Comment on essd-2022-137
Anonymous Referee #1

Referee comment on "An 8-day compositied 36 km SMAP soil moisture dataset from 1979 to 2015 produced using a random forest and historical CCI data" by Haoxuan Yang et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-137-RC1, 2022

The authors of this study tried to use a random forest model combined with historical CCI soil moisture estimates from multiple satellites and sensors, to incorporate the strength of SMAP soil moisture data from 2015, making an 8-day composite, 36 km SMAP-like soil moisture data set from 1979. Overall, this study has potential to provide an updated global soil moisture data set for a range of research and applications. However, there is an important assumption made in this proposed approach that changes in temporal variability in soil moisture estimates from CCI sol moisture data are similar to those of SMAP data. This should be fully investigated, because these are two different sources of information. The authors may want to seriously consider the spatiotemporal consistency among different data sets prior to conducting such an analysis. Second, there is no report on statistic metrics of comparison of the random forest-based SMAP estimates and in situ measurements. People may want to know if these estimates achieve a basic acceptable accuracy in terms of an overall error of less than 0.04 m$^3$ m$^{-3}$ for the volumetric soil moisture estimates. I hope the authors have a chance to rethink about this study, and have more time to redo such an analysis prior to resubmission.