

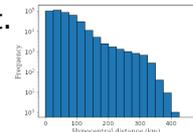
Reply on RC2

Megha Chakraborty et al.

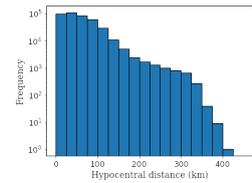
Author comment on "A study on the effect of input data length on a deep-learning-based magnitude classifier" by Megha Chakraborty et al., EGUsphere,
<https://doi.org/10.5194/egusphere-2022-4-AC2>, 2022

The authors thank Anonymous Referee #2 for his feedback. Listed below is our response to "specific comments":

- It is difficult to experiment with decision boundaries above 5 because the number of waveforms for such high magnitudes present in the dataset is severely limited. Although we experimented with decision boundaries of magnitude 3 and 4 and got similar results. (This is also explained in AC1 in this discussion - <https://doi.org/10.5194/egusphere-2022-4-AC1>).
- Currently no selection criteria are applied to the source-to-site distance. (see next comment)
- The authors thank Anonymous Referee #2 for suggesting the analysis of source-to-site distance of the training data. The figure for the distribution of source-to-site distances is shown below and will be added in the revised version of the manuscript.



- As discussed in AC1 (<https://doi.org/10.5194/egusphere-2022-4-AC1>) we have analyzed the model performance for different source-to-site distances and observed that the model is indeed capable of performing reliably over a wide range of hypocentral distances. In other words, no clear dependence between the model performance and hypocentral distance can be observed. Shown below is the relevant figure which can also be added in a revised version of the manuscript.



- Convolutional Neural Networks have often been found to be useful for seismological data analysis as they are capable of extracting patterns in the data (features) without any temporal dependence. When combined with LSTMs the temporal relations between these features can be obtained. In applications such as magnitude-based classification of earthquakes, this aids in the effective analysis of signal features as compared to the pre-signal background noise. The dropout layers are used to prevent the model from overfitting and the maxpooling layer is a method to reduce the data dimensionality so that only relevant features can be retained. The final layer is a softmax layer which outputs the probabilities corresponding to each of the three classes that the data is classified into. This description will be added to a revised version of the manuscript.