Reviewer Comments for Manuscript ESSD-2026-202

This manuscript presents a flatfile of strong-motion recordings for shallow crustal earthquakes in Colombia. The dataset comprises uniformly processed three-component accelerograms together with a set of intensity measures. The authors also perform validation and consistency checks by comparing the dataset against both a global and a regional Ground Motion Prediction Model (GMPM). The database has the potential to become a valuable resource for seismic hazard assessment, ground-motion modelling, engineering seismology, and future development of regional GMPEs. The manuscript is generally well organized. Although the dataset is potentially valuable, I identified several issues that should be addressed to ensure that the database can be readily understood and reliably reused by the community. My main comments are provided below.

Paper	
18-19: the Fourier Amplitude Spectra for each component is provided in the frequency range from 0.04 to 30 Hz	The data file contains frequencies up to 50 Hz (maybe more considering the discrepancies from frequencies - see comments for COL_FAS). Please clarify the actual frequency range provided in the dataset.
105: The predominant period of the site (T0)	Does this correspond to the Tn values reported in the Sites_Vs30_HV file? Please use consistent notation throughout the manuscript and the accompanying datasets.
Abbreviations	All abbreviations should be defined when first introduced. eg. SGC – for readers who are not familiar with the study region, the meaning of this abbreviation is not obvious.
COL_FAS	- the file contains 9330 components instead of the 8086 reported in the manuscript. Please explain this discrepancy. Are some components duplicated? - it is unclear what eqid (col. B) represents. If it is an identifier, it should not be stored in scientific notation. - date2 (col. B) contains valid dates only up to row 1693. From that point onward, the entries no longer appear to be dates. - col. E – H contain latitudes and longitudes values Please clarify whether these coordinates correspond to the earthquake epicenter or to the recording station. - the frequencies for FAS_NS, FAS_EW and EAS appear to be incorrect. For example, the frequency in column CQ is 18.789 Hz, while the adjacent column CR is 2.006 Hz, and column CX is 20.15 Hz. Please verify the frequency labels. - it would be more convenient to provide a separate row containing only the frequency values instead of including both the parameter name and the frequency in each column header. - the table headers should contain explicit and self-explanatory parameter names and measurement units
COL_GM_journal	- the table headers should contain explicit and self-explanatory parameter names - please clearly distinguish between station and event coordinates in the column headers - measurement units should be included in column headers - there is no distinction between spectral accelerations (SA) for the EW, NS, and vertical components. Furthermore, SA at 0.01 s should be very close to PGA. This appears to be true only for the NS component (assuming PGA is reported in cm/s² and SA in m/s²). Are spectral accelerations provided only for the NS component? The manuscript states that SA values are available for all components.
Sites_Vs30_HV	- the table headers should contain explicit and self-explanatory parameter names and measurement units - I would recommend starting the station numbering from 1 instead of 0 - the columns lat, lon, lat_y, and lon_y appear to contain similar coordinates. Please clarify the meaning of lat_y and lon_y. Similarly, please explain what is dist_lon and dist_lat. - is Tn equivalent to the predominant period T0 referred to in the manuscript? If so, please use a consistent notation throughout the paper and the supplementary files. - please define the parameter P*.