

This manuscript uses global vegetation models (GVMs) to project the likelihood of biome shifts in different climate change scenarios. The study shows that the used gvms are capable of reproducing current biome distribution as biome maps derived from model outputs agree with observation-based biome maps. The gvm biome projections show increasing risks of biomes shifts in higher emission scenarios and that higher latitudes were more likely to have biome shifts than lower latitudes, which the authors attribute primarily to temperature effects.

The manuscript is well written and generally clear, despite the methodology requiring some clarifications and the discussion being excessively long and descriptive in my opinion. I found the study results interesting and useful for scientists interested in climate change. My main criticism is related with how some hypotheses are tested. The authors use a classification based on expert knowledge of temperature limited and not-limited biomes and compare how these groups shift. The problem with this approach is that the actual temperature change in the grid cell is not considered, then it is impossible to claim if the biome really shifted because of temperature effects or some other reason (that is, even a very temperature limited biome still couldn't have shifted because of temperature if there was no temperature changes in that particular grid cell to being with). I believe a more rigorous test should involve assessing the relationship between biome shifts and the actual temperature change in the grid cell.

Thank you for the overall positive feedback and the suggestion for additional analyzes to relate biome change and temperature change. We agree that our classification into temperature limited and other biomes is expert-based and created by the authors. As commented, we did not relate biome shifts to actual temperature shifts projected by the climate models. We only referred to previous studies indicating that temperature change is projected to be stronger in higher latitudes and concluded on biome shifts.

However, our hypothesis that predicted biome shifts are more consistent in temperature-limited biomes was based on two observations: 1.) in particular cold-temperature limits are relatively well understood and, therefore, probably well and similarly represented in different GVMs 2.) Climate models simulate very consistent and very strong temperature changes in the high-latitudes (where cold-temperature limits are crucial for biome borders). In other words, the difference in temperature change is a mechanistic component of our hypothesis (see line 33). We think that correcting for temperature change to better understand the models would be beyond the scope of this study, also because reviewer 2 already commented that having the methodological advances in terms of biome classification and a focus on the future results might be too much for one paper. We will clarify these issues and explain our hypothesis better in the revised manuscript. Apparently, it was not fully clear in the current version.

Specific comments:

L24: Provide references for the claim that biomes are often used to assess vegetation change.

We will add appropriate references already cited elsewhere in the ms, such as Gonzales et al 2010, Huntley et al 2023, Bonannella et al 2022, Tobian et al 2024.

L35: I think this statement needs to be clarified. Model projections of precipitation are more uncertain than temperature projections but that doesn't mean precipitation regimes won't change in the future, only that model capacity to project it is currently more limited/varied between models. Therefore, I think it needs to be clarified that temperature limited biomes could respond more to current climate change projections, which does not necessarily mean they will be the most affected by climate change. In addition, it is important to consider that increased temperatures could also contribute to drought, through temperature effects on atmospheric vapor pressure deficit.

We fully agree about these points and will clarify them, which is partly related to the arguments for our hypothesis that we describe above. Of course, precipitation changes, in particular in combination with increasing temperatures and evaporative demand, can cause large-scale biome shifts, such as drought- and heat-related forest dieback. There is a large debate how changes in drought and drought-induced tree mortality may influence ecosystems under future climates. Yet, given the challenges and limitations of modeling plant hydraulics and tree mortality, as well as more consistent projections of temperature change in different GCMs, we argue that biome shifts related to temperature should be represented more accurately and consistently in GVMs and have larger effect size.

We will revise this statement as suggested, to be more nuanced and concise. We will make clear that current models should be more appropriate to model biome shifts related to temperature than to precipitation change, particularly drought and temperature effects on VPD.

L95: I don't fully understand the justification for using only increased CO2 scenarios. This was the combination that had the highest number of models available?

Yes, this combination had the highest number of models available. This maximizes comparability across models as all forcings are identical. We will state this more clearly in the revised ms.

L98: Please clarify if all these models have a dynamic vegetation component (e.g. PFTs competing for grid-cell space, Prentice et al 2007)? I think it should be better explained how each model handles vegetation dynamics and responses to climate (especially temperature) in supplementary material S1.

The analysis includes both models with (LPJ-GUESS, ORCHIDEE-DGVM, CARAIB) and without (ORCHIDEE, CLM) a dynamic vegetation component. We agree that it is important to state this clearly in the ms. We will clarify this in the main text and the supplement, and also describe how vegetation responds to climate and temperature change (e.g., by describing if bioclimatic limits related to temperature are included). Note that in ORCHIDEE and CLM, simulated changes of LAI of PFTs in response to climate change can also trigger biome changes, even if the PFT distribution is not dynamic.

L115: Is it really sensible to discuss biome shifts in anthropogenic areas (for example Agricultural lands)? You are assuming agricultural lands are currently a natural biome based on potentially minuscule PFT fractions on it? Wouldn't be more sensible to exclude these areas?

We share the concern that human impacts are highly relevant w.r.t. current biome distributions and potential future biome change. Therefore, we do not analyze biome shifts in anthropogenic areas but remove them by (1) removing anthropogenic PFTs from the Fischer et al biome maps, and (2) removing anthropogenic PFTs from GVM results using the scaling approach described in L 115. Thereby, our analyses only include natural biomes (from observation-based products) and natural PFTs (from GVMs). We will revise the introduction and methods for clarification. We think that some readers will find global potential natural biome shifts interesting, even though they are not very relevant in areas currently covered 100% by croplands and cities. Excluding such areas based on some kind of arbitrary threshold would not add much, we believe. Furthermore, it is not impossible that some agricultural areas will be reforested until the end of the century, in fact quite a few scenario studies suggest that the agricultural area might decline in the future in some areas.

L145: Why 2 PFT products? One current and one future? Please clarify

Both PFT products are for present-day vegetation, future maps are not available. The two products differ in the PFTs included the method to create the products and the spatial patterns of PFTs. To account for these differences, we included two different products for current conditions. This analysis allows us to evaluate how robust the GVM-based classification is, compared to classification of observation-based PFTs. We will clarify this in the text.

L187: That doesn't seem like a very rigorous test of the temperature effect.

This classification into temperature driven and not temperature-driven biomes was not based on an analysis but on our knowledge. We argue that this classification is not unique as, to some extent, all biomes are influenced by temperature, precipitation, soil, and other factors. Yet, some biomes such as boreal forests have well-established bioclimatic limits related to temperature (extreme cold and growing season length towards the north and chilling requirements towards the south, which have been well constrained already in the 1990ies last century). Our classification is based on such model assumptions.

L240: Please show graphically or at least describe these effects better. From the info in the previous paragraph, the difference between scenarios does not appear very large. P-values from traditional statistics are not always very informative when using very large samples that can be obtained from model outputs and will be highly "significant" even when the actual effect is very small.

We will add a table showing these results and describe them better in the text. Regarding p-values in the previous paragraph: those were not derived from statistics using the full global spatial information of biome maps and GVMs, but contain only 31 data points for each of the 31 biome maps. Spatial data were aggregated into one point per biome.

L270: Where do this p value comes from and what it refers to? The difference between coefficients of variation?

Yes, in this analysis we compare coefficients of variation as indicator for consistency of biome shift of temperature-limited and other biomes. We will reword this paragraph and make clear that the analysis well-described in the methods section.

L279: I do not think this claim was rigorously tested in your results. I suggest investigating the temperature changes in grid cells with different likelihood of biome shifts.

See our comments above. We do not claim to disentangle how strongly the biomes respond to a standardized temperature change.

L308: If fire is included in the GVMs you use and they still have low-agreement with data it shows that either the fire model is not working very well or that other factors not explicitly included in the models (such as herbivores) or even other aspects of the model structure or parametrization do not work well for the biomes here.

We agree that multiple processes may explain data-model disagreement, and that there is still high potential for model development and improvement. Yet, we think that it is beyond the scope of our study to dive deep into each GVM and understand which process explains disagreement. Our study can provide helpful insights for the different modeling groups where mismatches with respect to biomes occur and accordingly what biomes or regions should be prioritized. Those insights are based on a large number different classification models informed by the 31 biome maps. We will revise the discussion to make this clearer.

L333: Couldn't these acclimation effects also affect the vulnerability of higher altitude biomes? As far as I know they aren't included in GVMs.

We agree that acclimation effects may also be relevant in higher altitudes. We will revise the discussion to account for this point. In the meantime, several studies have addressed this issue with GVMs at the global scale, which we will discuss in our revision.

L385: Do any of the used GVMs use these empirical limits? That would make the high agreement with observation maps less impressive than if were the result of bottom-up vegetation dynamic models as suggested in the beginning of the discussion.

Yes, many GVMs use a set of bioclimatic limits, e.g., frost or minimum temperature thresholds, to constrain climatic niches of PFTs. Those bioclimatic limits are based on empirical evidence and implicitly represent eco-physiological processes that constrain growth once those limits are crossed. PFT defined by bioclimatic limits and biomes are not fully identical such that the models still produce variables (biomes) not directly used for model parametrization and calibration. So it is not a priori clear, that GVMs using bioclimatic limits, and the underlying ecosystem dynamics will produce LAI and PFT patterns that represent the biomes included in different biome products. Yet, we agree that that GVMs where bioclimatic limits emerge from ecological processes bottom-up, that would allow adjustments of bioclimatic limits in response to climatic and atmospheric change would be desirable.

In the revised ms, we will better describe bioclimatic limits included in the GVMs and how they were derived.

L394: It would be important to clarify which processes are affected by temperature in each model in the methods section.

We agree, and we will add more information on temperature-related processes and bioclimatic limits in the methods section (see also a previous response). Thereby, we will focus on the direct impacts through bioclimatic limits for the distribution of PFTs.

Fig. 4. Please define “Man” in the legend.

“Man” is managed land, we will add this in the legend.