Dear Reviewer #2,

Thank you very much for conducting a comprehensive review. I have addressed the issues raised to the best of my ability. Throughout this letter, given words are written in blue.

This manuscript aims to predict the global biome distribution using machine learning method based on climate characteristics and estimates its accuracy. Although CNN is a promising technology for image based vegetation classification, it remains limitedly underutilized for climate envelope modeling. I think it is good to see them being used here.

Unfortunately however, I have some concerns regarding the comparison of model performance. The authors need to clearly state that different type of input data was used to CNN and other models. The authors should also explain how these differences in data sources affected the results of the model performance comparisons.

I have four main concerns:
(1) The input data for training is not clear - The author explains that CNN uses graphical images of climate data as training data. I am not sure what graphical images of climate data meant. Is this RGB transformed climate data? If so, the authors need to clearly explain how they converted the climate data to graphical image data. The input data for RF, SVM, and NV also unclear. Are these models trained by climate data variable itself? If so, the authors also need to clearly state CNN and the other models were trained by different type of data. The authors should provide more detailed explanation about how the models used in this study trained.

Response 1-1:

The details of the method for converting climate data into graphical images are based on a previous publication (Sato & Ise, 2022). While this manuscript cited that paper, I will insert the following explanation on Line 137:

"The size of one graphical image is 256 × 256 pixels, and this image is divided into rectangular cells for as many data points as it represents, arranging tiles in each cell that express the values in grayscale. Before this visualization, climate variables were standardized to 0.01-1.00 with log transformation. The R code for drawing images is available in online open data."

If you find the above insufficient, I can also include examples of the generated images and further details of the image conversion in the supplemental information.

Response 1-2:

I will add an explanation as follows to clearly state that CNN and the other models were trained by different types of data.

Present manuscript (Lines 133):
"CNN algorithms are more complex than the others included in this study."

Revised manuscript:
"Although models except CNN were trained by climate data themselves, the application of CNN algorithms requires converting climate data."

Also, please refer to my response 3, which clarifies that models except CNN were trained by climate data itself. For your concern that the difference in accuracy or robustness between CNN and the other models would reflect the difference in input data, please refer to my response 4.

(2) One of my main concerns is that the fairness of model performance comparison. My understanding is that in this paper, CNN model was trained by graphical images of climate data while the other models were trained by climate data itself. Therefore, difference of accuracy or robustness among CNN and the other models seen to reflect not only model performance but also input data difference. Since model performance comparison generally aims to evaluate the performance of algorithms, I am not completely sure whether the model performance comparison in this study is truly meaningful or not. First, the author needs to clarify the reason why convert the climate data into graphical image. Second, the authors should also explain how these differences in data sources affected the results of the model performance comparisons.

Response 2-1:
CNNs are being used to evaluate graphical information. Sato & Ise (2022), the basis for this paper, developed and evaluated a method for coding numeric values in graphical information and then employed its classification with CNN. To clarify this method's advantages, I will explain as follows.

Present manuscript (Lines 136-137):
"I follow Sato and Ise (2022) in training our CNN with graphical images as input variables representing climatic conditions."

Revised manuscript:
"I follow Sato and Ise (2022) in training our CNN. This method represents climatic conditions using graphical images and employs them as training data for CNN models. This method can automatically extract non-linear seasonal patterns for climatic variables relevant to biome

classification."

Response 2-2:

Please refer to my response 4.

(3) P5 L129-132: Authors should explain the model setting so that other researchers can check the validity of their methods without checking code. I feel that the descriptions of the settings of machine learning methods other than CNN are insufficient. The authors should clarify more detail about the settings of machine learning methods, such as the information of the parameters or the type of kernel they employed.

Response 3

Sure! I will supplement the sentence in lines 129-131 as follows.

Previous:

"I used the default model parameters for simplicity and to prevent potential overfitting, i.e., training the model too closely to a particular dataset, thereby creating a model that might fail to fit additional data or reliably predict future observations."

Revised:

"More specifically, I utilized the commands *randomForest*(*VegNo*~., *Dataset_Train*), *ksvm*(*VegNo*~., *Dataset_Train*), and *naiveBayes*(*VegNo*~., *Dataset_Train*), where *Dataset_Train* represents the training dataset table, and *VegNo* is the name of the column within the table that holds the biome category. By opting for the default settings in these commands, I aimed to maintain simplicity and mitigate potential overfitting. Overfitting occurs when a model is trained too closely to a specific dataset, leading to a model that may perform poorly on new data or reliably predict future observations.

(4) P5 L129-132: I also have concern regarding the parameter optimization of ML. In this study, author used ML without optimizing parameters. I would suggest that the author try to optimizing parameters of ML using commonly used method such as grid search.

Response 4

Optimizing machine learning models requires significant amount of effort required. For instance, in the study by Sato & Ise (2022), which serves as the basis for this paper, optimizing and conducting sensitivity analysis of a single method, CNN, required adding 19 pages of Supplementary Information. Besides, optimization depends on data-sets for training. The purpose of our paper is to provide a quick perspective on the differences in performance and uncertainties among chosen methods when non-experts in machine learning construct models

for bioclimatic envelope construction using default settings. However, given that I employed each algorithm with its default settings, it prevents us from conclusively stating that CNN is the superior approach.

To address this point, I will insert the following clarification on Line 207:
"While it is necessary to consider that (1) default parameter settings are used for all methods adopted in this study, and (2) models except for CNN were trained with climate data themselves, while CNN employed graphically converted climate data, it prevents us from conclusively stating that which is the superior approach. However, "

(5) Minor points

P4 L101: Please check the format of references

Response 5

Sure! Thanks for noticing it.

Best,
Hisashi SATO