

Specific comments

Abstract:

“Measurements with SeaVision were quality controlled and validated by comparison with Spotter buoy data and WaveWatch III experiments.”

One of my main criticisms to this work is the use of wave modelling for validation of in-situ observations. It needs to be considered that numerical models provide an approximation of physical phenomena, and they need to be quality controlled against observations, it does not sound logical/scientific to go the other way around.

Section 2.1:

I suggest more elaboration of the expeditions in this section considering that the main purpose of this manuscripts is to present the dataset which has been collected during the expeditions. For example, information such as region, location (including approximate coordinates), date (including day of month), name of the vessel, distance sailed, start and end point, stations, instruments used, metocean parameters collected, general depth etc. should be included in the text.

Figure 1:

Explain what the reason is behind measurement gaps as depicted in panels (a) and (c). Briefly explain cruise numbers (50, 57, 58) in the text of the section. For instance, cruise 50 out of how many cruises per month/year aboard RV Academic Survey Vavilov? Indicate what RV stand for. For instance, by adding “research vessel (RV)” in the caption. The published dataset at <https://sail.ocean.ru/tilinina2021/> seems to be including the measurements at stations shown with orange dots only (Spotter and SeaVision). I wonder why the rest of the observations (SeaVision only measurements shown in green dots) have not been published?

Further, the number of stations shown orange dots are not consistent between this figure, Appendix A, and the published dataset. In total, 52 points are shown Figure 1. However, 50 stations are provided in Appendix A and observations at 48 stations have been published. From Cruise 57, station number 3836 is not included in the dataset but provided in Appendix A. From cruise 50, stations 2771 and 2792 are included in the Appendix but not in the dataset and station 2913 is in the dataset but not in the Appendix.

Line 115: “other regular deep ocean observations”

What are these “other regular deep ocean observations”? Elaborate more or provide examples.

Lines 116 to 119: “In particular, two cruises in the North Atlantic (Figure 1a,b) are related to regular deep ocean observations at the 59,5°N (Verezemskaya et al., 2021; Falina et al. 2007; Gladyshev et al. 2018, 2019; Sarafanov et al. 2008, 2018) and the Arctic expedition is a part of the IO RAS “Floating University of IORAS” program (Stepanova, 2018).”

This sentence is confusing. It is providing information of the first two expeditions versus the “program” which the third expedition is associated with. Also, it has already been mentioned at lines 115-116 that all three expeditions are carried out by IO RAS. So, what is the need to repeat it here about the “Arctic expedition”?

Line 118: “Arctic expedition is a part of the IO RAS”

What is “Arctic expedition” referring to? The expeditions have not been named previously.

Line 121: “During all “stations””

It is not clear what the word “stations” is referring to i.e., no information about any station has been previously defined in the text.

Lines 122 and 123: “to provide conditions for Spotter buoy wave observations in the free floating mode.”

Is “free floating” referring to the deployment of spotter buoy? If yes, it needs to be clarified in the text. Also, explain why this deployment option has been targeted.

Section 2.2:

I think this section can be improved by being organised in a more informative manner. I was expecting a thorough description of SeaVision system and its comparison and superiority against other widely used observational systems such as WaMoS-II as it is indicated in the conclusion that “the commercial oceanographic systems for the wind waves monitoring such as WaMoS II, SeaDarQ and WaveFinder already exist ... [this manuscript’s] main aim is to develop in the nearest future a low cost, independently operating and stable system that would allow to broaden observational network for the wind waves”.

Explain why the dataset only includes one dimensional frequency wave spectrum and no information is provided about the directionality of waves.

Underway observations and their file types deserve a much more comprehensive explanation considering the scope of ESSD journal.

Line 133: “For our purposes we used the shortest possible pulse length”

Explain what the reason is for setting the radar to short pulse length i.e., how the resolution of images may be affected etc.

Line 162: “this results into the three-dimensional spectrum $S(k_x, k_y, f)$ ”

Explain what k_x and k_y are.

Line 164 and 165: The correct expression of the linear dispersion relation is: $\omega = \sqrt{gk \tanh(kh)}$. The angular frequency has been previously defined as ω , why has it been changed to Ω in this equation? The unit of g (gravitational acceleration) should be corrected to $m*s^{-2}$ and H should be replaced with h which is the depth, not wave height!! Also, correct the equation in Figure 3.

Line 166: “the signal (Ω_{BGN}),”

What is this signal? What does “BGN” stand for?

Line 185 and 186: “wave period T_s estimated traditionally using the first moment of the spectrum.”

Provide the equation for the wave period calculations. I also suggest using the standard annotation of wave period based on the first moment “ T_{m1} ” or “ T_{m01} ” instead of “ T_s ”.

Line 191: “Table A1 provides a list of all locations where SeaVision+Spotter buoy (or SeaVision only) measurements were carried out.”

From the published dataset, I can see in all the stations provided in Table A1 both SeaVision and Spotter data exist. No data is available at stations with SeaVision only measurements.

Lines 200 and 201: “Analysis of the raw vertical and horizontal displacements recorded by buoy starts from the selecting in timeseries the “free floating” measurements”

Explain in the text how free-floating measurements have been identified from timeseries.

Line 202: “using common definitions (see Appendix in Raghukumar et al., 2019)”

I suggest adding an appendix that provides definition of all the parameters in the manuscript/published dataset.

Line 203: “in the frequency range of the wind waves.”

I wonder why you have limited yourself to wind waves only, both from the spotter buoy in this section as well as SeaVision and WaveWatch-III in other sections of the manuscript. What is

the reason behind disregarding swell observations? Couldn't it be included in the dataset/manuscript? If not, it needs to be clarified why.

Further, a range of 0.05 to 0.3 Hz seems to be underestimating frequency of wind waves. For example, from Semedo et al., (2009) frequency of wind sea in North Atlantic seems to be above 0.2 to 0.3 Hz in summer when IORAS expeditions have taken place (see Figure 7 at Semedo et al., (2009)). The frequency/ period ranges can also be confirmed from example wave spectrum in Figure 4 of current manuscript.

Please also confirm the frequency range applied to the published data (one dimensional frequency spectra) for calculation of integrated wave parameters provided in the "Global Attributes" of the NetCDF files i.e., Hs_radar , Ts_radar, Hs_buoy, Ts_buoy.

Figure 4:

It should be indicated that this spectrum is an example. Please also provide the station number where this spectrum has been recorded.

From the published dataset, I can see that only one-dimensional frequency wave spectra are provided. I am surprised to see the information about wave directions in the bottom right panel. Is there a point I am missing here?

It may be better to plot the wave spectrum in frequencies instead of periods.

Indicate what PSD stands for. Also, describe what θ_p and σ_{θ_p} are.

I suggest labelling different panels with (a), (b), (c), etc. This also applied to Figures 2 and 3.

Section 2.4:

I cannot find any information regarding the meteorological data in the published dataset at <https://sail.ocean.ru/tilinina2021/>. It would be beneficial to publish the meteorological data (after being quality controlled) together with the wave observations.

Section 2.5:

A WaveWatch-III model for this study need a much more extensive explanation than one paragraph only. Model physics and packages, setup, calibration, and validation should be comprehensively explained. I am surprised to see that no information is provided about model calibration and validation.

Lines 223 to 225: "We run WaveWatch III (WW3DG, version 6.07, WW3) spectral wave model with ERA5 reanalysis (Hersbach et al., 2020) as lateral boundary conditions with 0.1° spatial and 1 hourly temporal resolutions."

What does it mean to run the model with ERA5? The model's forcing and boundary conditions (parameters with their temporal and spatial resolution) need to be clearly explained. It is not clear whether "0.1° spatial and 1 hourly temporal resolutions" refers to the resolution of forcing and boundary conditions or the actual wave model.

Line 238 to 240: "In general, for the lower wind speeds SeaVision underestimates Hs by up to 50 cm and overestimates Hs for the higher wind speeds. This effect can be due to better ripples development on the ocean surface during higher winds affecting the signal to noise ratio (Formula 1)."

I am a bit confused here, from Figure 5(a) the difference between Hs observed with Spotter and SeaVision are less, and it increases for wind speeds more that 8 m/s. The text seems to be indicating the opposite while describing better ripple developments at high winds.

Line 247: "there are two stations (2901 and 2937 see Table A1) where this difference reaches almost 100 cm"

Show these points in Figure 5 by a label and/or different colour.

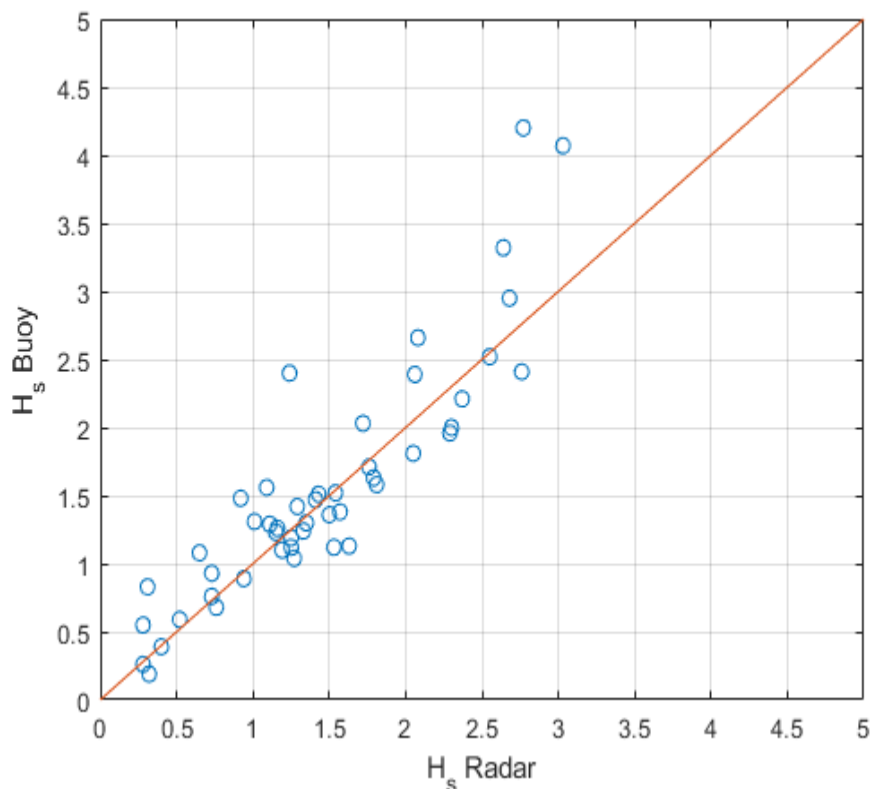
Figure 6 (and lines 253 to 259):

The fitted lines should cross the [0 0] points i.e., logically when Spotter record a wave height (or wave period) equal to zero, SeaVision should return zero, too (it is similarly true about WaveWatch-III estimates).

Overestimation of wave height and period by WaveWatch-III against Spotter measurements may be indicating that the model has not been properly calibrated.

I suggest including other error statistics such as Root Mean Square Error (RMSE) and Scatter Index (SI).

I wonder why the number of points in these plots (~32) is less than the number of stations (~50)? Surprisingly, I get different looking plots from the published dataset while reading the values from “Global Attributes” of each NetCDF file published. Below is the significant wave height from buoy versus radar, for example:



Lines 265 and 266: “waves direction (from)”

Does it mean the wave directions are in “coming from” convention? It needs to be further explained.

Figure 7:

This figure needs a more extensive explanation. The sources of errors and inconsistencies need to be described.

Similar to figure 4, I am surprised to see directional wave roses while no information about directionality of waves is provided in the published dataset, please explain.

Data Availability:

I suggest including a calendar as well as a metadata in the repository to provide information about the dataset each folder/link includes.

Parameters/information available in the dataset and their description (including an extensive description of Global Attributes and Variables in NetCDF files) are worth being added to the

manuscript, probably as an Appendix. Also, indicate the full length of time that the measurements are covering.

Technical corrections

The English language, specifically grammar and punctuations, need to be revised and corrected throughout the text.

Consistent symbols and annotations should be used in the text and figures. For example, the WaveWatch-III wave model is indicated by “WaveWatch III” in some parts and “WW3” in others. Other examples are:

Significant wave height being indicated by “H” and “Hs”, and Figure being indicated by “Fig.” and “Figure”.

Line 135: “SeaVision system (Fig. 2) connected to the radar via splitter, **it** digitizes and records directionally”

Replace with “SeaVision system (Fig. 2), connected to the radar via splitter, digitizes and records directionally”.

Line 154: “FFT-based”

It should be indicated what FFT stands for. Replace with “Fast Fourier Transform (FFT) based”.

Line 175: “to estimates **to**”

Replace with “to estimate”.

Line 176: “spectra power”

Replace with “spectral power”.

Line 121: “During all “stations””

Replace with “**At** all stations”.

Line 190: “waves observations”

Replace with “wave observations”.

Line 199: “we use”

Replace with “we used”.

Line 279: “Nowadays there is still exists gap”

Correct the sentence grammatically.

Line 279: “winds waves is”

Replace with “wind waves **are**”.

Line 280: “component”

Replace with “components”.

Line 283: “wave energy spectra”

Replace with “wave energy **spectrum**”.

Line 289: “2,5 s”

Replace with “2.5 s”.

Line 294: “already exist and successfully operating and providing”

Replace with “already exist and **are** successfully operating and proving”.

Lines 296 and 297: “ships navigating”

Replace with “ship navigations”.

Lines 250 to 252: “Further examination and methodology adjustment required together with more data collection during different conditions in the open ocean are required to investigate into these differences.”

Rewrite this sentence.

Line 255: “worsen”

Replace with “worse”

Line 266: “doesn’t”

Replace with “does not”.

References

Semedo, A., Sušelj, K. and Rutgeresson, A., 2009, September. Variability of wind sea and swell waves in the North Atlantic based on ERA-40 re-analysis. In *Proceedings of the 8th European Wave and Tidal Energy Conference, Uppsala, Sweden, 7-10 September* (pp. 119-129).