

Earth Syst. Sci. Data Discuss., author comment AC2
<https://doi.org/10.5194/essd-2022-16-AC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC2

Xunhe Zhang et al.

Author comment on "Mapping photovoltaic power plants in China using Landsat, random forest, and Google Earth Engine" by Xunhe Zhang et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-16-AC2>, 2022

In this manuscript, machine learning and visual interpretation methods were combined to map the PV power plants in China. The topic is very important, and the study results would be useful for developing PV industries in the future. However, there are some problems in the manuscript which should be solved before reconsideration for publication in Earth System Science Data:

Response: We thank referee #2 very much for the positive feedback.

1. The Introduction section should be rewritten. The novelty of this study compared to the previous studies should be highlighted. And the useless contents should be removed.

Response to comment 1: By following referee #2's comment, we have rewritten the Introduction. In the new Introduction, we introduced the advantage of cloud computing, machine learning, visual interpretation and freely available remote sensing imagery to map the PV power plants in China. We also highlight the advantages of our study compared with previous studies in the introduction part. The current Introduction has been largely refined.

2. The "Dunnett's dataset" was used as the basic training and validation samples, which means your model's ability was "limited" to this dataset. In other words, the PV power plants that cannot be identified in "Dunnett's dataset" may also be ignored in your model. Although you mentioned that you modified the dataset, how it was implemented and what is the difference between the modified dataset and the original dataset were not clearly stated.

Response to comment 2: We thank referee #2 for pointing out this. Suitable training samples are indeed crucial for an RF model's classification accuracy and stable performance. We modified the Dunnett's dataset as our training samples in this study, and found that the labelled PV power plants in Dunnett's dataset are rarely distributed in eastern China, which will limit our model's performance to identify the PV power plant in similar areas. So we further manually selected and edited the extent of different PV power plants that were not annotated in Dunnett's dataset to ensure the labelled data covered

most of the parameter space of PV power plants in China. The total area of the PV power plants in China is about 897 km² from primary Dunnett's dataset and the area of the modified training regions was 1121 km². We then randomly sampled points within the training with a balanced quantity from humid and arid regions in China. We have put the statements in the revised manuscript in Section 2.1.3 for clearer future readership.

3. Writing should be taken more seriously as there are many writing and grammatical errors in the manuscript. For example, in line 201, " we discovered that the EVI values of PV power plants in 2020 was strongly and positively linked with the that in 2013".

Response to comment 3: We have thoroughly gone through the manuscript to edit the writing and correct the grammatical errors.

4. The quality of the figures should be further improved. For example, the "North arrow" in Fig. 1 and Fig. 5 were missing. And in Fig. 2, what do 1-6 mean?

Response to comment 4: We have further improved all figures in the revised manuscript: we added the "North arrow" in Figure 1 and Figure 5. We are sorry for the unclear description of Figure 2. We refined the figure 2 with accurate description. We added the scale bar and coordinate in the Figure 2. The specific meanings of 1, 2, 3, 4, 5, and 6 are the 6 example sites to show the true-color images from Landsat-8, Sentinel-2, and Google Earth for visual interpretation.