Comment on bg-2021-37
Anonymous Referee #1

[I've reviewed the same manuscript before. As far as I can see (main conclusions are unchanged, figures are identical), the manuscript version under review at here is identical to the earlier version I have reviewed. Therefore, I am posting my previous report here again.]

This paper investigates trends in global leaf area index (LAI) and attributes them to drivers (climate, CO2, land use change) based on factorial simulations with a set of DGVMs and a fully coupled Earth System Model. This is basically a revisiting of a study published earlier (Zhu et al., 2016 NatCC) that applied the same approach (model-based attribution of drivers) and used the same LAI product (GIMMS3g, based on data from the AVHRR mission; Zhu et al. also used GLOMAP and GLASS to obtain more robust results).

Winkler et al. reach conclusions that have potentially high relevance for our understanding of global vegetation dynamics in response to climate change and (in particular) to CO2. Effects of rising CO2 remain a major uncertainty in Earth System Model projections, owing to challenges in observing and attributing effects. Hence, deriving new insights from available observational records is needed.

The paper by Winkler et al. is well written and display items are of high quality. The fact that their conclusions directly challenge findings by Zhu et al. (2016), although relying on largely the same method and data, caught my attention. Winkler et al. write in their abstract ("Our results do not support previously published accounts of dominant global-scale effects of CO2 fertilization.", l. 16) and in their conclusions ("A cause-and-effect relationship between CO2 fertilization and greening of other biomes could not be established. This finding questions the study by Zhu et al. (2016) that identified CO2 fertilization as the most dominant driver of the Earth's greening trend.", l. 722), and in the Key Points ("Most models underestimate the observed vegetation browning, which could be due to an excessive CO2 fertilization effect in the models.")

Strong conclusions require strong evidence. However, I have several strong concerns with how these conclusions were reached. In my view, the evidence presented here does not support this main conclusion (represented by the three citations I refer to above). Although I'm convinced that the analysis itself is diligently carried out and I consider that the paper offers a valuable discussion of the wide and relatively recent literature on the topic, I am concerned that the main conclusion will not meaningfully contribute to
advancing the field. My concerns revolve around two main points:

- Winkler et al. rely on a single LAI product to derive trends. Yet, several papers have documented inconsistencies between greening and browning trends between satellite data products. In particular, the product used here (LAI3g) is based on data from the AVHRR mission. It has been reported that respective data is affected by orbital drift of the satellite (Tian et al., 2015) and sensor degradation (Piao et al., 2019). The MODIS Collection 6 does not support the AVHRR-derived browning trends in several regions (see also Chen et al., 2019). This affects in particular North American boreal forests. The AVHRR-based browning seen in this region combined with the apparent failure of models to capture the same trends has been used by Winkler et al. (with a largely over-stretched logic) to argue that CO2 effects were inappropriately represented in models ("Thus, it is important to focus model development not only on a better representation of disturbances such as droughts and wildfires, but also on revising the implementation of processes associated with the physiological effect of CO2, which currently offsets browning induced by climatic changes.", l., 744). This is non-sense both in view of the known lack of robustness of apparent browning trends, and in in the logic of the argument itself. The failure of models to capture a browning trend may also be due to insufficiently sensitive responses to climatic drivers; and in this case (North America) is most likely due to inappropriate representations of disturbance in models (Anderegg et al., 2020 Science).

- One aspect that distinguishes the study by Winkler et al., from that of Zhu et al. (2016) is their probabilistic driver attribution. As the authors write, the method has been adopted from Pearl et al. (2009) and Marotzke et al. (2019) who applied it to attribute drivers of near-term climate change. However, I consider that the application of this method to investigate drivers of vegetation change is ill-conceived. The usefulness of probabilistic attributions is evident when dealing with systems that are characterized by a substantial inherent stochasticity (deterministically chaotic systems). In such cases, simulated variations are not necessarily forced, by may result from unforced internal variability. This is not the case for vegetation dynamics, where the (simulated) internal unforced variability is typically zero (except for some models, e.g., LPJ-GUESS, that simulate stochastic gap formation, or some relatively small internal variability arising from stand dynamics - which are actually not simulated explicitly at the individual/cohort level in TRENDY models). This actually facilitates driver attribution. All simulated trends are uniquely attributable to drivers using factorial analyses. As a consequence of relying on a probabilistic attribution, Winkler et al., find, e.g., no clear attribution in some semi-arid regions (Africa, South America, AUS) (l. 680-690) due to high interannual variability of green vegetation cover. As I read the paper by Winkler et al., such findings underlie their conclusions (e.g., "We find that CO2 fertilization is an important driver of greening in some biomes, but not dominant globally as suggested previously", l. 126). I would argue that the findings by Winkler et al., do not provide new insights that allow for a revision of findings by earlier studies (e.g., Zhu et al., 2016), but rather fail to identify drivers (including CO2 effects) due to their application of an inappropriate attribution method. In most other biomes, attributions made here are largely identical with attributions made by Zhu et al., 2016 and also summarised by Piao et al., 2019.

I regret that I cannot offer a more positive assessment of this manuscript. However, my review should not discourage authors to use their results for a revised manuscript, where more attention is paid to assessing robustness of greening/browning signals in the context of multiple satellite products, and where caution is applied when reaching conclusions based on absence of evidence following the attribution method ("Causal Counterfactual Theory") applied here, and claiming evidence for an overestimated CO2 effect in the
current generation of terrestrial biosphere models.

[Text in italic was not in my previous review report.]