We thank Kevin Martins for the detailed review and very thorough look at the dataset. The suggestions improved the dataset and accompanying manuscript. Below we provide our responses to the reviewer comments. Original reviewer comments are shown in **black font**. Our responses and modifications affected in response to reviewer comments are shown in **green font**. Modifications in the manuscript are shown in the **blue font**.

General/Major comments:

Overall, the manuscript is relatively well written and organised, but I found that the introduction could be improved. I feel that the authors fail to find the right balance between the specificity of the site, and how the dynamics at this location is representative and similar to other low-energy sites around the globe: in other words, is this dataset going to help find generic processes, which can in turn be useful elsewhere? Or instead, is this study more relevant to the North Sea context? There is a sort of review on "sheltered" beaches, whose definition is not completely clear (is this the role of this paper to define such morphological features anyway, I wonder?). Part of my feeling also comes from phrases jumping from the Dutch context to generalities, which are sometimes extrapolated. For instance, the first phrase: "Sheltered coastlines are traditionally defended by hard coastal structures made of concrete, asphalt or stones.". I am not a specialist but I do not think this actually the case everywhere.

We understand the reviewer's suggestions and questions on the introduction. We think the findings are applicable at many sheltered shorelines as the fetch limited conditions are often dominant at these beaches. We improved the introduction in which we removed some of the specificity of the site and explicitly mention that we believe that the dataset will lead to generic knowledge (e.g. on wave nonlinearly during these low energy, short period conditions). Publications on generic processes using this dataset are in the making. The revised introduction of the manuscript can be read in L13-79.

1 -The authors often refer to "sediment transport" as measured or, at least, that can be studied with the present dataset. But is this really the case? Turbidity measurements could not be transformed into SSC, so how can the dataset actually serve to validate sediment transport (not morphological changes)? Because of this, I would tone down the sediment transport part, and rather focus on a dataset capturing the morphological evolution of the site as well as describing the spatio-temporal evolution of sediment characteristics. It does not mean that the dataset is less useful to model the site morphodynamics, there just is less means to validate sediment transport directly.

The reviewer is right to point out that the initial ambition of the campaign, to measure both hydrodynamics and sediment concentrations, was not realized as we decided to not publish sediment concentrations derived from the optical backscatter data due to difficulties with flocs in the water column.

Apparently, the current formulation of the scope may leave the reader with the impression that sediment transport was measured. We take the reviewers comment at heart and adjusted the manuscript to make sure the scope of the dataset is clear to the reader. Therefore, we have reviewed all mentions of 'sediment transport' in the manuscript and made more specific references to the scope of the data that is published and presented. The list below shows all mentions of 'sediment transport' in the old manuscript and changes in the revised manuscript (~~old text~~ >> new text):

L8-10: The novelty of this data set lies in the detailed approach to resolve forcing conditions on a sheltered beach, where ~~sediment transport~~ <u>morphological evolution</u> is governed by a subtle interplay between tidal and wind-driven currents, waves and bed composition, primarily due to the low-energy (near-threshold) forcing.

L41-42: ~~This is particularly the case for the unresolved and therefore parameterized transport processes that are calibrated on field and lab observations.~~ (we removed this line)

L45-47: For example, tidal currents enhance wave-induced bed-shear stresses (Kleinhans and Grasmeijer, 2006) and may simultaneously alter the dominant transport direction from crossshore to alongshore (Héquette et al., 2008). (we left this sentence unchanged)

L51-52: These processes affect the strength and direction of wave radiation stresses, and thus the extent to which the wave field drives alongshore currents that could transport sediment (Feddersen, 2004). (we left this sentence unchanged)

L55-57: Third, the restorative capacity of wave-driven onshore transport by returning sediment high in the beach profile from lower parts after storm erosion (Hoefel and Elgar (2003)) is limited on low-energy beaches (Hegge et al., 1996; Jackson et al., 2002; Nordstrom and Jackson, 2012). (we left this sentence unchanged)

L61-63: ~~Finally~~ <u>At these sites</u>, the nourished sediment composition does not always match the natural gradation. Sediment heterogeneity, although widely recognized as an important control in beach development (e.g., Huisman et al., 2016; Bergillos et al., 2018), is generally inadequately accounted for or resolved in sediment transport models.

L68-70: Yet, how the local grain-size distribution and its spatial heterogeneity resulting from the implemented sediment mixture affects sediment pathways at a mixed-energy site requires more field data before transport models can be improved accordingly. (we left this sentence unchanged)

L71-72: The extent to which these aforementioned processes are ~~resolved~~ <u>captured</u> in ~~the transport equations of~~ engineering-type models requires validation.

L83-85: The overall aim of this study is to unravel ~~the drivers of mixed-sediment transport in a low-energy system~~ <u>the forcing mechanisms of sediment transport in this mixed-sediment, low-energy system. These data are usable for validating</u> ~~, and to be able to validate~~ model parameterization of unresolved processes (e.g., wave non-linearity, wave breaking, multi-fraction <u>sand dispersal</u> ~~transport~~) in engineering-type models at sheltered beaches.

L288-289: The cross-shore wave transformation governs where to expect wave-driven sediment transport and also determines whether this transport is onshore, offshore or alongshore directed. (we left this sentence unchanged)

L313-314: This data set allows in-depth analysis of the <u>forcing mechanisms of sediment transport in this mixed-sediment, low-energy system.</u> ~~driving mechanisms of mixed-sediment transport at sheltered beaches in a process-based approach.~~

L322-324: Altogether, this comprehensive data set provides new data <u>of a case study in the Wadden Sea (NL)</u> to improve our predictive understanding of morphodynamics in sheltered (semi-enclosed) systems, where ~~sediment transport~~ <u>morphological evolution</u> is governed by a subtle interplay between waves, tidal and wind-driven currents and bed composition due to low-energy (near-threshold) forcing.

2 – I have concerns over the processing of the bottom pressure, and the bulk wave parameters provided in the paper. My attention got caught especially in Figure 8, with the much larger Hm0 offshore compared to other sensors, without an obvious explanation. Reconstructing the free surface elevation from bottom pressure is a problem that I know well, and I am fully aware of the challenges in the present dataset for reconstructing a signal with such short waves, and sometimes in the presence of relatively strong non-linearities. In short, there is no ideal solution, but at least the problem should be acknowledged, and the related uncertainties into bulk parameters quantified. First, there is no mention of the cutoff frequency for the linear correction, except at line 209 where 1 Hz is noted but this is surely not the correct one. Here, I analysed several bursts of data at the OSSI L2C10 (similar behaviour is observed at C9, I have not checked at other sensors) as follows. 40-min of data were taken during 4 situations of "low" and "high" energy conditions (2 each), including the one chosen. Note that the behaviour I describe next could be found in most bursts I extracted and tested. Over this 4 bursts of data, I computed power spectra of the hydrostatic (ζ_{hyd} in figures)and linear (ζ_L) reconstructions of the free surface elevation from the detrented pressure signal using 60 Hann-tappered blocks of 64 s overlapping by 50% (effective d.o.f. are ~112). From these, the significant wave height H_{m0} and mean wave period T_{m02} are computed using a range of different cutoff frequencies. Lastly, the bicoherence b, a measure of the bound energy at a given frequency, was computed from ζ_{hyd} following Hagihira et al. (2001). From the provided plots, several important remarks can be made:

- Bulk wave parameters are extremely sensitive to the cutoff frequency, especially H_{m0}. Let us consider the energetic example of Figure 8 (19 September 2021 around 08:15): integrating up to 0.5 Hz gives Hm0 ~ 0.5 m, while integrating up to 0.8 Hz gives more than the double! In terms of energy, that is a factor four. Here, it should be noted that by looking at the hydrostatic reconstruction, choosing 0.8 Hz as a cutoff does not look incorrect, since it still corresponds to energy levels 2 orders of magnitude above noise level.
- The cutoff frequency chosen by the authors seems to vary in time, and sometimes even between parameters. In order to retrieve the values stored in the 'tailored' data, a cutoff frequency at 0.6 Hz should generally be applied, but it actually seems to vary between 0.4 Hz (13 October 00:15) to 0.9 Hz (1 October 02:45), and potentially less or more, respectively, since I only checked a limited amount of data bursts.
- Non-linearities seem relatively strong, and interactions around the peak frequency or between distinct bands of frequencies seem both intense and pretty common in the present dataset. Those explain large fractions of bound energy at high frequencies, leading to a potentially large overestimation of the pressure-to-surface transfer function (e.g., Martins et al., 2021), which affects the computation of bulk parameters. As shown in some of the examples provided, the blow-up of the linear reconstruction occurs rather "fast" in some cases, e.g. well below frequencies where noise start to dominate the pressure signal.

As mentioned above, I acknowledge the fact that there currently does not exist a method adapted to such wave conditions (non-linear reconstructions being currently limited to weakly or moderately dispersive regimes). Furthermore, the raw pressure data is provided, so that informed users will be able to decide themselves how to compute bulk parameters. However, a large fraction if not most users will not be fully aware of the issues, and will be mostly interested in bulk products as computed by the authors. In this context, and in my opinion, the issues related to the reconstruction of such short waves often undergoing non-linear processes should be acknowledged, and the uncertainty on

bulk quantities shoud be estimated and discussed. Here I chose bursts of data at high tide, where we expect minor influence from the current, but considering the wave periods measured here, this is definitely another aspect that need to be discussed as it unsure whether the authors account for it for computing wavenumbers.

The reviewer points out an important aspect that we did not pay sufficient attention to in the original manuscript. We thank the reviewer for raising this point. Indeed, we did not explain in enough detail how the reconstruction of sea-surface variance from the near-bed pressure signal was performed. In fact, the Transfer Function $Sw = cosh(kd)/cosh(kh)$ was computed with linear wave theory but capped at an amplification factor of 5 for amplification of the power spectrum, to prevent blow up of noise below 1 Hz in the computed wave statistics. We have modified the manuscript accordingly.

As mentioned by the reviewer and his accompanying figures, choosing one appropriate cutoff frequency can be difficult in case of a linear correction of the pressure data. We use an amplitude cap on the correction (rather than a frequency cutoff) and with this cap, the wave statistics Hm0 and Tm02 are almost insensitive to the choice of the cutoff frequency if the cut-off frequency is chosen above the sea-swell frequencies. This can be seen in the provided plots in Figure 1 (right panels, green lines). Here, we produced figures similar to the ones of the reviewer, using 10-minute bursts of data (as this was the chosen burst length for this dataset) at the similar moments in time as the reviewer chose. We show the effect of the capped Transfer Function Sw on the reconstruction of the sea-surface variance density spectrum in comparison with an unbound application of the Transfer Function. We also show how wave statistics Hm0 and Tm02 then vary with the upper integration limit (cut-off) frequency. Because the hydrostatic variance at frequencies above 0.6 Hz is not negligible for the short-crested waves we recorded, we believe our method with a capped transfer function is preferable over a shorter integration limit up to e.g. 0.6Hz.

Then, the reviewer points out that linear theory has it limits and we should acknowledge those. To that end we added a paragraph on consequences of bound wave energy on the reconstruction. The lowest frequency on which the cap operates determines the reconstruction and is therefore a measure of the reliability of the linear reconstruction. At L2C10SOLO, the cap on Sw onsets below 2fp for 38% of observations, see the attached Figure 2.

Lastly, the reviewer asks for a discussion of the influence of currents on the published wave statistics. We made an analysis of the effect of currents at station L2C10VEC. At this station, we expect the effect of currents to be largest, as currents are predominantly alongshore and at this deepest ADV station the waves are still most obliquely incident. The attached Figure 3 shows a 2D histogram of the current-corrected and the uncorrected wave length. The statistics from this comparison were included in the text of the manuscript.

All the above was incorporated into the manuscript as follows:

L208-240: Then, near-bed pressure was converted to sea-surface elevation using linear wave theory:

$$Pss = S\_w \; Phydrostatic \qquad (1)$$

$$Sw = cosh(kd)/cosh(kh) \qquad (2)$$

with Pss the variance density spectrum of the sea surface, Phydrostatic the variance density of the hydrostatic surface elevation, k the wave number, h the instrument height above the bed and d the water depth. Due to the short-crestedness of observed waves, the sea-surface reconstruction was performed over the frequency band [0.05-1] Hz. This band deviates from common choices such as [0.05, 0.33] Hz (e.g. Neumeier and Amos, 2006). It was chosen such that sea-surface spectra for short-crested wind sea waves (with peak frequencies up to 0.5 Hz) are represented in the dataset. To prevent blow-up of noise in the power spectrum within the integration range, the linear transfer function, Sw was capped to a maximum of Sw = max(Sw, 5). For example, for the deepest pressure sensor L2C10SOLO, on average 51% of the hydrostatic power density is present between [0.33, 1] Hz. This portion of sea surface variance density is on average amplified by a factor 260 in the linear reconstruction without a cap on Sw, but with the cap gets amplified by a factor 6. Using a cap in the sea surface reconstruction, the wave statistics are rather insensitive to the exact choice of the upper integration limit. For example, setting the upper integration limit to 0.8 Hz instead, reduces the significant wave height on average by 1% and increases the mean wave period Tm01 by 1.2%. Some sensitivity to the choice of the cap value on transfer function Sw however remains.

The cap on Sw mitigates the blow-up of noise, but does not solve the problem that waves at high kd are strongly attenuated at 15 cm above the bed. This results in unrealistically large Sw values, that are replaced by the cap value in our approach. The lowest frequency on which the cap operates determines the reconstruction and is therefore a measure of the reliability of the linear reconstruction. When interested in statistics on spectral shape, the user could discard those observations where the onset of the cap on Sw lies too close to the peak frequency fp. At L2C10SOLO, the cap on Sw onsets below 2fp for 38% of observations.

The effect of currents on the sea-surface reconstruction is not incorporated in the linear reconstruction. This was decided as the effect of currents for those sensors where both flow and pressure observations were present was seen to be small, and moreover not all pressure sensors were accompanied by a flow sensor. At the deepest ADV L2C10VEC, the effect of the current on the relative wavelength was computed with the method of Guo (2002). For this, block-averaged horizontal velocities were decomposed in a component in direction of wave propagation and one perpendicular to that. On average, the root mean square error in wave length when neglecting currents was RMSE=0.3 m (3%).

Lastly, applying a linear sea surface reconstruction occurs under the assumption of a random-phase sea state, thereby ignoring the effects of non-linear processes (e.g. triad- and quadruplet wave-wave interactions) that phase couple sea-swell frequencies to higher harmonics that do not adhere to the linear transfer function (e.g. Martins et al., 2021). This means that potentially, bound wave variance, if present, can get over represented in the reconstructed sea-surface signal. The user should be aware of this if higher-order statistics of the reconstructed sea surface (not part of the published tailored dataset) are computed from the raw data.

More specific comments:

#1 – L19: along the lines of my comment above on the introduction, what then makes them different than other sandy beaches? What about winds?

We restructured the introduction, also in line with the earlier comments. The revised introduction identifies the main aspects of sheltered beaches that sets them apart from exposed sand beaches: 1. Tidal and wind driven currents, 2. Isotropic young sea states, 3. Limited restorative capacity of wave-driven onshore transport. 4. Sediment heterogeneity in the case of nourished sheltered beaches. These are discussed in line L34-51.

#2 – The cross-shore and longshore referencing used is really hard to follow in the paper, as it changes between Figures. I found no information on the coordinate system used (only vertical referencing), and in this regard, maybe the mean coastline orientation could be provided so that users could convert easily between real and local coordinate systems (includent long- and cross-shore directions).

This could indeed have received some more attention in the manuscript. This information was and remains accounted for in the tailored netcdf's for ADV and ADCP instruments. We have altered the manuscript as follow:

L322-323: Positive onshore direction at this cross-section is 135 degrees counterclockwise from East, and positive alongshore direction is 225 degrees counterclockwise from East.

The local coordinate system that is used for figures was indeed not consistent. Now, we have made sure the figures all use the local coordinate system that is also used in the transects netcdf coordinate "d". Figure 3, 8 and 9 were updated. In Figure 9 we showed changes on transect #0 in the earlier manuscript, this is now updated to transect #1 that overlaps the densely measured cross-section.

3 – The unit of raw pressure was not intuitive (Pa around a mean Pa value relative to NAP). At this point, they could be provided directly in m (hydrostatic) relative to NAP or please clarify somewhere.

We agree that the description of the unit Pa+NAP might indeed not be the most intuitive. Therefore we have changed the unit of the pressure on Quality Controlled and tailored timeseries for the pressure sensors to m+NAP so that p indeed represents hydrostatic surface elevation in m referenced to NAP. For the ADV and ADCP instruments, this was already the case for the variable eta. To align ADV/ADCP and pressure sensors, we renamed the variable eta to p, with unit m+NAP on the revised dataset.

#4 – L214: why only resticting to such a narrow frequency band? First, it becomes really sensitive to the definition of the peak frequency, which is clearly hard to define in such wave conditions. Second, the bottom pressure "contains" what reaches the bottom in terms of non-linearities so that including frequencies up to the noise level will be much more representative of wave non-linearities.

This chosen frequency band, scaling with the peak frequency, was chosen in anticipation of a following piece of work in which we focus on sea-swell wave non-linearity at this and another sheltered site compared to more exposed sites. In that analysis, we observed phase coupling of pressure variance at the peak frequency with lower harmonics in our observations by analysis of bispectra. This bound variance had little influence on spectral wave statistics in the band [0.05, 1] Hz, but was observed to affect the velocity skewness, see e.g. the attached Figure 4 that shows the real and imaginary part of the bispectrum of September 13th 12:30 (corresponding to the same time period of the bicoherence plot the reviewer plotted). At more

exposed coasts, this bound variance is identified as infragravity and gets excluded in the integration between [0.05, 1] Hz. Given our short-crested waves, they are observed above 0.05 Hz and hence are included in the integration. This is what motivated us to adjust the frequency band adaptively.

Although we are eager to share a more in depth analysis of the impact of the frequency range we believe this data descriptor manuscript is not the right place to extensively describe this finding. Therefore, we followed the reviewers suggestion to publish velocity skewness and asymmetry in this dataset using the more conventional frequency band [0.05, 1]. With that the statistics are in line with standard processing protocols (e.g. Ruessink et al. 2012, Rocha et al. 2017). A more in-depth discussion on phase coupling in these short-crested waves is saved for another manuscript.

Mentioning of this in the manuscript is modified:

L250 and L268: in the frequency band [0.05-1] Hz

#5 – Sections 8.2 and 8.3 feel light. There is no mention of Figure 9 in section 8.2.

We expanded the paragraphs 8.2 a bit more to discuss the changes observed in Figure 9 and fixed the missing reference to Figure 9. Likewise, we expanded section 8.3. This section of the article was however only meant to compactly visualize the type of data available, not to analyse the observed signal in-depth.

#6 – All data files have different time origins. Is this the best choice for processing multiples files?

The reviewer makes a good point. Writing to netcdf was semi-automatically taken care of by the analysis tools we used (python/xarray) on the fly, and the time origin was therefore not a thought over choice. It makes sense indeed to take a stable time origin for ease of analysis using different tools. We have modified the time origin in the revised dataset for all instruments to 2021-09-01 00:00:00 and set the time axis in seconds from this origin.

#7 – The table indicates 30 min bursts for the ADV L2C10 while the data actually suggests it is 10 min. Please verify that the figures provided in the table are correct.

ADV's recorded raw data in bursts of 30 minutes indeed, as tabulated in Table 1. The reader was left confused because the decision to analyse in 10 minute blocks was only mentioned very briefly in Section 6. The explanation of this discrepancy is now more explicitly stated in Section 6:

L194-198: For all instruments apart from the ADCP, it was decided to perform analysis on 10 minute data blocks instead of the full 30 minute bursts. At this site with short crested waves, 10 minute blocks are sufficiently long for mean spectral statistics in the sea-swell band and 10 minute blocks increase the total duration of valid measurement blocks with submerged instruments. Therefore, the raw data folder consists of data structured in 30 minute bursts, and qc and tailored data folders contain data structured into 10 minute bursts.

#8 – In relation to #7: the Table indicates that the sampling frequency of L2C10SOLO is 8 Hz, while the data file indicates 10 Hz. By comparing with the pressure from the ADV, it

clearly is 8 Hz. Please verify that all netcdf files include the correct information. As reviewers, we cannot verify every individual files…

We thank the reviewer for finding this error, the sampling frequency of the SOLO pressure sensors was 8 Hz indeed. In a last modification of the processing scripts the sampling frequency of the SOLO sensors was mistakenly set to that of the OSSI sensors. We have corrected this now on the SOLO dataset, and this indeed also affects the wave statistics reported at the SOLO tailored datasets. All other instrument files were checked after identification of this error, and were found to be fine.

#9 – I felt that more information could be provided within the netcdf files, and fields could be more descriptive in general.

We have reviewed the attributes of variables and added units and long names to some variables where these were mistakenly missing. Thanks for pointing this out to us. This was for example the case for the coordinate N, the block local time, and the vertical coordinate z for the ADCP datasets. On top of that, we have added the campaign summary that is also part of the 4TU overview page to the general dataset attributes for each file in the hydrodynamics dataset.
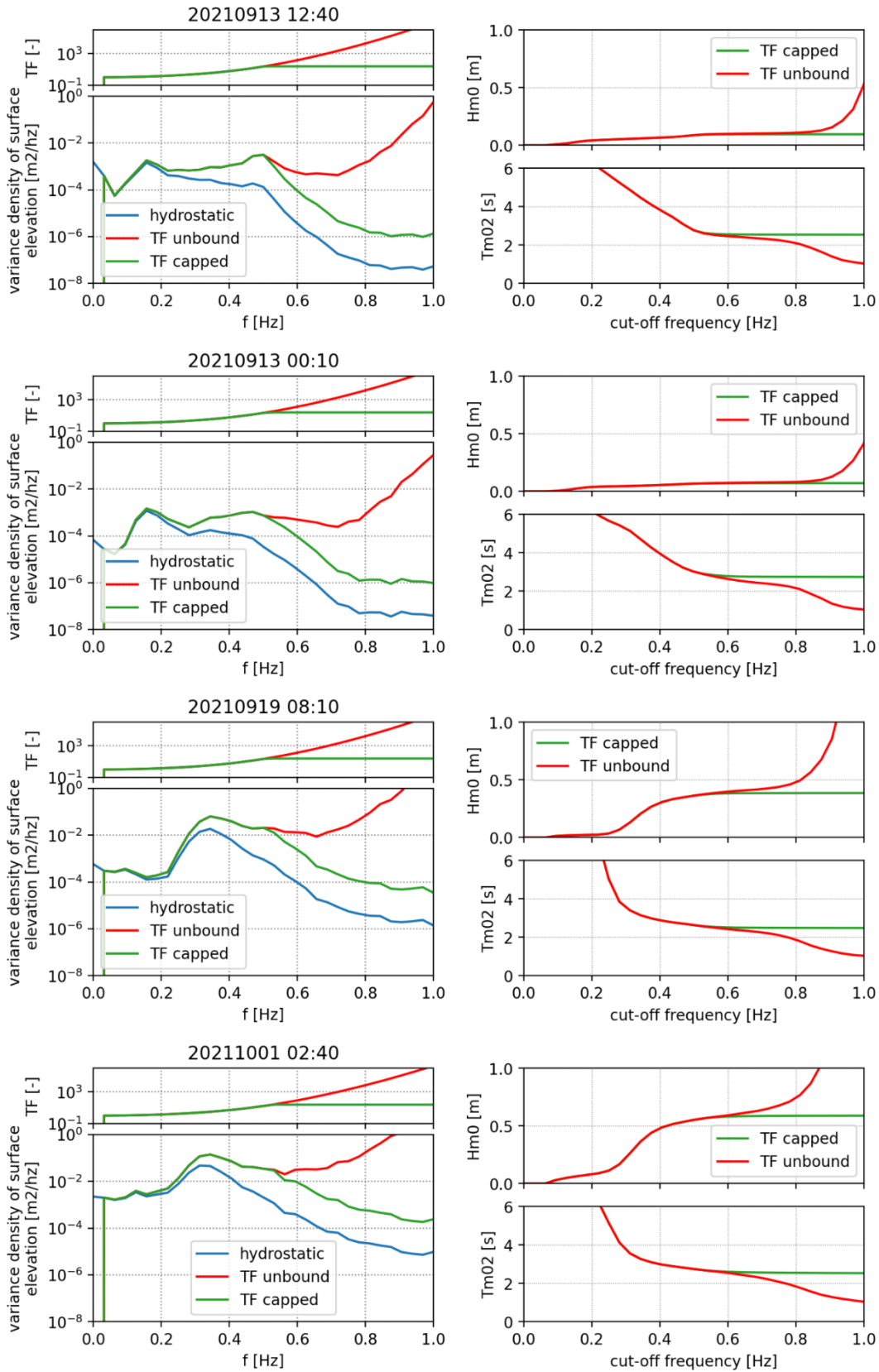
Figure 1 Effect of capping of transfer function on computed wave statistics for the four moments in time analysed by the reviewer. The left panel shows the transfer function with cap and without as a function of frequency, and the reconstructed variance density with and without cap. The right panels show the effect of the cap on the transfer function on the sensitivity of the cut-off frequency
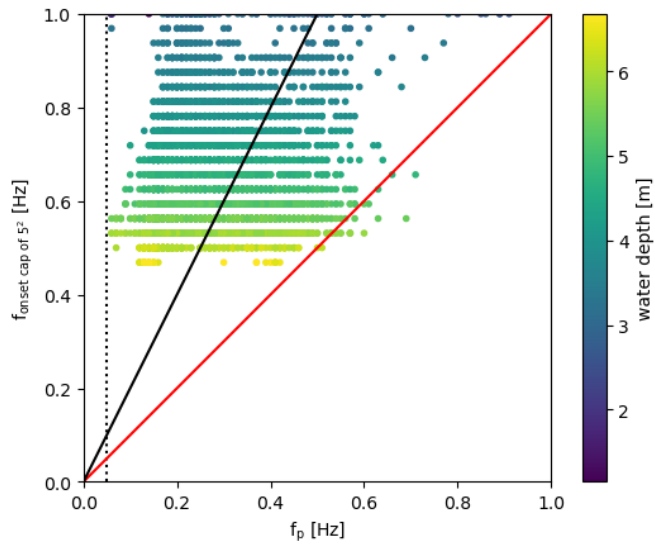
Figure 2 Effect of correcting for ambient current on wave length at ADV L2C10VEC. 2D histogram of current corrected wave length versus wave length without current correction. Dashed lines indicate mean value of wave length and current corrected wave length of entire timeseries.
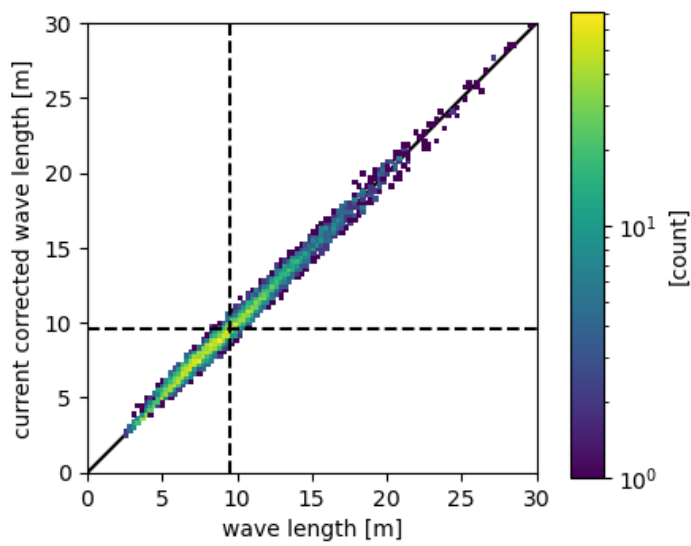

Figure 3 Effect of correcting for ambient current on wave length at ADV L2C10VEC. 2D histogram of current corrected wave length versus wave length without current correction. Dashed lines indicate mean value of wave length and current corrected wave length of entire timeseries.
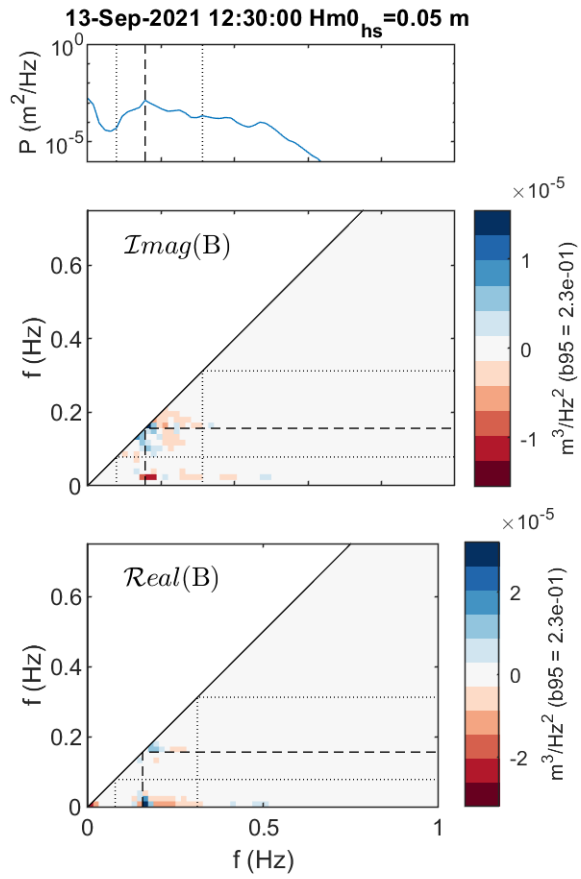
Figure 4 Bispectrum of calm burst on September 13th 12:30. Power spectrum in top panel, Imaginary and Real part of the Bispectrum plotted below. Legend of colorbar mentions statistical significance threshold of bicoherence (Bicoherence not shown here).