Multiparametric Sea State Fields from Synthetic Aperture Radar

Andrey Pleskachevsky[a], Björn Tings[a], Sven Jacobsen[a], Egbert Schwarz[b], Detmar Krause[b], Sergey Voinov[b]

[a] DLR, Earth Observation Center, Maritime Safety and Security Lab Bremen, Am Fallturm 9, 28359 Bremen, Germany
Andrey.Pleskachevsky@dlr.de
[b] DLR, Earth Observation Center, Maritime Safety and Security Lab Neustrelitz, Kalkhorstweg 53, 17235 Neustrelitz

Abstract

The study presents a method embedded in a software for estimating series of integrated sea state parameters from satellite-borne synthetic aperture radar (SAR), which allows processing of data from different satellites and modes in near real time (NRT). The developed Sea State Processor (SSP) estimates total significant wave height H_s, dominant and secondary swell and windsea wave heights, first, and second moment wave periods, mean wave period and period of wind sea. The algorithm was applied for the Sentinel-1 (S1) C-band Interferometric Wide Swath Mode (IW), Extra Wide (EW) and Wave Mode (WM) Level-1 (L1) products and also extended to the X-band TerraSAR-X (TS-X) StripMap (SM) mode. The wide scenes are processed in raster format and resulting in continuous sea state fields. For the S1 WV 20 km × 20 km imagettes averaged values for each sea state parameter are provided. Validated with worldwide data the reached RMSE for H_s are 0.25 m for S1 WV, ~35 cm for TS-X SM and ~60 cm for the coarser S1 IW and S1 EW.

1 Introduction

In recent years, the ongoing development of space-borne SARs together with corresponding data transfer and data processing infrastructures has made a series of oceanographic applications possible in near real time (NRT), e.g. [1], [2], [3]. Several minutes after image acquisition, the acquired scene can be processed and derived meteo-marine parameters are transferred to weather services that can use these products for forecast validations and to support marine traffic. In turn, sea state or sea ice information, for instance, can also be sent directly to a ship's bridge in order to optimise the ship routing. Oceanographic products, such as significant wave height fields and surface wind speed fields, ice coverage maps, oil spill locations, etc., can be processed in parallel from the same satellite image or from acquisitions from different satellites and combined to support Maritime Situational Awareness (MSA) by fusing data from various sources based on remote sensing, *in situ* measurements, forecast modelling or communication systems. One essential data source for MSA is high-resolution weather and sea state spatial information, which can be estimated only from remote sensing data. Further, the fusion of oceanographic parameters not only generates direct additional value for MSA, but can also be applied for forecast model assimilation and to gain insights into the underlying physics.

2 Subject and linear regression method

This work introduces an algorithm and processor for meteo-marine parameter estimation for near real time (NRT) applications. The main goal of the study was to improve the accuracy of existing algorithms [4], [6], expand the list of obtained sea state parameters and also apply the method to various satellites and modes.

In this study, the empirical approach was applied, as it is most suitable for direct estimation of sea state parameters form SAR features. The linear regression method delivers a first guess H_s, which is then supplemented using machine learning to increase accuracy of H_s estimation further. Correspondingly, the developed sea state processor (SSP) consists of two parts: CWAVE_EX (extended CWAVE) based on widely known approach using linear regression [4] and additional machine learning postprocessing.

The advantage of the empirical CWAVE approach is that an analytical solution exists. The function's coefficients can be obtained quickly using singular value matrix decomposition (SVM), an extensive machine learning training is not necessarily. As will be seen further in this work, although the empirical solution is inferior in accuracy to that obtained by adding machine learning, this solution is already stable at about 80,000 random samples (the distribution of wave heights in the worldwide tuning data is near to the normal distribution with ~70% of H_s<3 m, see percentage in **Table 2**): additional tuning data does not distinctly affect the obtained results from the point of view of total H_s.

In order to improve accuracy of the original CWAVE [4][6], series of modifications were considered and introduced:

- additional data preparation steps i.e., image artifact filtering, SAR-feature smoothing, etc.
- new SAR features [8] additionally to features used in [10].

In detail, five types of primary SAR-features are involved:

Type-1: Normalized radar cross section (NRCS) and NRCS statistics (variance, skewness, kurtosis, etc.).

Type-2: Geophysical parameters (wind speed using CMOD/XMOD algorithms).

Type-3: Grey Level Co-occurrence Matrix (GLCM) parameters (entropy, correlation, homogeneity, contrast, dissimilarity, energy, etc.).
Type-4: Spectral parameters based on image spectrum (ISP) integration of different wavelength domains (0-30 m, 30-100 m, 100-400 m, etc.) and spectral width parameters (Longuet-Higgins, Goda-parameter).
Type-5: Parameters defined in [4] based on products of ISP with orthonormal functions and cut-off wavelength estimated using autocorrelation function (ACF).

Secondary features are combinations of primary features.

The algorithm was tuned and validated using two independent global wave model hindcasts:

- WaveWatch3 (WW3, NOAA) [8] with a resolution of 0.5 degrees, spatially interpolated on 0.25 degrees (ca. 20–25 km grid cells, which corresponds to the S1 WV imagette size, data available for the entire S1 mission).
- CMEMS (Copernicus) [9] with a spatial resolution of 1/12 degrees (data available from April 2016 onwards).

The 3 h outputs of both models were interpolated temporally.

As *in situ* the National Data Buoy Center (NDBC) buoy measurements were used. The records (mainly 1 h interval) are also interpolated temporally. Data with a measurement time gap of over 6 h were excluded.

The cross validations carried out using CMEMS and WW3 shows: in comparison to with NDBC data, using only CMEMS ground truth resulted into an accuracy ~3 cm better than when the model function was tuned using WW3 data. The H_s comparison between CMEMS, WW3 and NDBC resulted into an RMSE=0.26 cm for CMEMS/NDBC and an RMSE=0.23 cm for CMEMS/WW3 at NDBC buoy locations. Generally, in terms of H_s, the ground truth noise can be assessed to an error of ~0.25 m.

The reached root mean squared errors (RMSE) for CWAVE_EX for the total H_s are 0.35 m for S1 WV and TS-X SM (pixel spacing ca. 1–4 m) and 0.60 m for low-resolution modes S1 IW (10 m pixel spacing) and S1 EW (40 m pixel) in comparison to CMEMS.

Additional to H_s, integrated sea state parameters were considered: dominant and secondary swell and windsea wave heights ($H_s^{swel\text{-}1}$, $H_s^{swel\text{-}2}$, H_s^{wind}), first and second moment wave periods, mean wave period and period of wind sea (T_{m0}, T_{m1}, T_{m2}, T^{wind}). The accuracies of the four studied periods are in the range of 0.45–0.90 s for all considered satellites and modes. Similarly, the dominant and secondary swell and wind sea wave height RMSEs are in the range of 0.35–0.80 m compared to CMEMS wave spectrum partitions. More details are given in **Table 1** and in [7].

3 Machine learning postprocessing

The postprocessing step uses machine learning. The support vector machine technique (SVM) [5] was applied. As input, the primary SAR-features estimated directly form the image (no features combinations) are used [11], extended by three additional features:

- first-guess H_s from linear regression solution.
- precise incidence angle (degree, third decimal place)
- flag identifying the satellite (0 for S1-A and 1 for S1-B).

For the training of hyperparameters, around 0.5 Mio model collocations were applied. The entire SVM-model training used ~1.5 Mio samples and took around 3 weeks. For the validation, data from the entire S1 WV archive Dec. 2014 - Feb. 2021 (around 15 Mio samples) was used. The resulting accuracy of H_s reaches an RMSE=0.25 m by SVM postprocessing of S1 WV. Comparison to 61 NDBC buoys, collocated at distances shorter than 50 km to S1 WV worldwide imagettes, result into an RMSE=0.41 m.

Table 1 RMSE of sea state integrated parameters using linear regression CWAVE_EX approach.

Parameter	Unit	Satellite mode			
		S1 IW	S1 EW	S1 WV (wv1/wv2)	TS-X SM/SL
H_s	m	0.57	0.61	0.34 / 0.38	0.36
T_{m0}	s	0.96	0.86	0.46 / 0.51	0.72
T_{m1}	s	0.97	0.85	0.51 / 0.56	0.59
T_{m2}	s	0.82	0.86	0.46 / 0.51	0.51
$H_s^{swell\text{-}1}$	m	0.68	0.63	0.42 / 0.47	0.33
$H_s^{swel\text{-}2}$	m	0.38	0.44	0.41 / 0.46	0.27
H_s^{wind}	m	0.77	0.66	0.40 / 0.46	0.37
T^{wind}	s	0.97	0.95	0.62 / 0.67	0.71

4 Sea state processor (SSP)

The Sea State Processor (SSP) is a C++ software package integrating series of operations and functions, e.g.:

- SAR image reading and calibration
- Land-masking and pre-filtering NRCS artefacts
- SAR feature estimation, feature filtering
- Empirical model functions for different sea state parameters parameters for different satellites and modes
- filtering and control of results
- writing of results

SSP was designed in a modular architecture for S1 IW, EW, WV and TS-X SM/SL modes. The DLR Ground Station “Neustrelitz” applies the SSP as part of a near real-time demonstrator service that involves a fully automated daily provision of surface wind and sea state parameters estimated from S1 IW images of the North and Baltic Sea. Due to implemented parallelization, a fine raster can be practical set: for example, S1 IW image with large coverage of 200 km × 250 km can be processed using a raster with 1 km grid cell (~50000 subscenes) during minutes. The maritime environment and ship detection products are combined in layers (see **Figure 1**). The applied series of technical innovations improved not only the total RMSE, but are also seen in the local effects: e.g. the antenna beam pattern that impacted the results in the old SSP version [7] is not more present in the new sea state products (see **Figure 2**). This pattern resulted from NRCS gradients at the beam borders in S1 IW and EW images. Also, the NRCS “black gaps” (NRSC≈0), a frequent error at image borders, does not affect the processed parameters anymore. **Figure 3** shows an example for additional parameters for the SAR image presented in **Figure 2**: swell wave height $H_s^{swell-1}$, wind-sea wave height H_s^{wind} and periods T_{m0}, T_{m1}, T_{m2}, T^{wind}.

5 Processing Sentinel-1 Wave Mode archive

Using the SSP, the complete archive of S1 WV from December 2014 until February 2021 with around 60 overflights/day, each including around 120 imagettes, was processed. The validation using the CMEMS model within latitudes of -60°<LAT<60° to avoid ice coverage resulted in an RMSE of 0.245/0.273 m for wv1/wv2 imagettes, respectively. Comparisons to 61 NDBC buoys, collocated at distances shorter than 50 km to S1 WV worldwide imagettes, result into an RMSE of 0.41 m.

The monthly estimated total RMSE (compared with CMEMS) varies form 0.22 m to 0.31 m. These RMSE fluctuations around the mean value are caused by different amounts of acquired storms in separate months. As high waves have a higher RMSE of around 52 cm, they increase the total RMSE when their relative percentage in a month is higher (see **Table 2**).

Table 2 Accuracy distribution for H_s using the machine learning approach SVM for four sea state domains and in total. Dataset: entire S1 WV archive with ~15 Mio samples

H_s domain, m	Sea state, %	RMSE (m)		
		wv1	wv2	Averaged wv1/wv2
0.0 - 1.5	11	0.280	0.342	0.311
1.5 - 3.0	62	0.196	0.227	0.2115
3.0 - 6.0	24	0.304	0.332	0.318
6.0 <	2	0.519	0.558	0.5385
Total	100	0.245	0.273	0.259

All processed S1 WV data including derived state parameters and imagette information (geo-location, time, ID, orbit number, etc.) are stored both as *ascii* and in *netcdf* format for convenient use. The data is made available to the public within the scope of ESA's climate change initiative CCI. **Figure 4** shows an example of Sentinel-1 Wave Mode WV archive processing for Feb. 2021.

6 Accuracy reached for different satellites and modes

In comparison to previous studies (e.g. [7]) a large amount of data was collected and applied. As many S1 IW, EW and TS-X images are acquired in coastal areas, only ~40% of all processed subscenes show ocean (wet points) and could be used for tuning/validation. The relatively coarse grid resolution of the used sea state models required an additional distance of 20 km from the coastline. The amount for resulting samples used for tuning and validations are shown in **Table 3**. The conducted studies and comparisons for S1 and TS-X show that using linear regression, the method tuned for different satellites and modes reaches an accuracy that depends more on the SAR image's pixel resolution rather than on the used frequency band (C-band for S1 and X-band for TS-X) or satellite altitude (ca. 700 km for S1 and ca. 500 km for TS-X). A RMSE of around 0.35 m was reached for both S1 WV and TS-X SM which have similar pixel resolution (1-4 m dependent on product and incidence angle), and an RMSE of around 0.60 m was achieved for the lower resolution S1 IW and S1 EW (pixel spacing of 10 m and 30 m, respectively). A limited validation case study conducted for TS-X ScanSAR (SC, pixel spacing 10 m) using 150 TS-X SC images also showed an RMSE of 0.60 m, which confirms this assumption.

For S1 WV almost all data were acquired in VV polarization (only a few months HH polarization data were acquired, e.g. May-July 2017), and since 2019 only in VV polarization. So, the S1 WV function was designed for VV polarization. The HH polarization data, validated separately, resulted into RMSE lower of around 10% than VV pol. data. The S1 IW and EW are acquired in both HH and VV polarizations. However, to estimate the sea state parameters VV polarization images were actually used. The reason is that tuning H_s using different polarizations results into better results for VV pol. images. Also, using combined HH and VV polarization images did not improve the results reached using only VV pol. [3].

Around 60% of all pooled TS-X SM data had dual polarization. As the function was tuned for each polarization individually, the dual-polarization products allowed to extend the number of samples for each polarization. The results of the study showed later that the data for both polarizations tuned independently resulted in the same accuracy for TS-X SM.

Table 3 Amount of data used for tuning and validation of S1 IW, EW and TS-X SM

Data		N ID	N collocated subscenes total	N sub-scenes tuning	N sub-scenes validation
S1-IW		1,062	517,289	300,000	217,289
S1-EW		2,093	1,162,492	800,000	362,492
TS-X-SM	HH	2,047	216,938	120,000	96,938
	VV		138,885	120,000	18,885

7 Conclusion

The purpose of this work is a consistent improvement of the method based on the classical CWAVE approach (involving both: SAR features and a solution using linear regression). The following was achieved:

- the series of additional operations and features significantly increase the accuracy in terms of H_s compared to the original CWAVE approach [4][6]. The studies with different options conducted in this work show that the cumulative improvement in RMSE reached around 15 cm (ca. 30% of RMSE~50 cm using the original CWAVE).
- an additional machine learning technique (SVM) further improves the H_s results by another ~10 cm compared to the linear regression solution and brings the results to the noise level of the ground truth data of around 25 cm (see ground truth comparisons in Section 2).

It was found that for a stable solution that is valid for worldwide applications, the linear regression approach needs around 80,000 random samples for tuning. Only when using this number of samples including any new data into the tuning does not improve the accuracy of the resulting model. For the machine learning (SVM) approach with a more complex function, this amount needs to be larger with around 600,000 samples. This is consequence of the H_s distribution in the worldwide acquired SAR data: H_s<3 m in ~75% of all cases, 3 m<H_s<6 m in around 24% and only around 1% of H_s>6 m and even less than 0.1% for H_s >10 m. For accurate reflection of the physical reality, the algorithms should be based on data covering all sea state domains with sufficient density.

8 Acknowledgements

This study was supported by the European Space Agency (ESA) within the CCI+ Sea State project. The authors are also grateful to Susanne Lehner for organizing the research group and providing initial scientific steps and the late Wolfgang Rosenthal for qualified advice concerning sea state, ship navigation and suggestions concerning algorithm developments.

9 Literature

[1] Schwarz, E., Krause, D., Berg, M., Daedelow, H. Maas, H.: Near Real Time Applications for Maritime Situational Awareness (2015). Proc. 36th Int. Symp. On Remote Sensing of Environment, Berlin, Germany, Remote Sensing and Spatial Information Sciences, vol. XL-7/W3, 6p.

[2] Pleskachevsky, A., W. Rosenthal, Lehner, S.: Meteo-Marine Parameters for Highly Variable Environment in Coastal Regions from Satellite Radar Images." JPRS, Vol. 119, pp. 464-484., 2016.

[3] Pleskachevsky, A., Jacobsen, S., Tings, B., Schwarz, E.: Estimation of sea state from Sentinel-1 Synthetic aperture radar imagery for maritime situation awareness. IJRS, Vol. 40-11, pp. 4104-4142., 2019.

[4] Schulz-Stellenfleth, J., König, Th., Lehner, S.: An empirical approach for the retrieval of integral ocean wave parameters from SAR data. JRL, Vol. 112, pp. 1-14., 2007

[5] Wen, Z., Shi, J., Li, Q., He, B., Chen, J., 2018.: ThunderSVM: A Fast SVM Library on GPUs and CPUs, Journal of Machine Learning Res. 19 (2018) 1-16.

[6] Stopa, J., Mouche, A.: Significant wave heights from Sentinel-1 SAR: Validation and Applications. JGR, Vol., 122, pp. 1827-1848., 2017.

[7] Pleskachevsky, A., Tings, B., Jacobsen, S.: Sea State from Sentinel-1 SAR Wave Mode Imagery for Maritime Situation Awareness. Proc. EUSAR-2019.

[8] http://marine.copernicus.eu

[9] https://www.ndbc.noaa.gov/

[10] Pleskachevsky, A., Tings, B., Jacobsen, S.: Multiparametric Sea State Fields from Synthetic Aperture Radar for Maritime Situational Awareness, under review, RSE.

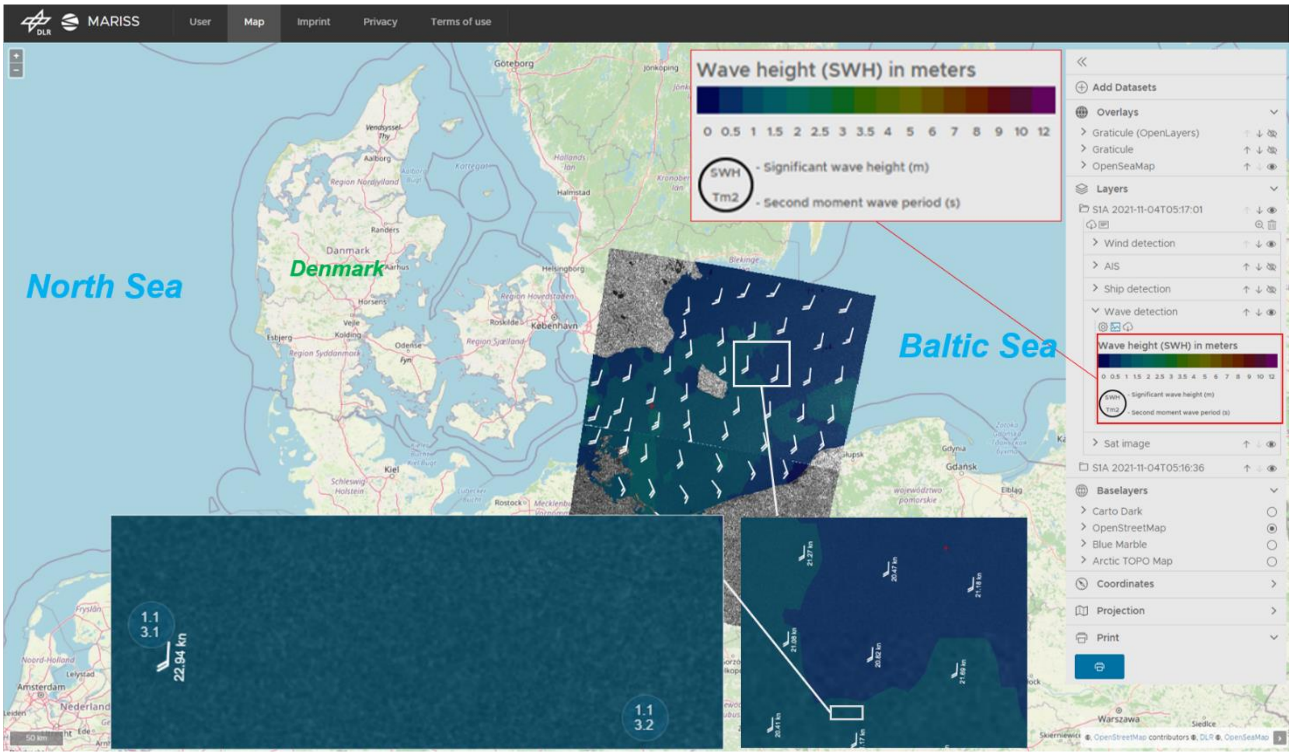

Figure 1 Screenshot of the demonstrator for NRT services at Ground Station Neustrelitz. The demonstrator runs daily for Sentinel-1 IW in Southern North Sea and Western Baltic Sea. The actual processing raster is 3 km (the processed subcenes cover an area of ~4 km × 4 km), the wave-detection layer shows wave height (colored) and period (in circles: H_s above, Tm_2 below). Data for all eight sea state parameters can be downloaded as google-earth ID*kmz file. The wind-detection level shows the wind speed estimated from SAR image.

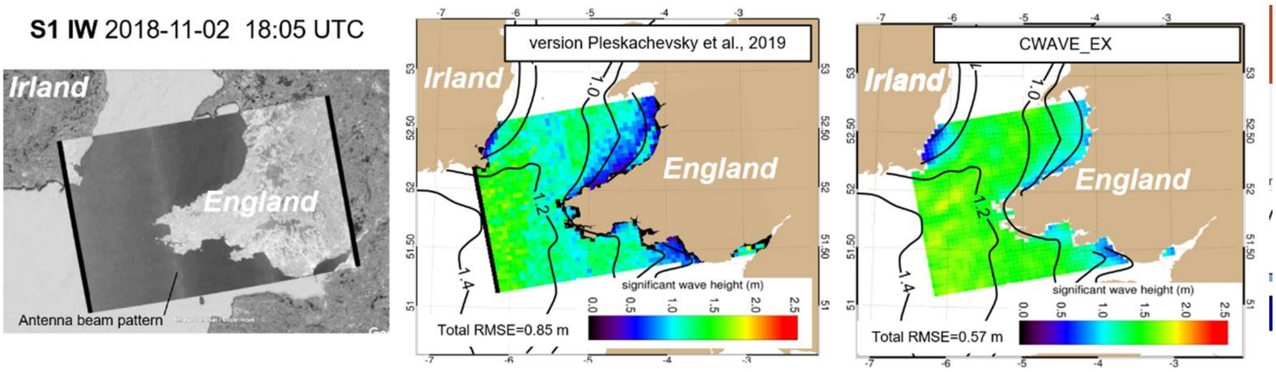

Figure 2 Significant wave height field estimated from S1 IW image acquired on 2018-11-02 at 18:05 UTC. The results for the previous algorithm version [2,7] and the results from this study are presented. The errors "black gap" (boundary effect, NRCS=0) and antenna beam pattern (middle of the SAR image, flight direction) are not more present. Isolines show the H_s from CMEMS on 2018-11-02 at 18:00 UTC.

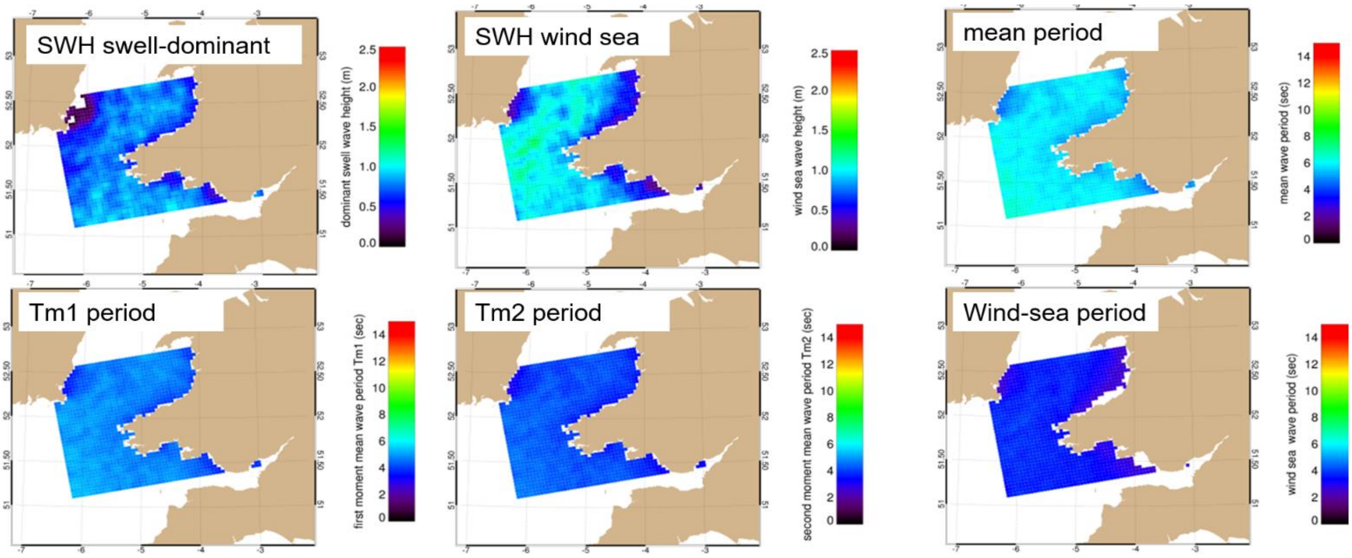

Figure 3 Example for integrated sea state parameters for Sentinel-1 IW image shown in **Figure 2**: swell and wind-sea wave heights, mean, first and second moment periods and period of the wind sea.

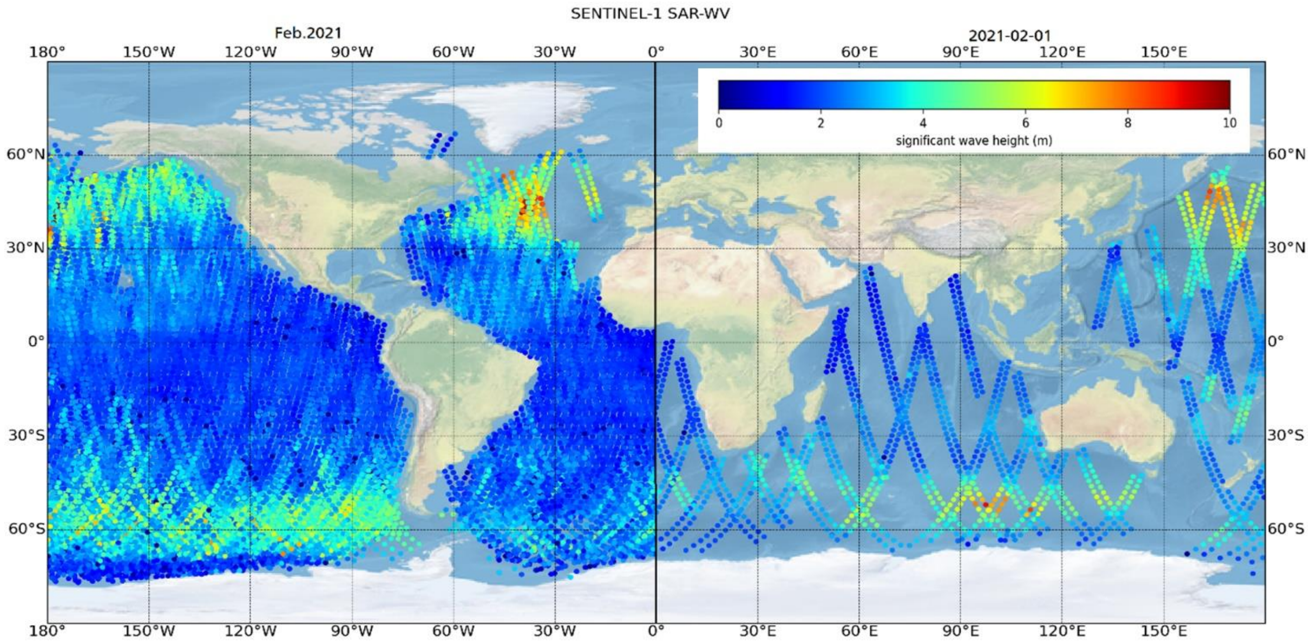

Figure 4 Example of Sentinel-1 Wave Mode WV archive processing. On the right half of the globe only one-day acquisitions are displayed, on the left half all data acquired during February 2021