Interactive comment on "An assessment of geographical distribution of different plant functional types over North America simulated using the CLASS-CTEM modelling framework" *by* Rudra K. Shrestha et al.

Anonymous Referee #2

Received and published: 5 June 2017

This paper evaluates the performance of the competition module of the CLASS-CTEM modeling framework at the fine spatial scale of 1 degree by comparing the geographical distribution of vegetation types as simulated by the model with observation data. The paper also proceeds to try and identify the main reasons why the simulated results deviate from the observed patterns. The research represented by this paper is very important to the scientific community as it (1) documents the geographical regions where the model does and does not perform well, (2) by seeking to establish the reasons for deviations from observation, sheds more light on inter-model (ESM) uncertainty

and (3) finds out the response of terrestrial vegetation to increasing atmospheric CO_2 and a changing climate. However, the paper failed to evaluate the performance of the competition module of the model on certain key aspects and did not do thorough job of assessing the reasons for model deviation from observations. I would suggest a substantial revision of the paper.

My main/general comments are:

1. Establishing reasonability: The primary objective of a model is to 'look' at the terrestrial biosphere for such conditions that can not be observed. For example, experimental manipulations of ecosystems, simulations for climatic conditions in the future. It would be highly improbable to have model outputs exactly match observational data. Two independent observational datasets have differences. Therefore, while assessing model performance, it is essential to establish quantitatively what 'reasonable performance' is. It could be a measure of deviation from multi-observational mean (ensemble) or a threshold for spatial correlation (R^2). Establishing this 'reasonability' will help users of this model to limit themselves to assessing simulation results for certain geographical regions and/or certain Plant Functional Types (PFTs).

2. More comprehensive model performance assessment: a) A critical element of model evaluation that is missing is comparison with other published DGVM outputs. DGVMs are similar in their basic structure but differ with respect to certain processes and/or parameterization. Thus comparison with other DGVM outputs will allow the authors to establish how the carbon cycle dynamics and/or terrestrial vegetation patterns of CLASS-CTEM compare with that of other models. This will further help to establish the 'reasonability' of CLASS-CTEM. I would suggest using the ensemble of DGVM outputs from the TRENDY project, which is a consortium of DGVM groups who set up a project to investigate further the spatial trends in Net Biome Production (NBP) and agreed to perform a factorial set of DGVM simulations over the historical period, 1901 – 2010 (see http://dgvm.ceh.ac.uk/). b) Since the primary objective is to compare model performance, it will be a good idea to compare the model GPP to a satellite based model

output like the MODIS MOD17 data (see http://www.ntsg.umt.edu/project/MOD17/ →
'Data Product'). Using this will also help in establishing observational uncertainty.

3. Regional analysis: The paper currently does perform a regional analysis, but it is
qualitative. The future users of CLASS-CTEM would want to know in which specific
regions of North America is the model performance reliable enough. I would thus
suggest using shapefiles to demarcate the different regions being studied and then
carry out statistical comparisons for each region. It would be helpful for the reader if
the maps of individual regions are explicitly shown in the paper.

4. Time series comparison: According to the authors, the primary goal of the paper is
to evaluate the performance of the competition module of CLASS-CTEM. While com-
paring current vegetation patterns and productivity with that of observational data is a
good way to evaluate competition, it is not comprehensive since it looks at only one
time slice and equilibrium conditions. Competition can be observed in reality when an
ecosystem is subjected to environmental changes and/or disturbance. The time-series
data of both CLASS-CTEM data and observational data both incorporate the response
of the terrestrial ecosystems to changes in the climate and atmospheric CO2 concen-
trations. Thus a better evaluation of the competition module should include equivalent
time series comparisons of model and observation.

5. Attributing deviations from observation: The authors admit that it is difficult to de-
termine whether the model limitations are due to biogeochemical parameterizations or
structural limitations (Pg 20, lines 578-580). If this is the case, then attributing devi-
ations from observations is mere speculation. While speculation backed by scientific
logic is okay for the discussions section, it shouldn't be present in the abstract since
this isn't something that the authors have assessed. That being said, model limita-
tions can be diagnosed by tuning parameters. The biogeochemical parameter that is
the main suspect causing deviations from observation should be identified and parallel
simulations should be run using varying values of the parameter and/or using values
which are a closer match to North American species.

6. Impact of change in spatial scale: The authors say that one of the goals of the paper
was to look at how the 1 degree version of the model changed results when compared
to the earlier study performed using the 3.75 degree version (Pg 23, lines 674-678).
The authors make qualitative statements and not any quantitative assessments. I was
expecting to see RMSD and R2 values for both model versions. This would help es-
tablish whether a higher spatial resolution helps to improve model performance or not.

Minor comment:

The paper would be of much more value to the climate science and/or carbon cycle
community if the authors used more of carbon cycle flux and/or stock terms/values
to assess CLASS-CTEM performance. While it is useful to have a detailed analysis
of the major PFTs, analyzing the minor PFTs doesn't add much value to the model
assessment. Instead grouping PFTs under broader categories (eg. Trees/grass and/or
needleleaf/broadleaf trees) adds value to the paper as it makes comparisons with other
models and datasets simpler.