Review of "An ensemble estimate of Australian soil organic carbon using machine learning and process-based modelling" by Wang et al. for consideration in *EGUsphere.*

Based on in situ observation data of SOC across Australia, the authors investigated the controlling factors of SOC using machine learning models and predict the gridded SOC map in Australia using trained machine learning models and process-based models. The topic of this study is interesting and important. The manuscript is well organized, and the conclusions of this study can be robustly supported by the analysis results. Nonetheless, a few more analysis and information should be added to further improve this paper.

We would like to thank the reviewers for their time and feedback on this manuscript. Please find our point-to-point responses below.

1. **Several soil databases and studies have provided gridded SOC map over Australia, I would suggest to compare the estimates of SOC in this study to previous estimates of SOC, and provide a brief explanation on the reasons of the differences between present and previous estimates.**

Thanks for the suggestion. Yes, there are many existing gridded SOC maps over Australia, and we have compared our results with some of those most widely used (e.g., Australian baseline map (Viscarra Rossel et al., 2014)) and newly released (e.g., Walden et al., 2023; Wadoux et al., 2023) SOC maps (see section 4.2 and 4.3). We aim to better highlight this comparison and improve our discussion by clarifying the potential reasons causing the difference between our estimates and other maps.

2. **I would suggest the authors to include the optimized values of MIMICS model in Table 1 or a new supplementary table. As the authors has conducted leave-out cross validation. The parameter values used for predicting SOC contents across Australia, and the range of each parameter obtained from the cross validation should be showed.**

Thanks for the suggestion. We'll present the optimized parameter values in the revised manuscript.

3. **In Fig. 2, the locations and numbers of the observation sites belonging to each PFT are showed. I suggest to add a similar map and box plot to show the locations, total numbers of the observation sites grouped in each of the 6 clusters based on the k-means algorithm.**

Thanks, we'll add this map to show the distribution of sites belonging to each environmental cluster in the revised manuscript.

4. **Did the authors tried to optimize and evaluate MIMICS model using data from all observation sites together, rather than optimize the model parameters for each PFT or cluster? What is the performance of MIMICS if it was not optimized for each PFT or cluster?**

We didn't optimize parameters using the whole dataset, and we are confident that this will result in poorer performance in MIMICS. Prior to determining our parameter optimization approach, we conducted an analysis of SOC observations, as well as climate and soil conditions across Australia. Our investigation revealed significant variations in both climate and soil characteristics across the continent, for example, mean annual temperature ranged from 4.75 to 29.15 °C, mean annual precipitation ranged from 107.9 to 5536.7 mm, and soil clay content ranged from 3% to 59%. Additionally, SOC observations exhibited a positive skew.

Given these diverse and nuanced environmental conditions, and the low dimensionality of MIMICS relative to real-world processes, we recognized the limitations of optimizing parameters using all observations collectively. Model results are typically poor when the optimized parameters are applied to regions where the environmental conditions deviate significantly from the average conditions used in parameter optimisation.

As described in the manuscript, we instead adopt a data-driven grouping of the observations based on plant functional types and environmental conditions. This improves parameter optimization and enhances the performance of MIMICS, particularly in regions with limited observations but comparable environmental conditions.

Specific comments:

5. **L35: The t/ha should be changed to t ha^{-1}. The form of other units used in this study should also be adapted like this.**

Thanks, we'll revise this in the revised manuscript.

6. **L362: What is the 'test data'?**

We randomly selected 70% of observations to train the model and the remaining 30% to validate the model. The remaining 30% are test data – we'll make this explicit in the revised manuscript.

7. **L361-364: Why do not train the models using all observation data first? Then using the parameters trained based on all observation data to simulate SOC stocks across Australia.**

We opted for not adopting the approach suggested based on preliminary testing and our previous experience running the model. Please see our response to (4) above for details.

8. **L446: greater àlarger**

Thanks, we'll change this word in the revised manuscript.

9. **L491-496: This discussion might be one-sided and give a wrong causal relationship. Many studies also suggested that SOM is one of the key controlling factors of soil bulk density (BD), as the density of SOM is smaller than soil mineral particles and stimulate the formation of aggregates and activities of insects and microbes. It is better to discuss the interaction between SOM and BD, rather than only the effect of BD on SOM.**

Yes, we agree with this comment. BD and SOM are intimately related. We will add discussion on these interactions in the revised manuscript.