Open Access Earth System Science Data Discussions

Earth Syst. Sci. Data Discuss., author comment AC2
https://doi.org/10.5194/essd-2021-134-AC2, 2021

Reply on RC2

Petra Zemunik et al.

Author comment on "Minute Sea-Level Analysis (MISELA): a high-frequency sea-level analysis global dataset" by Petra Zemunik et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-134-AC2, 2021

The authors propose a global dataset (331 stations) of high temporal sampling (1-min.) data recorded by tide gauges over the period 2004-2019. Their primary source of data (95%) is the IOC Sea level station monitoring facility (SLSMF) at VLIZ in Belgium, supplemented with data from two national agencies (17 stations), which do not contribute to this data assembly centre of the global sea level observing system (GLOSS) of IOC/UNESCO. The major added-value of the derived dataset is thus the quality control of the SLSMF data, and this is really a major outcome.

- We thank the reviewer for the constructive comments. Our responses to all comments are provided below.

Indeed, many users have misunderstood what is the goal of SLSMF data assembly centre, whose goal has never been delivering high-frequency data, but informing about the station real-time operational status (Is the station operational, or not). This status reporting goal has been achieved via the collection of the stations real-time data. It should be made a little clearer in the manuscript to underline the originality of what the authors propose. They can also quote one of the comments published by the GLOSS programme directors in J. Coastal Research in response to a clear misuse of the IOC SLSMF dataset for science or technology assessment (Aarup et al. 2019).

- We will emphasize in the text that IOC SLSMF is strictly focused on reporting of the station availability and performance and that any research use of this data require additional processing (e.g. quality control). Aarup et al. (2019) will be cited.

Many discussions have taken place within the GLOSS group of experts regarding its SLSMF dataset, in particular to take actions towards a data quality control, hence addressing the needs of scientific applications investigating high-frequency sea level phenomena. In this regard, the work undertaken by the authors is relevant, and clearly fills a gap. However, it raises several issues that need to be clarified: how this action and dataset articulates with GLOSS? What are the perspectives in terms of update and maintenance? Indeed, if the dataset is not meant for update and maintenance, its scope becomes rather limited

(2004-2019 and ca 300 stations).

- All of these questions are relevant, as it would be the worst-case scenario that our efforts will not be updated and expanded in the future. Still, this is beyond this paper and our efforts, which are largely the part of the PhD research of the first author. Our idea here was – similarly to the creation of the GESLA dataset (which is a project that "grew out of the interests of several people in learning more about the changes in the frequency and magnitude of extreme sea levels") – to allow for better research on high-frequency phenomena, which is an emerging topic in the sea-level research nowadays. So, we hope that our efforts will be recognized by GLOSS and sea-level data centres, which may invest more man power and resources to continue our work. Regardless of that, we plan to update the MISELA dataset in the future, to which students might be of a great help (one of co-authors here was engaged as master's degree student). We will add a few sentences on that in Section 5.

It is further limited by the filtering choice, which is clearly an application dependent restrictive feature. (In this respect the title is misleading.) In my opinion, this aspect should be left out to the scientists (what filter they want to apply, and why). The most invaluable aspect achieved here is the 1-minute sea level quality-controlled data. Filters are application-oriented, and often have pros- and cons (here, these are not discussed). In terms of quality-control, the manuscript does not develop the details of the datum and clock shifts. How are these flaws handled? Why the multiple sensor locations are not exploited to fill gaps, instead of using interpolation? Several other technical issues are commented below.

- For the choice of filter, please also refer to the response to the similar comment from Anonymous Referee #1. We have chosen Kaiser-Bessel filter because it is often use to obtain high-frequency sea-level signal. We will add references in the text in which the filter is described and in which it is used for obtaining high-frequency signal.
- Datum and clock shifts have not been treated, as requiring information which is not available at IOC SLSMF. Indeed, IOC SLSMF contain no levelling information as designed for operational purposes only, while clock shifts are hard to detect at a minute timescale. We are aware that these problems are present in the data, but we consider that they occur in a small percentage of the overall data (as coming from visual quality control, when going through all the data records).
- We choose not filling the gaps with other sensors at a station (where multiple exist, which is minority of stations) as it appeared (at some stations) that high-frequency sea-level signal is different between the sensors (at a minute timescale), due to the technology behind the sensors, and that may influence the statistics.
- We will change the title to "MISELA: High-frequency sea-level analysis global dataset".
- All above will be added and clarified in the text to highlight the limitations of the dataset. However, as both anonymous referees and Philip Woodworth (https://doi.org/10.5194/essd-2021-134-CC1) emphasized, despite to these shortcomings, the dataset might become an invaluable tool for the science of sea level at high frequencies and is filling the gap in available sea-level products. We hope that the future versions of the dataset will handle at least some of these problems.

To wrap up, the manuscript is overall well written, and organized. It can represent an invaluable contribution to the international sea level community and science, provided its current limitations are addressed (filtering, interpolation) and perspectives clarified (update, maintenance).

p.2, L43-44: The PSMSL requests their updated reference to be quoted. This data assembly centre has substantially extended its contents since 2003 with many useful developments

- Will be updated with reference Holgate et al. (2013, https://doi.org/10.2112/jcoastres-d-12-00175.1)

p.3, L82-83: To my understanding, the goal of the IOC SLSMF has never been delivering high-frequency data, but informing about the station real-time operational status (Is the station operational, or not). This goal is achieved via the successful collection (or not) of the real-time data. This aspect should be made crystal clear (see Aarup et al.) for the full appraisal of the added-value provided by MISELA. See general comment above.

- We will remove the first part of the sentence and change to: "The main objective of the facility is to inform users about the status of station availability and performance (Aarup et al., 2019). This includes displaying the tide gauge station metadata and regularly checking the operational status of all stations, as well as contacting operators regarding non-operating stations. Another important objective is a display service through which one can undertake quick visual inspection of the raw data in a selected half-daily, daily, weekly or monthly period during which the chosen station was operational (IOC, 2012). It is also possible to download the data for the whole operational period. However, any research use of these data would require additional processing (e.g. quality control), in order to properly prepare and involve data in statistical analyses and avoid misleading results and conclusions (Aarup et al., 2019)."

p.4, Fig.1: What are the problems illustrated here? I suggest to state them in the figure caption so that the reader makes a direct association.

- The illustrated problems will be stated in the figure caption: "… a) gaps, b) spikes, c) shifts, and d) spurious oscillations …"

p.5, L109: The rationale for the criteria (2-year long) should be developed, if there is any.

- MISELA is a research product for statistical analyses of high-frequency sea-level oscillations. For that reason, short series are not of use for any statistical analysis there. We will clarify this in the text as: "As the dataset is intended to be applicable for statistical analyses of high-frequency sea-level processes, we choose a length of 1.4 year (70% of 2 years) as a threshold, because short time series or those overly intermitted with data gaps would not significantly contribute to the research."

L111: what is "too many"? "incorrect records"? Need for objective criteria.

- "too many spikes, incorrect records" will be replaced with "spikes that are distributed throughout most of the time series and appear on an hourly or multi-hourly basis, obvious incorrect records like spurious oscillations produced by malfunctions of instruments". As is written, these are spotted by visual checking, therefore do not have

an exact numerical threshold. Later in the text we also explain that the visual control is subjective to a certain extent.

p.5, L117-118: I do not understand what it means: "or 30 cm differing from both neighbouring values (20 and 15 cm, respectively". Please, clarify.

- The sentence will be rephrased: "The automatic quality-control procedures included removing of out-of-range values, i.e. values 50 cm differing from one neighbouring value or 30 cm differing from both neighbouring values (in case of the FMI stations 20 cm differing from one or 15 cm differing from both neighbouring values).".

p.5, L119: how much is "deviate"? Needs clarification.

- Three standard deviations. Will be added in the text.

p.5, L128: "Not all problems" Can you give an example?

- g. spikes, artificial oscillations, stucks of instruments. Will be added in the text.

L130-131: How many users have expressed their interest within this atmospheric community? Did the authors make a survey, for instance within international programmes like GLOSS?

- No, we didn't perform any survey, but just coming from recent research activities on high-frequency sea-level phenomena (e.g. being spotted by the meteotsunami community at the First World Conference on Meteotsunamis, izor.hr/mts2019). We will add a few words and references on that: "The next step in creating the MISELA dataset was to exclude sea-level records observed during seismic tsunamis, since the applications are directed towards research on atmospherically-induced sea-level oscillations, which has been an emerging field during last decades (e.g. Pattiaratchi and Wijeratne, 2015; Vilibić et al., 2021).

p.6, L138: what is the interpolation rationale behind the duration of "one week"?

- We have filled these gaps to simplify our filtering procedure. However, interpolated data are, of course, non-reliable, and are thus marked with a flag.

p.8, Table: Is the 0.01-degree precision sufficient? At this spatial resolution gravity waves can be rather different in terms of amplitude and impact. For instance, in terms of exposure along a protected coastline (harbour, estuary…).

- Majority of stations in the MISELA dataset have higher precision (0.001- or 0.0001-degree), yet some stations provide 0.01-degree precision at the IOC SLSMF portal, e.g. Japanese stations.

p.9, L139: “acceptable”? Needs clarification. I suggest to remove this subjective term.

- Will be changed with a sentence: “Figure 4 shows that stations included in the MISELA dataset cover many of the World’s coasts”.

p.9, Last line: “excellent”: the regions in c) d) and f) have large gaps along long parts of the coastline; in other words, far from “excellent” in my opinion.

- “Excellent” will be replaced with “satisfactory”.

p.11, L203: June 2018 at the latest in a SLSMF? Needs explanation in the context of a real-time facility, something sounds weird (objectives, misunderstanding…).

- We will clarify in the text that we have obtained records from the IOC stations for the period from as early as 1 January 2006, when the portal started operating, up to 14 June 2018 at the latest, when we have last downloaded the data. Unfortunately, we have not downloaded sea-level time series after this date due to extensive time needed for performing quality control of the data. To emphasize, this work has largely been done as a part of the PhD research, with limited human resources, thus cannot provide up-to-date research-quality data in an operational way. We hope that the future developments of the MISELA dataset will be taken over by data centres, with much higher human capacities and resources to perform such a work.