

Drought responses of a Norway spruce forest on drained peat soil: combining sap-flow sensors, eddy-covariance, soil and UAV data

Pavel Konstantinovich Alekseychik, Mikko Peltoniemi, Raisa Mäkipää, Ville Tuominen, Tuomas Laurila, Hamlyn Jones, Mitro Müller, Evgeny Lopatin, Helena Rautakoski, Timo Vesala, and Samuli Launiainen

Response to Review No. 1

We sincerely thank the Reviewer for the valuable insights and constructive recommendations on how to improve our work. Our responses will be highlighted with blue.

General comments

Overall, the paper is suitable for publication in Biogeosciences, as it provides an important overview of the efficacy of various methodological approaches for investigating the responses of boreal forests to drought. However, given the many concepts and datasets presented in this paper, the overall structure needs improvement to provide more clarity for the reader. Some of the ideas presented are not well connected, or require clarification to contextualize them within the paper's premise and the scale of the results. Moreover, because the figures can be very data-heavy, captions and/or legends need to be improved so they can stand alone without requiring embedded supporting text in the manuscript. Nonetheless, the conclusions are clearly drawn and can support further research on the usability of the proposed methods to understand the onset and effects of drought on boreal forest productivity. - We are very pleased to receive such feedback on our work and once again thank the Reviewer for the positive and thorough assessment of this manuscript. We will clarify the text and improve the structure. Figure legends will be improved to make them stand alone.

Specific comments

The introduction needs some light reorganization to improve the flow of ideas. For example, in lines 51 to 56, the introduction of the Standardized Precipitation Evaporation Index (SPEI) feels disconnected from the ideas presented in the previous paragraph. I understand that this index is used later in the paper to explore regional drought characterization, but it is not well integrated here, nor does it support a strong segue to the next paragraph. A suggestion is to combine the paragraphs from lines 45 to 56 into one and make the connection between Norway spruce responses to VPD clearer in preparation for the next paragraph. – These paragraphs will be reworked to improve the flow, as advised.

In the introduction, the paper's main goal is first presented: to investigate the drought responses of two adjacent stands in a boreal forest in Finland (lines 98 and 99). However, in line 104, the statement about the difficulty of detecting drought responses in boreal forests serves as the main premise. Given the importance of this premise in the introduction and abstract, I suggest reframing the goals to clarify the paper's main question. If the paper aims to identify the utility of different methodologies to characterize drought and then explain how two adjacent stands behaved differently under drought, a reframing could benefit the reader. – [The aims and hypotheses will be reformulated. We understand that, in its present form, this part does not achieve the desired effect.](#)

In the methodology section, specifically in “meteorological and soil environmental conditions” (Section 2.2.2), there is mention of datasets that were collected but not incorporated into the discussion. I suggest removing mention of any datasets that are not presented in the results and discussion. – [Will be done](#)

I suggest renaming section 2.2.3, “meteorological drought assessment”, to better represent both the atmospheric and soil drought definitions. Moreover, a clearer distinction between these two definitions is needed, as their use in the results and discussion sections was often confusing (see other comments below). Moreover, this section does not seem to belong in the “field setup” description and would be more appropriate in another section describing derived indices and drought characterization. The same would be more appropriate for Section 2.2.4 on drought proxies. In this section, acronyms need to be more clearly defined and used consistently throughout the paper, and providing more context on why these proxies were chosen would probably help the reader. – [Thank you for this recommendation, we do believe it would indeed improve the structure. The sections 2.2.3 – 2.2.7 will be unified under a new subsection, “2.3 Proxies for drought assessment”. The acronyms definitions will be revised.](#)

The novel approach to measuring tree stomatal conductance (section 2.2.7) also does not appear to be well suited to the “field setup” section. Moreover, this seems like a pivotal contribution of this paper to the literature. It may be best to include it in another section of the methods for physiological stress characterization, for example, or within section 2.2.6 (e.g., 2.2.6.1). – [We agree, and this will be done, as per the previous comment,](#)

In the results section, I am unclear what is the goal of the monthly SPEI distribution for the entire country (Figure 2c). It is tied to national rainfall patterns, but this is not investigated enough to provide sufficient contextualization for the study site-specific results. I suggest this be reframed for greater clarity or removed entirely. Moreover, Figure 2b needs more clarity: the soil drought points correspond to both atmospheric and soil drought conditions (line 395). The figure should be comprehensible on its own, and the current caption and legend are not sufficiently self-explanatory. – [The Fig. 2b legend will be improved. The geographical distribution of SPEI over Finland is intended to allow the reader to evaluate how strong the drought conditions were at the](#)

measurement site (marked with the red dot) compared with the rest of the country, and we consider this a useful insight. This will be better explained in the main text and caption.

Please include a legend for Figure 2a. I assumed that the colors and line patterns corresponded to different blocks and years, as in other figures, but that should be clearer here. – Will be added

Lastly, I understand that Figure 2b is representative of the harvest block, likely due to the placement of the eddy-covariance tower. That is not well described in section 2.2.2, nor in the caption or the text. Make this clearer to the reader in one of these places. – Thank you for this note, we will clarify that.

Results for the Bowen Ratio in lines 425-428 are unclear because there is no frame of reference for the Bowen Ratio in the methods section. Please include contextualization for what it means when the Bowen ratio increases or decreases. Contextualization may also be needed for the Canopy Water Stress Index, since the averaged CWSI during drought periods is not presented relative to other values. – Thank you, we will rethink the presentation of the Bowen ratio and the CWSI, and add the missing interpretations.

In Figure 6a-6d, it is unclear why and how the VPD values were binned. What was the criterion for binning the values? Could a line have displayed that with either shaded standard deviations or a shaded confidence interval? Moreover, the legend is small and hard to read. Please increase the point size of the legend. – the binning was done for every 0.25 kPa step in VPD – This will be clarified. The font sizes will be increased.

Please consider reorganizing section 3.2 for greater clarity. This is a very data-heavy section and can be confusing for the reader. It might be beneficial to discuss ecosystem-level data first, as currently presented in Figures 4 and 6, and then use the block distinctions in Figure 6 to discuss differences between blocks, including Figure 5. The main issue identified here is the discussion of NEE that opens and closes this section, which leads to some repetition; perhaps bringing the two together could facilitate the reader's comprehension of the data. Moreover, a reorganization may help with the weather contextualization of this section (e.g., drought is often associated with more sunlight, weather conditions for EC measurements, etc.). – Thank you for the recommendation. Section 3.2 will be rearranged as proposed.

Throughout the results section, please consider being clearer about the atmospheric and soil drought. Currently, the term “drought” is used in both contexts with much explanation, which has hindered easier comprehension of the results. Perhaps some context can be introduced in the relevant methods section and reinforced here. – Thanks for important observation. The current ambiguity is for two primary reasons: First, it arises from our approach, where we attempt to address both types of drought and have not succeeded to be specific enough when atmospheric vs. soil drought are discussed. This will be revised in the next version of the manuscript for better clarity. Second, the concepts of atmospheric, soil and ecological drought are still not mature and

clear overall. Currently, the research community has no effective or unified way to describe drought. We will highlight this problem further and clarify the definitions as much as possible.

Two previous studies discussing the complexity of drought definitions will be cited, Crausbay et al. (2017), “Defining Ecological Drought for the Twenty-First Century”, and Wilhite et al. (1985), “The Drought Phenomenon: The Role of Definitions.”

In Figure 7, similar to Figure 6a-6d, it is unclear how the VPD values were binned. Please provide context for the specific criteria used to bin these values. – It was done by dividing the VPD range of [0...3] into eight intervals. This will be clarified and explained in the text.

In lines 511-512, it is noted that the harvest-block trees show higher sap flow than the control trees. Is that the result on average? Is it per range of tree diameter? How are these values contextualized? – This part refers to Figure 7, where these relations may be observed visually. What is meant is the mean ratio of sapflow at $VPD > 1$ for all trees in Harvest and Control. However, the text will benefit from a numerical expression of these relations, which we will provide.

In the UAV survey results (section 3.4), inconsistent use of abbreviations and symbols may hinder the reader’s ability to understand the results. Please consider using abbreviations and symbols more consistently, such as when discussing thermal stomatal resistance and stomatal conductance. Thank you for pointing this out - the quantity names will be harmonized, e.g. g_{sT}

The paragraph in lines 585-595 compares the median of the distributions from the tree stomatal conduction to the eddy covariance values, which is not immediately clear from the figure. An indication of the tree medians could be included to facilitate comparison with the EC lines. Moreover, rewording this paragraph may benefit the reader; lines 593-595 were confusing. – These sentences will be reformulated for clarity. This is correct: the paragraph talks about the median values and widths of the plant index distributions.

In Figure 12, tree height and NDVI values were binned. As with previous figures showing binned VPD values, it is unclear here why or how these variables were binned. Please provide some context. – Will be done. As before, the binning was done by dividing the x-axis range into equal intervals.

Again, in the discussion section, there’s confusion about which type of drought the argument concerns (atmospheric or soil). In the first paragraph of this section, for example, it is not immediately clear that changes in vegetation functioning with drought refer to the atmospheric drought. Improving clarity would greatly benefit the dissemination of this paper. – As we indicate in a response to another comment above, this lack of clarity in drought definitions is in fact one of the foci and premises for this work. We will try to present this conundrum in a clearer way.

The ideas in the paragraph on lines 654-657 should be distributed across the previous paragraph and the following paragraphs for greater clarity. – Will be done

In lines 661-665, it is unclear why bring up a previous study finding greater reductions in boreal forest productivity under low soil moisture contents if the last sentence is about the effects of atmospheric drought in the absence of soil drought. – Apologies for poor wording in this paragraph – the mentioned studies explored the conditions of drought, when high VPD were most often found to coincide with low soil water. These studies did not intend to explore the individual effects of air or soil drought, to my best understanding.

In the results section, it states that soil moisture was measured at 20 cm in addition to 5 cm. Why is that not shown in the paper, and why was it not used to explore additional stress from soil drought further? Moreover, how might the water table depth mediate the soil response to drought stress? How can the decoupling between soil moisture and water table depth help explain the effects of soil drought on boreal forest productivity? – The mention of the -20 cm measurement level will be removed, as it indeed is not used in the present study.

The -5 cm measurement level was chosen in order to highlight the short-term dynamics in topsoil moisture content. This level also has a closer link with the soil surface condition, which in turn directly affects the ground surface conductance and thus the site-average g_s , Bowen ratio and other drought proxies. Conversely, the -20 cm moisture time series are much smoother, concealing the short-term weather change effects. The trees in the site have shallow root systems so they too can be expected to be affected by topsoil (-5cm) moisture.

The discussion section follows a strong structure, from ecosystem-level responses to drought and block effects (thinning) to individual tree responses. It ends with a comparison of the EC data to the UAV approach, which is incredibly valuable to the community. - Thank you very much for this positive assessment!

Technical corrections

Line 234: include BR abbreviation in parentheses or brackets (consistent with other abbreviations). – Will be done

Line 239: include symbol and units for aerodynamic resistance to water vapour transport in parentheses or brackets (consistent with other symbols). – Will be done

Line 253: include symbol and units for surface conductance in parentheses or brackets (consistent with other symbols). – Will be done

Line 254: include Canopy Water Stress Index abbreviation in parentheses or brackets (consistent with other abbreviations). – Will be done

Line 263: include WUE abbreviation in parentheses or brackets (consistent with other abbreviations). – Will be done

Line 268: include LUE abbreviation in parentheses or brackets (consistent with other abbreviations). – Will be done

Line 435: Figure 4 needs a title before the description of the specific panels. – Will be done

Line 467: add a comma after PAR for greater clarity. – Will be done

Line 474: unsure whether there's a typo or if the sentence needs more clarification. Unsure what "at no drought on the c shows" means. – Will be clarified

Line 481: Figure 6 needs a title before the description of the specific panels. – Will be added

Line 529: Figure 9 needs a title before the description of the specific panels. – Will be added

Line 578: Figure 10 needs a title before the description of the specific panels. – Will be added

Line 669: correct Mirabel et al. (2023) citation. – Will be done

Line 709: Is Figure 12 the correct figure here? Shouldn't it be figure 9? – Sorry, you are correct, it should be Figure 9.

Response to Review No. 2

This study presents a commendable investigation of forest drought responses, integrating eddy covariance, sap flow, UAV remote sensing, and soil measurements across control and harvested blocks. The breadth of measurements spanning different scales and levels of plant functioning is a major strength of this work. The author presents results clearly for the most part, discusses limitations in considerable detail, and applies sound analytical techniques throughout. - We sincerely thank the Reviewer for all the motivating and constructive comments! Our responses will be given in blue text.

However, two issues require attention before the manuscript is suitable for publication:

1. Certain statements and conclusions lack specificity or contain internal inconsistencies that should be resolved. Key findings, particularly the discrepancy between sap flow and canopy-level indicators, should be more prominently communicated, including in the abstract. – We will address this need and sharpen these points in the manuscript.
2. Consistency in color schemes and legend readability should be improved across several figures. – Will be done according to the recommendations of the Reviewer.

I recommend this manuscript for publication, subject to the revisions detailed in my specific comments. Please note that the detailed comments below do not follow the order of major and minor revisions but instead follow the sequential structure of the paper for ease of reference.

Abstract:

The major conclusion of the paper that the sap-flow is more indicative of the drought at the stand level is somehow missing in the abstract section of the paper.

Also, the abstract could be streamlined to clearly convey the research's output.

For instance:

"Sap flow data showed that the CCF-harvested block shows greater tree resilience to high vapor pressure deficit (VPD) than the control block, likely due to the higher soil water availability during the rainless and hot period." - Thank you for this important remark. We will edit the relevant parts of the manuscript in order to highlight this idea.

Sap flow measurements revealed contrasting drought responses between the two blocks. While normalized sap flow in the control block was insensitive to soil drought conditions, the harvested block showed elevated sap flow during drought relative to non-drought periods (Figure 7). This divergence intensified during the hot, rainless June–July 2021 period, alongside rising VPD (Figure 8), though whether it reflects genuine drought resilience or

simply reduced competition for water among fewer trees in the harvested block remains unresolved. – We agree that these two concepts highlighted by the Reviewer form part of the complexity of the processes under scrutiny. The evidence presented in this work could be interpreted as the fusion of both: fewer trees means more water availability, which means higher ecosystem- or stand-level resilience. Whether it involves any physiological mechanisms enhancing the drought resilience on an individual tree scale could not be verified with the available monitoring data.

Especially when UAV retrievals show a different story: stomatal conductance derived from canopy temperature was lower in the harvested block than in the control (Figure 10). While the author thoroughly discusses this and other measurement limitations in the discussion section, which I appreciate, this particular statement, in my view, remains unfounded in the abstract section. – We will try to add this point to Abstract, as much as the word limit will allow it. The Reviewer is absolutely correct that this conflict between the ecosystem monitoring products is one of the interesting observations. Unfortunately, we the potential reasons for this discrepancy remain speculative and warrant future studies.

Introduction:

There is an internal inconsistency between lines 85 and 101. At line 85, the author states that thinned stands are generally expected to have less inter-tree competition for soil water, implying greater water availability per tree and, therefore, lower drought stress in the harvested block. However, at line 101, the author hypothesizes that trees in the harvested stand experienced greater drought stress than those in the control. In my view, these two statements are contradictory. Though this is just a hypothesis, it leaves the reader confused following the introduction. Do you also think so? – Indeed, there was a contradiction between these portions of text, thank you for pointing it out. on line 83 we say “It is generally expected that thinned stands have less inter-tree competition for soil water, particularly during a drought in both Norway spruce and Scots pine (Laurent et al. 2003; Gebhardt et al. 2014, Friedrichs et al. 2009, Martín-Benito et al. 2010, del Río et al. 2017)”. However, in the hypothesis on line 101 “We hypothesized that the trees in the selection harvested stand were more stressed than the trees in the unharvest control to the drought in 2021. ”We will revise the text, if necessary, so that no apparent unintended conflicts remain. Our hypothesis would naturally go along the lines of what had been observed previously.

Methods:

Since vegetation height, roughness length, and displacement height are required inputs for deriving turbulent fluxes from eddy covariance measurements, and given that the stands exhibit considerable heterogeneity in vegetation height, it is unclear which height value was used in these calculations. – The measurement height of 29 m and displacement height estimate of 10 m were used in the calculations (the canopy height is about 10 m, and in the Harvest block it is

rather sparse). These parameters did not vary with wind direction. The displacement height could be somewhat higher in the Control block, but the difference with Harvest block is difficult to estimate. The differences in turbulent field between the two blocks were described by the roughness length, whose measured 30-min (i.e. wind direction-dependent) values were fed to the footprint algorithm.

For consistency, does the SpaFhy model use the Penman-Monteith equation to calculate ET for SPEI derivation, given that CWSI is also computed within the Penman-Monteith framework? While I do not expect this to change the interpretation, it is worth clarifying that the point-based model often uses a gradient-based resistance scheme to compute ET. – SpaFHy model uses two-layer approach to predict daily ecosystem ET. Penman-Monteith equation is used separately for canopy transpiration and evaporation of intercepted rainfall from the canopy. The respective aerodynamic and surface conductance depends on meteorological conditions, as well as LAI, plant water use traits and seasonal phenology. Evaporation from forest floor follows radiation- and temperature-driven Priestley-Taylor equation, where available energy depends on LAI of overstory canopy and ET is reduced linearly in case the organic top soil layer, resembling moss and humus layer, moisture content is below a threshold value. Forest floor ET is set to zero when soil is snow-covered. For details, see Supplementary material in Launiainen et al., 2019.

The author selects 1 kPa VPD and 0.2 m³/m³ soil moisture as thresholds for atmospheric and soil drought, respectively, based on earlier research. While not mandatory, these boundaries could be further corroborated from a plant hydraulics perspective. For instance, by relating them to known water potential thresholds associated with stomatal closure or hydraulic failure in the studied species. This would provide a more physiologically grounded justification, if the author chose to do so. – We will try to present further corroboration of these first-guess limits. These thresholds are indeed intended to mark the initial stages of reaction, which most importantly consists of stomatal closure.

In Equation 6, substituting Equation 7 appears to yield a zero in the numerator, which would render the CWSI expression trivial. The correct form should likely be $(y^* - y) / (\tau + y^*)$, where the numerator captures the departure from the non-water-stressed baseline. Please check. It is unclear whether g_s was calculated at 15-minute intervals and subsequently averaged to a daily value, or whether daily mean VPD was used directly to compute a single daily g_s estimate. – Apologies for the typo; the correct expression is $\gamma^* = \gamma(1 + r_{sp} / r_a)$, where r_{sp} is the canopy resistance at potential evapotranspiration. However, since the r_{sp} term is typically not known precisely, it is assumed that $r_{sp} = 0$, which leads to $\gamma^* = \gamma$. Therefore, the equation becomes:

$$CWSI = \frac{\gamma(1 + r_{s,EC} r_a^{-1}) - \gamma}{\zeta + \gamma(1 + r_{s,EC} r_a^{-1})}$$

This is the actual form of the equation that was used in the present version of the manuscript.

Results

1. In Figure 3, the first row and others present variables aggregated at the ecosystem level, while the second row shows per-block comparisons. Rearranging the panels so that all ecosystem-level variables appear at the top and per-block comparisons at the bottom would improve readability. – Will be done
2. At line 451, the authors state that correlations with VPD were more pronounced in 2020 than in 2021, attributing this to the higher tree density before harvest and the stronger 2021 drought. While this interpretation is plausible, it is not directly evident from Figure 4 as currently presented. – It is better to reformulate this as “High VPD periods caused clearer reductions in transpiration-related drought proxies in 2020 than 2021”, as this is what we intended to suggest.
3. Lines 476–479 are unclear. In particular, the statement that higher sunlight during drought compensates for lower P_{max} , leading to similar integrated GPP, is difficult to follow. Since P_{max} is derived from the GPP–PAR relationship using a constant k over the growing season, any change in P_{max} is inherently a consequence of changes in GPP and/or PAR themselves. The authors should clarify the causal logic here, as the current phrasing risks circular reasoning. – This is about the integral of the GPP vs. PAR curve. A “lower” daily curve, i.e. one with a smaller P_{max} , might have the same integral as a higher curve, if the ecosystem experienced higher PAR values on that day. Adding a schematic in the Appendices might aid the understanding of this idea.
4. Lines 198–199, section 2.2.2, state that humidity, air temperature, and PAR were measured at nearby meteorological stations. It is therefore unclear how these variables are plotted separately per block, unless separate sensors were deployed in each block. The authors should clarify this in Figures 6 (e,f,i). – These measurements were located just a few meters distance from the EC tower, so, on the border between the Harvest and Control blocks.
5. Figure 6: Legend are too small to understand in the top row. – Will be edited
6. In my view, while discussing Figure 6, we could only agree that at very high VPD ($> 2\text{kPa}$), NEE appears to differ from low VPD conditions in both the control and harvested blocks. And NEE is the measured signal. However, the ecosystem respiration (R_e) estimate used here is solely temperature-dependent, omitting potentially important drivers such as soil moisture. This may introduce unaccounted-for uncertainty into R_e and, consequently, into the partitioned GPP. Furthermore, the partitioning approach presented here differs from the standard Reichstein (2003) method widely used in the flux community. Given these limitations, the authors should exercise caution in drawing strong conclusions from GPP and P_{max} , as these are derived quantities sensitive to the

choice of partitioning algorithm rather than directly measured fluxes. I thus encourage the author to moderate the conclusions based on GPP and Pmax throughout the paper, while still presenting them. For instance:

In line 677,

While the drought thus suppressed photosynthetic efficiency in 2021, as was reliably shown by CO₂ exchange in the Control, the mounting soil dryness and prolonged VPD stress did not seem to have had a similar effect in June-July 2021. Both Pmax and GPP increased throughout the 2021 drought and during the hotter period of 2020 (Figs. 4-6). - Absolutely, NEE is the directly measured quantity, and should be given the primary attention in the discussions of Fig. 6 and elsewhere. We agree that partitioning into GPP and Re involves aforementioned uncertainties, and will modify the discussion of these quantities as advised.

7. Line 504-507: Nice paragraph. I personally really liked Figures 5, 6 (top), 7, and 8 as their messages are more straightforward and succinct. – Thank you very much for the kind words!

8. In Figure 9, panel (e) is referenced in the caption but absent from the figure itself. – Sorry, that panel had been removed from the figure, but the caption was by mistake not updated.

9. In Figure 9 (c), since meteorological forcing is measured at a single tower, it is unclear how the separate Penman-Monteith estimates for the control and harvested blocks were derived. The forcing inputs should be identical for both, as they are measured by the same tower. The authors should clarify whether wind direction criteria, analogous to those used for EC flux attribution, were also applied to partition the meteorological forcing between blocks. If so, we might be looking at different times (which block is dominant) in Figure (c) and comparing them. Its same confusion as in point 4. – The potential evapotranspiration (PET_{PM}), representing here weather-depended reference ET assuming a uniform reference vegetation, is the same for both blocks, while the sap flow timeseries are coming from each block individually. This will be clarified.

10. In Figure 10, the colorbar for g_{sT} (panels c–d) is inverted relative to all other panels; higher values are shown in blue and lower values in red, while the reverse is true for the others. This creates confusion unless it's stated in the caption. – Thank you for pointing this out, it will be clarified in the text. Higher canopy temperature does imply lower transpiration and lower stomatal conductance. Since the reader already associates higher temperature and stress with red colours, the g_{sT} is coloured in the same manner to preserve the colour association, hence the inverted colourbar.

11. Furthermore, the figure 10 appears to show higher canopy temperature in the harvested block, alongside lower g_{sT} , which is physically consistent with the idea that higher

canopy temperature indicates greater water stress and reduced stomatal conductance. However, the accompanying text states that higher drought-period temperatures in the harvested block led to markedly higher g_{sT} estimates. (Line 565). This contradicts the figure in my view. Please clarify. – “Markedly higher r_{sT} estimates” were meant, but mistyped as g_{sT} . Thank you for this comment, it would be best to correct this as “markedly lower g_{sT} estimates”.

12. In Figure 11, for consistency, please use the same colors for the control and harvest blocks as in the rest of the paper (blue and red, respectively), rather than distinguishing them by line style here. This makes it difficult to follow. – will be done

Discussion

The discussion section is well-written. – Thank you very much for the positive evaluation! In section 4.2.2, the author could further elaborate on the temporal mismatch between UAV-derived stress indicators, particularly NDVI and temperature-based stomatal conductance. Water stress driven by VPD or soil moisture deficits manifests rapidly in stomatal conductance and is therefore detectable at the timescale of individual flight acquisitions. NDVI, however, reflects changes in canopy spectral properties tied to the structural and biochemical state of the photosystems, which respond to stress over days to weeks. Consequently, a single flight acquisition may capture stress responses in temperature-based conductance while missing the corresponding signal in NDVI entirely – We will develop this part of discussion as advised. This is an important note, thank you.

Response to Review No. 3

We thank the Reviewer for the time and effort invested in critical assessment of our work. The review makes a number of points that we find valid and useful and will gladly respond to them.

This study has used sapflow data, eddy covariance and UAV measurements to assess the response of a Norway spruce forest to drought over two years (2020 and 2021). In the meantime, the forest was affected by thinning. Throughout the paper it is not clear what exactly the research questions are (effect of drought on transpiration or carbon fluxes or both? Effect of thinning on carbon fluxes? How thinning changes response to drought? What is the question?). - Thanks for this important feedback on the ambiguity of the study aims. We believe the complexity follows from our broad goal to conduct a holistic analysis of potential drought responses on commercially managed drained peatland forest. To achieve that based on intensive multi-scale un-controlled field measurements, we adopted a combination of approaches to analyze the impacts of atmospheric (high VPD) and soil drought (limited water availability) on tree and ecosystem-scale water and carbon flux dynamics. Further, we tested different in-situ approaches to detect tree-scale drought proxies. This admittedly makes the paper heavier read, but was necessary in order to address each facet of the problem. While we agree further refinement of text is beneficial for overall clarity, the study aims are in our opinion already clearly stated in the lines 98-110:

“The objective of this study was to compare the drought responses of two adjacent stands, a recently selection-harvested and an untreated control stand, on drained peatland soil in southern Finland. The study was conducted during the non-drought year 2020, and during a long and intense drought period in 2021 likely representative of the future Boreal climate (Balting et al. 2021), which was also the year the site was selection harvested. We hypothesized that the trees in the selection harvested stand were more stressed than the trees in the unharvest control to the drought in 2021.

Overall challenge to detect (likely) non-linear responses to progressing drought in boreal forests serves as the primary premise for this work. Therefore, a diverse set of continuous in-situ monitoring data available at the Ränskälänkorpi peatland forest harvest experiment was used to examine the possible drought responses of trees and stands. The standard ecosystem carbon, water and energy balance measurements (eddy-covariance CO₂ and H₂O fluxes, surface energy balance, sap-flow sensors) are supplemented by campaign-wise thermal and multispectral UAV data. Based on these data, we conducted the analysis of how the drought progresses and trees respond to it, and compare and discuss the merits of different methods for observing drought stress of individual trees and tree stands.”

In fact, the Reviewer refers to the same lines in the comment on Line 107 below.

Because of the poor structure of the paper, it is not clear what new insights are gained here which we did not already have, based on considerably longer timeseries and a greater number of sites. - The reviewer does not give concrete hints as to what the issues with the structure are, so this is not very helpful. We found it most efficient to first represent each topic / use of dataset in Materials and Methods, presenting the combinations of different approaches in Results, and finally drawing a synthesis in Discussion. We will make our best to further streamline the text in the revised version of the manuscript. Our goal will be to improve the clarity overall, but specifically on the exact points identified by Reviewers. We are confident this will relax the overall concern of Reviewer 3. Please note that the two other reviewers did not find the paper particularly poorly organized or hard to follow.

The paper is written very poorly (results are provided often with a qualitative language) and basic statistical analysis is missing. - Extra care will be taken to ensure the correctness of language in the revised version.

Overall, we are against conducting statistical analysis just for the sake of it and using p-values as the god's word. The methods and representation has been selected according to their necessity and suitability and we consider them to be sufficient for the aims of the present work. The drought responses, where present, can be readily seen in the graphical form in the figures. , Also here, the lacking the specificity of the Reviewer's comments makes it difficult to provide a more detailed answer. We also note that other two reviewers did not find significant shortcomings in our methods, analysis or representation, except in terms of lack of clarity of some specific details (please see our responses to other reviewers).

Thus, what the authors present as "response" to drought might not hold at all. Also, I have reservations with how the authors defined drought (see details below). - The main takeaway of this work is perhaps that the responses of a conifer forest on drained peat soil are i) complex and ii) vary between trees and tree and ecosystem scale. Clearly, they also depend on the mode of recent forest management. Each type of data involved provides a unique insight into specific processes and scales involved. Thus, it is impossible to single out any single response or process of superior importance; in fact, there are many ways to observe and understand what the "drought response" is. However, the analysis framework is straightforward enough in our opinion – it consists of identifying soil and air drought based on in-situ soil and micrometeorological sensors, and tracking the potential tree responses in the flux and sap flow data, and looking for corresponding signals in the novel UAV data. For drought onset, we made reasonable first guesses regarding the point at which the drought responses could start to be observed - and these were indeed confirmed by the results.

Overall, this was a frustrating experience reviewing this paper. I do not believe that the study is mature enough for publication in Biogeosciences. I have listed my reasoning below.

Title: Rather than listing the methods, the title should indicate what “response” of the forest is the central point of this paper. - We regret that the reviewer has felt the paper that frustrating. Regards to title, it could be useful to modify the title as “Complex drought responses of a Norway spruce forest...” As suggested above, the observed responses are disparate enough to be collectively crammed into the title.

The Abstract is written very poorly. It is not clear what exactly the research questions are (not clear response to drought is assessed based on what function?) and it is a summary of methods and some results without an appropriate link made across the methods. - We will improve the abstract as suggested also by other reviewers, taking care of the word count limit. Especially the research question will be clarified as requested.

Only the results from each of the observation methods are reported. Judging by the Abstract alone, there is no novel finding or mechanistic understanding provided by the study. It is not clear to me what this paper adds to the existing body of literature.

We acknowledge the comment and will make our best to improve the Abstract in the next version. In short, the novelties are: - 1) The use of such a diverse multi-method and -scale dataset is a novelty in itself; 2) drought responses of conifer forests on drained peat soils are known poorly in general, yet these ecosystems are central for Northern European forest carbon balance and ecosystem services, particularly in Finland and European Russia where forestry drainage contains millions of hectares of boreal peatlands, 3) A site with two harvest treatments (thinned using selection harvesting and control) is seldom featured in literature.

Unfortunately, still by the end of the Introduction the objectives are not clear (is it management effect? Why is not the relevant literature assessed for this topic?) The Introduction is a collection of points without a clear link to form the overall objective of the study and the motivation of the study is lacking. - Please see the last lines of the Introduction as quoted above. We will try our best to account for this comment and improve the introduction.

Not enough details on EC flux calculation are given: e.g., what frequency are the high-resolution data, raw data filtering approaches and checks, coordinate rotation, time-lag detection etc. - We consider this beyond the scope of the present paper focusing on ecosystem behavior, not flux data processing. Common and established methods of the EC flux community are used throughout in measurements and data processing. We assure the Reviewer that the eddy-covariance measurements were carried out by a team of highly skilled specialists and the measurements were carried out to the highest current standards. Aubinet et al. (2012), already cited, summarize the current EC measurement standards and best practices, which we followed. Further details can be found in Laurila et al. (2021), which is already cited in the manuscript, but a reference will also be added in the paragraph on EC methods.

I understand that the study employs multiple different observation methods but rather than focusing on the method would be good to focus on the ecosystem properties these methods were used to assess. E.g., in Line 22-25: combining monitoring for what factor? Dataset of what? - The dataset in question is explained the Abstract, quoting Lines 24-25: “A dataset combining eddy-covariance (EC) fluxes, sap-flow sensors, UAV mapping, soil and weather data was collected”. We understand that the Reviewer would like to see the mention of the studied quantities, such as the turbulent CO₂ flux density and stomatal resistance. This will be clarified in the Abstract, as per the word count limits.

Is the research question about soil (highlighted in the abstract) effect on drought response? Or on management (thinning) effect on drought response? What is the question here? Not clear. The Introduction starts with the topic of management so maybe this is the key question? We don't know. - We refer here to our earlier response on study goals.

There is also a major issue with how “drought” in the soil is extracted. The authors have used absolute soil moisture content as a threshold for drought, and not even based on the local observations but based on another study! Soil water limitation depends on water potential, NOT water content and this link is highly dependent on soil texture. The studies that are used to determine the soil water content threshold for drought are on mineral soil and not relevant for this forest.

Using previous studies was necessary in order to come up with the first approximation of what the critical soil moisture level may be. Considering the porosity and shape of water retention curve in typical drained forest peatlands, it may be that our estimate of approximate threshold soil moisture is conservative. This will be checked and use of threshold moisture value is better motivated in the revised version.

Results are presented with qualitative language: ample, remarkable. They don't provide sufficient comparison and a statistical test to show if the responses are at all significantly different is completely missing from the paper. - We do not observe a major problem with our use of the words “ample” (1 instance) or “remarkable” (1 instance). In both cases, the meaning of these words tallies with what is shown in the figures. Otherwise, we are not ready to impoverish the language of the manuscript but will attempt to be more quantitative (e.g. providing magnitudes of relative impacts) whenever possible. We agree that this will strengthen the text.

It is not quite clear what the Reviewer means by “show if the responses are at all significantly different”. If they mean the difference between the treatments, these are very apparent in the graphical representation of the CO₂ fluxes (e.g. Fig. 5), sap-flow (Fig. 8), and temperature-based stomatal conductance (Figs. 10-11). If the reviewer is pointing to whether e.g. specific fluxes or resource-use efficiencies within treatment are statistically different is another question. Given that we analyze time series from field experiment under progressing drought and varying

meteorological conditions, addressing statistical significance would be very contextual to selected sub-setting of data.

Bare minimum standard writing and analysis for a scientific paper is lacking. - We kindly disagree with that statement and consider it unfounded, as reviewer does not provide specific backing to this claim.

Specific comments:

Abstract should already mention the timeframe of the study. – The timeframe will be specified

Line 26. “It’s temporal and spatial features are identified „, temporal features in what? What is meant by spatial features of the forest stand? – Temporal and spatial features of spruce reaction to drought are meant.

Line 32: Are you talking about an increase in temperature and decline in NDVI due to drought, which was higher in CCF plots? Not clear how this comparison reflects the effect of drought on the two blocks. – Correct, this is what is meant here. Since there is no space to delve into this any deeper in the Abstract, the reader has to proceed to the main text.

Line 37-39: the two statements are not linked. - In order to establish a link, the second sentence is rephrased as: “The Boreal managed forests might benefit from the enhancement of tree growth in the region due to higher air temperature and CO₂ concentration which is forecast for the coming decades (Danneyrolles et al. 2023, Wang et al. 2023)...”

Line 41: What exactly is the example reference supporting? – The sentence will be rephrased as: “only if water availability is not a limiting factor in new conditions (e.g. d’Orangeville et al. 2018).”

Line 46: given that low resilience is confirmed for Norway spruce, what is your study adding? - Our study was certainly not designed to confirm the basic fact of Norway spruce susceptibility to drought, but to explore its extent and timing in an environment that has so far been underrepresented, and management practices (stand thinning) that are suggested to reduce drought stress, using the tools that are seldom combined. As explained later on in the Introduction.

Line 51: the topic of drought identification is not linked to the previous statement. – I do not see how the lines 45-56 could be lacking this link, I think it is sufficiently clear; Norway spruce drought resilience being a major source of uncertainty for boreal forestry is the point.

Line 55: “atmospheric drought”, not air drought - While “air drought” is a valid term which is very common in other literature, it perhaps makes sense to harmonize the text and only use the term “atmospheric drought”, as suggested.

Line 61 and 62 could have been merged – Will be done

Line 77: provides a buffer rather than protection – In a sense, the microclimatic buffering can be understood as protection, as the small trees are literally sheltered from direct sunlight by the dominant trees, and grow in lower VPD conditions than the dominant trees. This will be revised.

Line 76 so the non-dominant trees and understory are growing in the dark?! – “More dark microclimate” is what is meant here. Will be rephrased as “lower PAR environment”.

Line 85: and? What does this “less competition” mean for their response to drought? What is the link? – This will be clarified. Lower inter-tree competition is beneficial at the times of lower availability of soil water because, on the one hand, the trees do not collectively drain the soil water reserves, and, on the other hand, the dominant trees do not leave the rest without access to water.

Line 87: What is even the point of this sentence here that UAV and RS is helpful or research?? – We find it logical and according to basic principles of a research article to first state the overall success and facility of UAV remote sensing tools and then highlight the current limitations and knowledge gaps. The latter might indeed be insufficiently represented in this specific portion of text, and will be given more space.

Line 91: this is not true. There are numerous studies using UAV to detect drought effect on a range of structural and physiological functional traits in forests. – Numerous is maybe an overstatement, but there are some. What the sentence in question is lacking, though, is and addition in the end – “especially in the Boreal region”.

Also, is this the place to justify a methodology? I am struggling to find the research questions. Yes, there are some background information given here and there (that the resilience of Norway spruce is tested before, effect of tree response on stand response) but what is the exact question in this study that addressed all these functions in a manner that has not been tested before? - To our best knowledge, UAV-based remote sensing combining both thermal and RGB imaging has not been conducted in a Boreal forest drought studies, which justifies its application. The research questions follow just below and the UAV data are mentioned. We consider adding a specific mention that leaf colour (~NDVI) and temperature are the primary parameters (physiological and functional proxies) provided by the UAV platform.

Line 98: if the objective was to compare for the effect of management, this topic should be the focus and properly addressed and related literature reviewed for it, and provide the basis for the hypothesis that is put forward.

- While the harvest-control differences are a circumstance that adds an important aspect of tree density effects on drought response, this does not warrant an in-depth review of all aspects this problem. A brief overview will be added in the Introduction.

Line 104: what causes this difficulty? Reading this sentence, I would conclude that you have developed methods that specifically address the challenge of detecting drought response? (a remote sensing-focused analysis). How does using multiple methods that are each very demanding and not available easily, provide a solution to the challenges of studying boreal forests? So again: What is the research question exactly?

- On the methodological side, the research question was: What added value does the combination of several tools for ecosystem monitoring provide for drought response studies? This will be clarified in Introduction and Discussion.. Another research question, related to tree physiology, was about the potential difference in transpiration between the control and harvest blocks, as stated in the Lines 102-104.

- This should be rephrased as “difficulty to harmonize the drought indications across the multitude of existing methodologies”, as this is one important aspect. – Will be done

Line 107: OK so finally there is a mention what is studied here: carbon, water and energy? This should have been made clear in the Abstract. Sapflow is clear but UAV or EC don't indicate what entity has been measured and what is the focus of this paper. – I will try to fit this into the Abstract

Line 23: this is about density. What about the contribution to stand basal area by each species? If inventory data was collected DBH should be available for the trees. – This information is actually available in Martinez-Garcia (2026), “Short-term effects of harvesting alternatives on soil nitrous oxide fluxes in a boreal drained peatland forest”. This will be cited in the manuscript.

Figure 1: footprint area around the tower should have been shown to link the sapflow measurements with the fluxes – We acknowledge this and consider adding a cumulative EC flux contribution isoline (e.g. 60%)for clarity.

Line 150: Standard flux terminology is ignored. Calculating fluxes is processing, gapfilling and partitioning is referred to as post-processing. - We will consider the comment and ensure text follows established terminology as closely as possible.

Line 193: nearby means how many m away? Ground level means on the ground!? What variables were measured on the ground? – Immediately near the EC tower, not more than 10 m away.

Section 2.2.2 The location of the GW measurements and soil profile should also have been shown on the map. – Will be added

Line 220-226: why was soil drought defined based on thresholds of SWC from soil that is not even organic? This undermines the definition of soil drought selected for this analysis. – We do not see a reason why this cannot be done. There is an extensive body of literature discussing specifically the soil water content in peat.

Line 274: stem growth sensors – This subtitle lists the measured quantities and not sensors, therefore the replacement of “stem diameter change” with “stem growth sensors” would be inappropriate.

Line 284-282 what is the reference that validates this approach? – The task of sapflow upscaling is not novel per se, reviewed in Čermák et al. (2004) “Sap flow measurements with some thermodynamic methods, flow integration within trees and scaling up from sample trees to entire forest stands”. The previous studies have explored sap flow relationship with the tree structural parameters, including tree height (e.g. Zhang et al. (2018), “The tree height-related spatial variances of tree sap flux in subtropical forests”). In the current study, tree height was chosen as the explanatory variable for the following reasons. First, the number of sap flow sensors per treatment (7 and 8 in control and harvest, respectively) allows having only one explanatory variable. Second, it is known that tree-specific sap flow is strongly dependent on tree size/, in the general sense. Third, out of all tree structural parameters, tree height is most readily obtained from the UAV photogrammetric digital elevation model. Fourth, the tree height shows a strong relationship with the sap flow in the 15 trees that have sap flow sensors installed (appendix 4). Hence our use of tree height as the sap flow driver in the upscaling exercise.

Same as for equation 10. This is a rather standard expression for tree-scale stomatal conductance derived from sap flow, e.g. Oren et al. (1999) “Survey and synthesis of intra- and interspecific variation in stomatal sensitivity to vapour pressure deficit”, Köstner et al. (1992) “Transpiration and canopy conductance in coniferous forests”. The sap flow utilized therein is normalized by sapwood area of the measured tree.

Figure 2: a map without coordinates is just a shape figure. – These panels are intended to show the spatial distribution of SPEI over the territory of Finland. In that sense, it is a shape figure indeed, and it serves its purpose. Moreover, the size of figure would not allow the addition of coordinate labels in a fashion that is readable and non-obstructive for the main colour layer.

Line 387: What supports this statement? When can the reader get this information? – if the Reviewer means the WTD drop, it is presented in the Figure 3 just below. A reference to the relevant figure and panel will be added.

Line 417: How much is “ample”? – The exact amount will be specified. This means that it was the period of strongest precipitation of the year.

Figure 4: ET (or LE) should have been plotted too – Will be added

Line 455: seasonality in what? Light? Phenology? What is this statement describing and how is it connected to the previous statement – Will be replaced with “phenology”

Figure 5 The difference between the lines (mean for each category plus the uncertainty around them) should have been tested to see if the differences are significantly large enough – This will be done

Line 468: but we don’t know if it was significantly lower? - Will be specified, as suggested above

Line 476: what is meant by mild weather? – Means “non-drought”, this will be clarified

Figure 6: why is the daytime 30-min fluxes called flux density? – In micrometeorology, the quantity which is commonly referred to as “30-min average flux” is, in fact, termed “turbulent flux density”. We will consider reformulating the sentence.

Figure 6 lower panel: this is describing the drier conditions and should have been mentioned earlier – These panels can be moved to Figure 2, where they would be a better fit indeed.

Figure 7 which of these differences are significant and which not? – The significance of differences between the red and blue bin-averages (drought/non-drought) is readily shown by the error bars plotted on the bin-average markers. As one can see, these differences are typically insignificant at VPD < 1 kPa, and significant beyond 1 kPa, although the exact threshold varies between the trees.

Line 506: if it is usual, then references should have been given to support this. – Will be added

Figure 8 seems like some panels (b, c) are cropped at top (y-axis limit should be increased) – There is only one tree that shows a much higher sap flow than the rest, and thus the present y-limits were selected to best illustrate typical behavior. The figure label will be improved to clarify this.

How was resilience defined (the abstract mentions higher resilience of thinned stands to high VPD. Also how do the authors know that this range of VPD is “too high”? – The definition of

resilience will be added and related text reformulated (based on Holling 1973, “Resilience and Stability of Ecological Systems”) We are not completely sure what the Reviewer refers to in the second half of this comment, but generally VPD exceeding 1 kPa is known to initiate plant physiological responses such as stomatal closure, which also emerges from the present study.

The manuscript should have been checked for language (many grammar mistake that I won’t list here) – An extra language check will be made before the revised manuscript is submitted.