Review of the manuscript Yang et al., "Real-Time Snow Depth Estimation and Historical Data Reconstruction Over China Based on a Random Forest Machine Learning Approach."

The authors have made significant improvements in the revised manuscript. However, some issues, as mentioned previously, remain. I am not convinced by some of the responses by the authors. These issues are as follows.

1. My primary concern was the selection of the RF algorithm for snow depth estimation. With the broader availability of deep neural networks for similar applications, I am not convinced by the utilization of RF for snow depth estimation. Therefore, the authors should either demonstrate that such advanced methods do not provide a significant improvement in accuracy over RF algorithm or re-do the experiments with these methods. Without this analysis or these experiments, we cannot ascertain the potential of the proposed method, particularly in practical applications.

   The experiment with ANN by the authors is appreciated. However, this is not at all convincing, since, we already know from the literature that for medium resolution or coarse resolution imagery, SVM or RF outperforms the ANN.

2. The authors have now discussed the impact of the diurnal changes in the snowpack geophysical properties, including the microstructure, as mentioned in my previous comment. However, this discussion is not comprehensive. Due to the absence of any experiments, this cannot be correlated to the errors observed. Thus, the analysis and the discussion part in this manuscript is still lacking.

   For example, the authors have considered in their RF algorithms inputs such as the latitude and the longitude. But why not any snowpack specific parameters. These can be obtained from the ERA5 data and used as inputs to their RF models. The authors then should check the discrepancy in the RMSE based on these parameters.

   Another analysis that I find missing is seasonal variation in accuracy, which is very important. The readers should know the difference in the performance of the algorithm in the off-peak snow seasons, i.e., during the early winter and the melt season.

3. Please mention the upscaling method for 1-km to 25-km LULC map. Additionally, I also mention which latitude and longitudes were used as inputs. I hope these were geographic ones and not projected ones.

4. I find that the majority of the discussion is based on the comparison between the results from the RF method and the WESTDC. While I believe there should be a better balance in the discussion for comparison with the in-situ measurements.

5. Another analysis to the manuscript that the authors should add is for the selection of trees. Presently, the authors have fixed this to 1000. However, is there any merit in using this large number? What is the performance for other smaller number of trees, for example, 500, 250, etc. Is the performance significant enough to use such a relatively large number of trees?

6. An important parameter in the discussion is the R-square that is missing?

7. In Table 5, the results are contradictory to the explanations given by the authors. The authors have mentioned that for shallow snow depth, the PMW could be insensitive. However, Table 5 shows the contradictory, i.e., rather some sensitivity in case of shallow snow depths as compared to deeper snow where there is nearly no correlation. I believe the inconsistency of the results requires a thorough analysis.

Minor issues

- Please check Figure 4 for correctness. Is the validation dataset correct? This is very confusing from the authors response.
- Some of the colors used by authors are very poor. Figure 6 the turquoise text, Figure 9, the orange legends and lines.
- Although I meant with my comment to include a discussion, the MEMLS simulation experiment is appreciated. However, the parameters used in the MEMLS simulation are not correct in the response by the authors. How often do we have a snowpack representing fresh snow characteristics? This is very seldom. Thus, there is no point showing an experiment with a snow density of 100 kg/cubic m corresponding to fresh snow. Instead, the authors should have selected something between 200-350 km/cubic m for the simulation, which corresponds to old snow.