Summary of replies to Rev1

All responses to the reviewers' comments, as well as the corresponding updates introduced in the manuscript, are highlighted in red to facilitate quick and easy tracking.

Main Comments

In the introduction, the authors rightly point out that long-term observations in the deep regions of the Mediterranean Sea are rare. They should also mention the various types of operational deep observatories that complement their own. Notably, these include the HYDROCHANGES network (Schroeder et al., 2013; https://ciesm.org/our-science/hydrochanges/), which has continuously monitored hydrological and hydrodynamic variability at various locations in the Mediterranean since 2002. This includes deep sites in the Adriatic and northwestern basins, as well as key locations such as the Corsican, Balearic, Sicilian and Gibraltar straits. Further results concerning the long-term evolution of hydrological conditions in the deep basins of the Mediterranean can be found in Chiggiato et al. (2023). It would be interesting to include a perspective on the convergence of measurements and the complementarity of the different datasets in the introduction or conclusion.

We thank the reviewer for this suggestion. The Introduction has been revised accordingly [lines 58–63 and 67–68] to improve clarity and provide additional context, and relevant references have been added.

Regarding the post-processing of current meter data, I was interested in the results in Figure 3, which compare current measurements obtained by an ADCP and a punctual current meter. According to Table 1, the acoustic punctual current meter (Falmouth 3D-ACM) operates at 2 Hz, while the ADCP (RDI WH 300 kHz) operates at 2.8×10^{-4} Hz (1 sample/hour), and the comparison is made using hourly data. The hourly average of 7200 measurements from the punctual current meter is then compared to the hourly measurement from the ADCP. Is the hourly ADCP measurement the result of a single ping, given that the standard deviation for a single ping is at least 2 cm/s? Or is it the result of averaging an ensemble of pings, in which case the standard deviation would be lower? Information on the acquisition modes specific to certain instruments and the standard deviation associated with the measurements for each instrument is important for making comparisons. Additionally, it would be preferable for this figure to plot the 1:1 line rather than the least-square regression line.

The Falmouth 3D-ACM current meter recorded velocities at 2 Hz. For comparison with the ADCP record, the time series was smoothed using a 1-hour moving average and sampled at hourly intervals to obtain an hourly estimate of the flow. The RDI Workhorse 300 kHz ADCP was configured with 100 pings per ensemble and a ping interval of 6 s, resulting in ensemble durations of approximately 10 min. Each reported velocity profile therefore represents an internal average of multiple pings, and the hourly ADCP measurements used in this study correspond to these ensemble-averaged profiles. In accordance with

reviewer's suggestion, Figure 3 has been revised to display the 1:1 line as well as the least-squares regression line. The text has been updated accordingly [lines 229-239].

Figure 4 illustrates the variation in data quality across different sensors and sites, categorising data as good (flag = 1), interesting/suspicious (flag = 3), or missing/removed (flag = 9). Firstly, the flag labels should be consistent between the figures and the legend. Furthermore, it is surprising that most of the data (between 50% and 90%) for the different sensors is considered interesting/suspicious. Could the authors provide details and explain why this percentage is so high after the data has undergone a quality check? Are these percentages associated with the raw or final processed data?

We thank the reviewer for this comment. The original figure contained an error in the color reference of the quality flags, which were inadvertently inverted. This has been corrected in the revised figure, where the labels are now properly aligned. As a result, most of the data correspond to the good quality flag, while the suspicious and rejected categories represent only a small fraction of the dataset. The figure and caption have been corrected accordingly in the revised manuscript [lines 293-294].

Figure 6 shows the quality control performed on a time series of current intensity, as measured by an ADCP, as well as the two-level time series. Why does the Hovmöller diagram not start in mid-December 2003, as it does for the time series? This would enable us to observe the filtering of the initial data, which appears very noisy in the upper part of the water column sampled by the ADCP. The depth and altitude relative to the bottom of the measurements should also be indicated systematically in the figure titles and legend for the time series to help guide the reader.

We originally displayed the Hovmöller diagram over a shorter time window to allow a clearer comparison between the raw and quality-controlled data, focusing on a representative month where the effect of the filtering could be easily visualized. However, following the reviewer's suggestion, the Hovmöller diagram has now been extended to cover the full time series, aligned with the temporal range shown in the time series panels below. This modification allows the reader to observe the initial noisy period in December 2003 and the effect of the quality control on those data. In addition, the labels have been revised to report the absolute depth of the measurements, improving readability and making the information on sensor position clearer in the figure titles and legends. Caption has been updated accordingly [lines 388-390].

Figure 8 shows time series for which missing data has been filled using a method that combines singular spectral analysis and optimal interpolation. It may be helpful to review the definitions of the K1 (lunisolar diurnal) and M2 (principal lunar semidiurnal) tide components, and the origin of the inertial frequency f. The sentences "As shown in Figure 8, the warming signal in the reconstructed time series confirms the robustness of the spectral estimate" (lines 443–446) are unclear to me. Could you please explain your reasoning more clearly?

We agree that the original wording could be misleading. The sentence was intended to indicate that the spectral estimates derived from the reconstructed time series are supported by the 95% confidence limits, within which the main spectral features are contained. We have therefore rephrased this part of the manuscript [lines 454–470] to clarify this point and improve the overall explanation.

Finally, it would be interesting if the authors could provide an example of the simultaneous variation of different variables, such as temperature, salinity, current, and turbidity / oxygen. This could be over a short period to illustrate an event for which high-frequency measurements are useful, or over a longer period to illustrate seasonal or longer-term variability.

We appreciate the reviewer's comment. Figure 7 is intended to summarize the variability observed in the temperature time series at the four sites, together with the hodographs relating measured velocities to water-mass densities, in order to highlight the main dynamical features of the observed variability. The figure was carefully designed to provide a comprehensive overview within a single panel, while avoiding the inclusion of additional plots that would largely repeat the same information and unnecessarily increase the length of the manuscript. We believe that this approach helps maintain clarity and conciseness while still conveying the key dynamical aspects of the dataset.

Summary of replies to Rev2

All responses to the reviewers' comments are provided in red to facilitate easy tracking, while references to revisions made in the manuscript are highlighted in blue to distinguish them from those introduced in response to the first review.

Main comments

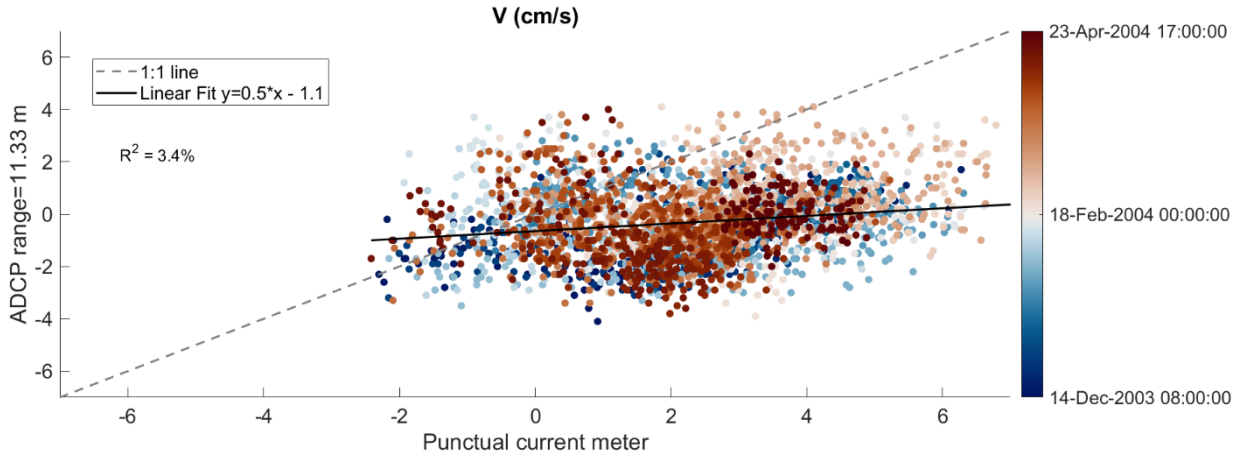
In L.242 it is mentioned *"cross-correlation techniques"* used, particularly between current meters and ADCPs. Could the author explain what are these techniques? Because in the following part of this paragraph the R^2 coefficient is used, which is a regression analysis coefficient considering no lag, while cross-correlation depends on it. If there was a confusion, I would suggest to just change *"cross-correlation techniques"* by *"regression analysis"*. I would also like to point out that the R^2=0.28 named as *"reasonable agreement"* is more a *"moderate agreement which requires particular attention"*. What is written in L.232, if depending only on R^2, should be subject to more caution: *"This consistency…"* could be rephrase to *"moderate consistency"*. Concerning properly the Fig.3, a sentence in the text explaining that the *"ADCP range=11.33"* represent the depth of the current meter (I think so at least) is necessary. There is a last element I was curious about: what were the result for *v*, the other component of the horizontal velocity?

1. **We agree that the use of the term "cross-correlation techniques" could be misleading, then, following the Reviewer's suggestion, we have replaced "cross-correlation techniques" with "regression analysis" to more accurately describe the methodology used. Regarding the interpretation of $R^2 = 0.28$, In the revised manuscript, this has been changed to "moderate agreement" better explaining the context of the measurement setup. These clarifications have been incorporated into the revised manuscript, as follows**:

"For instance, the horizontal velocity component (U) measured by the punctual current meter and by the ADCP deployed at GEOSTAR-SN3 during the 2003–2004 mission showed moderate agreement ($R^2 = 0.28$, Figure 3). This result should be interpreted in the light of the differences in functioning, and positioning: the ADCP samples the open water column, whereas the punctual current meter operates within the observatory frame. Also, the current-meter record, acquired at 2 Hz, was smoothed using a 1-hour moving average and sampled at hourly intervals to allow comparison with the ADCP data. "

Accordingly, L232 has also been revised to "moderate consistency".

For what concerns the other component, see attached Fig.3_V:



On the figure 4, I was wondering if the result is the final one or just after one step of QC because a large majority of *"Interesting/Suspicious Data"* is present. I would have also added a horizontal space between bars when moving to another site for readability (bigger horizontal space between bar 5 and bar 6, between bar 11 and bar 12, and between bar 16 and bar 17).

2. To be fair, also thanks to the reviewer1, we noticed that in the originally submitted version, the color labels were inadvertently reversed: Good data are represented in blue, not yellow. We have corrected this in the revised figure. Additionally, following the reviewer's suggestion, we have improved the overall visualization to better distinguish between the different datasets.

L.320: It is written that the rate of change (ROC) is *"equivalent to removing outliers, i.e. data over three standard deviations from the mean"*. However, up to what I read on the QUARTOD manual, the ROC is not demonstrated to be equivalent to that. ROC checks the temporal consistency of a time series by comparing changes between successive

observations while the mean+3 standard deviations is an outlier detection method relative to the mean. I kindly ask the author to explain from where the equivalence assertion comes from.

3. We agree that the Rate of Change (ROC) test and a classical outlier detection based on mean ± 3 standard deviations are conceptually different and should not be described as equivalent. The ROC test evaluates the difference between consecutive observations, requiring that the rate of change between sample $n-1$ and n remains below a threshold defined as a multiple of the standard deviation (e.g., $N\_DEV \times SD$). Importantly, this standard deviation is computed over a moving time window to account the natural variability. Both the number of standard deviations (N_DEV) and the time window (TIM_DEV) are operator-defined parameters. In this framework, the standard deviation is used to characterize the local variability of the time series, allowing the threshold to adapt to changing environmental conditions. Therefore, describing this approach as equivalent to a mean ± 3 SD method would be misleading, since the latter identifies outliers based on their deviation from a central value and does not account for temporal sequencing. We have therefore revised (in blue) the manuscript as follow:

Rate of change – This test inspects the time series for changes between consecutive observations that exceed a defined threshold. The observed quantities can change substantially over short periods in some locations, hindering the value of this test, so the thresholds have been chosen carefully and tailored based on the sensor's operational limits and location. The threshold is defined using the local variability of the signal (e.g., a multiple of the standard deviation computed over a moving time window). Unlike classical outlier detection methods based on deviation from a mean value, this test evaluates the temporal consistency of the measurements and flags, rather than removes, suspect data. In the case of dissolved oxygen measurements, because of its dynamic nature, this test does not involve outliers' removal but flagging.

L.330: Concerning the sensor tilt test, it is written *"this test can be replaced with a rate of change test performed along the vertical (i.e., along depth instead of time)"*. Here again, this requires further clarification. The vertical gradients may provide information, but the equivalence of the proposed method with a tilt detection test is not generally established.

4. **Thanks for this comments that one more time evidencing a possible misleading. Our intention was not to imply a strict equivalence between the sensor tilt test and a rate-of-change test applied along the vertical. In the QUARTOD framework, the tilt test is specifically designed to detect changes in sensor orientation that may affect current measurements. In our case, the sensors are mounted on a fixed and stable seafloor platform, where significant tilt variations are not expected under normal operating conditions. Therefore, the relevance of a dedicated tilt test is reduced. Instead, we proposed a complementary approach based on inspecting the vertical consistency of the velocity profiles (i.e., gradients along depth). While this does not directly measure sensor orientation, abrupt or non-physical vertical variations can indicate potential issues, including misalignment, platform disturbance, or measurement artifacts. We acknowledge that this approach is not equivalent to a tilt detection test and should not be interpreted as a replacement in a general sense. We have revised (in blued) the manuscript to clarify how this method is used:**

Sensor Tilt (for current meter and ADCP) - Current sensors must be aligned within an expected range of tilt angles to properly measure horizontal and vertical currents. For fixed-

mounted sensors, as in the case of seafloor observatories, significant variations in tilt are not expected, and the applicability of this test is therefore limited. In this context, we examined the vertical consistency of ADCP velocity profiles (i.e., gradients along depth), which can help identify non-physical variations potentially associated with instrumental issues or platform disturbances, although this approach does not represent a direct substitute for tilt measurements.

L.342: *"All tests were conducted within the valid measurement range for the instrument, as specified in Table 4"*. Table 4 provides min, max, mean, and standard deviation but the *"valid measurement range"* is nowhere to be found. I kindly ask the author to explain what they intended by it (manufacturer specifications?, QC thresholds?)

5. The sentence referred to a previous version of the manuscript that included a table reporting the operating limits of the sensors, which was subsequently removed. As a result, this sentence is meaningless and has simply been removed.

Figure 7: This figure is presenting the time series of potential temperature and hodographs in the four different sites, but, even if I like the idea of presenting a map with the results, is hardly readable. I am unable to read legend of time series figures, hodographs are too small, their ticks label, colorbar values and labels, title are far too small. Concerning hodographs, it is also written *"Eastward [Northward] velocity (km)"*, and a velocity unity is not km. Hodograph are usually presented with the North as 0° and angles increasing in the clockwise direction. I would suggest the author to refine the figure so a person that print the article can read it. I suggest the author the use of tools like Gimp to merge different figures and save the result in a high quality (300dpi). On this figure caption *"Temperature"* should be changed to *"Potential temperature"*, and a more accurate description of the temporal scale at which the *"variability [...] in the Tyrrhenian Sea"* is observed is necessary.

6. The hodographs were generated using the JLAB hodograph routine, part of a toolbox specifically developed for oceanographic applications, which plots the eastward velocity component against the northward component, i.e., a Cartesian U-V hodograph.

Figure 7 has been improved following reviewer suggestion.

L.398: *"in agreement with the global warming trend"* should be nuanced as the amount of data is lower than 5 years for 3 of the sites. Such trends could only reflect a regional or decadal variability rather than a long-term climate change. I kindly ask the author to temper this statement or clarify its limitations.

7. We thank the reviewer for this comment, and we understand the caution suggested that is needed when interpreting trends derived from relatively short time series. Our intention was not to claim a robust detection of long-term climate change, but rather to highlight that the observed positive temperature trends across all monitored deep sites are consistent in sign and in agreement with the broader warming patterns reported in the literature, including IPCC (2013). We acknowledge that, for several sites, the available time series are relatively shorter and therefore may reflect local or decadal variability, as discussed later in the text.

However, the coherence of the signal across multiple independent deep-sea observatories, combined with the known sensitivity of the Mediterranean basin as a climate hotspot, supports the relevance of these observations.

We have revised (in blue) the manuscript to better clarify these limitations and to temper the interpretation accordingly.

While the available time series are relatively short for each site and might therefore reflect regional or decadal variability, all temperature trends measured between 2001 and 2013 at all deep layers are coherently positive and broadly consistent with the warming patterns reported in the literature (IPCC, 2013).

L.404: *"providing a coherent and reliable estimate of a decadal scale warming process"*,

8. **fixed**

L.410: *"not a simple warming signature but a real change of the deep water masses"*,

9. **fixed**

L.417: *"[The shift in the hodograph] may be attributed to the alternating advection of water masses that the Ionian basin receives from the Adriatic or Aegean Sea, which could sustain the better known decadal reversals (BIOS)"*. I kindly ask the author to clarify whether the observed warming is really reflecting a long-term trend or more a variability associated with BIOS phases. In the case where a doubt subsist, reliability of the trend estimate can be debated.

10. We thank the reviewer for this important comment. It is important to clarify that this study does not aim to provide a comprehensive assessment of long-term climate change. Rather, its objective is to highlight the presence of a consistently positive temperature signal in the deep layer across multiple sites, and to emphasize its potential relevance. In this case, we distinguish between processes acting at different temporal scales. The changes highlighted by the hodograph, and the shift in current direction, reflect a reorganization of deep circulation, likely associated with BIOS-related variability. At the same time, the observed temperature increase represents a local long-term tendency, which, although derived from relatively short time series, is coherent across sites. These processes may be interconnected: modifications of deep thermohaline properties, can influence the density structure of water masses and potentially contribute to circulation changes at larger spatial and temporal scales.

The main concern I have is about the last part of the article and the use of SSA/OI. While using SSA/OI is reasonable in the case of gaps mitigation, some statements may be overstated there. L.430 *"SSA is a [...] method particularly suitable for time series with long and continuous gaps"* while Kondrashov and Ghil (2006) (cited in the article) concluded that *"The accuracy and reliability of the method depend on the pattern of missing data, the relative length of the gaps with respect to the total length of the data set"* ; the article presented in Fig.8 a time series of 3 year with a gap of more than a year, so the accuracy of SSA in this case is not proved. On L.431 it is also written *"does not require a priori*

assumptions", but this is not fully true as for instance the number of modes retained is important. Concerning the modes used for this analysis, the author used *"the dominant modes (6 mode)"* without giving any explanation for this choice. Could the author clarify this point in the text? L.441: *"ensuring temporal coherence and dynamical consistency"* This is I think may be overstated because SSA/OI will provide a statistical coherence, but it does not guarantee physical/dynamical consistency. L447-L.448: *"dominant energy at low frequency, reflecting the slow and persistent variability typical of long-term warming process evident in the reconstructed time series"* Here it is not that *"evident"* because the low frequency is not necessarily equal to the warming trend, it could be due to decadal variability, changes in the circulation, all the more as the time series chosen for applying the SSA/OI method are both less than 3 years short. The *"robustness of the spectral estimate"* (L.446) may be acceptable but it refers on a time series which is too short to be fully statistically significant. I suggest the author to refine the part concerning SSA/OI, avoiding the overstatement I mentionned above and providing a more balanced discussion with the limitations of this method. It was an interesting analysis which has still its place in this article and which could benefit comparison with models outputs, reference measurements in another article. Following this comment, the conclusion has to be modified in consequence; for instance in *"the reconstruction of longer and more continuous records increase the resolution and robustness of spectral estimates..."* the *"robustness"* may not guarantee by reconstruction especially with SSA/OI and in *"the application of advanced analytical methods [...] minimized spectral leakage and enhances the reliability of frequency domain diagnostics"*, the article did not show a leakage reduction quantitatively, neither a quantitative validation for the *"reliability"*.

11. **We thank the reviewer for this detailed and constructive comment. We agree that some statements in this section were overstated and have revised the manuscript to provide a more balanced description of the SSA/OI approach and its limitations. We acknowledge that the performance of SSA-based gap filling depends on the data structure and gap distribution, as highlighted by Kondrashov and Ghil (2006). In our case, the relatively short time series and the presence of a long gap do not allow for a rigorous assessment of reconstruction accuracy. The SSA/OI method was therefore not intended to provide a quantitatively validated reconstruction, but rather to obtain a continuous time series suitable for estimating large-scale, low-frequency variability and associated trends. This clarification has been added to the text. Regarding the statement that SSA "does not require a priori assumptions," we agree that this was imprecise. While SSA is non-parametric in the sense that it does not assume a predefined model, it does require methodological choices, such as the selection of the embedding dimension and the number of modes retained. We have revised the text accordingly. Concerning the selection of the six dominant modes, this choice was based on the fraction of explained variance, retaining the leading modes that capture the majority of the signal while excluding higher-order components dominated by noise. We now have explicitly stated this criterion in the manuscript and provided a reference to similar approaches (e.g., Fig. 8a in Marullo et al., 2011 https://doi.org/10.1175/2011JCLI3884.1) to support this choice.**

Secondary comments

L.188 : *"storage efficiency"* while on L.130 it is named *"data acquisition efficiency"* and in Table.1 *"sensor efficiency"*. It is necessary to use every time the same definition. The definition provided L.213-L.214 could be given earlier in the text (L.130 or L.188).

12. **changed**

Table 1: Change *"OXYGEN"* by *"DISSOLVED OXYGEN"*. Consider that in Table 4 it is referred as O_2 too.

13. **We clarify that in Table 1 the term "OXYGEN" refers to the sensor type and therefore cannot be replaced with "DISSOLVED OXYGEN," which denotes the measured variable. In Table 4, however, the notation has been revised, and "O_2" has been replaced with "DO" to ensure consistency when referring to the measured parameter.**

Figure 3: it is often written in the text *"horizontal velocity component (u)"* to refer to the meridional velocity component (u). Horizontal velocity component could be misleading as it is constituted by *(u,v).* I found occurrences in L.223, Fig.3 caption, and L.337. In L.341 it is mentioned more clearly *"eastward velocity component"*.

14. **fixed**

L.240: What does the author mean by *"stability range"*? This could be explained more carefully.

15. "Stability ranges" refer to the acceptable limits of variation in sensor measurements over time, ensuring that the sensor does not exhibit significant drift or degradation in performance. To avoid misunderstanding it has been better explained as follow (in blue):

Before any QC procedure can be meaningfully applied, however, rigorous Quality Assurance (QA) is essential to ensure that the sensors themselves provide measurements within their expected accuracy and long-term stability specifications (i.e., limited drift over time).

L.241: I would mention after *"CTD casts"* that they are performed by CTD instrument that have higher performance than the mooring CTDs (more likely it is SBE911 or an equivalent). The sentence is already good but adding this would improve it.

16. **done**

L.244: Would the author have a reference for *"even the most sophisticated QC procedures cannot fully compensate for poorly calibrated or improperly functioning sensors"*?

17. added

L.250: It is written "enabling prompt corrections that can minimize data gaps" but a correction will not minimize a gap as there is no data to correct. In this sentence "data gaps" should be removed.

18. **done**

Table 3 is missing at least one reference.

19. **added**

L.252: *"It involves comparing with reference datasets, historical records, or model outputs"*. For a comparison, models are ok. For a correction it is trickier. As the sentence before mentions about a "more comprehensive review to improve data accuracy" and after "correction of errors missed" is written, I would suggest the author to be cautious when mentioning models in a part concerning quality control, all the more as the study consider deep-sea measurements.

20. thanks for the suggestion. The reference to the models has been removed

L.266-L.269: To be sure, *"a QC protocol was customized [...] Custom thresholds"*, means that the QC was following QUARTOD with modification of some coefficient in the tests considering the different sites area? If so, it is, of course, not a problem as QUARTOD is internationally recognized.

21. **yes, QC protocol follows QUARTOD standards and was further implemented as a tailored protocol, with sensor-by-sensor adjustments of thresholds and parameters based on the specific environmental conditions at each site.**

L.290: I do not understand what the author is referring to when writing *"sensible"*? *"Sensible"* is too vague if not explained clearly.

22. changed with reliable

L.303 to L.306: the sentence *"For Teledyne [...] only one beam."* is a bit long and difficult to understand. Maybe the author could split it in two sentences and improve its clarity.

23. it has been change as follow (in blue):

For Teledyne RDI ADCPs, when the coordinate frame is not set to beam coordinates, the Percent Good test refers to the percentage of valid three- and four-beam solutions, i.e., the proportion of data for which no beams or only one beam has been rejected by the sensor.

L.321 and L.349: *"dynamic nature"* referring to the oxygen measurements is too vague. It should be explained more carefully (biological activity, physical mixing,...) and supported by a reference.

24. **added Fusi et al. 2023**

Figure 5: I really like the figure because it is visual and we clearly see the improvement given by the QC. I have a few suggestions. On the legend of the second subplot, *"Raw after step 1"* is ambiguous. It would be better to write *"Data after step 1"*. In the caption, change *"Quality control"* by *"QC"* as QC was widely used in the previous pages, *"(a) Raw data showing outliers"* could be replaced simply by *"(a) Raw data"*

25. **improved**

L.362: *"... where the measurements are noisier."* Here clarifying why the measurements are noisier would improve the sentence (surface wave effects, turbulence, or instrument limitations,...)

26. **sentence improved as follow:**

…particularly in the upper ranges where the measurements are noisier due to greater signal dispersion.

Table 4: *"W"* is not defined earlier in the text; the reader should be able to clearly understand it refers to vertical velocity component. Minimum and maximum are written on two lines which Is not visually pleasant.

27. **improved**

Figure 4: The author could have written in the text a sentence about the strong velocity difference present (at maximum 15cm/s) between 5.33m range and 20.33m range, which is an interesting aspect. On the caption, *"2004"* is missing after *"January"*

28. **Maybe reviewer was referring to Figure 6? Here the caption was already changed following also suggestion of first reviewer:**

Figure 6: Quality control of ADCP velocity magnitude data at GEOSTAR-SN3 (2003–2004). (a) Color map of current magnitude across depth: top panel = raw data; bottom = QC-checked. (b) Time series from two depth bins: range 5.33 m corresponding to -3314.67 m depth (on the top panel), and range 20.33 m corresponding to -3299.67 m (bottom panel). Black = raw data, green = quality-checked data, yellow = flagged.

The observed velocity peak corresponds to a volcanic ash fallout event that is currently under investigation. However, details on the specific processes are intentionally omitted, as their inclusion would go beyond the scope of the present study.

L.386 to L.389: *"After post-processing a QC validation routine has been performed, potential temperature and density anomaly data derived following TEOS-10 international standards (https://teos-10.org/), exhibit interesting variability at all sea-bottom sites monitored during these years (Figure 7)"* could be replaced with *"After post-processing, a QC validation routine has been performed, potential temperature (θ) and density anomaly (σ_{2000}) data were derived following TEOS-10 international standards (https://teos-10.org/, [date of last access]). They exhibit interesting variability at all sea-bottom sites monitored during these years (Figure 7)".* Maybe the definition of potential density anomaly given L.412 could be also included here for more clarity. Concerning the definition given L.412, σ^2 is the symbology used, but it could be interpreted to a variance. The author should consider using σ_{2000}.

29. **fixed**

L.397: *"measured between 2002 and 2013 at all monitored deep sites"* can be misleading as all the different sites do not have data for the period 2002-2013, only one have. A more accurate sentence should be written.

30. this sentence has been already changed following previous comment

L.411*: "This is evident looking at the current hodograph"* Here *"evident"* may be overstated as no information about dates are given on the hodograph. The sentence could be changed to *"This is evident looking at the current hodograph, also showing potential density anomaly, where velocities are found to be oriented toward …° with a mean potential density anomaly of …kg/m^3 on the period … while they are oriented toward …° with a mean potential density anomaly of …kg/m^3 for the other period…"*

31. **improved**

L.413 to L.415: *"four times bigger than the usual range of inter-annual variability expected at these depths in the Ionian bottom water"* Here a reference is missing to justify this assertion.

32. **fixed**

Other small comments (editorial corrections, rephrasing, missing reference in the list)

L.25: it would be better to replace *"high-resolution"* by *"high temporal resolution"*

33. **fixed**

L.35 an L.37: *"The unsteady state of the deep ocean is a quite new achieved knowledge […] that emphasizes the significance of abyssal processes in redistributing heat and energy, thereby influencing surface climate variability"*. Here, first, I would replace *"quite new"* by *"recent"*.

34. **fixed**

L.41: *"The deep sea is indeed a highly complex and interconnected environment"* is vague. I would suggest changing it by something like *"The deep sea is indeed governed by physical and biogeochemical processes occurring on a wide range of temporal and spatial scales, resulting in a complex system."*

35. **fixed**

Following my previous comment I would modify the sentence L.42 to *"It is influenced by the continuous vertical exchanges across the water column and at the air-sea interface, as well as by lateral exchange with surrounding ocean basins, guided by morphological dynamics"*

36. **fixed**

L.81: *"challenging […] conditions"* could be removed, it is already written in L.78 *"extreme conditions"*

37. We thank the reviewer for the suggestion; however, we prefer to retain this element, as it helps illustrate the types of issues that can affect this kind of system.

L.187: *"Preliminary analysis was always conducted, as the first step…"* Here *"first step"* is just a repetition as *"preliminary"* was mentioned before, so I would only keep *"preliminary"*

38. **fixed**

L.189: *"custom-designed software tailored"* is a bit heavy in my opinion, it would be better to keep it simple as the *"custom-designed software"* are not presented here. I would stay with *"software tailored"*

39. **fixed**

L.192-193: *"This process enhances the accuracy, reliability, and interpretability of the raw data, refining its quality and facilitating better interpretation"* Here *"refining its quality and facilitating better interpretation"* is not necessary as it is just repeating the information which is given in the first part of the sentence.

40. **fixed**

L.209-L.211: *"thorough […] meticulous"* Using both is redundant, it would be better to keep only one of them.

41. **fixed**

L.212: *"[…] providing a solid foundation for subsequent analysis and interpretation"* I do not find this necessary; it was already clear that we were talking about the first stage of post-processing.

42. **fixed**

L.215: "Marmara Sea (Table 2)" should be changed to "Marmara Sea (Table 1)"

43. **fixed**

L.215: *"very high"* Instead of this, giving a value would be more appreciated *"greater than 90%"*

44. **fixed**

L.233: I would replace *"height"* by *"depth"*

45. We thank the reviewer for the comment. This aspect actually refers to the different height relative to the punctual current meter; the sentence has therefore been revised accordingly, specifying the high from the seafloor

L.238: The use of *"however"* in the middle of the first sentence of the paragraph is confusing. Considering the sentence, it can simply be removed.

46. **done**

L.248: Maybe it is only my opinion but *"invaluable"* is a bit too much. *"necessary"* would have been enough.

47. **changed**

Table 3: Space missing between *"Tyrrhenian"* and *"Sea"*

48. **done**

L.289: *"timestamps requires"* should be changed to *"timestamps. It requires"*

49. **fixed**

L.296: *"The global range"* should not be replaced by *"The gross range"*?

50. **yes, changed**

L.301: *"the beams' percent good"* would be better written *"the percent good of the beams"*

51. **fixed**

L.314: I would change *"This test checks for single value spikes usually due to an electrical signal from the sensor, relative to adjacent data points"* to *"This test checks for single value spikes relative to adjacent data points, usually due to an electrical signal from the sensor."*

52. **fixed**

L.322: *"outliers' removal"* should be changed to *"the removal of outliers"*

53. **fixed**

L.324 and L.326: *"since"* is used two times in the same sentence. The second since could be replaced by *"which"*

54. **done**

L.358: *"This sensor measures current time series"* is not totally correct as ADCP measures acoustic Doppler shifts and then derives current velocity. The sentence could be simply modified to *"This sensor measures indirectly current time series"*

55. **done**

L.359: *"Consequently, visual inspection of the raw dataset is more difficult to interpret, as can be seen in Figure 6(a), noisy data is not immediately evident in the raw data"* could be modified to be clearer to *"Consequently, visual inspection of the raw dataset is more difficult*

to interpret and noisy data is not immediately evident in the raw data, as can be seen in Figure 6(a)"

56. **done**

L.384: *"EOVs"* should be explained *"EOVs (Essential Ocean Variables)"*

57. **Improved**

L.410 *"one decade is"* should be replaced by *"one decade, it is"*

58. **Improved**

Ghil et al., 2002 is missing in the references list

59. **added**