

Earth Syst. Sci. Data Discuss., referee comment RC1
<https://doi.org/10.5194/essd-2021-134-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on **essd-2021-134**

Anonymous Referee #1

Referee 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-RC1>, 2021

This paper describes the derivation of a global sea level dataset that is purpose-built for the study of non-seismic extreme events that act on very short timescales. The authors rightly point out that, for the most part, historical tide gauge datasets do not offer the temporal resolution that is required to study such events. This is largely due to the past limitations on data storage and telecommunications of earlier tide gauge technology and the lower computational power of data-processing systems that were available at that time. Faster sampling technology and near real time telecommunications systems, together with increased back-office computational power have meant that higher frequency data are now more prevalent. Even so, many of these high frequency datasets are only used for operational early warning purposes and there is certainly a need for high frequency research quality data. The MISELA dataset therefore will address a gap that exists amongst global data portals.

However, I can't help wondering why the authors tailored (and thereby restricted) the use of the dataset to studying only atmospherically-induced sea level oscillations? At some stage, prior to removing tsunamis from the dataset, the authors would have possessed a quality-controlled high frequency sea level dataset that could have multiple scientific applications, including the study of tides, tsunamis and meteotsunamis, seiches, storm surges etc. Sadly, that gap remains unaddressed.

Another drawback is that the dataset only uses tide gauges that afford 1 min sampling, as those that offer faster sampling are disregarded. A means of sub-sampling or averaging these data to one minute intervals would be desirable.

Nevertheless, this dataset is undoubtedly an improvement on current data provision and the allocation of a DOI is excellent. In addition, the processing methodology appears sensible. I recommend publication, subject to correction of a few discrepancies which are listed below:

Page 1, Line 20, The reference is missing from the list of references on page 15

Page 1, line 25, replace 'processes' with 'phenomena'

Page 1, Line 28, 'no quality-checked globalan hour'. This isn't quite true – BODC and GESLA datasets do offer some data at higher frequency than an hour. I'd rephrase to say 'no quality-checked global sea-level datasets afford sufficiently high temporal resolution to cover periods at which'

Page 2, line 45, BODC supplies data at the resolution provided by tide gauge operators, which is sometimes hourly intervals, but is generally higher, e.g. 15 minute for the UK and 6 minute for the USA. Also, BODC provides sea level data for a vast number of locations. The authors should be clear about which BODC dataset they are referring to. Is it the GLOSS Core Network or another?

Page 2, line 57 'with a minute resolution' should read 'with a minute or higher resolution'

Page 3, line 78, remove 'with'

Page 3, Line 79, The meaning of 'not completely operational...facility' is not clear. Are the authors saying that many stations do not afford high enough frequency data? Or are many stations offline i.e. in disrepair?

Page 3, line 84, use 'displaying' rather than 'controlling'

Page 3, line 89, use inverted commas for 'as received'

Page 3, line 90, should read 'quality control' not 'quality'

Page 3, line 95, should read 'each of the stations'

Page 3, line 96, replace 'and other' with 'etc'

Page 13, line 215, I don't think it is fair to say that the hazard has been underrated. Instead, the ability to study this hazard has until recently, been restricted by technological and computational limitations (as I described above).

Page 13, line 226. It is a bit of a 'leap' to suggest that the MISELA dataset might encourage operators to increase their sampling resolution or to install more tide gauges in observationally sparse regions. The determining factor for both of those issues is more likely to be financial.

Page 13, line 230, use 'enabled' instead of 'allowed for'

Page 13, line 234, should read 'The recent manual ...'

Page 13, various uses of 'quality-check' should really be 'quality control' or 'quality-checking'

Page 13, line 238, 'procedures' should be singular

Page 13, line 241, 'automatized' should be 'automated'

Page 14, line 248, 'with as much quality-check....' Is a bit clumsy. Maybe rephrase to 'to report, so far as possible, near real-time quality-controlled data.'

Page 14, line 2501 'evolution' not 'evolving'