Open Access Earth System Science Data Discussions

Earth Syst. Sci. Data Discuss., referee comment RC1
https://doi.org/10.5194/essd-2022-345-RC1, 2022

Comment on essd-2022-345

Anonymous Referee #1

Referee comment on "DL-RMD: a geophysically constrained electromagnetic resistivity model database (RMD) for deep learning (DL) applications" by Muhammad Rizwan Asif et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-345-RC1, 2022

This manuscript proposed a standardized dataset for deep learning-based electromagnetic methods. The database is geophysically constrained, which produces good accuracy performance and satisfactory generalization and consistency. Overall the paper is very well written, and the data shows its high readiness for the community. I would recommend it get published before some of my concerns are addressed.

Major:

- Evaluation (section 4): This study mainly employed the proposed dataset to train a deep-learning (DL) model and surrogate the forward modeling process and demonstrated that this dataset shows its great performance. The assessment method is rational; however, this section needs additional result comparison with previous relevant DL-based studies. By comparing this proposed data with other DL studies that used limited input data, the authors may demonstrate this proposed database can be treated as the benchmark. Otherwise, it is only another DL experiment for improving the computation efficiency.

Besides, please try to find some weaknesses in the training data utilized by previous DL studies and demonstrate your progress on it after comparison. For example, the introduction is well written but those training sets are not involved after then. Benchmark is a strong word that requires more comprehensive assessment and evidence.

- The assessment section (section 4) needs to provide additional quantified comparisons and descriptions for the figures rather than just some summaries.

Minor:

1) References in Table 1 should be Qin et al. (2019) rather than (Qin et al. 2019), and generally, the table caption is above the table.

2) the dataset is formatted as txt, which caused the code reading speed very slow.