weighting matrix for results

5.1 Altimeter

The Consortium has discussed offline and during the progress meetings on the possibility of defining an univocal weighting matrix to assess the results. The conclusion was that this is impossible at this stage of the project, given that the weighting would be highly dependent on the application of choice. For example, the weighting applied to the coastal statistics by a coastal modeller that want to validate the simulated wave data would be different from the one applied by an oceanographer that want to study open ocean storms. But while one algorithm could offer the best performances in a coastal environment, the same could perform worse than others in the detection of extreme wave heights.

It was decided that the Altimeter Algorithm Development Team and in particular PML and TUM, which are responsible for the Round Robin exercise, will bring the results of the Round Robin without applying any weighting matrix. Considering these, the Consortium will take a decision in collegiality during a progress meeting. The Consortium also reserves the possibility to propose more algorithms for production, in order to allow a wider Validation.

The decision can involve considerations based on the outcome of the User Requirement Document in order to identify the most important applications for the future users of the SS_cci dataset.

5.2 Synthetic Aperture Radar

The decision process for all algorithms selected for production will be identically held on a collegiality basis, as defined in the previous subsection for the Altimeter Round Robin.

6. Summary

The first version of the PVASR aimed at presenting the plan for the Round Robin exercise that will determine the algorithm(s) to be used in the generation of the official SS_cci product. The Round Robin exercise will be separated according to the sensors used to produce the SWH estimations: Synthetic Aperture Radars and Radar Altimeters. In the latter, a further differentiation will be made between LRM and DD Altimeters.

Key statistics that will be computed on the test datasets have been described in these documents and will focus on internal evaluation and comparison with buoys and model data. The results will be presented in the next version of this documents in the form of tables. Multiple statistics will be produced for sub-divisions of the dataset according to their geographical location (open ocean vs coastal ocean) and intensity of sea states (low, average, high and very SWHs).

Finally, this document contains the rules of participation for each internal and external participant to the Round Robin exercises.

7. References

Bidlot J.-R., D. J. Holmes, P. A. Wittmann, R. Lalbeharry, H. S. Chen, 2002: Intercomparison of the performance of operational ocean wave forecasting systems with buoy data. Wea. Forecasting. 17. 287-310.