Comment on essd-2021-349
Grzegorz Kwiatek (Referee)

Referee comment on "Moment tensor catalogue of microearthquakes in West Bohemia from 2008 to 2018" by Václav Vavryčuk et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-349-RC2, 2022

Manuscript essd-2021-349 presents development of seismic moment tensor catalog for Western Bohemia region. The Authors provide the vast catalog of moment tensors calculated between 2008 and 2018, as well as associate this catalog with pre-processes waveform data enabling the scientific community to perform a follow-up studies. I genuinely welcome such approach very much and appreciate

Regarding the core scientific part, i.e. the description of development of MT catalog, I have no comments. The Authors provide much more that I would expect regarding the moment tensor processing, and they also provide some more in-depth information on the peculiarities of the MT catalog. The core MT catalog is described sufficiently and it's easy to get deeper into the applied procedures by digging into the referenced studies.

1) The organization of data is slightly confusing or alternatively I don't understand fully how Authors want to share MT catalog. There is already associated data publication [1] containing seemingly full waveform data, locations, basic station information, and velocity model. This data publication is complete and seemingly consistent internally. The submitted manuscript provides additional supplementary information which contains already some overlapping data, i.e. velocity model and station information. There is also new information (figures and moment tensor catalogue) which both constitute the core of the supplementary information, and these are broadly described in the publication. However, I have a problem with remaining folders, i.e. waveforms and figures. These folders contain limited data (not for all MTs in the catalog), and these are certainly duplicated of what is in [1] already. I don't understand clearly (or maybe I missed something) why these limited set of waveforms is provided in the supplement. It makes to me more sense to provide only MT catalog and refer to [1] or even incorporate MT catalog in [1] as an update.

2) The data is provided mostly as text files with the waveforms already converted to actual ground motion. It could be argued whether this is the right way to do, but I personally welcomed that the authors aimed to simplify data accessibility by removing sensor response. An alternative would be to provide data in some recognized format (mseed?) and provide sensor responses *.resp files, but as I said, I am not insisting on that. However, I would clearly expect there is some more data provided detailing this step of data processing. More specifically, how sensor response was removed, what filters have
been applied etc. to obtain the signals that finally finished in waveform files. The sensor information is not sufficiently well described. The publication contains some basic sensor information, but the pre-processing of data is not described well.

3) The organization of various data types is fairly clean (i.e. clear directory structures, keeping some key information in file names and folders that allows to link various data altogether, and so on). However, the key moment tensor catalog DOES NOT contain any unique event identifier! This is actually quite surprising, because it forces the future data user to get the date of event from MT catalog, go to main catalog from data publication [1] (/locations), take the identifier from there (e.g. 'A1000') by comparing the dates and times, remember year on the way, and then, having all this information, try to build up path to the waveform data stored in [1] in another folder. This could be easily omitted by providing even the e.g. 'A1000' identifier in MT table, or actually, for example the relative path to the folder containing waveforms.

Refs: