We would like to express our gratitude to Anonymous Reviewer 2 for the considerate revision of the paper and the numerous helpful comments.

Below are detailed replies to all the issues raised in the comment.

A revised version of the manuscript has been prepared based on the reviews.

Reviewer 2 wrote:

The paper by Miesner et al presents a database of forest surveys distributed across a large area characterized by ecoclimatic gradients in northeastern Siberia. This geographic region has important influences on global carbon and energy dynamics due to its large area and sensitivity to climate change. Despite this importance there is a relative paucity of freely available and easily accessible data that can be used to inform observational and modeling studies. From this perspective, this manuscript describes an important dataset that is worthy of publication and dissemination. Some revision is required before the paper and data can be considered further for publication.

The overall structure of the manuscript is appropriate. The data are described reasonably well, and the comparison with various gridded data products is useful for understanding the utility and limitations of the data set. There are a number of areas where additional detail and/or discussion are warranted. Some of these are noted in my specific comments below, but in general the discussion seems a bit superficial. In particular, it would be useful to have a deeper discussion of errors associated with the use of height as the primary unifying measurement, as well as the visual estimates of height. To my mind, DBH is a more common and useful metric than height, and seems an easier measure than the several required to triangulate height using a clinometer. Related, it seems that there is a high potential for error, that is hard to quantify, associated with the visual estimates of height. More critical discussion here would be nice.

Our response:

We agree that DBH is a more common leading variable than height, and it is often a more robust predictor. As mentioned in the Methodology section, our experience with sporadic control measurements showed that the error of the method is small, although we never quantified it in a systematic way. But when the error propagates, it may still be non-negligible. A reason for choosing height as main variable was that the survey protocol was developed in the tundra-taiga ecotone, where trees are often small in height and have low branches, so that height is not only easier to measure, but also more meaningful. In the more densely forested taiga, this protocol was kept, so that the data would be consistent. We elaborated on this in the Discussion.

Reviewer 2 wrote:

Regarding the data, I found the files a bit unwieldy to work with. There is a lot of awkwardly structured metadata at the top of each file, before the actual data, making it difficult to read the files into a program like R. There also seems to be some redundant data here, in terms of site names, campaign, PI, etc. It may be more appropriate to have separate metadata files in order to make the data more user friendly/analysis ready.

Our response:

The data formatting is part of the PANGAEA data publication and cannot be changed by the authors. There are several ways for dealing with this structure: When ignoring everything surrounded by the signs "/* … */", the table becomes more readable.

Alternatively, the data can be downloaded with "pangear" (an R client for the PANGAEA database (https://github.com/ropensci/pangaear)) or pangaeapy (a python client for the PANGAEA database (https://github.com/pangaea-data-publisher/pangaeapy)).

Reviewer 2 wrote:

These are primarily suggestions - having the data described and available is important, and this paper accomplishes that. The edits/revisions I suggest would improve the utility of the data set.

Below are a number of specific/minor comments, more editorial in nature.

L120: Perhaps Gridded Data Products would be a more appropriate term here. The CHELSA data is downscaled reanalysis/climate data, not a remote sensing data set.

Our response:

We changed it to just "*Data products*" in the heading, and specified in the following sentence:

...we used several gridded, mostly remote sensing derived data products ...

Reviewer 2 wrote:

L161: What variable is suitable for comparison with biomass? A little more detail/information here would be nice.

Our response:

Explanation added :

... like stem volume

Reviewer 2 wrote:

L242-3: Are the field data consistent with this? Are vegetation conditions consistent with recent disturbance?

Our response:

For both sources, in 50% of the cases where disturbance was detected, the observed vegetation showed clear signs of recent disturbance, while for the rest of the allegedly disturbed sites, it was not clear. We added:

We encountered clear signs of recent disturbance in the vegetation at only 50% of the sites where either of the data products detected forest loss.

Reviewer 2 wrote:

L246: Areas with tree loss in the Hansen data set hold more standing dead than those without? Please clarify.

Our response:

See response below

Reviewer 2 wrote:

L247: Plots indicated as having forest loss do not have any disturbed trees? Please clarify.

Our response:

Regarding both above comments, we re-formulated the entire sentence to:

The average quotient of basal area of living trees to overall basal area is higher for the sites without disturbance than for the sites with forest loss according to the Hansen et al. (2013) data set, which shows that there is more standing deadwood on sites with forest loss.

Hopefully this makes it more understandable.

Reviewer 2 wrote:

L256: It would be good to discuss in a bit more depth the error implications of visual height estimates. Also, since DBH is a common measure that is often allometrically related to height, biomass, LAI and other ecologically important processes it would be good to discuss the tradeoffs associated with using height instead.

Our response:

We added some sentences stating that we are aware that DBH is the more common measure, but we chose height because it is easier, and to keep the protocol consistent. We are aware that it generates an error which then propagates into other variables, but we did not quantify this error. The paragraph now reads:

The field work was carried out according to scientific standards. Tree height was chosen as the leading variable because it is easy to overview in sparse stands and it generally correlates well with other variables (stem diameter, biomass). Diameter at breast height (DBH), even though it is more commonly used as a predictor, is more laborious to determine for trees in sparse stands with low crowns. With frequent clinometer measurements, we assured precise height estimations, and the remaining errors can be expected to average out over the high number of observations, which were easily obtained due to the efficiency of the method. Drawbacks coming with this method are: Since the diameter is only predicted from height, errors from this prediction propagate into derived variables like basal area and stem volume. And the initial measurement error, even if small, propagates along the same way. This error was not quantified systematically.

Reviewer 2 wrote:

L261-4: These sentences are almost too vague to be helpful. What does it mean that the plots are not weighed accordingly? I'm not sure what the last sentence is supposed to mean.

Our response:

"*Weighed accordingly"* meant "*according to the occurrence of the vegetation type they represent*". We changed it in the text. What we meant to clarify with these sentences is: If we find e.g. a negative correlation between January temperature and basal area, this may be systematic throughout Eastern Siberia, or it may be due to our choice of plots.

Reviewer 2 wrote:

L270: Are there any patterns here, geographic or otherwise?

Our response:

Thank you for raising this very good question.

A quick glance into the data suggests that the sites towards the south are a bit more irregular in their height distributions then the more northerly sites. And indeed, the Gini coefficient for height is negatively correlated with latitude (compare Fig.5), albeit only slightly significantly and with low explanatory value (R^2 = 0,033).

We cannot tell, though, if this is due to the generally larger height range on the southern sites, or if it is indeed caused by the disturbance regime, which is known to be characterized by shorter fire return intervals in the south. Still, we added this sentence to the results section:

The Gini coefficients are negatively correlated with the geographic latitude of the plot (Figure 5), but significance and explanatory value of the linear correlation are not high (p-Value 0.021, $R^2 = 0.33$).

Reviewer 2 wrote:

L271: Are sites recently affected by fire indicated in the database?

Our response:

Unfortunately, no, as this was not thought of at the time of the data publication.

Reviewer 2 wrote:

L288-9: What about variables produced at 30m resolution? How to explain the mismatch for these?

Our response:

Even if a 30m survey plot may not lie exactly inside the pixel, we aimed to take a representative vegetation type for a larger area, so that a scale mismatch should not be the problem here. However, we mentioned other reasons for the mismatch in the following sentences.

Reviewer 2 wrote:

L311: Extend should be extent

Our response:

Corrected accordingly.

Reviewer 2 wrote:

L325: See also papers by Kropp et al and Walker et al for evidence of drought stress in Siberian larch.

Our response:

Thank you for these suggestions. Two of those were added and referred to in an additional sentence, and another one in the sentence before:

...growth has been diminished by drought stress and extreme events, which are increasing under climate warming, like the 2020 Siberian heat wave (Collow et al. 2022). Kropp et al. 2017 and Walker et al. 2021 support that water availability is a limiting factor for Larix cajanderi.

Collow, A. B. M., Thomas, N. P., Bosilovich, M. G., Lim, Y.-K., Schubert, S. D., and Koster, R. D.: Seasonal Variability in the Mechanisms Behind the 2020 Siberian Heatwaves, Journal of Climate 35.10 (2022): 3075-3090. https://doi.org/10.1175/JCLI-D-21-0432.1

Reviewer 2 wrote:

L349: I didn't think the WorldClim data set was used in this study, please correct/clarify.

Our response:

Thank you for spotting this. It should have been CHELSA; WorldClim was used in an earlier version. Corrected accordingly.

Reviewer 2 wrote:

L355: I'm not sure this conclusion warrants a stand alone paragraph.

Our response:

It is true that the paragraph does not convey any new information, but we liked it as a concise closing statement. It is therefore merged into the preceding paragraph.

Reviewer 2 wrote:

Figure 2 - panels should be labeled a, b, c, etc. and referred to as such in the text.

Our response:

Suggestion adopted.

Reviewer 2 wrote:

Figure 4 - it would be helpful if the figure capture indicated that these specific plots were selected to show examples of different size class distributions.

Our response:

We added to the caption:

...which were chosen as examples for differing height distributions.

Reviewer 2 wrote:

Figure 7 - are all of these results significant, and if so to what level?

Our response:

All correlations except the top left (mean tree height ~ T01) are significant, most are striongly significant. We added p-values below the R^2-values for each graph.

Reviewer 2 wrote:

Kropp, H., Loranty, M., Alexander, H. D., Berner, L. T., Natali, S. M., & Spawn, S. A. (2017). Environmental constraints on transpiration and stomatal conductance in a Siberian Arctic boreal forest. *Journal of Geophysical Research: Biogeosciences*, *122*(3), 487–497. https://doi.org/10.1002/2016JG003709

Kropp, H., Loranty, M. M., Natali, S. M., Kholodov, A. L., Alexander, H. D., Zimov, N. S., Mack, M. C., & Spawn, S. A. (2019). Tree density influences ecohydrological drivers of plant–water relations in a larch boreal forest in Siberia. *Ecohydrology*, *12*(7), e2132. https://doi.org/10.1002/eco.2132

Walker, X., Alexander, H. D., Berner, L., Boyd, M. A., Loranty, M. M., Natali, S., & Mack, M. C. (2021). Positive response of tree productivity to warming is reversed by increased tree density at the Arctic tundra-taiga ecotone. *Canadian Journal of Forest Research*, cjfr-2020-0466. https://doi.org/10.1139/cjfr-2020-0466