Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2021-271-RC1, 2021

Comment on gmd-2021-271

Anonymous Referee #1

Referee comment on "Predicting Global Terrestrial Biomes with Convolutional Neural Network" by Hisashi Sato and Takeshi Ise, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-271-RC1, 2021

In this manuscript, the authors applied a data-driven or machine learning approach, the conventional neural network (CCN), to estimate global distribution of the potential vegetation types. They evaluated accuracy of retrieving the present state, and then estimated future shifts under projected climate change. Finally, they discussed the merits and limitations of the empirical approach.

General comments

My first impression of the manuscript is that this is a mixture of old problem and new technique. Such revisiting is sometimes effective, but only if the new technique provides deeper insights and/or apparently higher comprehensiveness than those in precedented studies. In my view, regretfully, I could not find enough advancements in this study; it looks like an exercise of the CCN. In other words, I am unsure whether this manuscript falls within the scope of Geoscientific Model Development.

The manuscript is short and well-focused but need more methodological descriptions and insightful discussion. The manuscript starts from several statements about the Holdridge Life Zone, but I think this part is unnecessary. On the other hand, the authors gave few words on remote sensing of vegetation, even for validation of the estimation result.

As the authors discussed, the data-driven approach has limitations. The model may not be applicable to the states outside the range of trained data, and the present CCN model used only temperature and precipitation as input data. Namely, it did not account for the effects of atmospheric CO2, nutrient, and disturbance, each of which is hot issues in the study area and so needs further discussions.

I agree with the meaning of examining the potential vegetation, because natural disturbances and human impacts (e.g., land-use) are too complicated to discuss climatic impacts on global-scale vegetation. In this regard, the study is one of a few attempts to apply the machine-learning method to capture the potential vegetation. However, because of critical limitations and deficiencies, I cannot recommend accepting the manuscript for

publication.

Minor points

- Introduction: As mentioned in my general comments, Introduction starts from classic studies. I recommend putting more focus on modern and recent studies.
- Line 72: I am quite unsure why the ISLSCP2 data were selected as benchmark data of potential vegetation and why any remote sensing data were not used.
- Line 84: Please give references to NCEP/NCAR, HadGEM2-ES, and MIROC-ESM.
- Line 102: Did you used daily temperature? Or, monthly?
- Line 129: The computational times should depend on machine ability.
- Line 172: I could not find explanation about the 'certainty' of the CCN output in the method sections.
- Line 190: 'quantity' may be removed.
- Line 195: Table S9 should be moved to main body. Otherwise, you may rewrite this sentence.
- Line 203: Did you mean stand-replacing disturbances such as wildfire and wind throw? It may be better to provide several examples.
- Line 213: I could not understand the sentence. Why the model should have better performance than you showed, if the CRU dataset had high efficiency irrespective of grain size?
- Line 249–253: Here, you mentioned about the limitation associated with the atmospheric CO2 concentration. Indeed, atmospheric CO2 levels in 2100 are largely different between RCP2.6 and RCP8.5. So, you should make more discussions about this limitation.
- Line 275: At this very last part of the manuscript, you first mentioned about the hardware issue (NVIDIA DIGITS 6.0).