Author response to referee comments (cp-2023-16)

Referee #1:

Referee: „The authors compare pollen-based reconstructions of Holocene vegetation (REVEALS) with simulated vegetation from an Earth System model (ESM) and a dynamic global vegetation model (DGVM). They find that the simulated decline in Holocene forest cover is too abrupt and late in both models and attribute this to errors in the land-use history applied in the models. The general underestimation of forest fraction in the ESM is attributed to the parameterisation of disturbance.

This is interesting study with an important finding which is very relevant for this journal and which is strengthened by the use of two very different modelling approaches. This and the level of detail is a strength. My main comment (see below) is that the paper is overly long and the authors should aim to substantially shorten the text in several places in order to make it more accessible."

Authors: We thank Referee#1 for his/her constructive and encouraging comments that have helped to strengthen the manuscript. We substantially shortened the text according to his/her suggestions.

Main comments:

Referee: I appreciate that this is a technical paper probably mostly aimed at those within the field, but I found the text in places hard to digest. In the methods I would recommend that you include some summary paragraphs at the start of the section 2 to very briefly explain the differences between MPI-ESM, JSBACH, LPJ-GUESS and REVEALS. More generally I believe that you can make the text more concise, especially section 2 (currently ~ 6 pages), section 3.1 (2 1/2 pages), 4.1 (3 pages) and section 4.5 (2 pages). Some of the detail in section 2 could also optionally be moved to an appendix.

Authors: Thank you for this comment. We insert the following short summary of the used models to section 2: *„In this study, we use DGVMs with very different modelling approaches. JSBACH is the land-surface component of the comprehensive Earth System model MPI-ESM, and thus calculates interactively the grid-cell cover fractions of different plant functional types (PFTs) in line with the simulated climate (Reick et al., 2021). In contrast, LPJ-GUESS is a so-called "gap" model that calculates the growth of individuals or cohorts in patches, according to a prescribed climatic forcing (Smith et al., 2001). While we explore one of the few worldwide existing transient Holocene Earth System model simulations (MPI-ESM), the high spatial resolution time-slice experiments in LPJ-GUESS have been conducted in this study. LPJ-GUESS is used here because it is well-tested for the European domain (e.g. Hickler et al., 2012). The results of both models are compared to the pollen-based REVEALS reconstructions of plant cover for Europe from Marquer et al. (2017). These reconstructions were chosen before others based on other methods such as MAT (e.g. Davis et al., 2015) or pseudobiomization (e.g. Fyfe et al., 2010) because the REVEALS model has been shown to be the best approach to produce reliable quantitative reconstructions of forest cover and the only method available to date to reconstruct the cover of individual plant taxa (e.g. Hellman et al., 2008; Roberts et al., 2018; see also Introduction above). The methodological strategy used in this study is summarised in the flow chart of Figure 1."*

We shortened the method section and section 4.5 and moved one paragraph of section 4.5 to the method part of the REVEALS model, according to the suggestions of Referee#1 and #2. However, since climate biases in the models are commonly used as explanation for model-data mismatches, the climate comparison and the resulting exclusion of this argument for the main model-data differences in this study, are an essential part of our manuscript. Therefore, we kept the detailed discussion of the climate (sec.4.1) and description of our results (sec. 3.1).

Referee: Can the authors comment on whether other parameterisations besides the ones explored (disturbance and windthrow) could contribute the biases in the forest cover simulations?

Authors: All parameterisations can have a greater or lesser effect on the simulated climate and vegetation, but it is difficult to say how strong the effects are without running sensitivity simulations for individual parameters. In this study and for determining the causes of differences in forest cover estimates, the prescribed disturbances were the most obvious place to start because, besides natural mortality, the disturbances have a direct impact on simulated forest cover by reducing (or removing) the tree PFTs. In addition, some areas of Europe are often exposed to strong winds. In these regions, the wind-throw is particularly effective.

Minor Comments

Referee: Line 19: specify that LPJ-GUESS is the dynamic vegetation model.

Authors: We further specify the type of models.

Referee: Line 297: I could not see a figure with these results, please can you check this?

Authors: The REVEALS reconstructions are displayed as dots on top of the modelling results. We now specify in the text that these results are shown in Fig. 3.

Referee: Figure 5: could you add a key to this figure to make it easier to read?

Authors: We added labels to the bars on the figure.

Referee: Lines 588-600: It is probably worth clarifying here that the KK10 reconstruction has substantially earlier and more widespread land-use than LUH2.

Authors: Thank you for this remark. Nevertheless, we decide to keep the text as it is, because we already mention in the previous sentence (starting L594) that the land-use products differ substantially: *„However, KK10, HYDE and other ALCCs exhibit large discrepancies in their estimates of the starting time, spatial pattern and intensity of anthropogenic land-cover change, making it a challenge to simulate the effect of human-induced vegetation changes with DGVMs (Kaplan et al., 2017; Gaillard et al., 2010).”*

Referee: Line 615: I cannot comment on how topography is treated in REVEALS. As written this sounds a bit uncertain. Can you confirm that topography is not accounted for and how this would influence the model results. At the moment it sounds as if you're unsure whether it does or not.

Authors: The REVEALS model does not take into account topography. See discussion of this issue in Marquer et al. (2020). The authors conclude from their pilot study in the Pyrenees that the REVEALS model does not work well in mountains unless the number and location of sites capture pollen assemblages from various altitudes, below, at and above the tree line. The latter is mostly not realized in the grid-cell based REVEALS reconstructions we are using, therefore the REVEALS estimates in e.g. the Scandinavian mountains should be considered as less reliable.

We now write: "*Because topography is not accounted for in the REVEALS model and reliable reconstructions in mountainous regions require a large number of pollen records from large lakes representing the major altitudinal vegetation zones (Marquer et al., 2020), the REVEALS estimates used here need to be considered as uncertain in the Scandinavian mountains and the Alps, which may explain discrepancies with DGVMs at the grid-cell scale level in these areas (Fig.3)*"

Referee: Line 705: This seems to contradict the sensitivity test with increased disturbance you present earlier, can you comment?

Authors: We are not sure if we understand the referee's concern. In the sensitivity study, we basically analyse the effect of reduced wind throw. To further clarify our statement, we modified the text as follows: „*However, the mid-Holocene natural forests were probably much more stable and less sensitive to disturbances than the heavily human-altered present-day forests. This would justify and require the use of a lower storm vulnerability of forest in JSBACH for this period. Thus, whether the modern climate-derived model parameter values may be valid for the entire transient simulation, is questionable.*"

Line 717: This seems to echo results by Kaplan et al 2017 (*Land*). Is that correct? If so, can comment on this here and earlier in the text?

Authors: Kaplan et al. (2017) evaluated the performance of two commonly used anthropogenic land-cover change models (ALCC scenarios) by comparing their estimates with REVEALS reconstructions of open land cover, and discussed the discrepancies in the light of human-habitation-related unknowns in model parametrization and spatial allocation of deforestation. Here, we discuss a different problem, the observed discrepancies in forest fraction estimates of two vegetation models and REVEALS in Europe. What is said in our conclusion has been suggested by previous papers. But no earlier analyses have looked into that specifically.
To further clarify this, we modified our conclusions as follows:
"*Our study highlights the fact that model settings that are tuned for present-day conditions may be inappropriate for palaeo-simulations and complicate model-data comparisons with additional challenges. Moreover, our analysis identifies land use as the main driver of the decrease in forest cover in Europe during the mid- and Late Holocene, as has been suggested by pollen studies and the more recent efforts to quantify pollen-inferred changes in plant cover (e.g. Robert et al., 2018) and the various scenarios of anthropogenic land-cover change developed over the last ca. 20 years (e.g. HYDE and KK10, see synthesis in Gaillard et al., 2010). Mid- and Late Holocene changes in climate have only a minor effect on forest cover, although changes in cover and distribution of individual plant taxa depend on both land use and climate (Marquer et al., 2017).*"