Earth Syst. Sci. Data Discuss., referee comment RC1
https://doi.org/10.5194/essd-2022-177-RC1, 2022

Comment on essd-2022-177

Anonymous Referee #1

Referee comment on "A 1-km daily soil moisture dataset of China based on in-situ measurement using machine learning" by Qingliang Li et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-177-RC1, 2022

Apparently, this paper touches an important research topic on the generation of high spatiotemporal soil moisture data using machine learning (ML) technique and various sources of datasets. Moreover, compared with previous efforts, this work comprehensively used the quality-controlled 1,789 in-situ soil moisture observations in continental China and performed detailed training and testing, especially for 10 soil layers between the 10 cm and 100 cm deep. However, to improve the current version, the authors are encouraged to carefully consider the following questions and suggestions.

- In this study, only the Random Forest method was applied to derive the upscaled soil moisture data. Is it possible to try more MLs, such as CatBoost, XgBoost, and NeuralNetwork, to test how consistent or different the resulted products are?
- Most of the source datasets cover the period before year 2010. Is there any special reason why the new soil moisture only covers the period 2010-2020? Is it possible to extend the present time period to year 2000-2020?
- The “Materials and Methods” read too long, and the authors may try to shorten the text and put some figures into the Supplementary Material.
- For the “Results”, they seem to be a combination of results analysis and short discussion. Please move relevant discussion content to the “Discussion” part.
- The “Discussion” really needs to be reorganized and improved; the current one does not provide deep thoughts on the new soil moisture products, in terms of their differences/similarities/uniqueness compared to previous products/work and implications for the soil moisture modeling and detection and attribution .
- Grammar mistakes can be noticed in many places, for example, for the sentences between lines 82-87, 91-06, and 112-114 among others. The authors are suggested to get help from native English speakers and thoroughly check the whole manuscript before the next submission.