

Dear Referee #1

Thank you for taking the time to evaluate our manuscript. We feel your comments are pretty constructive for increasing the quality of the expected revised manuscript. Following is our one-to-one response to your concerns. Throughout this letter, your comments are written in blue color and are numbered.

(1) 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.

Response:

Yes, it is an accurate abstract of our work.

(2) 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 precedent studies. In my view, regretfully, I could not find enough advancements in this study; it looks like an exercise of the CCN.

Response:

Our approach has higher comprehensiveness than previous studies: it automatically extracts non-linear seasonal patterns for climatic variables relevant to biome classification.

The Holdridge Life Zone only considers annual climate means, and hence it cannot account for seasonal patterns of climatic condition, which affect biome distribution. Accordingly, many subsequent studies of biome mapping tried to incorporate seasonal patterns by assuming environmental constraints (such as tolerance of drought) for each plant group (such as plant functional types, biomes, or vegetation types). These approaches require absolute physiological limits for each plant group. However, there is no straightforward way to estimate such limits because plant groups contain a large number of species. By taking advantage of CNN, our approach can provide an easy, efficient, accurate, and comprehensive solution for this issue.

To clarify and emphasize this issue, we will insert the following sentence in the expected revised manuscript's abstract (line 14).

-- Unlike previous approaches, which require assumption(s) of environmental constrain for each biome, this method automatically extracts non-linear seasonal patterns of climatic variables that are relevant in biome classification.

Also, we will insert the following sentence in conclusion (line 283)

-- Reconstruction of global biome distribution substantially improved when climate seasonality was taken into consideration, demonstrating that the method successfully extracted seasonal patterns of climatic variables that are relevant in biome classification.

(3) In other words, I am unsure whether this manuscript falls within the scope of Geoscientific Model Development.

Response:

The authors' instruction of the Geoscientific Model Development (<https://www.geoscientific-model-development.net/>) defines scopes of manuscript types for considering peer-reviewed publication. Our manuscript satisfies the following items, and hence we are sure this manuscript falls within the scopes of Geoscientific Model Development.

-- Geoscientific model descriptions, from statistical models to box models to GCMs
-- Model experiment descriptions, including experimental details and project protocols

(4) The manuscript is short and well-focused but need more methodological descriptions and insightful discussion.

Response:

Description concerning machine specificity and software environment will be moved from supplementary information 1 (lines 23-26) to the main body (line 122). To keep the main body short, we hope to stay parameter setting descriptions at the supplementary information 1. Concerning discussion, we are happy to add more text if you specify what is not enough.

(5) The manuscript starts from several statements about the Holdridge Life Zone, but I think this part is unnecessary.

Response:

As you mentioned, there are several statements about the Holdridge Life Zone in the introduction. As our approach is a kind of an extension of The Holdridge Life Zone, we hope these statements to be maintained.

(6) On the other hand, the authors gave few words on remote sensing of vegetation, even for validation of the estimation result.

Response:

Please refer our response on the item (9).

(7) 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.

Response:

As you pointed out, our approach ignores atmospheric CO₂, nutrients, and disturbances like other equilibrium and niche models. Besides, it also ignores other mechanisms that can impact real-world responses and vegetative state transitions (such as reproduction times, dispersal abilities/limitations, and geographical barriers to migration). Nevertheless, our approach quickly assesses the degree to which potential natural vegetation (PNV) states are projected to persist or shift under climate change globally. Our approach provides one of the few applications of CNN at a global assessment of spatiotemporal dynamics among PNV using standardized, empirical, and ecologically relevant climate information.

Indeed, after submitting our manuscript, Elsen *et al.* (2021) published an article where they adapted the Holdridge life zone for evaluating how changing climate shifts terrestrial life zone. Our approach has a clear advantage to the study in considering seasonal patterns of the climatic condition by applying CNN.

Elsen, P. R., et al. (2021). "Accelerated shifts in terrestrial life zones under rapid climate change." *Global Change Biology*.

Your criticism is reasonable of cause, but it describes general limitations of whole studies employing the so-called "climatic envelope approach" not specific to our particular study. Besides, process-based approaches are also unreliable options, as explained in our manuscript (line 228-233, line 254-260). At this moment, it cannot be said which is the better approach for projecting global PNV distribution under changing climate.

For showing an example that a climatic envelope approach is used as a vital option for projecting biome map, we will add the following sentence, which refers to a recent study (at line 26).

-- For example, Elsen et al. (2021) applied historical climatologies and climate projections to the HLZ system for determining potential changes in global life zone distributions under changing climates.

Minor points

(8) Introduction: As mentioned in my general comments, Introduction starts from classic studies. I recommend putting more focus on modern and recent studies.

Response:

Please refer our response on the item (7).

(9) 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.

Response:

Following is a part of the description of the ISLSCP II Potential Natural Vegetation Cover dataset (https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=961), and it addresses your concern.

"The geographic distribution of contemporary land cover types can be derived from remotely-sensed data. However, humans now dominate much of the world and there is little evidence of the pre-human-settlement natural vegetation or Potential Natural Vegetation (PNV). PNV, as defined here, does not necessarily represent the world's natural pre-human-disturbance vegetation. Rather, our definition of PNV represents the world's vegetation cover that would most likely exist now in equilibrium with

present-day climate and natural disturbance, in the absence of human activities."

To clarify the nature of the dataset, and to clarify the aim of our study, we will insert the following sentences in line 74.

-- The ISLSCP2 dataset represents the world's vegetation cover that would most likely exist now in equilibrium with present-day climate and natural disturbance in the absence of human activities.

Besides, the "ISLSCP2" in line72 will be replaced by "ISLSCP2 Potential Natural Vegetation Cover" because ISLSCP2 is just the name of a project name, not a name of a dataset.

(10) Line 84: Please give references to NCEP/NCAR, HadGEM2-ES, and MIROC-ESM.

Response:

We will!

(11) Line 102: Did you used daily temperature? Or, monthly?

Response:

It's monthly. Thanks for finding our missing description!

(12) Line 129: The computational times should depend on machine ability.

Response:

Right, it primarily depends on a graphics card. The following description, which explained machine specificity, will be moved from supplementary information 1 (lines 23-26) to the main body (line 122). Information about computational time would be helpful for readers as it provides rough estimates of computation cost.

"The computer employed to execute the learning had Ubuntu 16.04 LTS installed as the operating system and was equipped with an Intel Core i7-8700 CPU, 16 GB of RAM, and an NVIDIA GeForce GTX1080Ti graphics card, which accelerates the learning procedure. On the computer, the NVIDIA DIGITS 6.0.0 software (Caffe version: 0.15.14) served as the basis for CNN execution, and LeNet was employed to train the CNN via the TensorFlow library."

(13) Line 172: I could not find explanation about the "certainty" of the CCN output in the method sections.

Response:

We will insert the following phrase (Line 172).

-- ", which is the probability (in %) of the classification judged by the CNN."

(14) Line 190: "quantity" may be removed.

Response:

In this section, "allocation disagreement" and "quantity disagreement" are distinguished. Your mentioned sentence describes quantity disagreement, so we cannot remove "quantity."

(15) Line 195: Table S9 should be moved to main body. Otherwise, you may rewrite this sentence.

Response:

As you suggested, Table S9 will be moved to the main body as Table 1. With this change, Tables S10 and S11 in the previous manuscript will be renumbered as tables S9 and S10, respectively.

(16) Line 203: Did you mean stand-replacing disturbances such as wildfire and wind throw? It may be better to provide several examples.

Response:

Here, we intended anomalous climate events that have catastrophic influences on plant mortality, and hence we will insert the following sentence in line 204.

-- For example, in response to anomalous drought during 2002-2003, regional-scale die-off of overstory woody plants was observed across southwestern North American woodlands (Breshears, *et al.*, 2005).

(17) 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?

Response:

We will replace the mentioned sentence as follows.

Previous: "Therefore, our validation method underestimates the actual performance of the models, and performance is much better than we demonstrated in this

manuscript."

New: "Therefore, our validation method, which suffers from the systematic differences among climate datasets, should underestimate the actual performance of the models, and performance would be much better than we demonstrated in this manuscript.

(18) 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.

Response:

We will insert the following phrases at the end of line 253.

-- Besides, projections of atmospheric CO₂ have significant divergence among socioeconomic scenarios from 421 ppm (RCP2.6) to 936 ppm (RCP8.5) at the end of the 21st century.

(19) Line 275: At this very last part of the manuscript, you first mentioned about the hardware issue (NVIDIA DIGITS 6.0).

Response:

For introducing the NVIDIA DIGITS 6.0 before this sentence, the following description in supplementary information 1 (lines 23-26) will be moved to the main text (line 122).

"The computer employed to execute the learning had Ubuntu 16.04 LTS installed as the operating system and was equipped with an Intel Core i7-8700 CPU, 16 GB of RAM, and an NVIDIA GeForce GTX1080Ti graphics card, which accelerates the learning procedure. On the computer, the NVIDIA DIGITS 6.0.0 software (Caffe version: 0.15.14) served as the basis for CNN execution, and LeNet was employed to train the CNN via the TensorFlow library."

Best,

Hisashi SATO (on the behalf of all co-authors)