

Response to referees' comments on the manuscript "*SHELDA: Sub-hourly European Quality Controlled Sea Level Dataset*" by Balić and Šepić, submitted to Earth System Science Data (essd-2025-767)

We sincerely thank the anonymous reviewer and Ivan Haigh for their valuable comments and suggestions, which improved the quality of the manuscript and the related SHELDA dataset.

Reply to RC1: 'Comment on essd-2025-767' by Anonymous Referee #1:

The paper describes quality control performed on 257 high-frequency sea level records (sampled at intervals from 1 to 15 minutes) along European coastlines. The manuscript is generally well organized and clearly written, with a few shortcomings noted below.

The dataset, as described, has the potential to make a significant contribution to the scientific community, particularly for researchers investigating sea level processes on shorter timescales. However, the dataset available via the provided link currently contains only 30 station records, rather than the expected 257.

In light of this discrepancy, I provide comments on both the manuscript and the available subset of 30 records, and I recommend rejection with the possibility of resubmission once the complete dataset is made accessible.

- **We thank the reviewer for the thorough review and positive feedback on the organization and clarity of the manuscript. We greatly appreciate your recognition of the dataset's potential contribution to sea level research on shorter timescales.**
- **As for the dataset availability, we apologize. We originally uploaded the entire dataset and were not aware that it is not fully available via the provided link. Thank you for noticing this. Following the requests, we have implemented the required changes to the NetCDF files and updated the dataset. The VLIZ team has been notified of the previous issues, and we have ensured that all records are now fully available for download.**

Comments on the paper

L15: Please specify the license under which the data are distributed.

- **The license information is added.**

L17: "rigorous quality control procedures": Please briefly outline the main steps undertaken, as well as those not applied (e.g. corrections of datum shifts and drifts). Also include in the Abstract the minimum and maximum lengths of the available records.

- **We included the requested information on quality control steps and record lengths in the abstract.**

L20: Please use a consistent name for Turkey throughout the manuscript (see also Fig. 10).

- **We revised Fig. 10 accordingly and ensured that the Türkiye is used consistently throughout the manuscript.**

L49: Please rephrase “some kind of averaging procedure”, as this is not appropriate. Also, I doubt you would measure wind waves with a tide gauge instrument typically installed on the coastline (as visible from your maps).

- **Rephrased.**

L58: ‘reduced’ > higher?

- **Changed.**

L91–94: Please rephrase for clarity; it is not clear whether MISELA includes only the high-frequency signal or both high-frequency and total sea levels.

- **We rephrased to make it clear that MISELA includes only the high-frequency sea-level signal.**

Figure 1: Can you offer a better solution for records shorter than 3 years, as the white circles on the grey background are not visible?

- **We revised Figure 1 by changing the colour of the circle.**

Sections 2.2.1 and 2.2.2: These sections would benefit from clearer organization. The current structure with multiple (i), (ii), (iii). points, some connected, others not, makes them difficult to follow.

- **We reorganized Sections 2.2.1 and 2.2.2 to improve clarity, as suggested.**

Figure 6: (%) > [%]

- **Changed.**

L211: I would refer to all data from one station (regardless of how many segments with different sampling frequencies it contains) as a “record,” as the current usage is confusing (e.g. the number changes from 257 to 275 and then back to 257).

- **Thank you for pointing out the potential confusion in terminology. To clarify, we strictly define a "record" as the complete time series from a single station (totalling 257 records). However, because different sampling frequencies require different pre-processing and QC thresholds, the 17 multi-frequency records had to be temporarily partitioned into 35 "segments" of continuous frequency. We revised the paragraph accordingly. The distinction between a "record" and a "segment" is now clear.**

L220: I understand that correcting datum shifts is not possible without additional data. However, users would benefit from knowing which stations are affected, so that they can exclude them if necessary. I recommend including this information in the manuscript and in Tables A1–A3 (and also in NetCDF files). Furthermore, please specify in the paper what you consider to be a shift, a datum shift, and a drift (including possible causes).

- **We added information about types of erroneous data found at each station record in both Tables A1–A3 and in NetCDF files. Furthermore, we defined what we consider a shift, a datum shift, and a drift (including possible causes).**

Figure 3: The bathymetry color bar is missing.

- **The bathymetry colour bar has been added. The updated figure uses a narrower colour bar range to show the bathymetry in more detail.**

L331: Please clarify what specific issues were identified in the German data that led to their exclusion from SHELDA.

- **Clarified.**

Table 1: See comments on NetCDF files below.

- **All taken into account.**

Figures 7, 8, 9: Please define what is meant by a step sensor.

- **Defined in the revised manuscript.**

L380: In some places it is unclear whether you are referring to the original or processed data. For example, is the 1-minute interval the most common in the original data or after processing? Please clarify.

- **In this sentence we refer to the 1-minute interval in the processed data, but the statement also holds true for the original data, where 1-minute sampling is the most common. We clarified this in the revised manuscript to avoid ambiguity.**

L380–387: It would be useful to include a brief discussion explaining the observed distribution of sampling frequencies (e.g. related to dynamics?) and sensor types (e.g. infrastructure constraints?).

- **We revised the manuscript by providing a more detailed explanation of: (1) the relationship between sampling frequency distribution and coastal dynamics and monitoring objectives (e.g., 15-minute sampling for tide and storm surge monitoring versus 1–10 minute sampling in regions prone to tsunamis, seiches, and meteotsunamis), and (2) how the distribution of sensor types reflects infrastructure constraints, funding availability, and the development of national monitoring networks across Europe.**

L394: Please include the data license.

- **Added.**

L471: ‘article’ > ‘manuscript’

- **Replaced.**

L19, L133, L397: In my opinion, the SHELDA dataset is most suitable for short-period processes, particularly since drifts and datum shifts were identified but not corrected. Therefore, the term “multi-yearly” appears too strong.

- **Omitted.**

Comments on NetCDF files

For each station, please include information on the data owner or the institution responsible for maintaining the station.

- **We added dedicated metadata fields in the netCDF files that clearly identify the original data providers for each station.**

2. The current qc_flags variable (0 or 1) provides limited added value, as the presence or absence of data can already be inferred directly from the dataset. It would be beneficial to enhance this by distinguishing between data that are originally missing (i.e., not available from data centers) and data that have been removed during quality control procedures. As a further improvement, the authors

might also consider introducing an additional flag to indicate data points that are present but potentially questionable or of lower quality (as noted in the manuscript, where some data were identified as doubtful).

- **We expanded the flags to distinguish between: (i) high quality data; (ii) data originally missing (i.e., not available from data centers), (iii) data removed during quality control procedures (e.g., 1 = High quality data, 2 = Missing data, 3 = Removed data)**
- **Specific data which are doubtful or potentially of lower quality were not systematically flagged at the time of quality control. However general issues at each station (e.g., out of range values, drifts, spikes, vertical shifts) were added to NetCDF files.**

3. Users would benefit from additional metadata indicating how many segments are contained within each file (i.e., whether a file consists of one segment or multiple segments).

- **We added metadata specifying the number of segments contained in each file.**

Comments on data (see attached Figures)

1. Check whether interpolation was applied across missing data in the first segment (Fig. bodr.png).

- **Interpolation was not applied across the missing data between two segments. The apparent interpolation observed in the original figure was solely a plotting artefact. In the initial implementation, records containing multiple segments were created by concatenating the segments into a single record, which caused the plotting routine to connect the end of one segment with the beginning of the next. In the revised dataset, each segment is stored separately together with its associated variables (1 NetCDF file with 2 segments), so gaps between segments are explicitly retained and no artificial connection is displayed. Corrected NetCDF files are uploaded.**

2. It appears that some stations still exhibit constant sea levels over time (see figures in Constant_SL).

- **Corrected time series are uploaded.**

3. Please check station batr (see figures in Suspect_data; ignore the red stars). Some high-frequency episodes appear to be “cut off” from above. Is this expected behaviour?

- **Corrected time series are uploaded.**

Reply to RC2: 'Comments - manuscript essd-2025-767' by Ivan Haigh

Overall, I consider this to be a very strong paper based on an excellent and valuable dataset. The authors should be commended for the substantial effort involved in compiling and quality-controlling the data; this represents a meaningful contribution to the sea-level research community. I particularly welcome the fact that the dataset has undergone explicit quality control, which significantly enhances its usefulness. The manuscript itself is detailed, clearly written, and generally well structured.

That said, I have two major comments that I believe need to be addressed, along with a small number of minor issues. If these points are satisfactorily resolved, I would fully support acceptance of the paper for publication in Earth System Science Data.

- **We thank the reviewer for the very positive comments and for recognizing the effort we put into compiling and quality-controlling the SHELDA dataset. We are glad that the reviewer finds the paper strong and the dataset useful for the sea-level research community. We carefully addressed all major and minor points raised and revised the manuscript accordingly.**

Major comments

Relationship to existing datasets (MISELA): A key question that remains insufficiently addressed is how SHELDA differs from, and adds value beyond, the existing MISELA dataset. MISELA already provides high-frequency (1-minute), quality-controlled sea-level data from 331 sites. In the Introduction (particularly around page 3), the authors should explain much more clearly why a new dataset is required, what its specific advantages are relative to MISELA, and what scientific or practical gap it fills. Related to this, it is not clear why SHELDA is not presented as a sub-component or extension of MISELA. Clarifying this distinction would greatly strengthen the motivation for the dataset. In addition, the rationale for focusing exclusively on Europe should be explained more explicitly, as this choice is not currently well justified.

- **Although both MISELA and SHELDA are based on data available at the IOC SLSMF website, these two datasets differ in the contained data. MISELA dataset contains only high-frequency component of sea level time series, i.e. it contains high-pass filtered series (with a cut-off period of 2 h). Quality control was done on high-pass filtered series only. The SHELDA dataset, on the other hand, contains quality controlled original time series. Thus, MISELA allows only for study of short-period sea level oscillations ($T < 2$ h), whereas SHELDA allows for study of both short and long period processes, and contribution of short-period oscillations to extreme sea levels. We have updated the Introduction to better highlight these distinctions and provide a stronger rationale for the formation of the SHELDA dataset.**
- **The regional focus to Europe is justified by: (i) the high effort needed to do quality control of 1-minute sea level data – doing quality control of this extent for the entire World would be too demanding given our resources; (ii) along with the northern America, the European coasts represent a continent for which most long-term 1-15 minute sea level data is available; (iii) sea level variability along the European coasts is very diverse – this refers to variability of tidal ranges, tidal types, ranges of low-frequency oscillations (e.g., planetary scale variability, synoptic scale variability, storm surges) ranges of high-frequency oscillations (e.g., long ocean waves, tsunamis, seiches), heights of extremes, exposure to seismic tsunamis and meteotsunamis, etc. We have added a corresponding discussion to the Introduction to justify this focus on the European coasts.**

Temporal coverage of the dataset: The dataset ends in December 2021, meaning the most recent data are now almost five years old. While I fully appreciate the significant effort required to compile and quality-control datasets of this kind, it would substantially increase the value of SHELDA if more recent data could be included. If this is not feasible, the authors should provide a clearer explanation of why the dataset currently only extends to the end of 2021, for example due to data availability, access restrictions, or quality-control constraints.

- **We thank the reviewer for acknowledging the effort involved in compiling and quality-controlling this dataset. We fully recognize that more recent data would enhance its utility. SHELDA dataset is limited to December 2020 due to the extensive time and effort needed for performing quality control of the data. We plan to extend the dataset in the future as resources allow. We further discuss alternative quality control methodologies, including AI based ones, as these might speed and improve the process of quality control. Accordingly, we have expanded our discussion on the temporal coverage and future plans within the Discussion and Conclusions sections of the revised manuscript.**

Minor comments

Line 95– While it is true that IOC data are updated most frequently, it would be appropriate to note that data quality is often poor and inconsistent, which limits its suitability for many scientific applications.

- **We briefly mentioned this point here, and we emphasized that this aspect is explained more clearly and in greater detail at the beginning of the section 2.2.**

Line 74– It is unclear which specific BODC datasets are being referred to here. The URL provided links to the general archive, and BODC no longer actively maintains a global sea-level dataset. I would recommend removing BODC from this list. Instead, the list would be clearer and more accurate if it included only PSMSL, UHSLC, IOC, GESLA, SONEL, and MISELA.

- **BODC has been removed from the list of global sea level data sources.**

Line 203– The text states that the dataset ends in December 2020, which appears inconsistent with other parts of the manuscript. Please clarify how 2021 data are treated.

- **To clarify, no data from the 2021 were included or analyzed in this study. The final data point in our time series is recorded precisely at 00:00 UTC on 1 January 2021. We updated the text. Consistency across the manuscript has been ensured.**

Line 228– The criteria used to remove values (a 50 cm difference from a neighbouring value, and a 30 cm difference from both neighbouring values) would benefit from a scientific justification. On what basis were these thresholds chosen? Were they informed by instrumental characteristics, tidal variability, or sensitivity testing?

- **These values were chosen to flag physically implausible spikes while preserving genuine short-duration sea-level variability, consistent with the high temporal resolution of 1-minute data. Inspection of data done by authors of original papers using this threshold (Vilibić and Šepić, 2017, Zemunik et al, 2021; in our reference list) revealed that these criteria remove only unphysical spikes. However not all of unphysical spikes are removed by the method, so additional automatic and visual controls are needed as well. We have added more details to the revised manuscript.**

Table 1: I strongly feel that the original data providers should be explicitly acknowledged in the netcdf files, not just IOC – this is to ensure each national tide gauge network gets the credit.

- **We added the original data providers to NetCDF file.**

Table A1: It would also be good to list the actual data providers for each site here. The data providers run the networks that go into IOC, so I strongly think, they should be named and get the credit – not just IOC.

- **We added information on the actual data providers to Table A1 and to the other relevant tables (Table A2 and Table A3).**