UPFLOW data paper - rebuttal

September 19, 2025

1 Editor's comments

Two reviewers have provided comments on your manuscript, including "It provides some very compelling data examples that highlight the quality and potential utility of this dataset" and "The paper definitely fits the scope of Seismica and should be considered for publication". They agree that your work needs substantive revision/clarification and provide numerous suggestions for improvement. Their reviews are attached.

We thank the editor and the reviewers for their detailed and constructive comments. We addressed them all as explained in the grey boxes below.

In addition, the following minor, non-comprehensive items caught my eye.
a. Fig 6, top has no red lines.
b. Fig. 9 does not indicate what channel's spectrum is shown.
c. Fig S15d, orientation vs. Backazimuth. The scatter in the orientation estimates for the +-marked events disappeared, while it appears preserved for the circle-marked events.

We updated Fig. 6 and added the CTBTO hydrophone noise models to the top panel. We expanded the caption of Fig. 9 to explain that vertical component data are used. The plot in bottom right of Fig. S15, which is now Fig. S14 in the revised supplemental materials (referred to as Fig S15d by the reviewer) does not show the poor quality measurements by design (poor quality measurements are indicated by crosses). It would visually clutter the plot without providing added value, as most of the bad measurements are nearly random, and the point of this plot is to show the consistency of accepted individual event measurements. The point of the other three plots is to justify the thresholds chosen for accepting measurements (of course, in actual practice we considered similar plots for all stations before deciding on the threshold).
We have amended the caption to point out that poor quality measurements (crosses) were left out on purpose (see caption of Fig. S14 in the revised supplemental mate-

rial).

S16d is also weird compared to a-c. Is there a plotting error?

None of the individual event measurements from this station were in fact accepted; therefore no circles are plotted, and also none of the horizontal lines showing mean and median values of the P wave measurements could be plotted, as there is nothing to average. The thin grey horizontal line indicates the measurement from the surface wave method for comparison; this important information was inadvertently omitted from the caption (the grey line is also present in the old Fig. S15, which is now Fig. S14 in the revised supplemental materials, but is fully covered by the body wave related lines). The point of showing this figure is that even though none of the measurements got accepted, as the few high SNR measurements failed rectilinearity or incidence angle thresholds, the high signal-to-noise-ratio (SNR) measurements are still consistent with the orientation estimate from the surface waves (and those of marginal SNR are wrong by $\sim 180°$). The caption has been amended to explain the absence of measurements in the bottom right plot, and to explain the meaning of the thin grey line (see Fig. S15 in the revised supplemental materials).

d. Typo in Section S9 title, ”propability”

Fixed (see the title of section S8 in the revised supplemental materials).

2 Reviewer A’s comments

Thanks for giving me the chance to review the paper titled “Performance of the 2021-2022 UPFLOW large ocean bottom seismometer array in the Azores-Madeira-Canary Islands region, Atlantic Ocean”.

In this paper, the authors introduced the readers to their OBS array “Upward Mantle Flow from Novel Seismic Observations” (UPFLOW), by analyzing the state-of-health (SOH) data and the noise level. More importantly, they determined the horizontal orientation by analyzing the P wave and Rayleigh wave polarization. They also calculated the transfer function of the compliance and the tilt. Finally, the authors showed some data examples and discussed the implications regarding future experiments. The paper definitely fits the scope of Seismica and should be considered for publication, but I do have two major concerns.

First, the paper needs to be more clear and concise to reduce possible confusion. For example, the authors do not need to justify their clock skew correction, when there has been one published paper (Cabieces et al., 2024) specifically discussing it. The downsample procedure is mentioned twice at different places with different sampling rates. Also, the authors do not need to repeat their noise level comparison conclusions several times. Figure 10 is plotted differently compared with the statement in the main text. Many places related to existing studies are not provided with enough references.

We thank the reviewer for their constructive comments and suggestions.
We revised the manuscript for clarity and conciseness, following the reviewer's suggestions. Specifically, we ensured consistent descriptions of the downsampling procedure, streamlined repetitive conclusions about noise levels, corrected the mismatch between Figure 10 and the main text, added missing references where previous studies were discussed, and corrected spelling mistakes.
In response to the reviewer's main comment on the possible benefits of tilt and compliance corrections for orientation estimation, we calculated the SNR for over 3,700 source-receiver pairs from 149 earthquakes, and compared waveforms with and without these corrections. The results, shown in the new Supplementary Figure S27, indicate that the SNR gain is generally small, with improvements in some cases and degradation in others. We now include these findings and their implications in the revised manuscript (lines 374–385).
All these changes are detailed in our responses below.

Second, the authors determined the OBS orientation first, then removed the seafloor compliance and corrected the tilt. They also mentioned that some earthquake data is not of high quality to allow them determine the OBS orientation. It is known that seafloor compliance and tilt can containment the long period earthquake recordings considerably (e.g., Bell et al., 2005). Therefore, I suggest the authors switch the sequence of their processing, by removing the seafloor compliance and tilt effect first. After that, they can proceed to determine the OBS orientation. Besides, it is also important to show how seafloor compliance removal and tilt correction can help improve the signal-noise-ratio of the earthquake records. Only showing daily power spectra comparison (Figure S21) is not enough.

We thank the reviewer for these suggestions. Irrespective of the order described in the manuscript, both processing steps (namely determining the orientation and removing the seafloor compliance and tilt correction) were carried out in parallel and mostly completed during the UPFLOW recovery cruise. An emphasis was placed on making these estimates available to data users quickly. We agree that the compliance correction improves the vertical signal quality, and thus applying the orientation determination procedure to compliance corrected data might improve the quality of the orientation estimates. Being pragmatic, the question is whether the expected improvement justifies the effort of reprocessing. Generally, noise levels at the longer periods used for the orientation estimates are much larger on the horizontal component (see Fig. 4, 5 in the manuscript). Specifically:

- For the P polarisation method, an improved vertical component will only have a marginal effect. Theoretically, the correction could lead to more linear particle motion, and thus more events to pass the QC thresholds. But the dominant criterion for which data are excluded is the SNR anyway, which is controlled by the horizontals. For the measurement itself, only the horizontal components are important. Therefore, there would be only a minor effect from improved verticals.

- For the surface wave method, the vertical component directly enters the measurement process, so there is more of a concern. The compliance noise on the vertical will degrade the correlation between corrected radial and phase-shifted vertical. But unless compliance noise is very large, the position of the peak of the correlation function will only shift very marginally, even though the maximum correlation value might be a little bit lower. Any shift of peak position, which we expect to be very minor in any case, will be random, so be further reduced by averaging individual measurements. Regarding the tilt correction, the situation is less clear. The tilt correction applies a transfer function to the horizontal components and subtracts it from the vertical. This transfer function is tuned to the current induced tilting noise. When an earthquake at good SNR on the horizontals is recorded, the earthquake signal rather than tilt dominates the horizontal time series. The tilt correction will thus project some of the Rayleigh wave energy (specifically the component perpendicular to the tilting axis onto the vertical), thus potentially biasing the orientation measurement in a systematic way. So we are rather sceptical about whether application of the tilt correction would improve the estimates, as time series are specifically pre-selected for good SNR on the horizontal components. Further investigation of this question with synthetic and data examples would of course be very interesting, but is well beyond the scope of the current manuscript.

Following the reviewer's comments, we computed SNR for over 3,700 vertical component recordings of the 149 earthquakes considered in our study (Tables S1-S3) for waveforms with and without tilt and compliance corrections. The SNR was defined as the root-mean-square (RMS) of the Rayleigh wave window divided by the RMS of the preceding noise window, using exactly the same filtering as in our horizontal orientation analysis. We built a new figure that was added to the supplementary information (Fig. S23) showing illustrative waveform examples as well as the distribution of the SNR gain, which is the difference between the SNR of the waveforms with and without tilt and compliance corrections. Fig. S23 shows that in most cases the SNR gain is small and close to zero. While in some cases the tilt and compliance corrections improve the SNR of the earthquake waveforms, in other cases it degrades the signal. We added explanations along these lines to lines 376–386 of the revised manuscript. These results support our expectations explained above that a re-determination of OBS orientation with UPFLOW data corrected for tilt and compliance will only yield a very marginal benefit (if at all). For these reasons as well as for pragmatic reasons given limited time resources, we thus refrain from re-determining OBS horizontal orientations with tilt and compliance corrected data.

I also have some detailed questions, listed as below.

Section 1.1 – Could the authors switch the first two paragraphs? I do not know the specific requirement of Seismica for this kind of paper, but I am more used to a start of the background geological setting and then talk about the aim of the experiment. For readers like me, I even do not have a clear clue how the Azores-Madeira-Canary islands region is connected to the mantle plumes.

We understand the reviewer's preference for introducing the geological setting first, as is common in traditional research articles. However, since this is a data paper (and please note that we now changed our manuscript from "Article" to "Report" in Seismica), our primary goal is to present and describe the dataset. Hence, we prefer to begin the manuscript with the motivation and objectives of the data collection.

Line 22 – It can be great if the authors can give the full spelling of "UP-WELL" here, which is the, just like they do in line 37-38.

Following the reviewer's suggestion, We added the full spelling of "UPFLOW" to the abstract (see lines 24–25 of the revised manuscript).

Line 55 – Please consider adding more references here, as I anticipate the word "many" means more than two. The two following references could be good choices: Montelli, R., Nolet, G., Dahlen, F. A., Masters, G., Engdahl, E. R., & Hung, S. H. (2004). Finite-frequency tomography reveals a variety of plumes in the mantle. Science, 303(5656), 338-343. Yang, T., Shen, Y., van der Lee, S., Solomon, S. C., & Hung, S. H. (2006). Upper mantle structure beneath the Azores hotspot from finite-frequency seismic tomography. Earth and Planetary Science Letters, 250(1-2), 11-26.

We added the reviewer's suggested references to the text (see lines 65-66 of the revised manuscript).

Line 56 – Why do the authors think "the situation is less clear beneath the Madeira and the Azores archipelagos"? More context needs to be given.

We added the following explanatory text to the manuscript (see lines 67–71):

> For Madeira, the proximity to the Canary Islands (∼300 km) and the limited resolution of existing tomography models have hindered efforts to probe the existence of deep plume structures in the area. Regarding the Azores, some tomography models suggest a lower mantle plume beneath the region, whereas others only show low velocity anomalies confined to the upper mantle (see, e.g., the study of Marignier et al., 2020, which used multiple tomography models).

Line 60 – References are needed to provide context for the debate.

Following the reviewer's suggestion, we added relevant references to the text (see line 77 of the revised manuscript).

Line 92 – Please add (100 s) after the 0.01 Hz at the end of this line.

We added the recommend alteration in line 112 of the revised manuscript.

Line 93 – Is it K.U.M. or KUM? There are two different spellings in the main text and in Table 2. Please unify them. Besides, it is not clear to me what KUM stands for. Is that a company name?

K.U.M. stands for "Umwelt- und Meerestechnik Kiel" in German. It is a known company that is developing customised subsea technology, like the LOBSTER and Nammu OBSs, amongst others. Both spellings "K.U.M." and "KUM" are valid, but for consistency, we modified the manuscript to use "K.U.M." throughout.

Line 128 – I think the burn wire mechanism is the main stream releasing mechanism for US OBSs as well.

That is correct, we updated the text (see lines 148–149 of the revised manuscript). See also our response to the other reviewer.

Line 148 – It is not clear to me what the authors are referring to as the "disparities in temperature recordings". Should you refer to Figure S3-S6 here?

Besides, this sentence should be within the last paragraph, instead of as the start of a new paragraph.

We updated our manuscript according to the reviewer's suggestion:

Original text

> Disparities in temperature recordings between stations could be due to local ocean currents/water masses, local hydrothermal activity, and/or be related to internal processes (recorder/seismometer).

Updated text (lines 172–174)

> Disparities in temperature recordings between stations could be due to local ocean currents/water masses, local hydrothermal activity, and/or be related to internal processes (Fig. S2-S6, 2nd panel; recorder: blue curves; seismometer: orange curves).

Line 150 – Please allow my ignorance, but I am not sure what the mass position and static sensor acceleration stands for and how to understand them. Maybe some explanations are needed, as some of the readers do not check the SOH information of their seismometers.

Indeed, SOH information is often overlooked or not easily accessible in many OBS experiments. In our case, we had the opportunity to extract and analyse these data, which can offer valuable insights on instrument's behaviour throughout the experiment.
Regarding the terminology, the term "static sensor acceleration" is used by Nanometrics (the manufacturer of the majority of the seismometers used in the UPFLOW experiment) and refers to the signal that results from the combined response of the seismic mass and the electronic feedback system. As for the mass position, it refers to the position of the seismometer's internal mass (which ideally should stay constant throughout the experiment). We added some text to the manuscript to clarify this terminology (see lines 174–177 of the revised manuscript):

Original text

> Finally, the third and fourth panels show the mass position (voltage) and static sensor acceleration, respectively, which are, again, quite stable for UP06 overall.

Updated text

> Finally, the third and fourth panels show the mass position (voltage) and static sensor acceleration, respectively, which are, again, quite stable for

UP06 overall. The term “static sensor acceleration” refers to the signal that results from the combined response of the seismic mass and the electronic feedback system.

Line 177 – Could the authors refer to the figures from smaller numbers to larger numbers, that is, S5, S6, S9, and S10?

Done. See also the response to the other reviewer.

Line 201-206 – I think the authors should remove these sentences along with Figure S13. There is no doubt that a clock skew of 12 s needs to be correct on the OBSs UP05. However, how to exactly correct it through the whole recording period cannot be as easy as adding a fixed skew value but assume the clock drift to be linear or nonlinear. The authors can directly refer to and introduce the work of Cabieces et al. (2024), just as they did in the next paragraph. The clock skew correction part is unnecessary but makes readers, like me, confused with the term of clock drift correction.

As suggested by the reviewer, we removed Figure S13 from the revised manuscript as well as the text around it. We kept the text about the typical errors in body wave teleseismic tomography for readers who may not be familiar with the latter.

Line 218 – Please consider spelling out samples-per-second when the abbreviation ”sps” appears for the first time.

Done (see line 253 of the revised manuscript).

Line 217-218 – I am confused here. In line 393-395, the authors stated that broadband has a sampling rate of 50 Hz, and high broadband has sampling rate 100 Hz. Then why do you want to downsample them so greatly here to 20 Hz? I think the authors need to make the paper self-consistent.

We used the 20 Hz resampled data for some on board processing but we uploaded LH channel data to EIDA together with the original channel sampling rate data. In order to avoid confusion we removed the text ”and Broadband BH with 20 sps” (see lines 250–251 of the revised manuscript).

Line 231 – Please directly write out "a grid search approach with 1 increments".

Fixed and some further rephrasing was applied to this sentence for clarity (see line 266 of the revised manuscript).

Line 242 – The rest of what is invalid? Measurements?

We meant that for the rest of the stations no valid estimate could be extracted, equivalent to our phrasing for the body waves. We rephrased this line to remove the ambiguity and now list numbers of stations with invalid estimates for both P wave and Rayleigh wave methods (see lines 278–279 of the revised manuscript).

Line 252-253 – The phase shift of what? P wave or Rayleigh wave? Please be clear. Also refer to Figure S18 if needed.

We now state that this statement refers to the P wave (see line 290 of the revised manuscript).

Line 304 – I am not sure if such a conversion is really meaningful. The hydrophones are recording pressures instead of ground motions. Even the authors switched to comparing the noise level with the CTBTO high and low noise models later in Figure 5. However, I am also fine if the authors keep this part as it is now.

Indeed, hydrophones measure pressure and are not typically used for standard ground motion analyses. Our intention in presenting this conversion is primarily illustrative. We aim to demonstrate that hydrophone data, although often underutilised, can be made more accessible and interpretable by converting them into a more familiar form for readers who are accustomed to working with ground motion data.
We agree that pressure data carry valuable information in its own right. This is why, in subsequent figures such as Figure 5, we chose to represent the hydrophone data in pressure units and compare it directly to the CTBTO noise models, preserving its physical meaning. Our goal was not to suggest that displacement conversion is the preferred way to treat hydrophone data, but rather to highlight the potential of this additional channel and encourage its broader consideration in seismological analysis. We are happy to keep this part in its current form.

Line 312 – I have a question about the phrase "the CTBTO high and low noise hydrophone models". The models are models describing the noise levels, right? What does it mean by referring to "hydrophone models"? I think the correct usage should be "the CTBTO high and low noise models," just as the authors wrote at the end of line 313. Similar issues also appear in the caption of Figure 5.

The expression "CTBTO high and low noise hydrophone models" was indeed imprecise. These are noise level models that use hydrophone data rather than models of the hydrophones themselves. We have revised the text and the caption of Fig. 5 to consistently refer to the "CTBTO high and low noise models" (see lines 357–359 of the revised manuscript).

Line 316 – I think this conclusion has been given once in line 309. Please consider improving the writing here to make the paper concise.

We have re-written this text in a more concise way (see lines 343–344 of the revised manuscript).

Line 321 – This is a very important piece of information. The authors should declare it at the beginning of this section, not in the latter part of the section. At least I wanted to ask about it when I started to read this section. Please move it to the first paragraph of section 5.1.

Done - we moved this information to the last sentence at the beginning of section 5, just before the start of section 5.1 (see lines 336–337 of the revised manuscript).

Line 324-326 – Please add some references to support your hypothesis.

As suggested by the reviewer, we added references to the text (see line 375 of the revised manuscript).

Line 333 – Please refer to Figure 1 after mentioning the Gloria Fault region. Please also mark the earthquake location in Figure 1.

Done (see lines 395–396 of the revised manuscript).

Line 336-338 – This is an interesting observation for someone like me, who is not familiar with normal modes. Do the authors think they should add a reference to indicate what kind of spectral analyses they used to create their observations (Figure 9)?

Done (see line 401 of the revised manuscript).

Line 349 – Should not you refer to Figure 10a at the beginning of this paragraph? Please check if this is a typo. Besides, in the caption of Figure 10, the station is UP06, but in the main text it is referred to as station UP27.

Line 349-350 – The authors mentioned the spectrograms of vertical and hydrophone components twice, but I can only find the spectrograms of the vertical seismic component in Figure 10. Where are the hydrophone spectrograms?

We thank the reviewer for spotting these typos and indeed we should refer to Figure 10a, to station UP06 and only to vertical component spectrograms (not hydrophone). We fixed all these typos (see lines 423–424 of the revised manuscript).

Line 361 – Do the authors think Figure 11a should be a better reference of the figure?

Line 361-363 – Just curious, why I cannot see the resonance peaks from the PPSDs in Figure 11b?

Line 366 – Please refer it to "Figure 11b," instead of "Fig. 11; bottom".

We thank the reviewer for spotting these glitches, which have all now been fixed (see lines 444 and 449 of the revised manuscript). The resonance peaks are visible as an elevated median PPSD curve for all GEOMAR OBS (yellow line) compared to the other LOBSTER OBS (blue line). We added some text to the caption of Figure 11 to explain this.

Line 395-397 – I agree with the author but I do not know if the tilt information could be useful for future studies, such as location earthquake phase picking. Do you think you should also list the tilt degree as a separate table in the supporting information? So that people can process their high sampling rate data themselves, if needed.

The dip and azimuth (i.e., tilt) information is fully included in the StationXML metadata files, which are deposited in the EIDA repository alongside the waveform data and will be open-access after UPFLOW's data embargo period. Users interested in reorientation, can access and use this information directly from the standard metadata format. Thus, we believe a separate table in the supporting information is unnecessary, especially as this manuscript and the supplemental information are already quite long.

Line 403 – What is mass? Do you mean mass position?

Yes, we have replaced "mass" by "mass position" (see text in line 486 of the revised manuscript).

Line 401 – I am not sure if it is proper to state that "the auxiliary data that are usually not shown or discussed in the literature".

Line 404-405 – I am not sure if the authors should speak so highly of being the first to extensively examine the SOH information. I guess a lot of engineers must have their SOH report when designing and evaluating the performance of their OBSs. Also, if an OBS deployment ends up so smooth, there will be no need to report the SOH information as it is all normal.

We addressed these two points together. We fully agree that engineers routinely examine SOH data during the design, deployment, and evaluation of OBS systems. However, our intention was to highlight that such analyses are rarely shared with the broader scientific community or presented in scientific papers. Nevertheless, we agree to tone down our statements and have removed the sentence "To our knowledge, no extensive examination of SOH information has been published." and we specified that SOH information is not often reported in the **scientific** literature (see line 484 of the revised manuscript).

Figure 1 – The current color scheme is not easy for the readers to see the stations. Have the authors considered plotting the submarine bathymetry in a gray scale? Besides, what does the red star near the Gloria Fault mean? Is that the location of the local earthquakes you showed in Figure 8b? Please explain.

We explored several alternatives, including grayscale bathymetry and varying transparency levels, to improve the visibility of the stations. After testing different configurations, we chose the current version as it provided the best balance between background context and readability for the key features we wanted to highlight. However, we agree that visibility could still be improved and have adjusted the contrast slightly in the revised figure to enhance station visibility further. Regarding the red star near the Gloria Fault: yes, this marker corresponds to the location of the local earthquakes shown in Fig. 8b. We have updated the caption of Fig. 1 to clearly explain its meaning and connection to Fig. 8b to avoid confusion for the reader (see the revised figure and caption in the revised manuscript).

Figure 4 – The citation of An et al., 2017 is problematic.

This has been fixed (see the revised caption of Figure 4).

Figure 6 – I personally feel the green dots are confusing, despite that I understand the authors wanted to address the fact that there are outliers in their dataset. This is because I do not know what the portion of the outliers is, compared with the total data. Just curious if the authors tried the violin plot?

Indeed, we believe that the boxplots that we show nicely summarise a range of information, clearly showing the outliers and key statistical information such as the first and third quartiles. We did also try to build violin plots, but we found that essentially it conveyed the same results as shown by the boxplot. Thus we prefer to keep the boxplots.

Table 1 – The GMRT depth may not be accurate enough for the OBSs. I am wondering if there are any pressure gauge that can provide a more accurate depth.

Indeed, more accurate depth estimates should have been obtained from onboard dedicated ship-based bathymetric surveys. Unfortunately, during the UPFLOW cruises, no high-resolution bathymetric survey was conducted, as the research vessel was not yet equipped with this capability. Therefore, we report GMRT depths in Table 1. Yet, we note that we do not use these depth values for any substantial analyses.

Figure S4 – The figure title "UP19" is very small and hard to see. Please make it as large as the titles in other figures (Figure S2-S3 and S5-S6). I also do not understand what the legends mean in the second panel. The three mass voltages are also mixed together. I think this figure needs to be replotted to be the same quality as other figures.

We updated the figure; a previous figure version had sneaked into the final version of the manuscript.

Figure S8 – The last sentence of the caption only has three words, and seems to be incomplete.

We thank the reviewer for spotting this glitch. We now replaced the previously incomplete sentence by: ”Higher noise levels correspond to darker shaded colours and can be tracked across neighbouring stations.”

Figure S8-S12 – I am not sure if I can really understand how strong the noise level actually is, because there is no colorbar plotted.

We added the colorbar to Figs. S8-S12. See also our response to the other reviewer.

Figure S13 – One pair of additional brackets is found for the reference.

As explained previously in our response to the reviewer’s comments about the clock drift section, we now removed this figure from the manuscript.

Figure S14 – A typo is found in the caption: It should be “Section Section S 7” instead of “Section Section S 7”.

This has been fixed (see Fig. S13 in the revised supplemental materials).

Figure S17-S18 – I think the figures can be great if the authors can label P waves on the waveforms. It is not as easy to see the arrival in Figure S17B.

We updated the figures following the reviewer’s suggestion; please see Figs. S16-S17 in the the revised manuscript.

Table S4 – The preferred orientation is one of the most important results of this study. I am not sure if the authors should move that column to Table 1 in the main text.

We updated Table 1 to include the horizontal orientation results.

3 Reviewer B's comments

Overview

This paper presents analysis of data and metadata of the OBS seismic dataset collected by the UPFLOW project. It discusses data availability and typical noise levels, documents data loss and reasons for loss, and discusses preprocessing analyses relevant to OBS data (tilt/compliance noise correction, horizontal orientation). It provides some very compelling data examples that highlight the quality and potential utility of this dataset. It is well written and mostly at an appropriate level of detail, doing a nice job of going into specifics where warranted but also taking a high-level perspective necessitated by the quantity of the data presented.

Medium issues:

- An omission from this study at present is a comprehensive depiction of exactly which stations/channels have usable data and when. This could save practitioners considerable time in querying/attempting to use the data. The information is very coarsely provided in figure 1/table 1 (as a single number for the entire duration for the hydrophone and seismometer) and somewhat more informatively provided in figures S7, and S8-S12. However, those figures only display the information graphically (which is not very helpful for automation) and also only coarsely, basically amounting to uptime/downtime plots. One addition the authors should seriously consider is a table (in .csv form) that displays whether or not a channel is available at a given station on each day of the deployment. The authors could go a bit further and flag days that DO have data as having mostly normal data or something anomalous (could do this on the basis of some spectral quantity from the PSDs). This would flag up serious power issues, or relevelling events, or whatever. The supplement of Eilon et al., 2021 provides an example of this.

We thank the reviewer for this constructive suggestion.
While issues like power failure, re-levelling events, and severe anomalies were a major concern in the deployment described by Eilon et al. (2021), we found that the UPFLOW dataset is generally more stable. Most spectrograms across the deployment period do not show days that are clearly or extremely anomalous. Noisy days certainly exist, but re-levelling events tend to be brief and rarely affect an entire day's data. Power-related problems or more serious instrument issues are captured in the

SOH logs.
Nevertheless, we agree that a tabulated overview of data availability and basic quality indicators would be helpful for users. We have therefore constructed four CSV files similar to the example provided by Eilon et al. (2021), and added a description of the approach to Section S 5 of the Supplementary Materials and line 630 - 632 of the manuscript, along with the new illustrative Figs. S29–S30.
In brief, for each day, station, and component, we compute the daily power spectral density (PSD) and derive a single RMS value as a summary of spectral energy. We then calculate the log-ratio of this daily RMS to the average RMS over the deployment $log_{10}(RMS_{day}/RMS_{avg})$. To determine whether a day is "anomalous", we examine the distribution of these log-ratios and apply a data-driven threshold: we flag days that fall outside ± 3 standard deviations of the distribution.
This method allows us to systematically and reproducibly flag days with unusually high or low energy without relying on assumptions from continental noise models. That said, we note that many users of OBS data prefer to evaluate data quality themselves and will typically download the entire dataset for their own purposes. Thus, while we believe this CSV file will be interesting for users, it should be seen as an additional aid rather than a strict classification.

- I was a bit surprised to see reference to an "in prep" manuscript concerning the logistics of the experiment. To the extent that those logistics bear on the dataset, that information should really be in this paper. To the extent that they do not, I am not sure why that manuscript is necessary, but I suppose that is the concern of whichever journal it is submitted to. Can the authors clarify that no information about the logistics is relevant to the experiment performance, or add such information into this manuscript?

The referenced "in preparation" manuscript refers to an overview paper submitted to the Royal Astronomical Society magazine, Astronomy and Geophysics. It presents the broad scientific goals of the UPFLOW project and its associated OBS experiment, highlights examples of observations, and documents the logistics and planning of the deployment and recovery of the 50 OBS stations. That manuscript stands independently from this data-focused paper. The present manuscript includes all logistical details that directly impact the data quality, performance, or interpretation, for example, metadata on deployment locations, clock corrections, data availability, and instrument health. Any remaining logistical aspects in the companion paper do not influence the dataset itself but rather serve as documentation of the operational effort behind the deployment, as well as the extensive science outreach work that accompanied it. We hope this clarifies the distinction between the two manuscripts.

- Can the authors clarify how the metric of "data completeness" accounts for four different channels per station (e.g. if three channels operated perfectly for the duration and one totally failed, does this station receive a “75%”?). Since the data use of the vertical seismic, two horizontal seismic, and hydrophone channels are each somewhat distinct, it might be useful to break the data completeness on each apart separately if relevant.

Yes, that is correct, it would be 75%. To quantify the data completeness, we calculated the theoretically expected number of samples against the actually recorded samples for all channels. We added some text to the manuscript explaining that data completeness is calculated over all channels (see the caption of Table 1 and line 218).

- The “quality” on

- Tilt and compliance corrections (Section 4.2): The authors should state in which order they computed the tilt and compliance corrections (hopefully tilt first, per Bell et al., 2015) and whether they re-computed corrected transfer/coherence functions between tilt-corrected vertical and pressure for compliance corrections in the process of this sequential noise removal.

Yes, we first correct for tilt and then subtract compliance. We re-compute corrected transfer/coherence functions between tilt-corrected vertical and pressure for compliance corrections, following the cascade procedure of Crawford and Webb, 2000. We added some text to the manuscript to explain this (see lines 323–325 of the revised manuscript).

- Line 297: The authors could have provided more information about how they calculated PSD functions, including data/sample window lengths, frequency bins, number of samples in spectral windows, taper, and method for PSD estimations. This could go in the supplement, but would be important for another researcher to replicate this work.

We added this information to Table S 5 in the Supplemental Information and we also added explanatory text to the caption of the PPSD Fig. S25. The PPSD estimates were calculated using the ObsPy software with its default parameters. Since this is a well-established and widely used method, we believe that referring to the software documentation and the original publication is sufficient for reproducibility.

Updated Text in the caption of the PPSD Fig. S22.

> PPSDs were calculated using obspy (Krischer et al., 2015) using default values (window length 3600.0 s, window overlap 0.5% and PSD values on

the frequency axis are measured every 1/8 of an octave), as explained in Table S5. Original sampling rates were used during the processing.

Minor Issues:

- Line 22, recommend altering last words to "global mantle dynamics." to avoid the repetition.

Done (see lines 25–26 of the revised manuscript).

- Line 47, not sure if this is just the formatting of the review version, but it would be better to avoid a single line at the bottom of a page below a fig+caption.

This is the formatting of the review version and will not show up in the final version.

- There is some repetition between **lines 47 and 48** regarding "upward/vertical mantle flow"

We have modified the text to avoid this repetition (see lines 55–60 of the revised manuscript).

- Line 88, for clarity, I recommend altering "and was deployed for the first time in a long- term deployment." to "and this experiment marks its first long-term deployment."

Done (see line 106 of the revised manuscript).

- The ordering of stations in Table 1 is confusing and in fact not apparent to me, since it does not appear to be linear in any of the columns. Since most users will search first by station name, I recommend sorting by this (first) column. For the global (read US-based) audience, it might be good to at least note the date convention of DD.MM.YY in the caption.

The original sorting was based on recovery date, which indeed is not the most apparent way of sorting the table. The sorting is now alphabetical based on the first column. Additionally, we added the final horizontal component rotation results from the appendix to the column based on a previous comment from the other reviewer. Please see the revised manuscript for the updated table.

- Line 128: It might be worth noting that the burn wire release mechanism is also standard for all broadband US-based/OBSIC instruments.

Done (see lines 147–148 of the revised manuscript).

- Line 143: the increase in voltage in the latter few weeks is not significant, and so I recommend the authors simply state that the voltage was essentially constant throughout to avoid confusing a reader.

Done (see lines 163–165 of the revised manuscript).

- Line 162: perhaps also refer to figures S3-S6 when discussing instrument power issues, since the first subaxes of those figures shows that.

Done (see line 173 of the revised manuscript).

- Line 168: should probably only refer to Fig. S4 (not "S4-S6") here, since that is the one germane to UP19 which is the station being discussed.

Done (see line 180 of the revised manuscript).

- Line 170-171: I think you mean "orange curve" for seismometer and "blue curve" for datalogger.

Thank you for pointing this out. This has been corrected; see lines 173–174 of the revised manuscript.

- Line 177: The cited figures are slightly off. Should read "(Figs. S5, S10 and S6, S11)”,

Done (see lines 190–191 of the revised manuscript).

- **Data availability**: I recommend collecting response files for all the sensors (or, better, station-by-station) used in this study and attaching them as a zip-file to this manuscript as a supplementary file. That way future users will be able to find them easily.

All the metadata files with instrument responses and other station information are in EIDA and will be openly available after the UPFLOW data embargo period. Thus, we believe a separate zip file in the supporting information is unnecessary, especially as this manuscript and the supplemental information are already quite long.

- Figure S4 seems to have inscrutably small axes labels, even smaller than the similar figures around it. It also seems to have perhaps outdated axes compared to the multi- axis versions of the figures around it (e.g. S2, S3, S5, S6), and the line legends in subaxis 2 are unclear.

Good catch, that figure was built with an old version of SOH plotting code; we now replaced it by the correct figure.

- Overall, the figures in the supplement have axis labels and tick values are very small and hard to read. I am not sure it is worth remaking them all if that is a challenge, but if it is easy to crank through them all and make them more readable, I’m sure people would appreciate that.

- All figures in the main paper could do with slightly larger font sizes on axis labels/ticks. They are mostly legible, if small, on the page, but would be completely illegible if any of the figures were displayed via slides in talks (and some of them are nice/useful enough that folks might want to!)

We address these two comments together. For the main paper, we have updated the figures where we felt axis labels and tick sizes could be improved, and we believe the revised versions offer better readability without compromising the layout.
Regarding the supplementary figures, many of them, such as the PPSD overview plots, are also made available in an online repository upon publication. There, they can be viewed at full resolution, with easier readability. Given the large number of

supplemental figures and their availability in a more accessible format online, we have opted not to regenerate all of them, but only those where it was highly beneficial.

- **Figures S8-S12** seem to have no color bars, which would be helpful to add to each panel/figure.

We added a colorbar to Figures S8–S12 and adjusted the figure captions.

- **Section S8**: I recommend adding panels to figure S21 to show the inter-channel coherence spectra in this time frame, which will indicate how applicable the noise corrections are at different frequencies

We added an extra figure to the supplemental materials - Fig. S21 - with the inter-channel coherence spectra in the same time frame as in the old Fig. S21, which is now Fig. S22.

- Figure 7: Make sure it is clear that this is for the vertical component. Also, I'm not sure it makes that much sense/is that useful to show this averaged across the entire array; Janiszewski et al., 2023 showed that the compliance noise (via the P-Z coherence) is very depth dependent, and so presumably your correction also is. Perhaps you could remake this figure with a few subplots, binning instruments in each subplot by different depth (d) ranges (where you could bin by sqrt(d) since the high-f cutoff for compliance goes as $1/\sqrt(d)$). You could also put this theoretical cutoff on the plot if that made sense.

We have clarified in the figure caption that the PSD comparison shown is for the vertical component, as suggested. While we agree that plotting the correction effect by depth bins could offer more targeted insight, we prefer to keep the figure as it is for a simpler, first order analysis of the effect of tilt and compliance corrections.
We also note that Figure S22 provides individual PSD comparisons before and after tilt and compliance correction, offering more detail than the array-wide average in Fig. 7.

- Figure 9 made me go “woah”! Figure 10 is similarly spectacular. Very cool. Figure 1 is also particularly beautiful and data rich.

Thank you for this encouraging comment.