



EGUsphere, referee comment RC1
<https://doi.org/10.5194/egusphere-2022-4-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-4

Filippo Gatti (Referee)

Referee comment on "A study on the effect of input data length on a deep-learning-based magnitude classifier" by Megha Chakraborty et al., EGU sphere,
<https://doi.org/10.5194/egusphere-2022-4-RC1>, 2022

The authors of the manuscript entitled "A study on the effect of input data length on deep learning-based magnitude classifier" present a very interesting contribution into addressing some crucial aspects of the automated techniques (powered by deep-learning) to estimate the earthquake magnitude. Despite the large perspectives that deep-learning techniques provide to the EEW framework, the choice of the right neural architecture, the dataset preparation and interpretation of results is crucial into determine their success. The authors have addresses many of the concerns that this sensitive deep-learning application has, showcasing a powerful neural classifier that can discriminate between noise, low-magnitude and high-magnitude earthquakes. Still, some open questions remain, which is why this reviewer suggested a major revision, so to address them more into detail and so to promote the scientific debate on it.

As metnioned, the manuscript is very interesting and it raises many questions about the decisions the authors made to perform their classification benchmark. The list of comments below is non-exhaustive: please, refer to the attached reviewed manuscript for further comments.

- The decision boundary between low-magnitude and high-magnitude is rather arbitrary, as stated by the reviewers and as shown in Fig.7a. Plus, It highly depends on the seismic context of interest and the risk assessment and vulnerability policies of each country/region. It would be interesting to test at least one another decision boundary or at least test somehow the sensitivity of the classifier to this choice.
- The effect of the source-to-site distance seems to have been disregarded. Maybe, separating the waveforms in different bins, based on source-to-site distance, could unveil some interesting aspects of the classification performance. Some comments on it would be beneficial to the manuscript overall clarity.
- Sometimes, it's rather useful to analyze the waveforms at stake in the Fourier's spectrum domain. The corner frequency is strictly related to the source spectrum, which mostly determines the magnitude (along with the distance) In this case, this

reviewer suggests to check the spectrogram of the classified waveforms, so to verify that the duration is compatible with the associated frequency corner value for the correspondent moment magnitude (see the statistical relationship between corner frequency and moment magnitude presented by *Courboulex F, Vallée M, Causse M, Chounet A (2016) Stress-drop variability of shallow earthquakes extracted from a global database of source time functions. Seismol Res Lett 87(4):912–918*

- Have the authors considered the earthquake type when preparing the dataset? A comment on this aspect would be very interesting.

There are no technical corrections, besides the need to detach the unity of measure from the number (ex. 10 s and not 10s)

Please also note the supplement to this comment:

<https://egusphere.copernicus.org/preprints/2022/egusphere-2022-4/egusphere-2022-4-R C1-supplement.pdf>