This document presents a review of the manuscript titled as “Ground motion prediction maps using seismic microzonation data and machine learning” by Mori et al. The main objective of the work is to apply the machine learning approach to provide ground motion maps predicted based on stratigraphic and morphological parameters. This is a somewhat novel idea when the validity of the outcome could be well-established.

The use of English language is generally fine. However, there are repetitive misuses of some particular words/phrases that make the reading tedious and confusing. There are also some misuses of punctuations (e.g. comma). Some examples have been given below. This reviewer strongly recommends a proofread of the article by a better English user.

The organization of the article is generally fine. However, the abstract does not seem to be a good summary of the work. It gives us the impression that the current work deals with near-real-time prediction of damage, which is not the case. The last sentence refers to a conclusion regarding the effects of short-distance variability, which has not been extensively addressed or proven by this article. The introduction segment could be improved. The texts between Line no. (L) 74 and 95 seem irrelevant. They should be presented in a different segment under a different title. The major weakness of this article is that the method section remains completely underdeveloped. As machine learning is at the core of this work, the method of evaluating the training dataset and validating the outputs should be clearly presented. At the current state, reproducibility of the results of this work remains under doubt. In the following sections, the interpretation of results could also be improved. The last segment related to the spatial correlation of the ground motion predictions seems completely underdeveloped. If the authors intend to present it in a future article (as mentioned in L 348-349), then there is no point addressing this issue and drawing some concrete conclusions without any explanation within this article.

Therefore, this reviewer recommends a major improvement of the article before being considered for the publication. The main comments are as below:

Scientific Comments:

- The presentation of the datasets as maps would be helpful. Median and percentiles could be indicated on Figure 1 and 2 for a better comprehension.
- The method section must be improved. How are the outputs estimated from the input data. Which parameters have been used and how? Which proxy performed the best? What are the differences among different ML prediction models (Table 2) and why have they been adopted? How come the SD values are reduced by 45-60% (L217)? It’s not clear how Openquake was used to determine the IMs – which site and fault parameters have been used for the calculation?
- The reviewer disagrees with the statement in L 264-365. It’s very difficult to visually determine if the maps in Figure 4 and 5 are coherent with geological and geomorphological

characteristics. No such maps of the area have been presented. For comparison, the existing ground motion estimation for this area (e.g. from the shake map) and/or microzonation map could be presented on a same color scale.

- The estimation of uncertainty presented in Figure 5 should be elaborated. What kinds of uncertainties have been considered and how have they been estimated?
- The validation of the results seems insufficient. It uses the results from only one earthquake, for one spectral period and only 3 cross-sections. How are the Sa profiles (in Figure 6) obtained and modified from Giallini et al.? What about the ground motions at other spectral periods? Do they compare well? Other than these 3 cross-sections from two sites, how many more sites have been used to validate the results? How long are the cross-sections?
- Segment 4.2 should be either eliminated or better explained. Results shown in Figure 7 and Table 5 should be interpreted in plain terms. The fact that the estimated short-scale variability is related to site conditions should be validated in terms of physical parameters. Which input parameters are causing these variations and to what extent? How close it is to the reality (example from a site)?

Editorial Comments:

- The phrases ‘referred to’ or ‘with reference to’ have been heavily repeated and misused. All of them must be replaced by appropriate vocabulary or the sentences should be rewritten. The phrase ‘bearing in mind’ has been tediously repeated.
- The resolution of the map mentioned as ‘50 X 50 m’ should be corrected as ‘50 m X 50 m’ for the entire article.
- For all the large numbers used in the article, the thousand separator should be corrected fro ‘XXX’000’ to ‘XXX,000’ In English, comma separates the thousand.
- Figure 2 has not been mentioned or explained anywhere. The cross-section labels on Figure 5 are not readable. The long web-address in the title could be replaced by a reference. The Figure 6 title should mention that the cross-sections are presented from Figure 5.
- L 108-112 should be rewritten. The current form is grammatically incoherent.
- In L 150 the phrase ‘here adopted’ should be corrected as ‘adopted here’.
- In L 232 ‘gauge the IMs’ sounds unnatural. It could be replaced by ‘estimate the IMs’.
- The webpage links should be referred to as a reference with last access date. It is sufficient to cite them as references in the text rather than citing the link everywhere.
- L 376 and 385 are not expressed as full sentences. Either they should be turned into paragraph subtitles by using a colon (:) after ‘following’ in L375 or they should be re-written as full sentences. Similar remark for L 397-403.