AGU ADVANCING EARTH AND SPACE SCIENCE

Earth and Space Science

RESEARCH ARTICLE
10.1029/2021EA002008

Special Section:
China-France Oceanography Satellite(CFOSAT): Scientific Applications

Global Validation of SWIM/CFOSAT Wind Waves Against Voluntary Observing Ship Data

V. G. Grigorieva[1], S. I. Badulin[1,2], and S. K. Gulev[1]

[1]Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia, [2]Skolkovo Institute of Science and Technology, Moscow, Russia

Key Points:

- For the first time ever two independent sea state data sources: visual observations and CFOSAT measurements are compared on equal terms
- Global statistics, annual distributions, and instantaneous individual measurements of wave and wind characteristics are analyzed
- Emphasis is made on discriminating wave systems by joint analysis of visually observed wind sea and swell and wave partitions from SWIM

Correspondence to:
V. G. Grigorieva,
vika@sail.msk.ru

Citation:
Grigorieva, V. G., Badulin, S. I., & Gulev, S. K. (2022). Global validation of SWIM/CFOSAT wind waves against Voluntary Observing Ship data. *Earth and Space Science*, *9*, e2021EA002008. https://doi.org/10.1029/2021EA002008

Received 9 SEP 2021
Accepted 6 JAN 2022

Author Contributions:
Conceptualization: V. G. Grigorieva, S. I. Badulin
Data curation: S. I. Badulin
Formal analysis: V. G. Grigorieva
Funding acquisition: S. K. Gulev
Investigation: V. G. Grigorieva
Methodology: V. G. Grigorieva, S. K. Gulev
Software: V. G. Grigorieva
Supervision: S. K. Gulev
Validation: V. G. Grigorieva
Visualization: V. G. Grigorieva

Abstract The article presents a joint analysis of wind and wave characteristics derived from Voluntary Observing Ship (VOS) data and measurements of the innovative Ku-band radar SWIM (Surface Waves Investigation and Monitoring) carried by Chinese-French Ocean SATellite (CFOSAT). Global distributions of significant wave height and wind speed in both data sets demonstrate good qualitative and quantitative agreement, especially in regions with a high spatio-temporal density of visual observations. A particular focus is made on discriminating wave systems by joint analysis of separately observed wind sea and swell characteristics in VOS and the partitions of wave spectra measured by SWIM. It is shown that three wave partitions from SWIM cannot be clearly attributed to wind sea, first, and secondary swell systems. The first partition aggregates both wind sea and swell in the operational wavelength range of the SWIM radar, while the second and third partitions fit neither wind sea, nor swell. A comparison of VOS and SWIM data within a 50 km radius and a 30 min time lag shows a very close match for most parameters in terms of mean values, yet with relatively high dispersion of individual measurements.

Plain Language Summary In this study, we analyze waves and winds derived from two independent sources: Voluntary Observing Ship (VOS) data and the innovative Ku-band radar SWIM (Surface Waves Investigation and Monitoring) carried by Chinese-French Ocean SATellite (CFOSAT). SWIM provides measurements of the directional spectra of ocean waves with further partitioning of multiple wave systems. In VOS, the separation is performed visually: wind seas are associated with local wind, and swells (first and secondary, if present) are not dependent on it. Significant wave height (measured by SWIM and calculated in VOS) is traditionally defined as the mean height of the highest one-third of the waves, or four times the square root of the zeroth-order moment of the wave spectrum. Satellite wind speed is evaluated using a parametric approach; VOS has high-precision anemometer measurements. Significant wave height and wind speed in VOS and SWIM are consistent in terms of average statistics, but show a large dispersion in individual measurements. The three wave partitions retrieved from SWIM measurements cannot be simply regarded as wind sea, first, and secondary swell systems. The first partition aggregates both wind sea and swell, while the second and third partitions look like residuals and fit neither wind sea, nor swell.

1. Introduction

Monitoring, analysis, and forecast of global sea state characteristics lay the foundation for various aspects of human marine activity, including natural hazard mitigation and the safety of navigation and off-shore constructions. Developing novel experimental tools is an essential part of progress for the exploration and research of the World Ocean. A recent example, the CFOSAT mission designed under the aegis of Chinese (CNSA) and French (CNES) space agencies is aimed at monitoring ocean surface waves and winds (http://smsc.cnes.fr/CFOSAT/index.htm). An entirely new concept of space observations is implemented with the real aperture Ku-band radar SWIM (Surface Waves Investigation and Monitoring). The innovation yields global measurements of the directional spectra of ocean waves in addition to the conventional altimetry parameters: significant wave height, H_s, and normalized radar cross-section, σ_0. Further wave spectrum partitioning allows for estimating significant wave heights, dominant wavelengths, and directions of multiple wave systems (up to three). Thus, CFOSAT provides new scientific products that cover a wide range of today's marine activity needs: from operational oceanography to fundamental academic studies of surface processes linked to wind and waves (Hauser et al., 2017, 2020).

Despite the relatively small operational period of the CFOSAT mission – merely 2 years – the data is still regarded as high in quality (Hauser et al., 2020) and relevant to global sea wave monitoring (Aouf, 2020; Aouf

Writing – original draft: V. G. Grigorieva
Writing – review & editing: S. I. Badulin

et al., 2019, 2021). The SWIM data was carefully validated against ocean buoy measurements (Jiang et al., 2021), the products of Copernicus Marine Service (CMEMS, von Schuckmann et al., 2020), and model simulations (Aouf et al., 2021; Le Merle et al., 2021).

There is, however, an alternative data set that has not yet been considered for comparison with the CFOSAT products but could be insightful for future research – the Voluntary Observing Ship (hereinafter VOS) data. Under the framework of the World Meteorological Organization (WMO), VOS unites thousands of international vessels with marine mates taking visual observations of numerous ocean and atmosphere characteristics all over the globe. The collected data consolidated in the ICOADS (International Comprehensive Ocean-Atmosphere Data set) make an invaluable contribution to sea navigation, operational forecast, and global climate studies (Freeman et al., 2017; Kent et al., 2019; Smith et al., 2019). The main advantage of visual observations in VOS is that different wave systems are estimated separately: wind seas (or wind-driven waves) are associated with local wind, and swell is wind-independent (Sverdrup & Munk, 1947). The observed parameters enable the estimation of significant wave height and dominant period as the most common characteristics of sea state (Gulev et al., 2003). The term "significant wave height" was proposed by Walter Munk (1944) as a new statistical quantity to be compatible with visual estimates of trained observers. Since then, significant wave height is defined as the mean height of the highest one-third of the waves or four times the square root of the zeroth-order moment of the wave spectrum.

VOS data was successfully used for validating satellite altimetry (Grigorieva & Badulin, 2016) and model simulations (Grigorieva et al., 2020; Gulev et al., 2003). The comparison showed a good agreement for significant wave height and mean period on seasonal, annual, and climatological scales. At the same time, wind sea and swell components derived from model simulations demonstrate significant quantitative differences with VOS, though their spatial distribution patterns and directional steadiness are very close on global scales (Grigorieva et al., 2020). Here we face the common difficulties of comparing results obtained by different approaches. Wave system discrimination in VOS is highly affected by the skill and experience of marine mates, so it can be hard to formalize the observations and relate them to instrumental measurements and numerical model outcomes. Similarly, the physical and technical constraints of remote measurements (e.g., SWIM) hinder the assimilation of informal yet useful information provided by VOS officers. Despite this, attempts to establish a correspondence between visual and instrumental observations in related fields, for example, in cloudiness measurements, prove to be successful (Krinitskiy et al., 2021).

In this study, visual observations from VOS were used for global validation of the SWIM data from CFOSAT, with a particular focus on the partitioning of wave systems.

The paper is organized as follows. Section 2 describes both data sets, namely Level 2 SWIM/CFOSAT data and VOS records from the ICOADS. Section 3 presents a comprehensive analysis and verification of VOS and SWIM wave and wind parameters in terms of conventional statistics and global annual distributions, followed by a comparison of collocated measurements. Discussion (Section 4) is focused on a physical interpretation of the obtained results, including an analysis of SWIM 2-year measurements and climatological VOS estimates. Conclusive Section 5 finalizes the paper by listing the main results and perspectives of VOS data for validating satellite measurements.

2. Data

2.1. SWIM Measurements From CFOSAT

CFOSAT was launched on 29 October 2018, with the purpose of permanent and global monitoring of wind and waves by two state-of-the-art devices: a "wave scatterometer" SWIM and a wind scatterometer SCAT.

A real aperture Ku-band radar SWIM, developed by CNES (Centre National d'Etudes Spatiales) operates at near-nadir incidences: 0° (nadir), 2°, 4°, 6°, 8°, and 10°. Nadir sounding of the sea surface with 0.2 s time intervals (5 Hz) ensures obtaining high-quality significant wave height (H_s) and normalized radar cross section (σ_0). Near-surface wind speed is calculated as a function of H_s and σ_0 using a parametric approach (Hauser et al., 2020; Gourrion et al., 2002; Ren et al., 2021). The accuracy of H_s measurements is better than 10% (or 0.4 m). In turn, the accuracy of σ_0 allows for assessment of the wind speed better than 2 m/s or 10% (whichever is greater) in a 4–24 m/s range (Hauser et al., 2020).

Off-nadir measurements provide two-dimensional surface ocean wave spectra, though in rather large spatial boxes of about 70 × 90 km (Hauser et al., 2017, 2020). A wave spectrum describes the distribution of total wave variance over direction and wavelength (spatial frequency). It represents a combination of individual wave systems, or partitions, likely originating from different meteorological events. For the analysis of the SWIM spectra, the watershed method is applied to identify up to three wave partitions (Hanson & Phillips, 2001; Hauser et al., 2020) and estimate their characteristics: significant wave height, dominant wavelength, and direction. The nominal accuracy for partial wave heights is 10% (or 0.5 m) and 10%–20% for partial wavelengths. In the current data version (V5.1.2), wave directions can be retrieved with an ambiguity of ±180° and a resolution of 15°. The wave partitions are estimated in a wavelength range of 30–500 m and are ranked (in decreasing order) by significant wave height (CFOSAT L2PBOX Product Handbook, Issue: 1.1, 07/09/2021).

Off-nadir SWIM parameters are also complemented by nadir wind speed averaged along track over the length of the box. Thus, waves and winds are derived from measurements of the same instrument at the same coordinates providing inherently consistent and robust estimates. Moreover, the used 70 × 90 km wave cell size corresponds to the standard spatial scales of Cal/Val procedures, which is 50 km for wave measurements.

In this study, we used the SWIM/CFOSAT data (version V5.1.2) for the period of 25.04.2019–31.03.2021. From nadir measurements, we operate with 1 Hz significant wave height (H_s, variable "nadir_swh_1 Hz") and the near-surface wind speed (variable "nadir_wind_1 Hz"). Off-nadir measurements availed ourselves of wave height, wavelength, and direction of each partition (variable "wave_param_part_combined"), as well as box wind speed (variable "nadir_wind_box"). Additionally, the integral wave height (H_{int}) was calculated as the square root of the sum of squares of partitioning wave heights. Only complete data records containing all of the above wind-wave characteristics with "good" quality flags were taken into consideration.

2.2. Voluntary Observing Ship Data

Visual observations of the sea state have been carried out by sailors for centuries since the launch of the very first ship. This practice still remains an important means of wave studies today despite the advent of buoys, satellites, and the development of wave theory and computational models. Consistency is the key feature of visual wave data from VOS collection: the observational practice was laid out in 1853 (Maury, 1854), which makes for the longest record of ocean waves.

All VOS data is assimilated in the ICOADS archive (Freeman et al., 2017) as a set of individual records of more than 100 observed atmosphere and ocean parameters at given coordinates and UTC (Coordinated Universal Time). Besides ship data, ICOADS includes wave measurements from buoys, rigs, light vessels, and so on. In this study, only visual ship-based observations are used to guarantee the homogeneity of each comparing data set. All collected data undergo initial quality control before being included in the ICOADS and are adjusted to an advanced International Maritime Meteorological Archive (IMMA) format (Smith et al., 2016). Thus, the data can be considered homogeneous and uniform.

As mentioned above, the original VOS data provides separate estimates of wind sea and swell characteristics, contrary to buoys, models, and altimeters that require additional processing procedures to identify different wave systems - which are, in a sense, merely a secondary product yielded by wave spectra decomposition. This concept urges us to consider visual wave observations as the "ground truth" for the verification and validation of wave systems derived from all alternative data sets. Code figures of the reported parameters (0.5 m for wave height, 1 s for wave period, and 10° for directions) can be accepted as characteristics of accuracy, which puts VOS and CFOSAT data on equal precision terms.

For a comprehensive comparison of SWIM measurements with visual observations, VOS significant wave height (SWH) and dominant wave period were evaluated from individual estimates of wind sea and swell parameters. Since VOS does not operate with wave spectra, alternative techniques are required to estimate SWH. In this research, we use two standard methods: (a) SWH1 is calculated as the square root of the sum of squares of wind sea and swell heights, an approach that is theoretically well-grounded by the estimation of spectral moments (Hogben, 1988); and (b) SWH2 is evaluated as the highest wave among simultaneously observed wind sea and swell heights (Wilkerson & Earle, 1990). The second approach is not as substantiated, but it works better for the off-shore regions and the subtropics, providing the least biased results (Gulev et al., 2003). Dominant wave period within visual observations is regarded as the period reported for the highest wave of both components (wind sea

and swell), which corresponds to the definition of the zero-up crossing period in the instrumental data (Srokosz & Challenor, 1987).

We selected complete VOS records which contain wind sea and swell heights, periods, and directions of propagation accompanied by wind parameters. To eliminate spurious errors and correct the occurring uncertainties in VOS data, special multistage quality control was applied (Gulev et al., 2003). Then SWH1, SWH2, and dominant wave period were calculated.

The details of data access are given in Data Availability Statement section.

3. Results

3.1. Wavelengths (Periods) of SWIM and VOS Wave Partitions

To compare wave characteristics derived from two fundamentally different sources, in our case, remote measurements and visual observations, we first need to consider the variability ranges of the parameters in each data set.

SWIM-derived wave partitions have a declared wavelength range of 30–500 m (~4.4–18 s periods; CFOSAT L2PBOX Product Handbook, Issue: 1.1, 07/09/2021). In the VOS coding system, wave periods have a range of 0–30 s (Smith et al., 2016), which implies a maximal wavelength of about 1,400 m. The actual number of these extreme waves is very low, especially after the data undergoes quality control. To further understand the distribution of VOS observations in the given range, a statistical analysis was performed, taking data from the past 50 years (1970–2020). Findings show that wind sea periods of 2–10 s (wavelengths of ~6–150 m) constitute more than 90% of observations. For swell periods, that percentage falls between 4 and 15 s, that is, ~25–350 m wavelengths.

Figure 1 shows the histograms of dominant periods/wavelengths from VOS compared to the periods/wavelengths of the first and second SWIM partitions. The VOS parameter choice was motivated by the fact that the dominant period can represent either wind sea or swell—whichever of the two is higher (see Section 2.2). In other words, this characteristic is able to accumulate all observed wave systems, portraying the full range of period/wavelength variability for further analysis against the SWIM partition ranges. The histogram bins (here and for all analyzed parameters below) were chosen twice larger than the accuracy of measurements according to the Kotel'nikov (Nyquist–Shannon) theorem (Kotel'nikov, 2006).

In general, the histograms demonstrate a poor agreement for the whole waveband. Both SWIM partitions fully capture waves of an 80–160 m mid-range that, apparently, include both wind sea and swell (Figures 1a and 1b). Short waves (0–80 m) primarily associated with wind seas are significantly underestimated by CFOSAT because SWIM simply does not "see" them. The underestimation of long swells (>160 m) by VOS, despite the presence of 30 s period waves, might result from poor observational density in the Southern Ocean and swell pool areas (Chen et al., 2002). To continue the analogy, VOS does not "see" vast ocean areas where long swell dominates. For wave periods, the dissimilarity of histogram shapes (Figures 1c and 1d) caused by the lack of short waves in the SWIM measurements is more pronounced which is explained by the dispersion relation ($T^2 \sim \lambda$, where T is wave period and λ is wavelength).

The performed statistical analysis implies that off-nadir SWIM measurements tend to underestimate visually observed short waves and overestimate long waves.

3.2. Heights of SWIM and VOS Wave Partitions

The above wavelength-period analysis does not allow us to confidently relate wave partitions from SWIM to VOS wind sea or swell. So we attempted to perform a cross-comparison of wave heights of the two wave partitions from SWIM and two wave systems from VOS.

The global histogram of the first partition heights coincides with the distribution of VOS wind sea heights almost perfectly (Figure 2a) and does not match the VOS swell distribution (Figure 2c). The second partition captures wind sea and swell heights only for a 1–2 m range (Figures 2b and 2d), that is, for the most probable values. Both partitions underestimate VOS high swells and overestimate small swell heights (0–1 m; Figures 2c and 2d). This

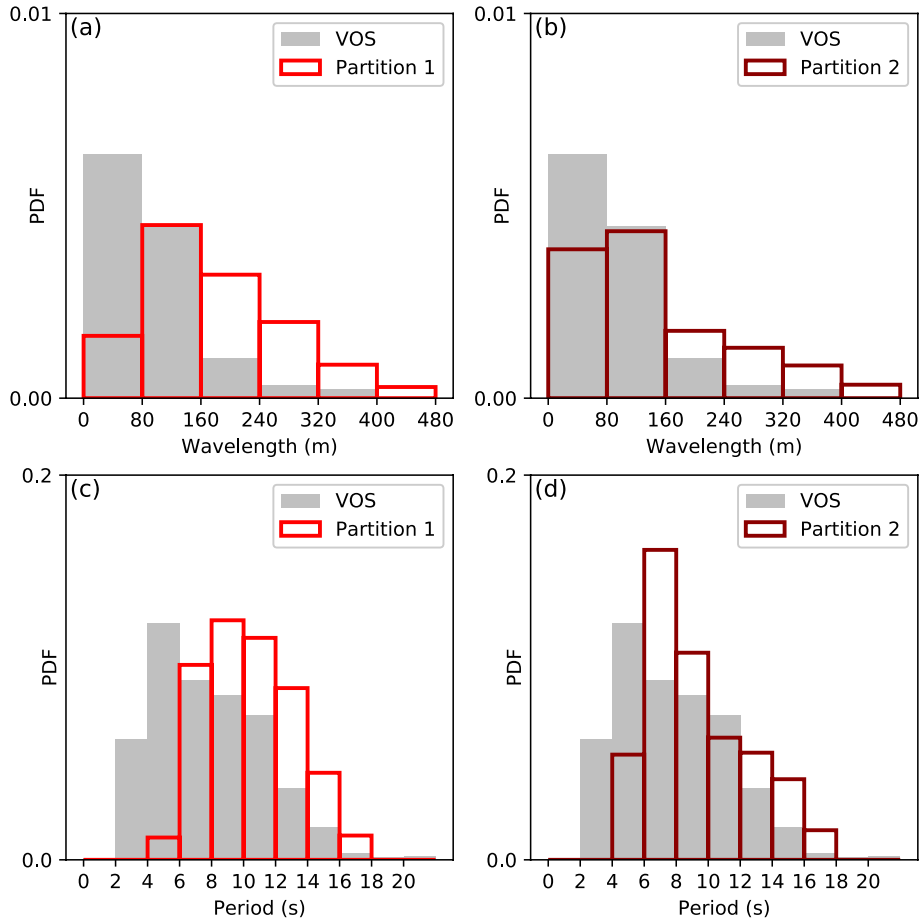

Figure 1. Global histograms of dominant wavelengths and periods from VOS (gray color) in comparison with the first partition (red lines) and second partition (brown lines) from SWIM, 2019–2021. (a) VOS—first partition wavelengths; (b) VOS—second partition wavelengths; (c) VOS—first partition periods; (d) VOS—second partition periods.

may be associated with the high accuracy of SWIM wave height measurements as well as with the rounding of VOS wave heights to half-meter values (due to the code figures in the VOS).

To supplement the statistical examination, global annual distributions of wave heights of the two VOS wave systems (wind sea and swell) and the two wave partitions of SWIM were mapped (Figure 3). The pattern of the first partition heights (Figure 3a) can be regarded as a combination of wind sea and swell, corresponding better with the VOS swell distribution in the Southern Ocean (Figure 3d) and with VOS wind sea in the stormy regions of both hemispheres (Figure 3c). Annual wave heights of the second partition substantially look like a minor replica of the first one with smaller magnitudes (Figure 3b), which is a direct consequence of ranking wave partitions by significant wave height in the decreasing order (L2PBOX Product Handbook, Issue: 1.1, 07/09/2021).

The most challenging part of wave partitions validation was to investigate their directional steadiness. A similar comparison was performed for wave directions derived from VOS and the third-generation spectral wave model WAVEWATCH III hindcast. The ship-based and model directional steadiness of wind sea and swell have shown a perfect consistency, thereby implying the high quality of VOS data (Grigorieva et al., 2020).

Since the determination of the SWIM wave directions at this stage is ambiguous, we carried out a prefatory experiment using the data we possessed. The first partition was related to VOS wind-driven seas based on the similarity of their global statistics (Figure 2a), and the second partition was regarded as a counterpart of VOS swells. Then, the meridional and zonal components of mean wave directions in both data sets ($2° \times 2°$ grid) were calculated. Finally, if the signs of the VOS and SWIM components in a given two-degree box were different, the

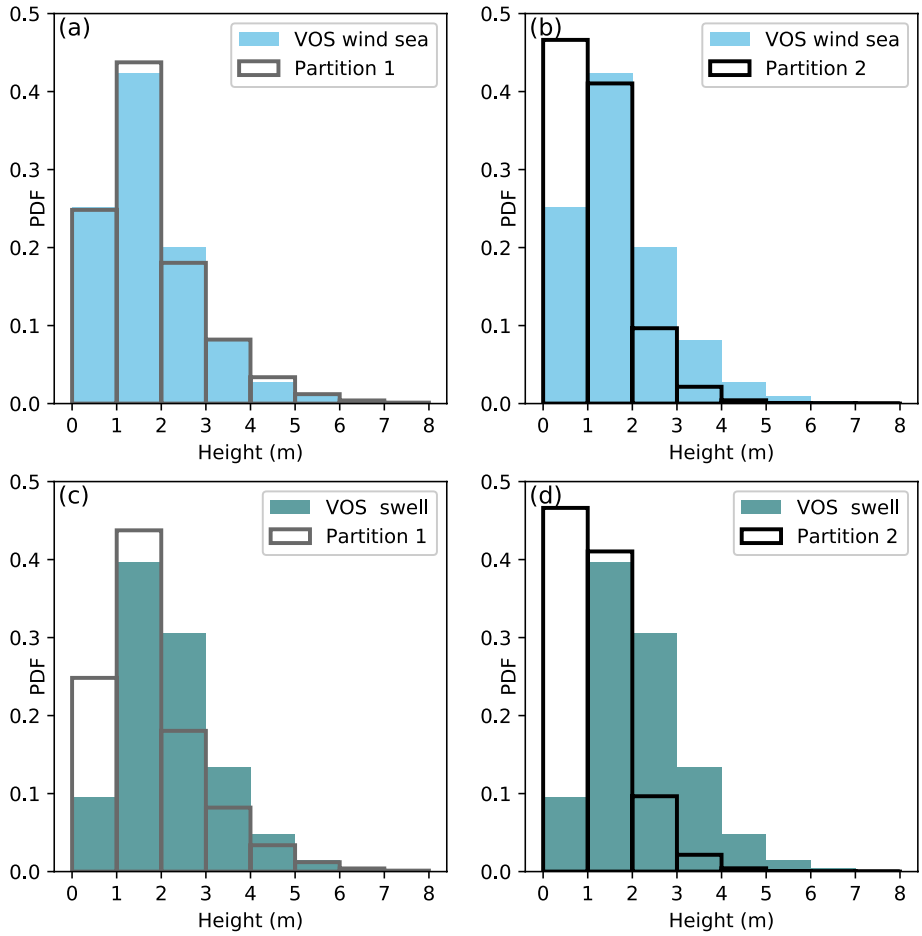

Figure 2. Cross-comparison of VOS wind sea (blue) and swell (cyan) heights with the first (gray lines) and second partition (black lines) heights from SWIM, 2019–2021. (a) VOS wind sea height—first partition; (b) VOS wind sea height—second partition; (c) VOS swell height—first partition; (d) VOS swell height—second partition.

VOS sign was taken as the primary one. Overall, signs were changed in about 50% of cases, which corresponds to the ambiguity ±180° of wave direction measurements in SWIM data.

The estimates of directional steadiness obtained from SWIM show rather trivial distributions, contrary to more sophisticated directional patterns from VOS. The restrictions on wavelengths measured by SWIM and a deficiency of today's methods of wave field reconstruction of directionality are the likely reasons for the discrepancies between these two types of data. The forthcoming development of the SWIM data processing algorithms aimed to remove the directional ambiguity could change the situation.

Summarizing the results of this section, we can conclude that SWIM wave partitioning is less reliable than the separation of wind sea and swell by the experienced observers in VOS. A joint analysis of wind and wave directions measured by SWIM and SCAT could provide a more accurate partitioning of wave systems because wind directions mostly coincide with wind sea (~95% of cases) while swell does not (~80% of cases of different directions according to VOS statistics).

3.3. Significant Wave Heights

Our next step of SWIM and VOS data validation dealt with significant wave height as the most common parameter of the sea state. Statistical analysis was performed for nadir significant wave height (H_s), the integral sum of all off-nadir partitions (H_{int}), and two significant wave heights (SWH1 and SWH2) from VOS (Figure 4). The distribution of nadir H_s is almost identical to SWH1 VOS (Figure 4a) while the distribution of H_{int} shows

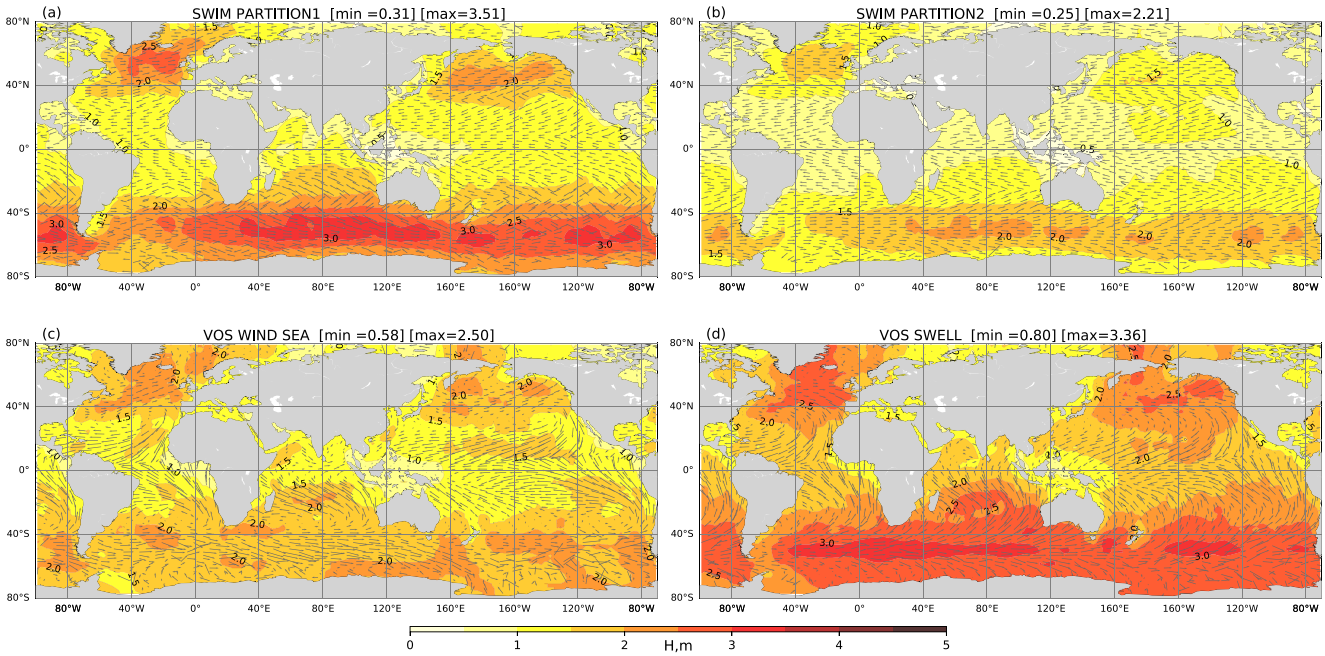

Figure 3. Global maps of annual wave heights (2° × 2°) of the first (a) and second (b) partitions from off-nadir SWIM measurements in comparison with annual wind sea (c) and swell (d) heights from VOS, 2020. Gray arrows show the directional steadiness. The length of the arrows is proportional to the wave phase velocity.

a good agreement with SWH2 VOS (Figure 4b). The results can be justified for both pairs "SWH1–H_s" and "SWH2–H_{int}." Nadir measurements H_s show surface elevation magnitudes and fit the SWH1 definition better. VOS SWH2 is the maximal height of two observed components (wind sea and swell), and in ~80% of cases that represents swell. The other 20% relate to prevailing wind sea, generally, mature or fully developed waves (Komen et al., 1984). The high coincidence of SWH2 with integral partitions height H_{int} (Figure 4b) can be explained by the SWIM operational wavelength range capturing long seas and swells by off-nadir measurements.

Statistical distributions are confirmed by annual global maps (2° × 2°) of significant wave heights (Figure 5). The largest differences for both pairs "SWH1–H_s" and "SWH2–H_{int}" (up to 1.4 m) are observed in the Southern Ocean where visual wave heights are smaller than satellite ones. Small negative differences (<0.5 m) are also found in

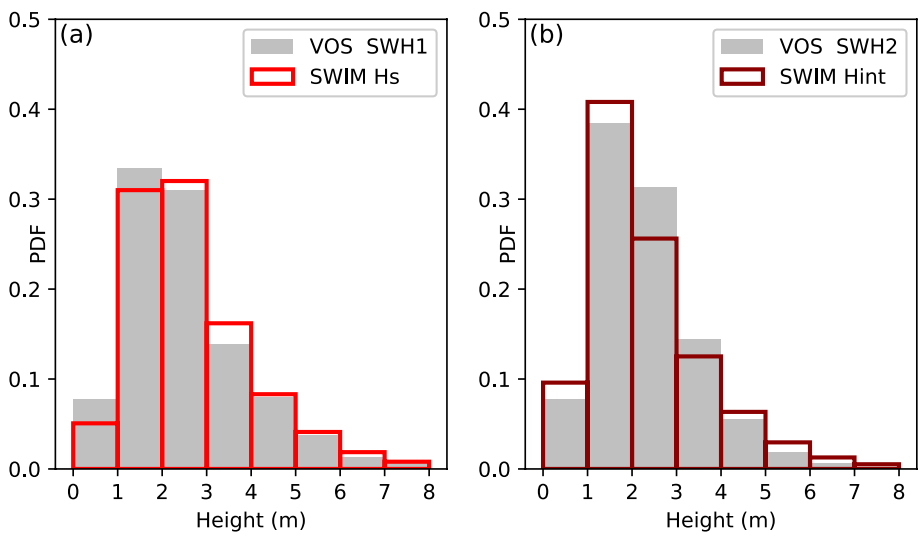

Figure 4. Global histograms of SWIM and VOS significant wave heights for 2019–2021. (a) VOS SWH1 (gray) and SWIM H_s (red lines); (b) VOS SWH2 (gray) and SWIM H_{int} (brown lines).

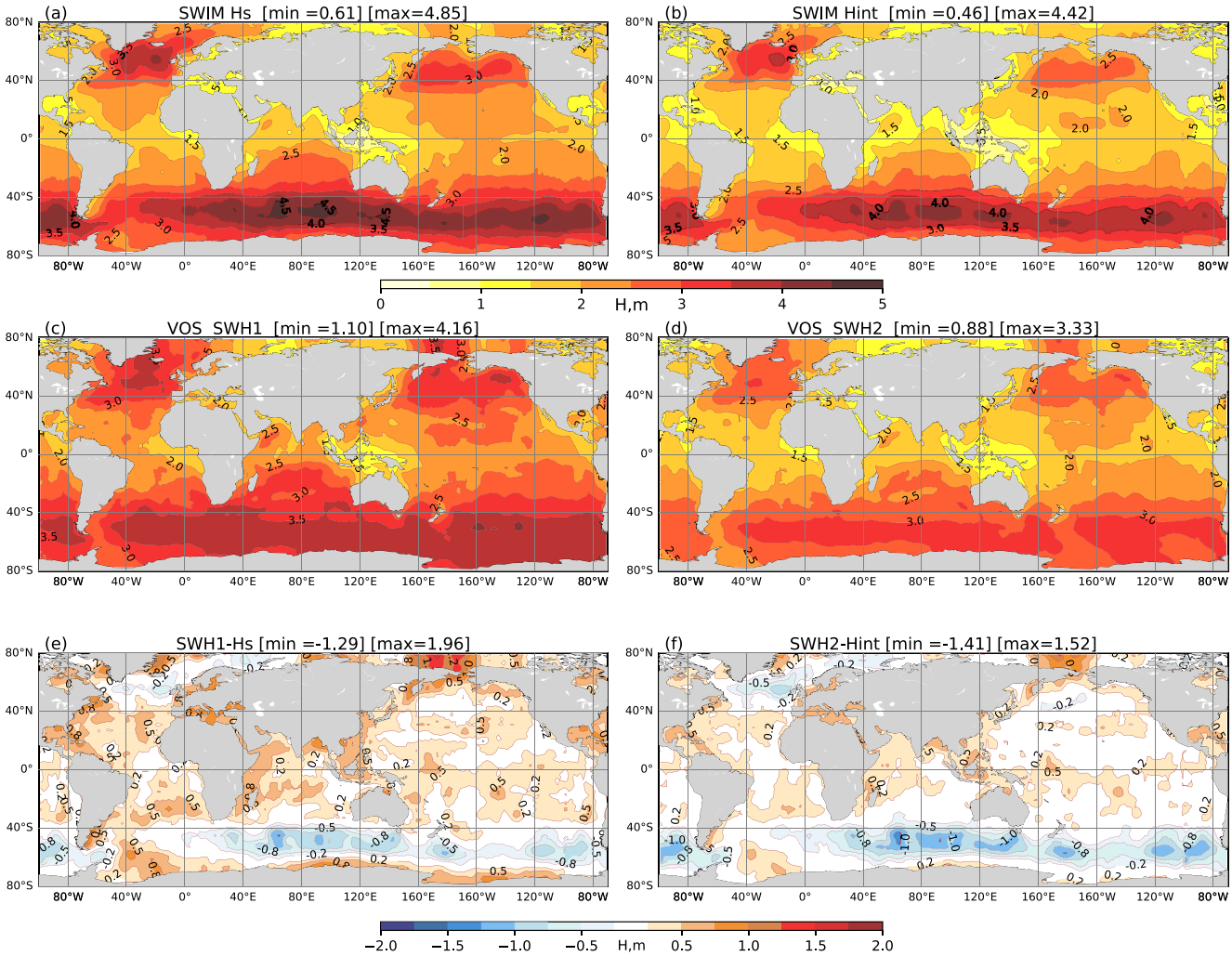

Figure 5. Global annual maps of SWIM and VOS significant wave heights (m) and their differences for 2020 (2° × 2°). (a) SWIM H_s; (b) SWIM H_{int}; (c) VOS SWH1; (d) VOS SWH2; (e) the difference between SWH1 and H_s, (f) the difference between SWH2 and H_{int}.

the stormy regions of the Northern Atlantic. That underestimation results from poor observational density and the tendency of ships to avoid storms (Gulev et al., 2003).

VOS SWH1 are notably greater than H_s in the coastal areas (differences reach 0.8 m). Similar tendencies were also revealed when comparing VOS observations with other satellite missions (Grigorieva & Badulin, 2016) and with the third-generation spectral wave model WAVEWATCH III hindcast (Grigorieva et al., 2020). We can tentatively suppose that these differences might stem from the better accuracy of significant wave height estimates in VOS, which is prompted by a high observational density in the coastal areas. Another relevant factor is the arguable accuracy of nadir SWIM measurements in the near-shore area. To come to a definite conclusion, additional research with a focus on near-shore observations is required for both data sets. VOS SWH2 and integral wave heights, H_{int} are more consistent, with differences of less than 0.2 m in most areas. Wave height observations in the Polar Regions should be considered with caution because of the sparse presence of VOS data and possible corruption of SWIM measurements due to ice concentration (International Altimetry team, 2021; Lebedev, 2016). In general, the differences in the annual significant wave height fields do not exceed 0.5 m, which falls within the accuracy of wave height measurements in both data sets.

It is worth noting that the consistency between significant wave heights in VOS and SWIM in terms of statistical and geographical distributions was expected. The previous comparison of visual wave observations with various

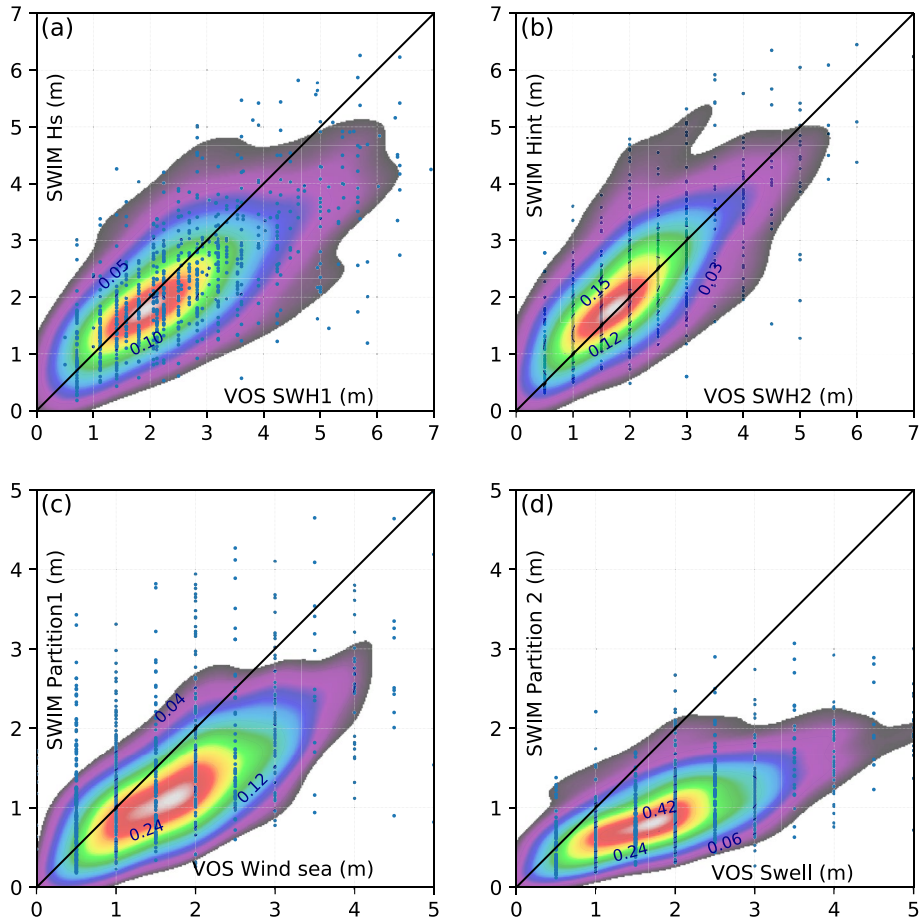

Figure 6. Two-dimensional probability density functions for wave heights from VOS and SWIM within a 50 km radius and a 30 min time lag. (a) VOS SWH1 and SWIM H_s; (b) VOS SWH2 and SWIM H_{int}; (c) VOS wind sea heights and SWIM first partition; (d) VOS swell heights and SWIM second partition. Probability density functions were estimated using 2D Gaussian kernels.

satellite missions (Grigorieva & Badulin, 2016) and model simulations (Grigorieva et al., 2020) has shown similar results on a global scale.

3.4. Wave Heights in Collocated Points

Comparison of any vast data sets in terms of statistics of mean values at global scales, generally, shows a good agreement. For a more detailed validation, we considered VOS and SWIM data in collocated points within two radii of 25 and 50 km and a time lag of less than 30 min. Figure 6 shows scattering plots and the resulting estimates of two-dimensional probability density function (PDF) for pairs of significant wave heights (Figures 6a and 6b), the first partition and wind sea heights (Figure 6c), the second partition and swell heights (Figure 6d) for a 50 km radius (801 pairs for 2020). All compared parameters demonstrate high dispersion and quite low correlation coefficients of about 0.5–0.7. Moreover, a shorter radius of collocation (25 km) does not dramatically change this result (not shown in the figures).

The distributions of "SWH1–H_s" (Figure 6a) and "SWH2–H_{int}" (Figure 6b) are elongated along the diagonal $y = x$, which indicates a good agreement of the mean statistics. The maximal probability density of the first partition is shifted toward the VOS axis which points to an underestimation of small waves of 1–2 m heights by SWIM. Similar tendencies are observed for global histograms (Figure 2a). The second partition, being less than 3 m for all points, significantly underestimates visually observed swell heights as it was found for global statistics (Figures 2d and Figures 3b, 3d).

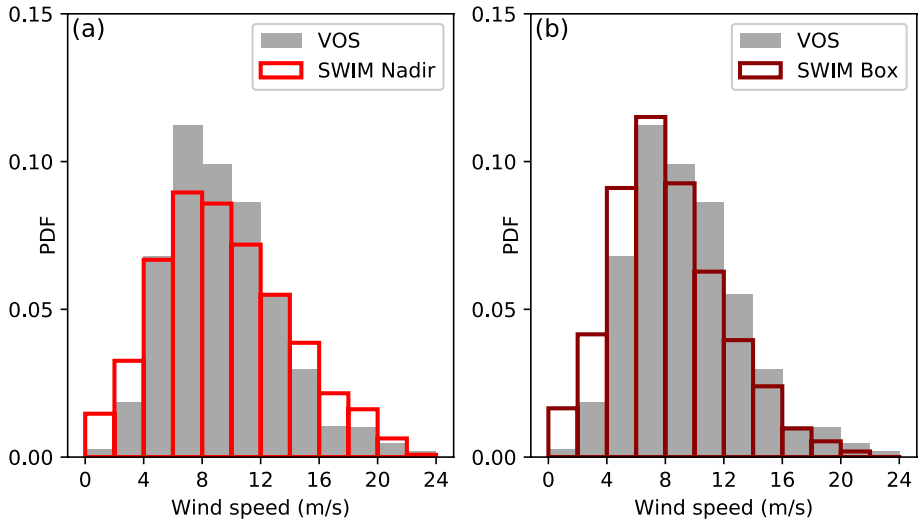

Figure 7. Global histograms of wind speed from VOS and SWIM, 2019–2021. (a) VOS wind speed (gray) and SWIM nadir wind speed (red lines); (b) VOS wind speed (gray) and SWIM box wind speed (brown lines).

In general, the results of the comparison of instantaneous individual wave height measurements in collocated points once again confirm our assumption that wave partitions of SWIM cannot be directly attributed to wind sea or swell. At the same time, all significant wave height estimates (SWH1, H_s, SWH2, H_{int}) representing, each in its own way, integral sea surface elevation, agree well. A possible reason for the large dispersion of analyzed pairs is the discreteness of visual observations (0.5 m) which manifests itself both for the directly observed wind sea and swell, and for the calculated significant wave heights.

3.5. Wind Speed

The above validation of SWIM/CFOSAT measurements relied on visually observed wave parameters. Yet, along with many other ocean and atmosphere characteristics, VOS also provides wind speed and direction data at a given point, which could be applicable to our research. According to the ICOADS IMMA format (https://icoads.noaa.gov/e-doc/imma/R3.0-imma1_short.pdf), "wind speed is stored in tenths of a meter per second (to retain adequate precision for winds converted from knots, or high-resolution data)." The use of high-precision anemometers over the last decades guarantees the homogeneity of VOS wind speed measurements, though some uncertainties associated with anemometer type, calibration, and location still remain (Gulev, 1999; Kent et al., 1993; Smith et al., 1999). In this study, wind characteristics did not undergo any special adjustment procedures. For the experiment with global wind statistics, we used VOS records that had passed general quality control for wave parameters.

Figure 7 represents histograms of VOS wind speed in comparison with nadir and box SWIM data. VOS reproduces moderate nadir-derived winds (4–14 m/s) well but underestimates low (0–4 m/s) and strong (14–20 m/s) winds (Figure 7a). A similar underestimation of low nadir winds by the ECMWF (European Centre for Medium-Range Weather Forecasts) model was recently found by Hauser et al. (2020). But contrary to VOS, ECMWF tends to overestimate high nadir winds (Hauser et al., 2020).

Box winds fit the most probable VOS values, notably overestimating small and underestimating high VOS winds (Figure 7b). The discrepancies between VOS and SWIM histograms might stem both from the different spatial resolutions of nadir and box winds and from the spatial inhomogeneity of pointwise VOS measurements.

To get more insight, global annual wind speed fields (2° × 2°) were mapped. Spatial patterns for the VOS, nadir, and box winds are very close. Figure 8 illustrates the annual distributions of the nadir SWIM measurements and VOS data for 2020 along with their differences. The found underestimation of high winds in VOS (Figure 7a) can definitely be associated with the spatial inhomogeneity of visual observations. The "VOS–nadir" differences reach −2.8 m/s in the Southern Ocean. However, in general, the distributions are similar and the differences are well within the accuracy of SWIM measurements (<2 m/s; Figure 8c).

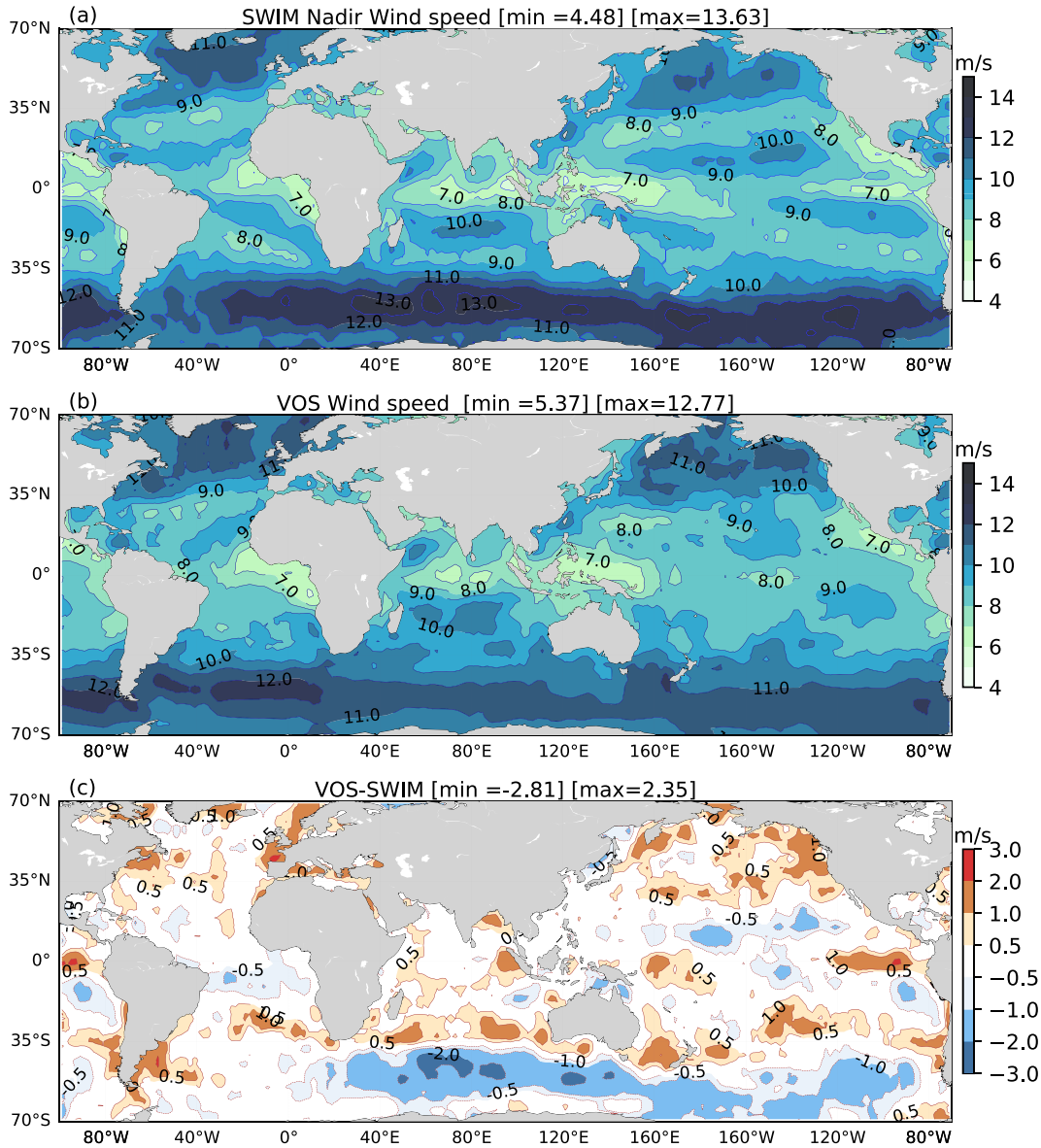

Figure 8. Global annual maps of nadir SWIM and VOS wind speed (m/s) and their differences (2º × 2°) for 2020. (a) Nadir SWIM wind speed; (b) VOS wind speed; (c) the difference between VOS and SWIM.

The comparison of wind speed in collocated points (50 and 25 km radii) also shows a good agreement in terms of mean statistics. The box measurements (Figure 9b) demonstrate a lower dispersion than the nadir-derived winds (Figure 9a). The minor shift of the probability density function toward the VOS axis indicates the underestimation of SWIM box winds that also is found for global histograms for 2019–2021 (Figure 7b). Correlation coefficients are 0.57 and 0.75 (0.59 and 0.77 for a 25 km radius) for nadir and box measurements, correspondingly.

In general, the revealed consistency of wind speed estimates can be seen as a success for both data sets and convincing evidence of a satisfactory accuracy of these types of measurements.

4. Discussion

We have compared two global independent data sets—satellite measurements from the innovative SWIM instrument and visual observations from VOS—on the same time scales: April 25, 2019 to March 31, 2021. The number of SWIM records is more than four times larger than the number of VOS observations ($2.7*10^6$ vs. $0.6*10^6$).

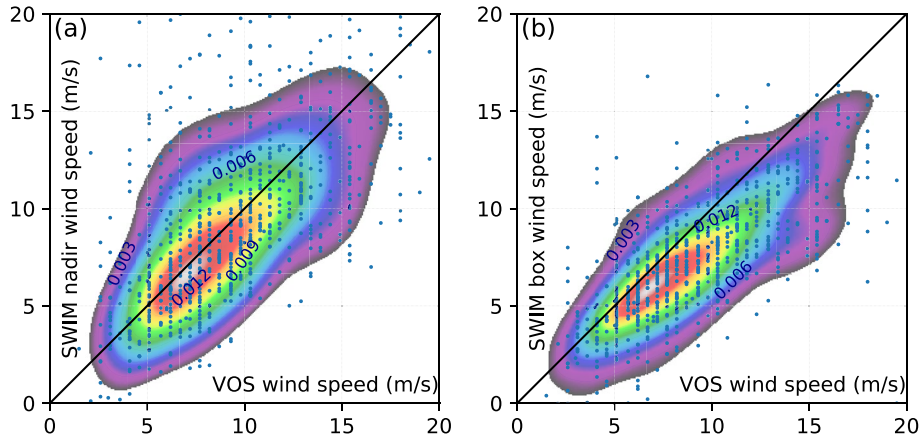

Figure 9. Two-dimensional probability density functions for wind speed from VOS and SWIM within a 50-km radius and a 30 min time lag. (a) VOS and nadir wind speed; (b) VOS and box wind speed. The probability density function was estimated using 2D Gaussian kernels.

Both numbers guarantee robust statistics according to the LLN (Law of Large Numbers), but VOS data still is at a disadvantage, mostly from spatial inhomogeneity, which might distort global estimates (Gulev et al., 2003).

To check the reliability of our statistical estimates, we have supplemented the performed validation of 2-year data with a comparison of the SWIM measurements and climatological VOS statistics for 40 years (1982–2021). Extending the time period has also allowed us to add visually observed secondary swell to the list of compared parameters and check its conformity with the SWIM wave partitions. Generally, there is one primary swell system. However, at times two (or even more) independent swells generated by different weather events can be observed. The number of VOS records containing secondary swell height is only 35,000 for 2019–2021, and $1.1*10^6$ for 1982–2021. The number of wind sea, first swell, and wind speed records for 1982–2021 exceeds $20*10^6$. Figure 10 presents global histograms of SWIM wave parameters and the climatological ones from VOS.

The distributions of climatological VOS SWH1 and SWIM nadir H_s (Figure 10a) demonstrate a nearly perfect agreement, quite similar to 2019–2021 (Figure 4a). VOS climatological wind sea heights slightly overestimate 0–1 m waves of the first partition and underestimate its 3–6 m waves (Figure 10b). For 2019–2021, they coincide much better (Figure 2a). VOS climatological first swell heights also agree well with the first partition, especially for waves higher than 3 m (Figure 10c). The same cannot be said for VOS swell and the first partition heights for 2019–2021 (Figure 2b). No similarity was found for the second partition heights and visually observed secondary swell (Figure 10d), as well as for the third partition (not shown in the figures).

When it comes to wind speed, we see almost identical distributions of climatological VOS observations ($>20*10^6$ records) and the 2-year box SWIM measurements (Figures 11b). VOS data perfectly captures moderate and strong winds with a slight underestimation of winds of less than 4 m/s. The correlation between VOS and nadir winds has not dramatically changed when moving to the climatological scale (Figures 11a). VOS overestimates moderate winds of 2–10 m/s and underestimates strong winds of 12–20 m/s.

The comparison of SWIM statistics with climatological VOS distributions confirms our assumption that spatial inhomogeneity of visual observations affects the distribution of wave heights, especially for wind sea and swell systems. Moreover, the dearth of observations probably causes the discrepancies in the SWIM box and VOS winds. Another reason for the underestimation of wave and wind characteristics in VOS, especially the high values, is connected with the tendency of ships to avoid stormy conditions, the so-called, “fair weather bias” (Gulev et al., 2003). Nevertheless, this experiment clearly shows the robustness of significant wave height statistics regardless of VOS data features and time scales (whether for several years or on climatological scales).

The analysis of wave partitions from the SWIM in relation to visually observed wind sea and swell characteristics is more sophisticated, first of all, because of the radically different approaches of wave partitioning used. Wave system discrimination in VOS is inherently based on joint analysis of wind and wave fields. The observer's estimation is also dependent on surrounding parameters (say, place of observation, ship, stationary point, etc.). Even

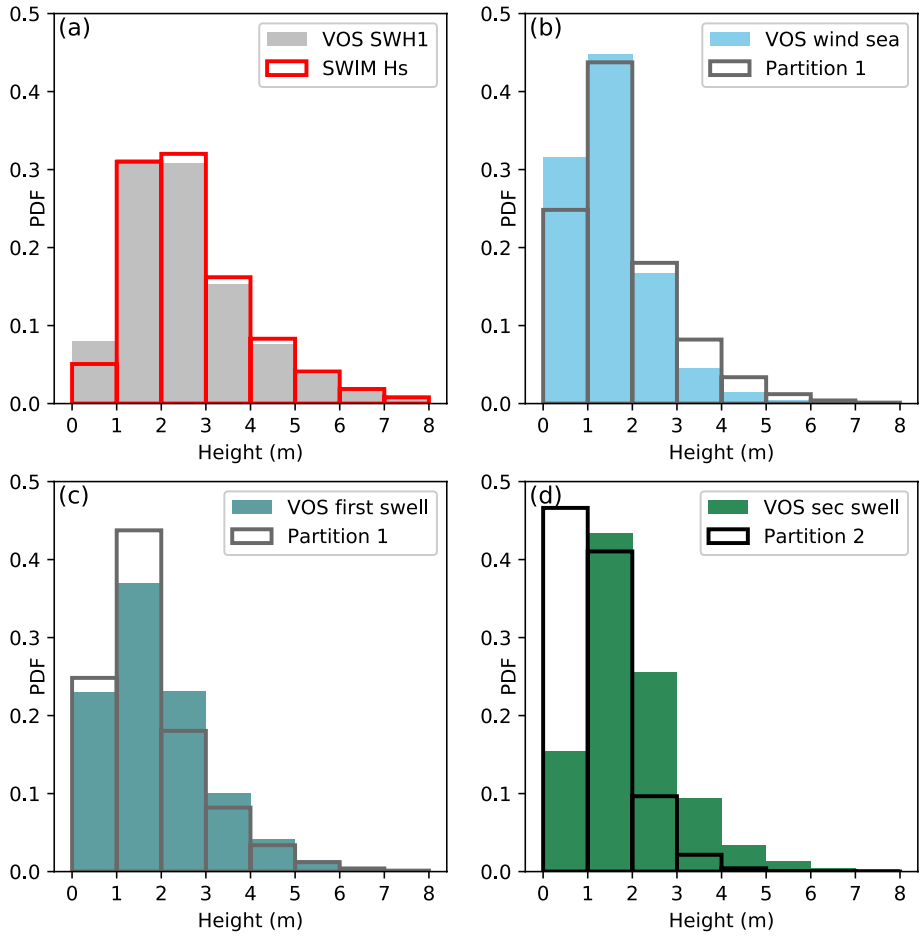

Figure 10. Global histograms of SWIM wave heights for 2019–2021 and VOS wave heights for 1982–2021. (a) VOS SWH1 (gray) and H_s (red lines); (b) VOS wind sea height (blue) and the first partition (gray lines); (c) VOS first swell height (cyan) and the first partition (gray lines); (d) VOS secondary swell height (dark green) and the second partition (black lines).

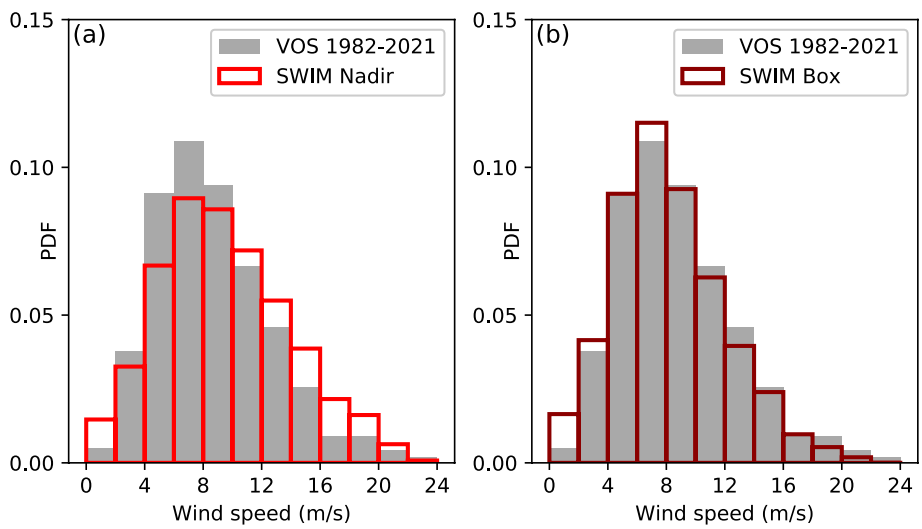

Figure 11. Global histograms of wind speed from VOS (1982–2021) and SWIM (2019–2021). (a) VOS (gray) and nadir SWIM (red lines) wind speed; (b) VOS (gray) and box SWIM wind speed (brown lines).

though the approach is essentially qualitative, it is possible to integrate all of these factors into a comprehensive observation. In contrast, the SWIM wave spectra partitioning follows formal magnitude criteria that can merge close but still physically different wave systems or, in some cases, break apart a single pattern. The ongoing problem of the wave directions ambiguity in SWIM data can enhance these effects of the formal algorithm of partitioning. Addressing this problem in the near future (Le Merle et al., 2021) will help develop advanced approaches that take wind direction into account and allow for combining physically related wave systems (Hasselmann et al., 1994; Hanson & Phillips, 2001).

5. Conclusions

For the first time ever wave and wind parameters derived from CFOSAT measurements have been validated against visual observations from VOS.

Significant wave height as the number one sea state parameter, in general, tends to demonstrate a high coincidence between all data sets, whether they are model simulations, altimetry, or buoy measurements. The VOS and SWIM data are not an exception when it comes to global mean statistics. But the comparison of individual instantaneous measurements in collocated points indicates a large dispersion of significant wave height values and a relatively low correlation coefficient between VOS and SWIM data sets. Presumably, the reasoning for these results lies in the features of visual observations when the significant wave height is calculated from separately estimated wind sea and swell heights with a rather low 0.5 m discreteness.

Cross-comparison of the SWIM wave partitions with wind sea and two swell systems of VOS was the most intriguing part of the validation. CFOSAT is the first satellite mission providing information on three wave partitions due to the novel SWIM instrument. Visual observations are the only source of direct estimates of wind sea and swell characteristics and, therefore, can be treated as the "ground truth" for this comparison. The comprehensive analysis showed that none of the partitions can be confidently attributed to wind-driven waves or swells, despite the good agreement found for the first partition with both wind sea heights for 2019–2021 (Figure 2a) and climatological swell heights (Figure 10c). Most likely, the agreement stems from two factors that compensate each other when assessing the global mean statistics. The first factor is the inability of the SWIM to "see" short waves, that is, young growing wind seas. And the second one is the spatial inhomogeneity of VOS data and systematic underestimation of visual wave heights due to several reasons indicated above. We believe the first partition aggregates both swells and mature wind seas which fall into the SWIM operational wavelength range. This assumption is confirmed by a coincidence of the SWIM integral wave height (the sum of the partitions) and SWH2 (Figures 4b, 5f and 6b), which is determined as the maximum of the observed wind sea and swell heights. Regional analysis in the areas with pronounced wave regimes (e.g., stormy regions and swell pools) will probably justify these findings. Promising results of similar nature have been reported recently (Aouf et al., 2021).

We found no noticeable agreement between the second and third partition heights with VOS wave systems, including secondary swell observations.

The analysis of the CFOSAT directional steadiness showed inconclusive results most likely due to imperfect methods of data preprocessing. We believe that the SWIM wave system partitioning can be improved if validated against visual observations, even despite their existing limitations, which can be alleviated by using wind wave data from other sources as well (models, reanalyses, and buoys).

A comparison of the VOS wind data with the SWIM winds, both nadir and box, showed a reasonable agreement for global statistics and the highest correlation of data sets in collocated points. This result, perhaps, can be regarded as unexpected, since VOS winds have never been considered as a reliable data source for validation.

To finalize this article, we would like to stress that one of the major problems that ocean studies continue to face is the scarcity of data from vast areas of the World Ocean required for the needs of basic weather forecasting, the provision of marine services, and wind wave climate research. The new generation of satellites and the CFOSAT mission, in particular, can help us overcome these problems. Visual observations continue to play an important role in validating results from other data sources.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Original VOS data from International Comprehensive Ocean-Atmosphere Data Set (ICOADS) Release 3, Individual Observations are available by courtesy of NOAA/NCEI at https://rda.ucar.edu/datasets/ds548.0/. Explicit reference for ICOADS data set is: "Research Data Archive/Computational and Information Systems Laboratory/National Center for Atmospheric Research/University Corporation for Atmospheric Research, Physical Sciences Laboratory/Earth System Research Laboratory/OAR/NOAA/U.S. Department of Commerce, Cooperative Institute for Research in Environmental Sciences/University of Colorado, National Oceanography Centre/University of Southampton, Met Office/Ministry of Defence/United Kingdom, Deutscher Wetterdienst (German Meteorological Service)/Germany, Department of Atmospheric Science/University of Washington, Center for Ocean-Atmospheric Prediction Studies/Florida State University, and National Centers for Environmental Information/NESDIS/NOAA/U.S. Department of Commerce (2016), International Comprehensive Ocean-Atmosphere Data Set (ICOADS) Release 3, Individual Observations, https://doi.org/10.5065/D6ZS2TR3, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO." All CFOSAT products, including wind/wave data are provided by courtesy of CNSA and CNES (available for download at http://ftp-access.aviso.altimetry.fr/cfosat). Data are also available at re3data.org repository (https://www.re3data.org/repository/r3d100011694) with the explicit reference: AVISO+; editing status 2021-10-08; re3data.org - Registry of Research Data Repositories, http://doi.org/10.17616/R3B332, last accessed: 2021-12-17.

Acknowledgments
All CFOSAT products are provided by courtesy of CNSA and CNES. We appreciate ICOADS team and personally Eric Freeman for the long-term maintaining the VOS data and making it available to all researches. SB was supported by RSF grant #19-72-30028. VG and SG benefited from the RSF grant #20-17-00139 (developments of validation algorithms). SG also benefited from RFBR grant 20-55-75002.

References

Aouf, L. (2020). *Quality information document for global ocean waves analysis and forecasting product GLOBAL_ANALYSIS_FORECAST_WAV_001_027* 2. Copernicus Marine Environment Monitoring Service. Issue 1.

Aouf, L., Dalphinet, A., Hauser, D., Delaye, L., Tison, C., & Chapron, B. (2019). On the assimilation of CFOSAT wave data in the wave model MFWAM: Verification phase. In *Proceedings of IGARSS 2019 – IEEE international geoscience and remote sensing symposium* (pp. 7959–7961). https://doi.org/10.1109/IGARSS.2019.8900180

Aouf, L., Hauser, D., Chapron, B., Toffoli, A., Tourrain, C., & Peureux, C. (2021). New directional wave satellite observations: Towards improved wave forecasts and climate description in Southern Ocean. *Geophysical Research Letters*, *48*, e2020GL091187. https://doi.org/10.1029/2020GL091187

Chen, G., Chapron, B., Ezraty, R., & Vandemark, D. (2002). A global view of swell and wind sea climate in the ocean by satellite altimeter and scatterometer. *Journal of Atmospheric and Oceanic Technology, CO*, *19*(11), 1849–1859. https://doi.org/10.1175/1520-0426(2002)019<1849:agvosa>2.0.co;2

Freeman, E., Woodruff, S. D., Worley, S. J., Lubker, S. J., Kent, E. C., Angel, W. E., et al. (2017). ICOADS release 3.0: A major update to the historical marine climate record. *International Journal of Climatology*, *37*(5), 2211–2232. https://doi.org/10.1002/joc.4775

Gourrion, J., Vandemark, D., Bailey, S., Chapron, B., Gommenginger, C. P., Challenor, P. G., & Srokosz, M. A. (2002). A two-parameter wind speed algorithm for Ku-band altimeters. *Journal of Atmospheric and Oceanic Technology*, *19*, 20302–22048. https://doi.org/10.1175/1520-0426(2002)019<2030:atpwsa>2.0.co;2

Grigorieva, V. G., & Badulin, S. I. (2016). Wind wave characteristics based on visual observations and satellite altimetry. *Oceanology*, *56*(1), 19–24. https://doi.org/10.1134/S0001437016010045

Grigorieva, V. G., Gulev, S. K., & Sharmar, V. D. (2020). Validating ocean wind wave global hindcast with visual observations from VOS. *Oceanology*, *60*(1), 9–19. https://doi.org/10.1134/S0001437020010130

Gulev, S. K. (1999). Comparison of COADS release 1a winds with instrumental measurements in the North-west Atlantic. *Journal of Atmospheric and Oceanic Technology*, *16*, 133–145. https://doi.org/10.1175/1520-0426(1999)016<0133:cocrww>2.0.co;2

Gulev, S. K., Grigorieva, V., Sterl, A., & Woolf, D. (2003). Assessment of the reliability of wave observations from voluntary observing ships: Insights from the validation of a global wind wave climatology based on voluntary observing ship data. *Journal of Geophysical Research: Oceans*, *108*(C7), 3236. https://doi.org/10.1029/2002JC001437

Hanson, J. L., & Phillips, O. M. (2001). Automated analysis of ocean surface directional wave spectra. *Journal of Atmospheric and Oceanic Technology CO*, *18*(2), 272–293. https://doi.org/10.1175/1520-0426(2001)018<0277:aaoosd>2.0.co;2

Hasselmann, S., Hasselmann, K., & Bruning, C. (1994). In Extraction of wave spectra from SAR image spectra. In G. J. Komen, et al. (Eds.), *Dynamics and modelling of ocean waves*. Cambridge University Press (ISBN: 0-521-47047-1), (pp. 391–401).

Hauser, D., Tison, C., Amiot, T., Delaye, L., Corcoral, N., & Castillan, P. (2017). Swim: The first spaceborne wave scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(5), 3000–3014. https://doi.org/10.1109/TGRS.2017.2658672

Hauser, D., Tourain, C., Hermozo, L., Alraddawi, D., Aouf, L., Chapron, B., et al. (2020). New observations from the SWIM radar on-board CFOSAT: Instrument validation and ocean wave measurement assessment. *IEEE Transactions on Geoscience and Remote Sensing*, 1–22. https://doi.org/10.1109/TGRS.2020.2994372

Hogben, N. (1988). Experience from compilation of global wave statistics. *Ocean Engineering*, *15*(1), 1–31. https://doi.org/10.1016/0029-8018(88)90017-0

International Altimetry team. (2021). Altimetry for the future: Building on 25 years of progress. *Advances in Space Research*, *68*(2), 319–363. https://doi.org/10.1016/j.asr.2021.01.022

Jiang, H., Mironov, A. S., Ren, L., Babanin, A. V., & Mu, L. (2021). Validation of wave spectral partitions from SWIM instrument on-board CFOSAT against in-situ data (p. 22). *Earth and Space Science Open Archive*. https://doi.org/10.1002/essoar.10506895.1

Kent, E. C., Rayner, N. A., Berry, D. I., Eastman, R., Grigorieva, V. G., Huang, B., et al. (2019). Observing requirements for long-term climate records at the ocean surface. *Frontiers in Marine Science*, *6*, 441. https://doi.org/10.3389/fmars.2019.00441

Kent, E. C., Taylor, P. K., Truscott, B. S., & Hopkins, J. S. (1993). The accuracy of voluntary observing ships' meteorological observations-results of the VSOP-NA. *Journal of Atmospheric and Oceanic Technology CO*, *10*(4), 591–608. https://doi.org/10.1175/1520-0426(1993)010<0591:taovos>2.0.co;2

Komen, G. J., Hasselmann, S., & Hasselmann, K. (1984). On the existence of a fully developed wind-sea spectrum. *Journal of Physical Oceanography*, *14*, 1271–1285. https://doi.org/10.1175/1520-0485(1984)014<1271:oteoaf>2.0.co;2

Kotel'nikov, V. A. (2006). On the transmission capacity of 'ether' and wire in electric communications. *Physics-Uspekhi*, *49*(7), 736. https://doi.org/10.1070/PU2006v049n07ABEH006160

Krinitskiy, M., Aleksandrova, M., Verezemskaya, P., Gulev, S., Sinitsyn, A., Kovaleva, N., & Gavrikov, A. (2021). On the generalization ability of data-driven models in the problem of total cloud cover retrieval. *Remote Sensing*, *13*(2). https://doi.org/10.3390/rs13020326

Le Merle, E., Hauser, D., Peureux, C., Aouf, L., Schippers, P., Dufour, C., & Dalphinet, A. (2021). Directional and frequency spread of surface ocean waves from SWIM measurements. *Journal of Geophysical Research: Oceans*, *126*, e2021JC017220. https://doi.org/10.1029/2021JC017220

Lebedev, S. A. (2016). Technique for processing satellite altimetry data for the waters of the White, Barents and Kara seas. *MDPI Remote Sensing Earth Space*, *13*, 203–223. https://doi.org/10.21046/2070-7401-2016-13-6-203-223

Maury, M. F. (1854). Maritime Conference held at Brussels for devising a uniform system of meteorological observations at sea, August and September, 1853. In *Proceeding of the maritime conference explanations and sailing directions to accompany the wind and current charts* (6th ed., pp. 54–96). Biddle.

Munk, W. H. (1944). Proposed uniform procedure for observing waves and interpreting instrument records. *Scripps Institution of Oceanography, Wave Project Report*, *26*, 22.

Ren, L., Yang, J., Xu, Y., Zhang, Y., Zheng, G., Wang, J., et al. (2021). Ocean surface wind speed dependence and retrieval from off-nadir CFOSAT SWIM data. *Earth and Space Science*, *8*(6). https://doi.org/10.1029/2020EA001505

Smith, S. R., Alory, G., Andersson, A., Asher, W., Baker, A., Berry, D. I., et al. (2019). Ship-based contributions to global ocean, weather, and climate observing systems. *Frontiers in Marine Science*, *6*, 434. https://doi.org/10.3389/fmars.2019.00434

Smith, S. R., Bourassa, M. A., & Sharp, R. J. (1999). Establishing more truth in true winds. *Journal of Atmospheric and Oceanic Technology*, *16*, 9392–9952. https://doi.org/10.1175/1520-0426(1999)016<0939:emtitw>2.0.co;2

Smith, S. R., Freeman, J. E., Lubker, S. J., Woodruff, S. D., Worley, S. J., & Angel, W. E. (2016). The International Maritime Meteorological Archive (IMMA) format. Retrieved from http://icoads.noaa.gov/e-doc/imma/R3.0-imma1.pdf

Srokosz, M. A., & Challenor, P. G. (1987). Joint distribution of wave height and period: A critical comparison. *Ocean Engineering*, *14*(4), 295–311. https://doi.org/10.1016/0029-8018(87)90029-1

Sverdrup, H. V., & Munk, W. H. (1947). *Wind, Sea, and Swell: Theory of Relations for Forecasting* (Vol. 60). Hydrographic Office Pub. US Navy.

Von Schuckmann, K., Le Traon, P.-Y. N., Smith, A., Pascual, S., Djavidnia, J.-P., Gattuso, M., et al. (2020). Copernicus marine service ocean state report. *Journal of Operational Oceanography*, (Vol. 13, pp. s1–s172). (sup1). https://doi.org/10.1080/1755876X.2020.1785097

Wilkerson, J. C., & Earle, M. D. (1990). A study of differences between environmental reports by ships in the voluntary observing program and measurements from NOAA buoys. *Journal of Geophysical Research, C: Oceans Atmosphere*, *95*(3), 3373–3385. https://doi.org/10.1029/JC095iC03p03373