

Interactive comment on “Gap-Free Global Annual Soil Moisture: 15km Grids for 1991–2016” by Mario Guevara et al.

Anonymous Referee #3

Received and published: 25 November 2019

General comments: While seeking a higher resolution global soil moisture product is certainly a laudable goal, I find the methods used in this paper lack a credible connection between ground measurements and the remotely sensed data. The authors use machine learning to generate regressions between multiple topographic variables thought to be related to soil moisture that are available at higher resolution with coarse resolution satellite data. It is difficult for the reader to discern if actual new information results from the downscaling because there is not a clear connection made between in-situ soil moisture measurements, their physical connection to the chosen topographic variables used, and resultant satellite measurements. To do so, requires first demonstrating the rigor of the methodology over a much smaller and better measured area, such as the area of the International Soil Moisture Network. Extending the algorithm to

C1

areas that are hydrogeomorphically and climatically distinct from the existing soil moisture measurement networks cannot be result in credible data. I do not recommend publication of this work.

Major comment: Much more information is needed about the regression methods and parameter selection. It would be helpful to discuss what each parameter is mathematically and why it can be useful in a soil moisture prediction context. Can it be demonstrated over smaller areas that there is a valid argument for using these variables, some or all of them. Explain in detail the cross validation process used and why. It is unclear if there is a single model developed and applied to all years or separate models for each year. Are all topographic parameters used in the model in all years? What is their weighting?

Major comment: Extending the results to areas without soil moisture measurements (whether in situ or remotely sensed values) without any validation is not an improvement over current spatial data. It is highly suspect and could lead to inappropriate applications using what is in effect non-data.

Moderate comment: I found the paper very repetitive and in need of detailed editing.

Interactive comment on Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2019-191>, 2019.

C2