Response to Reviewers for Minor Revisions on *Measurements of Nearshore Ocean-Surface Kinematics through Coherent Arrays of Free-Drifting Buoys*

Reviewer Comments in Red
Author Responses in Black

Dear Reviewer #2 and Dr. Marc Pezerat,

Thank you for reviewing this manuscript again, and we appreciate your feedback. Each of your comments is addressed below, and we hope that this will further improve the manuscript.

Reviewer #2:

Overview & Big Picture:
I appreciate the authors changes made to the manuscript based on my and others reviews comments. I believe the manuscript is improved by zooming out a bit and not attempting to provide a rigorous comparison between microSwifts and current meter wave statistics. The addition of the thought provoking Section 4 also adds to the manuscript. Although, I believe the manuscript is improved, and should be published after addressing a few concerns (outlined below) one could argue that less rigor results in a weaker manuscript. For instance, the most rigorous investigation into the quality of the microSwift data centers around Fig 9 and lines 270-297. At line 297, the authors conclude, "The agreement in significant wave height and scalar energy density spectra supports that the Level 2 data are useful for investigating wave spectra and statistics."
This is the authors main scientific finding: for some journals, this would not be enough to warrant publication, but I leave it to the editor of this journal to decide whether this is enough for this journal (as I'm not that familiar with this journal).
We agree that this finding is not a major scientific result; rather, we are presenting a dataset to the community. We are following the mission of this journal, which includes the following, “Earth System Science Data (ESSD) is an international, interdisciplinary journal for the publication of articles on original research data (sets), furthering the reuse of high-quality data of benefit to Earth system sciences. The editors encourage submissions on original data or data collections which are of sufficient quality and have the potential to contribute to these aims.

Articles in the data section may pertain to the planning, instrumentation, and execution of experiments or collection of data. Any interpretation of data is outside the scope of regular articles. Articles on methods describe nontrivial statistical and other methods employed (e.g. to filter, normalize, or convert raw data to primary published data) as well as nontrivial instrumentation or operational methods. Any comparison to other methods is beyond the scope of regular articles.”

Following this mission statement, we focus on the methods used to develop the instruments (microSWIFTs), the field experiment and data collection, the methods in which these data are processed, and how they are organized for future use. We plan to use these data in further studies that will specifically focus on scientific results from the dataset and will not include all of the details of the instruments and dataset as a whole. With this in mind, we believe that the findings presented in this study are appropriate for publication in this journal.

"Bigger Points"

This leads me to bigger point 1 that concerns Fig 9. Note that, "agreement in significant wave height" is based on Fig 9, where Hsig_mS vs Hsig_AWAC is scattered. The best fit line of the data is Hsig_mS = 0.61 Hsig_AWAC. This is quite poor agreement in my opinion, especially as the rms error is .37 m, a large fraction of the average Hsig (approximate 1.75)! The smaller Hsig_mS is attributed to shadowing by the pier. What is the slope if the shadowed values are omitted (gray dots in Fig 9)? The assertion that shadowing gives rise to the <1 slope should be tested. Shadowing can also be tested by investigating whether proximity to the pier matters (by coloring the dots in Fig 9 by distance to the pier?). Also, does it matter whether the waves are from the north or the south and the resulting direction of the mS? This should be explored in more detail and differences between Hsig_AWAC and Hsig_mS explored in more detail.
Shadowing only occurs when waves come from an oblique angle, and the microSWIFTs are on the opposing side of the pier from the direction of the waves. Therefore, the metric that we present in Figure 9 includes an investigation of the proximity to the pier (within the shadow must be within 200 meters of the pier) as described in lines 300-302, “Being in the pier ‘shadow’ is defined here as missions when the average location of the microSWIFTs during a mission is within 200 meters of the pier, and waves are coming from the other side of the pier based on the mean wave direction from the 8-meter array (furthest offshore sensor).”

If the shadowed values are omitted, the slope of the linear regression is 0.53, which suggests worse agreement; however, omitting all of the shadowed points reduces our sample size to 71% of the available data, which could lead to weaker statistics. We also do not expect perfect agreement between the microSWIFT arrays and the AWAC measurements since they measure at different locations, and the microSWIFT arrays measure waves in many different depths since we use data only outside of the approximated surf zone. The comparison should rather capture the general trends, microSWIFTs measure large waves in large wave conditions and small waves in small wave conditions, primarily what we see in this comparison. Further offshore, the microSWIFTs have shown much stronger agreement with nearby instruments for the significant wave height, as shown in Figure 5 of Thomson et al. 2023 (In Review, https://github.com/SASlabgroup/SWIFT-codes/blob/master/Documents/microSWIFTs_CEJrevision_6Sep2023.pdf).

The authors also state that the difference is due to a short time series. I suspect not, as addressed later in this review. Also, the AWAC is in about 4.5 m of water, which according to gamma=0.35 in the manuscripts means that waves bigger than 1.5 m are breaking at this AWAC. I think only Hsig when BOTH instruments are outside the surfzone should be compared. Hsig_mS in a region of breaking waves will not be reliable (as mentioned by the authors) as mS surf broken waves. Thus, the "good" part of the plot, Hsig_AWAC < 2 m, the relationship between mS and 4.5 m AWAC isn't very good. It doesn't make sense to include data when the mS might be surfing. If the mS might be surfing, that data should not be included in this

comparison. Unfortunately, the more I stare at Fig 9, the more I'm not so sure that microSwifts can tell me anything accurate about wave heights. If they can't get this statistic really well, what does that mean for other higher order statistics?
The choice of gamma=0.35 is a conservative estimate for gamma, suggesting that the surf zone may be narrower than our surf one edge estimate, and the location of the 4.5 meter AWAC should still be in a zone of intermittent breaking even under larger wave conditions. This was confirmed by visual observations during data collection. The caption associated with Figure 9 is now revised as , "Comparison of the estimated significant wave heights from the microSWIFT arrays, 6-meter AWAC, and 8-meter pressure sensor array (6-meter AWAC and 8-meter array have been corrected for shoaling) to the estimates from the 4.5 m AWAC. While the microSWIFT arrays are not in the same water depth as the 4.5 m AWAC, we see that the microSWIFT values are similar to the 4.5 m AWAC values. The gray bars indicate 95% confidence intervals around each of the significant wave height estimates, computed using a bootstrap method from the distributions of wave heights. The colors of the estimates depict if the microSWIFT array is in the 'shadow' of the pier, where we expect a reduction in wave energy. For significant wave heights greater than 2 meters, intermittent breaking may be occurring at the 4.5 meter isobath, leading to worse agreement between the AWAC and microSWIFT measurements."

We agree that the data from the microSWIFTs may not be well-suited to investigating higher-order statistics. This is why the manuscript has been restructured to focus on the kinematics of the surface and what we can learn from the Level 1 data with examples given in section 4 of the manuscript, including investigating 'surfing' transport of buoyant objects, surface kinematics, and spatial variability of breaking waves. We do not claim that the microSWIFT arrays are a precise tool for measuring wave statistics. Instead, we suggest that microSWIFT wave heights are useful contextual data when investigating breaking/surfing kinematics using the raw motion data.

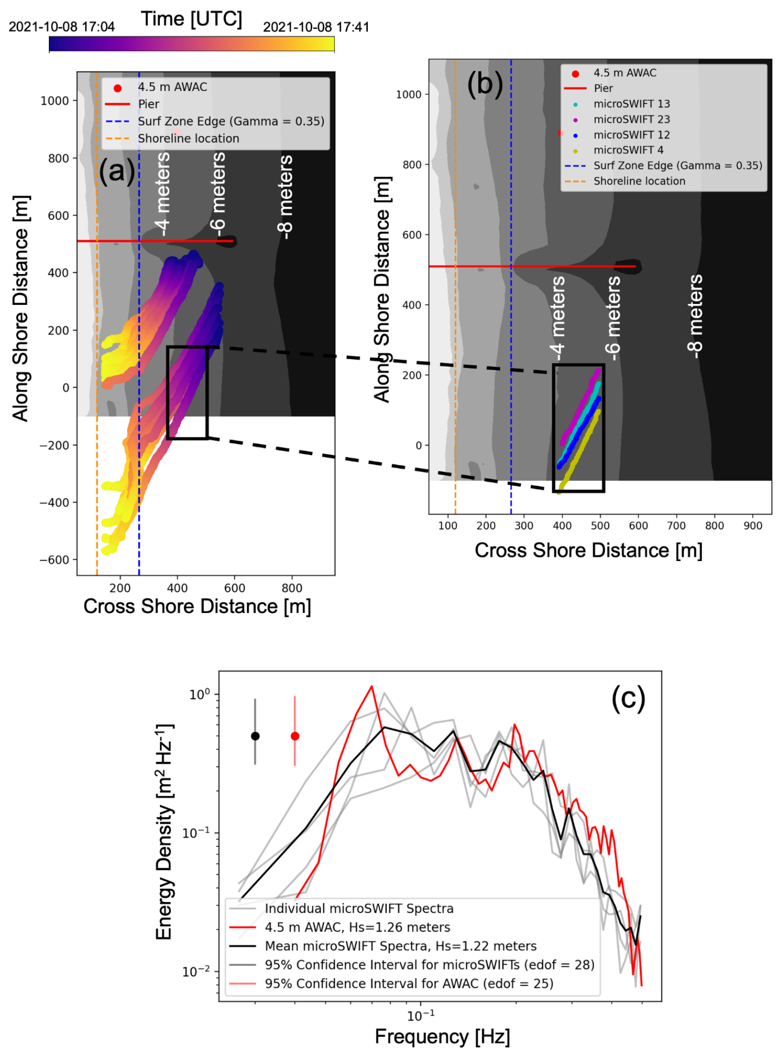
Although this doesn't affect the quality of the paper, because the spectra are only used qualitatively, I still contend that the EFD (effective degrees of freedom) isn't correct for the spectra calculated in this MS. And it should be done correctly. First, notice that in Fig 7 c, the variability of each spectral estimate is bigger than the 95% bars. This means that each peak in the spectra is real, which I doubt. I believe that the 95% bar should be longer more consistent with the variability within the spectra. I.e. the EDF used is too large. The authors state that for the microSwift the EDF is given by equation (2) in the MS,
EDF = (8/3) N/M
(from Thomson and Emery, 2014 table 5.5, but this is actually from Priestley 1981) where N is the number of points in the time series, and M is the 1/2 width. They use N=7200 (600 s x 12 Hz) and M=1800. Then 5 frequencies are averaged. So
EDF = (8/3) * (7200/1800) * 5 = 53. I don't think (8/3) N/M is being correctly used. Either that, or the formula itself is incorrect. I'm not sure, because these formula in T&E are not derived so it isn't clear where N and M come from. Regardless, for a spectra with 3 non overlapping blocks of data, the dof=6 and overlapping the blocks of data reduces the number of degrees of freedom. Thus for 3 blocks of data, the MAXIMUM EDF = 6 and then averaging 5 frequencies would yield

30 dof. The authors can not have more than 30 dof. T&E hint at this on page 476... " Spectra are then computed for each of the K segments and the spectral values for each frequency band then block averaged to form the final spectral estimates for each frequency band. If there is no overlap between segments, the resulting DoF for the composite spectrum will be 2K. This assumes that the individual sample spectra have not been windowed and that each spectral estimate is a chi-squared variable with two DoF."

I believe T&E can be confusing, especially table 5.5. EDF are considered in a variety of places. EDFs are derived in
http://pordlabs.ucsd.edu/sgille/sioc221a/lecture11_notes.pdf
and I highly suggest looking at this doc. Here, it is clearly shown that the EDF only depends on the number of blocks (aka segments or chunks ==Nb, "K" in T&E) averaged to make the spectra (which does not depend on N the number of samples within the block). With no overlap, EDF = 2*Nb (as outlined above), but since there is overlap, and a Hanning window is used, 2 becomes 1.9 and
EDF = 1.9 * Nb = 1.9 * 3 = 5.7
but 5 frequencies are averaged so EDF = 5.7 * 5 ~= 29.
The 53 stated in the MS is not the number of degrees of freedom for this spectra. Note, if N/M = 2, i.e. the window is 1/2 the length of the entire time series, which it is, then 8/3 *2 = 5.33 which is similar to the 5.7 above. Also note that in the above linked pdf it is stated that, "So what of the other texts? The 2014 edition of Thomson and Emery is as misleading as the earlier editions." I believe this is in reference specifically to Table 5.5 of T&E, so according to Gille, who I trust, maybe table 5.5, where the 8/3 N/M comes from, isn't the best reference regarding spectra dof? However, T&E state, "Nuttall and Carter (1980) report that 92% of the maximum number of equivalent degrees of freedom (EDoF) can be achieved for a Hanning window, which uses 50% overlap." I.e. 6 becomes 6*.92 = 5.52, and 5.52*5 = 28 dof. One less than the Gille formula.

For the AWAC, assuming a Hanning window,
EDF = 1.9 * 13 ~= 25 dof
not 42. Again, the larges EDF for the AWAC is 2*13 = 26. but the overlapping blocks result in slightly less.
This has been corrected. The following shows the adjusted Figure 7 with the equivalent degrees of freedom for the microSWIFTs equal to 28 and the equivalent degrees of freedom for the AWAC equal to 25. Lines 242-252 have been adjusted accordingly and are now the following. “The microSWIFT spectra are computed using Welch’s method, with five-minute (3600 sample) Hanning windows and 50% overlap between adjacent windows. The energy in each of the five adjacent frequencies is band-averaged to improve the statistical robustness of each estimate. The equivalent degrees of freedom for each microSWIFT spectrum is 28. This is based on the ten-minute time series (7200 samples at a 12 Hz sampling rate) used for each spectral estimate with three five-minute windows (50% overlap). Each window contributes 2 degrees-of-freedom and band-averaging the five adjacent frequencies increases the effective degrees of freedom by a factor of five. Due to the 50% overlap of the Hanning windows, the equivalent degrees of freedom are reduced to 92% of the maximum degrees of freedom (Nuttall and Carter, 1980). Therefore, the equivalent degrees-of-freedom for the microSWIFT spectra is 28 (3 windows * 2

degrees-of-freedom * 5 frequency bands * 0.92 = 28). The AWAC measurements consist of a 34-minute record with a sample rate of 2 Hz, and spectra are computed with 13 50%-overlapping windows (512 points per window) and no band-averaging, leading to approximately 25 degrees-of-freedom, comparable to that of the microSWIFTs (Christou et al., 2011)."



"Smaller Points"

Line 290: "We also expect that the microSWIFT arrays may under-predict some significant wave heights as the sampling windows are shorter than the AWAC, potentially not measuring the largest and least likely waves in the distribution and times that the microSWIFTs are within the 'shadow' of the pier."

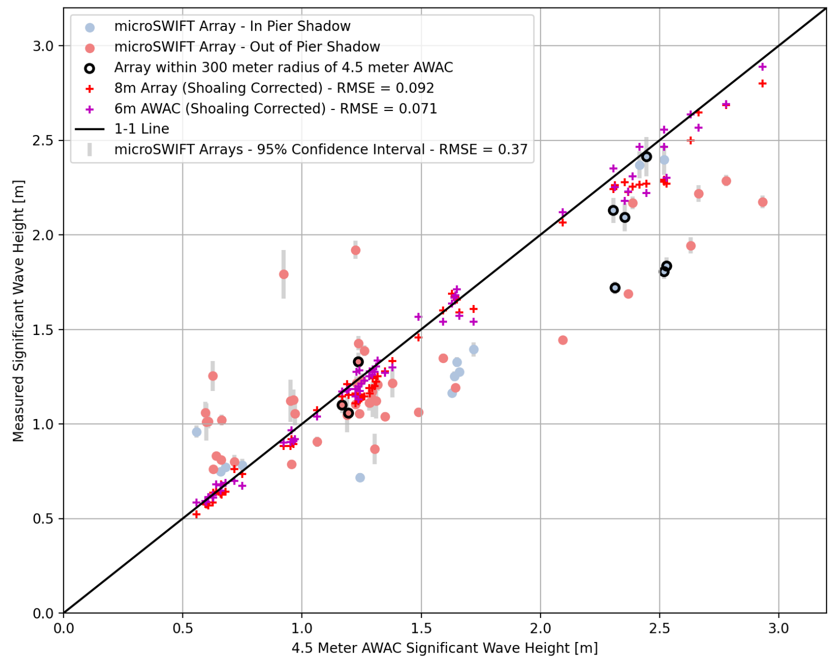
This seems a bit misleading. The shortness of the time series doesn't bias the difference between mS and AWAC Hsig, it just creates more variability. The authors could have also said,

"We also expect that the microSWIFT arrays may over-predict some significant wave heights as the sampling windows are shorter than the AWAC, potentially over representing the largest and least likely waves in the distribution and not measuring enough of the smaller waves."
This has been corrected. Lines 297-300 are now the following. “We also expect that the microSWIFT arrays have more variability in their significant wave height estimates since the sampling windows are shorter than the AWAC, potentially over-representing or under-representing the largest and least likely waves in the distribution. Further underestimation could be due to the microSWIFTs being within the ‘shadow’ of the pier.”

Line 323: Fig 11a. Is this GPS u? Or the Kalman filtered u? Hopefully the Kalman filtered u. If GPS velocities, are they de-spiked? Fig 11b. Is the acceleration in the vertical reference frame? If not it should be as the "body frame of reference" is not as obviously useful. The Kalman filtered velocities and accelerations should be used in this figure.
The velocities shown in Figure 11a are GPS velocities that have been de-spiked. The GPS velocities are already measured in the Earth reference frame, so we chose to use these velocities with less associated processing. Additionally, Figure 11b shows the vertical accelerations in the “body frame of reference” and have also been de-spiked. Again, following the same argument as for the GPS velocities, this is a less processed measurement (i.e., Level 1 data). These choices align with the reframing of the paper as a whole to focus more on the lightly processed measurements and what can be learned from them.

Line 273. 1.416 should be 1.414 as it is 2^{1/2}.
This has been corrected. Lines 279-280 are now the following. “The significant wave height is computed by first computing the root-mean-square of the wave heights and then multiplying by a factor of 1.414 to convert to significant wave height for a Rayleigh distribution (Dean and Dalrymple, 1991).”

Also, does your boot strapping method yield
Hs [1 - 0.41 / N^(1/2)] < Hs < Hs [1 + 0.41/N^(1/2)]
for the 95% confidence limits? where N is the number of waves in the estimate. I believe these are the 95% confidence limits of Hs based on a Rayleigh distribution and should be confirmed by bootstapping. In my opinion, it is better to use a derivable formula than just say, "we got this number by bootstrapping" as there is no way for a reader to determine if that number is correct.
The boot-strapping method does not exactly yield these confidence intervals, though we agree that they should converge as the square root of the number of realizations. We show the alternative version of Figure 9 below, using the equation from the reviewer. These intervals are smaller than the boot-strap method, presumably because the empirical boot-strapping includes other sources of uncertainty inherent to the data. We prefer to use these wider intervals, rather than rely purely upon the theoretical Rayleigh distribution (though the shape does seem to match our aggregate data, as shown in Figure 8e). In either case, the statistical uncertainty in each Hs estimate from the microSWIFT buoy arrays is insufficient to fully account for the mis-match with the AWAC observations. Again, we attribute these differences to the lack of colocation and alongshore variability.


microSWIFT Array - In Pier Shadow
microSWIFT Array - Out of Pier Shadow
Array within 300 meter radius of 4.5 meter AWAC
8m Array (Shoaling Corrected) - RMSE = 0.092
6m AWAC (Shoaling Corrected) - RMSE = 0.071
1-1 Line
microSWIFT Arrays - 95% Confidence Interval - RMSE = 0.37
Measured Significant Wave Height [m]
4.5 Meter AWAC Significant Wave Height [m]
0.0
0.5
1.0
1.5
2.0
2.5
3.0

Dr. Marc Pezerat Comments:

l.5 “the buoy’s global position” -> why global ?

Here, we wanted to be precise that the positions are from the GPS receiver and are, therefore, global positions. However, this may be unnecessary to specify, so it has been removed from the manuscript, and the line is now the following. “The microSWIFT is a small buoy equipped with a GPS module to measure the buoy’s position and horizontal velocities and Inertial Measurement Unit (IMU) to directly measure the buoy's rotation rates, accelerations, and heading.”

l.5 and 6 “horizontal velocities and accelerations” -> You should precise that you consider horizontal velocities obtained from GPS measurements not integrated from accelerations (according to l. 319).

This has been corrected in conjunction with the previous correction. See the comment above.

l.32 to 40 Why referring to modelling, it is a bit misleading, as beyond the scope of this study.

This section is included as a transition between the background and how the dataset can be used in the future. We believe there are still many questions regarding how phase-resolved wave processes are included in phase-averaged models and that the data presented in this paper may be useful in answering those questions. While this is not a modeling study, this section is a useful primer for where this data can be used in the future.

Somewhere around l. 50 It would be nice to distinguish free-drifting and moored buoy’s measurements.

A further distinction between free-drifting and moored buoy’s measurements was added between lines 49-55. The adjusted lines are the following. “As a complement to the fixed sensors and remote sensing methods, wave buoys are another option for obtaining direct measurements of the surface kinematics in various sea states. Wave buoys can be either free-drifting or moored. Moored buoys are effectively Eulerian wave measurements, with some movement due to the scope of the mooring, while free-drifting wave buoys are closer to Lagrangian measurements but move as a result of the wind, currents, wave-induced drift (Stokes drift), and surfing on broken waves. Free drifting buoys are essential for understanding how buoyant objects move in the nearshore (Spydell et al., 2007; Schmidt et al., 2003). Free-drifting buoys tend to move through the surf zone very quickly; prior studies have reported buoys reaching approximately 50 cm s^{-1} as a mean drift velocity.”

l. 68-69 “they are the only tool…” -> not the only one.

This has been corrected, and the sentence is now the following. “While buoys have inherent challenges in measuring nearshore waves, including distortion of surface elevation from accelerometer measurements (Magnusson et al., 1999) and inability to resolve second-order non-linearity (Forristall, 2000), they are one of the few tools that can be used to obtain direct measurements of the kinematics of the surface”

l.158-159 “Using this definition of γ_s , the variable H_s represents the offshore significant wave height (will use measurements from the 8-meter pressure gauge array, location in Figure 1,

panel (c))" -> Perhaps "… the variable Hs represents the offshore significant wave height, here measured from the 8 m-pressure gauge array…"
This sentence has been adjusted in the manuscript. The line is now the following. "Using this definition of γs, the variable Hs represents the offshore significant wave height, here measured from the 8 m-pressure gauge array (location in Figure 1, panel (c)), and the variable d represents the water depth during the mission."

Section 3. Could you introduce subsections to ease the reading?
Four subsections have been added to section 3. The added subsections are titled Data Cleaning and Level 1 Data, Level 2 Data, Spectral Exploration of microSWIFT Data, and Zero-Crossing Exploration of microSWIFT Data.

l. 271 "The significant wave height is computed from aggregated wave height measurements outside the approximate surf zone" -> So, If I understand correctly Fig. 8e is built from elevation timeseries when buoys are located seaward of the surf zone edge?
This is correct. Figure 8e is built from wave realizations when buoys are located seaward of the surf zone edge estimate. To clarify this in the manuscript, the caption is now the following. "Example of steps in processing each mission. Panel (a) shows the drift tracks of the microSWIFTs from mission 19 plotted over the surveyed bathymetry DEM. Panel (b) shows the same drift tracks as Panel (a) but shows each microSWIFT as a different color. Panel (c) shows the time series of computed sea surface elevation, with one time series being highlighted as an example. Panel (d) is a zoomed-in portion of the overall time series showing the locations of zero crossings and how we define the height of an individual wave in a time series. Panel (e) is the probability density of all wave heights, seaward of the approximate surf zone edge, where the colors show the contribution from each microSWIFT with the corresponding color. The probability density distribution fits a Rayleigh distribution. The vertical line shows the computed significant wave height for this distribution and the 95% confidence interval of the estimate."

l. 306 "These measurements can [help] investigate…"
This has been corrected, and line 306 is now the following: "These measurements can help investigate buoyant particles' cross and along-shore transport under various forcing conditions."