Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-149-RC3, 2020

Biogeosciences Discussions

Interactive comment on "Improving the representation of high-latitude vegetation in Dynamic Global Vegetation Models" *by* Peter Horvath et al.

Anonymous Referee #3

Received and published: 25 August 2020

The manuscript "Improving the representation of high-latitude vegetation in Dynamic Global Vegetation Models" by Horvath et al analyses the performance of three different vegetation modeling approaches with regard to the spatial distribution and relative abundance of plant functional types (PFT) in Norway. The modeling approaches include a dynamic global vegetation model (DGVM), remote sensing (RM), and a statistical distribution model (DM), which relates occurrences of vegetation types to multiple environmental variables. The authors found that both RM and DM showed a better performance than the DGVM when compared to observational data from a range of field sites. They then tested if it was possible to use the DM to improve the predictions of the DGVM with regard to PFT composition and distribution. It was found that,

through inclusion of three further bioclimatic constraints based on the analysis of the DM, the performance of the DGVM could be improved. The authors recommend DM as a complementary tool for the assessment and improvement of DGVMs.

The manuscript is well written and easy to understand in general. The research topic (assessing and improving DGVMs at high latitudes) is certainly relevant, and the chosen approach is original and seems useful to me. However, the description of the methods needs to be improved, with regard to the chosen statistical approaches, and also the motivation to carry out certain analyses. It often becomes clear only later in the manuscript why a certain method was applied. I therefore recommend minor revisions before a new version of the manuscript may be submitted.

Comments:

L 28 While the term 'DGVM' is explained at the beginning of the abstract, the term 'distribution model (DM)' is used in this sentence without previous explanation. Please explain shortly in the abstract what a DM is and how it differs from a DGVM, since some readers may not be familiar with the concept.

L 58 Please define or explain in more detail what you mean by 'thematic resolution'. Furthermore, it should be mentioned that recently, specific high-latitude PFTs, such as mosses, for instance, have been added to a number of DGVMs, e.g. Jules (Chadburn et al, 2015, The Cryosphere), JSBACH (Porada et al 2016, The Cryosphere), or ORCHIDEE (Druel et al 2017, Geoscientific Model Development) and several more.

L 60 Three examples are given for the difficulties of DGVMs to simulate extents of high-latitude PFTs correctly. However, I do not see how the underestimation of forest carbon storage by DGVMs relates to this, since this is rather a consequence, and not a reason for the incorrectly predicted extent. Please explain in more detail.

L 71 Please add a short statement to describe in which regard the RS products are not consistent.

L 83 At least one study (Druel et al 2017, Geoscientific Model Development), uses site data to assess the DGVM's performance with regard to plant traits. Please be more specific in this regard, and explain what exactly is new in the validation method.

L 121 I do not understand this sentence: If one plot is 0.9 km2 large, then 1081 plots are around 1000 km2, but 18x18 km are only 324 km2. Also, the plots are distributed throughout Norway, so the 18x18 km area has to mean something else. Is it the distance between the plots on a grid which covers Norway? Please explain.

L 129 To me it seems that low values of temperature and precipitation are underrepresented in the 20 selected plots compared to the full data set. This should be mentioned here briefly and then considered later in the Discussion section.

L 156ff By using the default surface parameter values for CLM, the DGVM may miss some relevant information to correctly predict PFT distribution, compared to RS and DM. Furthermore, by using climate forcing from 1980-2010 and running the DGVM into a steady state with regard to this period, historical climatic effects, which may influence today's PFT distribution are not considered. These points should be mentioned in the Discussion section of the manuscript.

L 162 Why was the CORDEX data not also used for the DM method? This should be briefly mentioned here.

L 175 Please explain 'supervised' and 'unsupervised' in more detail.

L 182 the number of explanatory variables (116) is rather high. It should be shortly explained what these are, and why such a large number is necessary for the regression. Even if this information is provided in Horvath et al (2019), it should be summarized here.

L 183 It would be good to add a short summary of the evaluation method for the DM here, so the reader can assess the DM better.

L 186 I wonder if, by discarding all other VT except the most probable one, biases in the

distribution of the VTs are introduced. Let us assume the logistic regression predicts a certain VT always with a slightly higher probability than a second one; according to the description, only the first VT would occur in the predicted map at all pixels, and all observations of the second one would be discarded, although this VT occurs quite frequently in reality. Please explain this in more detail.

L 200 I don't understand why an aggregated PFT profile is needed, I thought that the comparison of the 3 modeling approaches and the AR data is done for each of the 20 plots?

L 208ff This sounds like one comparison was done with the aggregated profiles (one for each method, aggregated over all 20 plots), using the chi-square test. Then, for each of the 20 plots the profiles were compared regarding their dissimilarity. It is not clear to me, why two different statistical methods were used to compare the models (DM,RS, DGVM) to AR.

L 222 I thought the dissimilarity index was used to assess the similarity between the 3 modeling approaches and the AR data. Why is it then necessary to do a pairwise Wilcoxon-Mann-Whitney test in addition? Please explain the reasons for the chosen statistical approach in a more detailed way.

L 230ff As mentioned above (L200), by aggregating the PFT profiles of the 20 plots, differences in profiles between plots are lost. Hence, it is not possible to evaluate the 3 models with respect to the correct prediction of differences in profiles between individual plots. Also, while the AR data (for each plot) can be interpreted as a random sample, it is not clear to me how the model approaches can be consistently included in this Chi-square test. Moreover, the number of elements (6 PFTs) is actually too small for a Chi-square test. The authors need to justify this better, or change their testing approach.

L 249 If I understand Fig. 2 correctly, the lines which connect the dots denote the individual plots, which means that for one method (e.g. DGVM), the dissimilarity can

be high (1.0), while for another method (e.g. RS) it can be much lower. The result that the goodness of the fit between a given method and AR data depends on the set of chosen plots may point to some underlying systematic deficiencies of each method and should be discussed later.

L 252 The statement in this sentence is not evident to me in Fig. 3, because this figure simply shows the profiles for each plot (which is a good way of illustrating the results, in my opinion). Wrong reference?

L 254 Please see also my comment to L 222; I assume that the authors use the Wilcoxon test to assess if the median values of the dissimilarity indices for the 3 models are significantly different from each other. However, I think it is more relevant how the models differ to each other with regard to the AR data. This information is contained in the values of the dissimilarity index, and it should be reported more clearly here. The pairwise comparison of the 3 models seems to me of secondary importance to assess the goodness of the fit to AR data.

L 262ff The visual comparison of the 3 models in Fig.3 and the associated description is more helpful to assess the modeling approaches than the statistical methods described before.

L 279ff This belongs into the Methods section. Explaining the sensitivity analysis earlier also makes it much easier to understand the goal of the overall approach.

L 287 The term 'precipitation seasonality' should be better described, in particular since it is found later that it is important to improve DGVM parameterization.

L 379ff The point about 'good' and 'poor' DMs is not clear to me. Why should poor DMs be used at all? Please explain, and also consider my comment above (L 186).

L 411 It may not be clear to readers why the lack of a shade-intolerant birch-PFT in the DGVM leads to the over-representation of NET in plots 17 and 18. The birch-PFT should rather have an advantage in mountainous regions compared to NET, which is

currently lacking in DGVMs. Please clarify.

L 450 Please check the literature for the recent progress in including high-latitude vegetation types into the PFT scheme of DGMVs, and add this to the discussion.

L 467 This sentence is hard to understand, please reformulate.

L 475 It should be mentioned if increased seasonality promotes or impedes growth of NET.

Supplement:

L 40 missing reference L 51 missing reference L 52 missing reference

L 55 The PFTs for this study are not in bold font, but shaded grey, please make this consistent.

L 56 The caption of Tab. S6 should be a bit more detailed: Is zbot the bottom height of the canopy (11.5 m above ground)? How is the coefficient of variation in precipitation seasonality computed?

L 90 The cover fractions in plots 801,2108,4268 are clearly not in a steady state. Please check if this significantly affects the results (e.g. by extrapolating the trends in cover), and repeat the DGVM runs, if necessary.

L 122 missing reference

Comments on style:

L 42 I think 'an' is not needed here.

L 55 'DGVMs' instead of 'the DGVMs'

L 60 'at high latitudes' instead of 'in the high latitude'

L 66 'in' not necessary

L 138 the second "of the" is not necessary

L 373 add ’the’ before ’reason’

L 401 ’differ’ instead of ’differs’

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-149, 2020.