Earth System
Science
Data

Open Access

Discussions

ESSDD

Interactive
comment

Interactive comment on "Mapping the yields of lignocellulosic bioenergy crops from observations at the global scale" *by* Wei Li et al.

Anonymous Referee #1

Received and published: 27 October 2019

The authors use a machine learning technique (random forest-RF) to develop an up-scaled global (0.5 x 0.5 degrees) yield data set for five bioenergy crops. To justify how realistic this empirically-derived global bioenergy yield map, the authors further compare their product with the yield map used by the Integrated Assessment Models (IAM). In general, I agree with the authors that this dataset can become potentially a useful product for either benchmarking the global crop models (e.g. LPJ alike models) or being as input to IAMs. However, I think the method and results of this manuscript suffer from the following major weaknesses, which cannot make me convinced that this is a reliable product. 1. The authors disregard the details of temporal resolution and coverage of training data sets. 2. The authors haven't provided good reasoning for how they decided the training data sets. The temperature dataset in CRUNCEP is similar

to the CRU data set, which is based on observations, but precipitation has less good reliability. Also, why do authors choose satellite-based short-wave radiation? Does the median value in the high-resolution dataset have any advantages over the 0.5-degree data set (e.g. the CRU sunshine hours)? The water available index is a model-derived data set, but actually, there should be some satellite-based dataset to indicate soil moisture. In a word, I think the authors should give strong reasoning on why they have chosen their training data sets. 3. Given the big deviation shown between the yield map used by IAMs and the yield map derived by the authors, it is difficult to convince me of the reliability of the yield map generated by the random forest approach. I also wonder why the authors don't compare this product with their model estimates (Li et al., 2018b). Because the ORCHIDEE model has also been calibrated based on the same global bioenergy crop yield data set in Li et al. (2018a), it would be more logical to compare the derived product with the ORCHIDEE model estimate in the spatial scale.