

Associate editor decision: Reconsider after major revisions

Public justification (visible to the public if the article is accepted and published):

Dear authors

Both referees recommended another round of major revisions, and based on their comments, I agree with their criteria. Specifically, referee 2, who also reviewed the previous version of your manuscript, has pointed out some methodological issues that I believe need resolving before we can recommend the acceptance of the manuscript. They have to do with inconsistencies between results and the text describing those results, and the propagation of uncertainty. I hope these comments help you improve your manuscript. Looking forward to receiving a revised version of the manuscript.

Response. We thank the editor and both referees for a constructive second round. Referee 2 raised two primary concerns: the main text overstated what the breakpoint simulations supported, and uncertainty was not propagated beyond breakpoint detection. We have addressed both concerns through new analyses rather than rewording alone.

The simulation results are now reported in the Results section, as requested by the referee, and the Monte Carlo framework in Appendix A now extends through segmented regressions and mixed models. The pooled estimates are reported in three new tables (S20, S21, and S22) and in the new Table A2 in Appendix A. Every slope contrast in the paper retains its sign and significance class after uncertainty propagation. One estimate, the shift in intercept between periods in Norway spruce, does not, and we have withdrawn its interpretation from both the Results and the Discussion.

We have also reviewed the Introduction, Methods, Results, Discussion and Conclusions for the specific inconsistencies identified by Referee 2 between the reported findings and the statistics in Tables 3 and 4, and we have corrected each one. Referee 3's comments on the resolution of the climate products, biotic disturbance, the rationale for selecting the three climate variables, and the definition of estimated marginal means have been addressed in Methods and Discussion.

Referee 2

I thank the authors for their careful consideration of my previous comments and detailed responses. Appendix A in particular is a careful, well-reasoned response to my initial major concern and a very nice addition to the manuscript. I commend the authors for these significant enhancements to their already valuable work. However, I still have two primary concerns along with several specific comments, noted below, before this work is ready for publication.

Response. We are grateful for the referee's continued engagement with this manuscript. We have addressed both primary concerns below with new analyses and have acted on every specific comment.

Primary concerns

1 Breakpoint detection Appendix A reveals substantially lower breakpoint detection rates than the main text suggests. I suggest to move key simulation percentages from Appendix A into sections 3.1 and 3.3, and revise your wording accordingly. Specifically, Appendix A reports:

Pine (mineral soils): only 9/1000 simulations detected significant breakpoints

Pine (organic soils): ~30% detection rate (vast majority returned no breakpoints)

Norway spruce (organic soils): ~45% detection rate (majority found no break-points at all)

Consequently, current main text claims need recalibration. Specifically, section 3.3 states breakpoints were “detected for all examined variables” in spruce, and section 3.1 claims “post-change trends remained robust,” but these statements don’t reflect the low detection frequencies reported in Appendix A. The 2009 pine signal appears driven almost entirely by site-level BAI (n=6 plots?) with minimal regional GI support.

Response. We agree, and the simulation results are now in the main text. Reporting them there turned out to sharpen the species contrast rather than weaken it.

The 1000 GI realisations clearly distinguish the two species. For Norway spruce, a regional breakpoint is a near-certain feature of the signal. It is detected in every realisation on mineral soils (median 2016.3, 95% range 2016.0 to 2016.7) and in 90.3% of realisations on organic soils. For Scots pine, no regional breakpoint is supported. A breakpoint is detected in only 0.9% of realisations on mineral soils and 32.0% on organic soils. In the latter case, the detected breakpoints cluster near 2019 rather than 2009. The probability that a breakpoint exists is therefore 1.00 and 0.90 for Norway spruce on mineral and organic soils, respectively, compared with 0.01 and 0.32 for Scots pine.

Four changes were made.

First, we corrected Appendix A itself (L834 to L843). The submitted text reported “one or two breakpoints in 45% of simulations, and none in 10%” for Norway spruce on organic soils, which leaves roughly 45% of realisations unaccounted for. Based on the same 1000 realisations, the correct partition is 9.7% with no breakpoint, 43.6% with one breakpoint, and 46.7% with two breakpoints. The 43.6% figure represents only the share with one breakpoint, not the share with “one or two,” and the previously unexplained category was the share with two breakpoints.

Second, the detection rate for each species and soil combination is now reported in the Results, together with the probability that a breakpoint exists at all. For Scots pine, these results are reported in Section 3.1 at L412 to L415 and Section 3.3 at L484 to L489. For Norway spruce, they are reported in Section 3.1 at L423 to L426 and Section 3.3 at L494 to L496.

Third, we disclose a property of the submitted code that the referee’s comment brought to light (L844 to L852). The breakpoint search stopped after two breakpoints, so realisations recorded as having two breakpoints were right-censored rather than terminated because no further breakpoint was significant. Re-running the same realisations without this limit shows that 33.0% of the Norway spruce organic-soil realisations with a detected breakpoint support three or more breakpoints, and that 98.7% of those that reached the limit have three or more once the limit is removed. Detection in the observed NFI series is unaffected. Without the limit, the procedure returns exactly one breakpoint for both Norway spruce series and none for either Scots pine series, identical to Table S19. We also note that the original simulation was run without a fixed random seed, so the reported percentages were derived from the preserved realisations rather than regenerated. A seed has now been added to the archived code.

Fourth, we revised the two sentences identified by the referee. Section 3.3 no longer states without qualification that breakpoints were detected “for all examined variables.” It now reports the number of plots at which each variable yielded a breakpoint (L484 to L489). Section 3.1 no longer describes the post-change trends as robust. Instead, it states where the regional decline is resolved by the NFI sample and where it is not (L427 to L433).

We also agree with the referee’s interpretation of the 2009 Scots pine signal, and the text now states it directly. The 2009 partition is based on breakpoints detected in the site-level BAI of three of six pine plots, with essentially no support from the regional GI. It is now described as an imposed comparison window rather than a detected regional change point (L412 to L415 and L494 to L496). The Abstract

and Conclusions have been revised accordingly and no longer date a regional decline in Scots pine to 2009.

2 Uncertainty propagation into down stream analyses Uncertainty quantification stops at break-point detection. The subsequent segmented regression and dEMM analyses treat the detected break-point year as known (conditioning on the minority of simulations that found breakpoints), ignoring the substantial detection uncertainty documented in Appendix A. The existing Monte Carlo framework should propagate this uncertainty through the full pipeline rather than fixing the break-point year arbitrarily. This is especially important given my above comment.

Response. We agree, and we have implemented this change. The Monte Carlo framework in Appendix A now extends through the segmented regressions and mixed models rather than stopping at breakpoint detection. The procedure is described in Section 2.4.3 (L378 to L392), and the results are reported in Section 3.4 (L540 to L545), Section 4.1 (L592 to L594), and Appendix A (L862 to L879).

For each of the 1000 GI realisations, we redraw the split year for Eq. (6) from that realisation's own breakpoint, refit the model, and pool the fixed effects across realisations using Rubin's rules with the Barnard and Rubin (1999) small-sample degrees-of-freedom adjustment. Realisations in which no breakpoint is detected are not discarded. Instead, they are fitted with the corresponding no-change model (response $\sim$ cSPEI + cVPD, with the same random intercept and AR(1) structure). Its three change parameters are structurally absent and therefore enter the pooling with an estimate and sampling variance of zero. The pooled estimates are model averages across the "change" and "no change" models rather than results conditional on detection. We report the probability that a breakpoint exists, which is 1.00 for Norway spruce on mineral soils, 0.90 on organic soils, and 0.95 when pooled, alongside the pooled coefficients and the proportion of realisations in which each term was individually significant. The two measures are not interchangeable. The pooled p-value summarises total uncertainty, whereas the proportion indicates how much of that uncertainty arises from the breakpoint location.

The analysis requested by the referee confirms the paper's conclusions and identifies the one estimate that does not remain robust. Every slope, both before and after the breakpoint for both species and response variables, retains its sign and significance class. For Norway spruce, the SPEI effect on BAI remains negative and significant in 100% of realisations (pooled estimate = -0.036, 95% CI = -0.065 to -0.008, $p = 0.013$), and the VPD effect on $\Delta^{13}\text{C}$ remains significant in 98% (pooled estimate = -1.84, 95% CI = -3.08 to -0.60, $p = 0.004$). For Scots pine, the VPD effects on both responses remain significant in 100% of realisations. The species contrast reported in the paper is therefore not an artefact of fixing the split at 2016.

The exception is the shift in intercept between periods. For Norway spruce BAI, it changes sign from +0.033 with the split fixed in 2016 to -0.010 when the breakpoint is redrawn. It is significant in only 30% of realisations, and 66% of its total variance is attributable to uncertainty in the breakpoint location rather than residual error. For Norway spruce $\Delta^{13}\text{C}$, the effect weakens from $p = 0.002$ to $p = 0.045$ when conditional on detection and to $p = 0.079$ under the full model average. We have therefore withdrawn the interpretation of the intercept shift from the Results and Discussion and now interpret only the slope contrasts, which are the comparisons on which the paper's argument rests.

Three new tables report these results. Table S21 presents every fixed effect in Eq. (6), both with the split fixed in 2016 and with uncertainty propagated, using conditional and unconditional pooling. Table S22 presents a fixed-split sensitivity analysis for 2015, 2016, and 2017, which reaches the same conclusion. The pre-period slopes retain their sign and significance class at every split, whereas the Norway spruce BAI intercept shift changes sign between the 2015 and 2017 splits. A new Table A2 in

Appendix A presents the pooled segmented-regression slopes for the growth index by species and soil type.

We note one limitation inherent in the requested analysis, which we now state in Section 2.4.3 (L390 to L394). Because the breakpoint varies among realisations, “pre” and “post” refer to slightly different periods in different realisations. Therefore, part of the between-realisation variance reflects variation in period definition rather than uncertainty in the slope.

Finally, we clarify the scope of the uncertainty propagation. It now covers breakpoint detection, the segmented regressions underlying Figure 4 and Table S19, and the mixed models in Eq. (6) underlying Tables 3 and 4. The estimated marginal means comparisons in Tables S13 to S18 remain conditional on the harmonised change years of 2009 for Scots pine and 2016 for Norway spruce, and we now state this explicitly in Section 2.4.2 (L331 to L334). We also note that only Tables S13 and S14 could incorporate NFI sampling uncertainty. The BAI and iWUE comparisons in Tables S15 to S18 are site-level quantities estimated from the tree-ring plots and are not affected by the NFI sampling design.

Specific comments

The specific comments below include apparent contradictions between reported findings and presented statistics, missing or incorrect citations, and typographical or content errors in figures.

Line 248 BAI units and the stand-structure confound. Section 2.3.5 defines BAI as $\text{mm}^2 \text{ tree}^{-1} \text{ yr}^{-1}$, but every subsequent reference is in $\text{m}^2 \text{ ha}^{-1} \text{ yr}^{-1}$ which is a stand-level quantity. The conversion or scaling is never described, please clarify.

Response. Corrected. BAI was calculated for each tree in $\text{mm}^2 \text{ tree}^{-1} \text{ yr}^{-1}$ from annual diameter increments and then converted to the stand level in $\text{m}^2 \text{ ha}^{-1} \text{ yr}^{-1}$ before analysis using the measured stem density of each plot (Td, Table 1). This scaling step was omitted from the submitted Methods and is now described in Section 2.3.5 (L252 to L253). Only stand-level values are reported in the Results, figures, and tables.

Line 498 “a statistically significant negative slope before 2016 that weakens thereafter.” Table 3 gives $\beta_{\text{pre}} = -0.035$ and $\beta_{\text{post}} - \beta_{\text{pre}} = -0.052$, i.e. $\beta_{\text{post}} = -0.087$ — the negative SPEI effect strengthens, this is also apparent in the bottom left panel of Figure 6. A strengthening negative SPEI slope in spruce apparently refutes the water limitation interpretation presented and is never discussed.

Response. The referee is correct about the arithmetic and its implications, and we have rewritten the passage in the Results (Section 3.4, L531 to L533) and added a paragraph to the Discussion (Section 4.1, L595 to L600).

The negative SPEI slope for Norway spruce BAI strengthens rather than weakens after 2016 ($\beta_{\text{pre}} = -0.035$, 95% CI = -0.059 to -0.011, $p = 0.005$; $\beta_{\text{post}} = -0.088$, 95% CI = -0.142 to -0.034, $p = 0.002$), although the difference between periods is only marginally significant ($\Delta\beta = -0.052$, $p = 0.081$).

We had not discussed this result, and we agree that it requires discussion. The text now reports it together with the corresponding response of $\Delta^{13}\text{C}$, which is positive in both periods ($\beta_{\text{pre}} = +0.134$, $p = 0.004$; $\beta_{\text{post}} = +0.228$, $p = 0.027$). Growth and carbon isotope discrimination in Norway spruce therefore respond to the summer water balance in opposite directions. Thus, the negative relationship between BAI and SPEI cannot be interpreted as evidence of soil water limitation.

We deliberately do not propose a mechanism for this relationship. All Norway spruce plots in this study are located on drained organic soils, where growth is influenced by drainage and water table dynamics as well as by the summer water balance. Separating these influences would require water table measurements, which we did not make. We have therefore based our interpretation of atmospheric drought in Norway spruce on VPD, the variable that shows a significant constraint on growth after 2016, and we no longer present the relationship between SPEI and BAI as evidence of water limitation.

Line 496-497 “the difference between periods is marginally significant” for spruce BAI * VPD. Table 3 gives $p = 0.310$. The marginal (“.”) flag is apparently on the SPEI interaction, not VPD. $p=0.310$ is well above your chosen significance threshold.

Response. Corrected (Section 3.4, L527 to L531). The marginal flag in Table 3 applies to the SPEI interaction for Norway spruce BAI ($\Delta\beta = -0.052$, $p = 0.081$), not to the VPD interaction, for which $p = 0.310$.

We have also separated two statements that were conflated in the submitted text. The post-2016 VPD slope for Norway spruce BAI differs significantly from zero ($\gamma_{\text{post}} = -0.459$, $SE = 0.162$, 95% CI = -0.779 to -0.140, $p = 0.005$; Table 4), whereas the change in this slope between periods does not ($\Delta\gamma = -0.232$, $p = 0.310$; Table 3). Both statements are true, but they are not equivalent. The revised text therefore identifies a significant VPD constraint on growth during the post-2016 period, but it does not claim a demonstrated increase in the strength of that constraint. We checked the remainder of the Results and Discussion for the same conflation and corrected it in the closing sentence of Section 3.4 and in the Abstract, both of which previously asserted that the constraint had strengthened.

Line 495: “significant changes in intercepts indicate different baselines in BAI and $\delta^{13}\text{C}$ between periods.” For BAI, $\alpha_{\text{post}} - \alpha_{\text{pre}} = 0.033$, $p = 0.520$, only $\delta^{13}\text{C}$ is significant. This statement is a direct contribution to the statistical evidence provided in Table 4. Also, I wonder if you meant to reference Figure 6 here instead of Figure 4.

Response. Corrected on both points (Section 3.4, L526 to L527). Only the $\Delta^{13}\text{C}$ intercept shift is significant ($\Delta\alpha = 0.455$, $SE = 0.146$, $p = 0.002$); for BAI, $\Delta\alpha = 0.033$, $SE = 0.051$, $p = 0.520$. The cross-reference now points to Figure 6.

In addition, as explained in our response to the referee’s second primary concern, neither intercept shift is identifiable once uncertainty in the breakpoint location is propagated. We therefore no longer interpret either shift (Section 3.4, L540 to L545).

Figure 6 caption indicates these are “standardized effect sizes” directly contradicts section 2.4.3; I think they are only mean-centered? Importantly, SPEI is standardized by construction (1 unit change means 1SD), complicating direct visual and quantitative comparisons with VPD (1 unit change is 1 kPa and is roughly an order of magnitude larger than the interannual variability). Therefore comparing VPD and SPEI effect sizes is challenging. This figure’s caption also says blue/red points while the new version of the figure uses blue/orange instead. Figure 6 is also missing the A,B,C, D sub-panel labels referenced in the caption & main text.

Response. Corrected on all four points. The second correction also improved one of our conclusions.

The description of “standardized effect sizes” was incorrect and contradicted Section 2.4.3. The predictors were mean-centred rather than standardized. The caption has been rewritten accordingly.

The referee is correct that slopes expressed in their original units are not comparable between the two predictors. The problem is asymmetric because SPEI is standardized by construction, whereas VPD is expressed in kPa. We have therefore rescaled every slope to represent a one-standard-deviation change in each predictor, and Figure 6 now shows these rescaled slopes. The procedure is described in Section 2.4.3 (L363 to L367), and the numerical estimates are given in the new Table S20. The standard deviations were calculated separately for each species over the entire analysis period, using pooled observations across that species’ plots for which both predictors were available ($n = 144$ plot-years per species). For Scots pine, the standard deviations were 0.759 for MJJ SPEI-3 and 0.100 kPa for MJJ VPD. For Norway spruce, they were 0.741 and 0.089 kPa, respectively. A one-unit change in VPD therefore corresponds to approximately ten interannual standard deviations in Scots pine and eleven in Norway spruce, a change that was never observed in these data. The rescaling is a linear reparameterization. It is exact and does not change any test statistic or p-value, as verified against the unscaled fit.

This correction changed one of our conclusions, and we are grateful to the referee for raising the issue. In the original units, the VPD slope exceeds the SPEI slope by factors ranging from 2.5 to 107, as described in the submitted text. When expressed per standard deviation, the ratio decreases to 0.30 to 14.1, and the species show distinct patterns. In Scots pine, VPD remains the stronger predictor for both responses in both periods. In Norway spruce, SPEI is stronger in three of the four combinations of period and response (BAI: 0.78 before and 0.63 after 2016; $\Delta^{13}\text{C}$: 1.73 before and 0.30 after 2016). The statement in Section 4.1 that “SPEI exerted consistently negligible effects relative to VPD” was an artefact of the units and has been replaced with a species-specific statement (L632 to L639): atmospheric dryness dominates the response of Scots pine, whereas Norway spruce responds to atmospheric demand and the summer water balance to a similar degree. The corresponding statements in the Abstract and Conclusions, which also identified VPD as the dominant driver across both species, have been revised accordingly.

The caption colours have been corrected from blue and red to blue and orange, and the panel labels A to D have been added to the figure.

Line 622: Norway spruce showed more stable $\delta^{13}\text{C}$ values over time than Scots pine, with the latter showing declining values.” Section 3.3 reports significant $\delta^{13}\text{C}$ breakpoints (shifting to negative trends) only in spruce, and none in pine. Please reconcile these two statements.

Response. We thank the referee for identifying this inconsistency. We agree that the original text incorrectly linked temporal stability in $\Delta^{13}\text{C}$ to physiological sensitivity. Correcting this issue strengthens the species contrast rather than complicating it. The passage has been rewritten (Section 4.2, L683 to L695).

Breakpoint detection and climatic sensitivity are different quantities, and the two species differ in both measures in the same direction. Scots pine showed no significant $\Delta^{13}\text{C}$ breakpoints at any plot, which is consistent with its relatively continuous increase in iWUE over time and its stronger $\Delta^{13}\text{C}$ response to VPD (-4.53‰ and -4.89‰ per unit VPD before and after 2016, compared with -1.92‰ and -0.57‰ in Norway spruce; Table 4). The response is gradual and persistent, so there is no structural break to detect. Norway spruce showed significant $\Delta^{13}\text{C}$ breakpoints at three of six plots because $\Delta^{13}\text{C}$ remained relatively stable for much of the record before shifting around 2015 to 2016. The breakpoint therefore

detects a change in trajectory rather than a larger overall physiological response. The revised text separates the two concepts and aligns the discussion with Section 3.3 and Table S19.

Line 137–139: I am unfamiliar with the CRU TS 4.08 data product, but a quick internet search tells me it is a monthly product, not a daily one

(<https://www.nature.com/articles/s41597-020-0453-3>). This should be corrected in the text.

Response. Corrected. CRU TS 4.08 is a monthly product, and all variables derived from CRU, including VPD and SPEI, were computed at monthly resolution (Section 2.1.2, L137 to L140).

Figure 1 caption covers 1980–2021 while section 2.1.2 states the same data were obtained for 1990–2023. Please clarify / reconcile.

Response. Corrected. Figure 1 has been updated, and its caption now states 1980 to 2023. The two periods differ by design rather than by error. Figure 1 shows the regional climate over the longer period from 1980 to 2023 to place the study period in context, whereas the CRU variables used in the analyses were extracted for 1990 to 2023 (Section 2.1.2, L137 to L140).

Isotope record start date Figure 6's caption implies $\delta^{13}\text{C}$ begins in 1999, and Figs. 3C–D and A1 start at 1999–2000, but the Methods never state this (maybe I am missing this). Please state the isotope period explicitly in the text. Also, could you please clarify (at least in a response to this comment) why Fig. A1/A3 begin in 1999 while Fig. 2 GI begins in 1990.

Response. The $\delta^{13}\text{C}$ series cover 1999 to 2022 at all 12 plots, providing 24 annual values per plot. This information is now included in Section 2.3.6 (L261 to L262).

Regarding the second question, the isotope record sets the analysis window. Carbon isotope measurements begin in 1999, so all change-point analyses were conducted over the common period from 1999 to 2022. The GI and BAI series, which extend further back, were truncated to match this period so that breakpoint estimates are directly comparable among GI, BAI, $\Delta^{13}\text{C}$, and iWUE. This is now stated in Section 2.4.2 (L327 to L328). Figure 2 shows the growth index over its full available period, from 1990 to 2022, because it is a descriptive figure and is not constrained by the isotope record.

Study period is variously defined as either being 1990–2022 (abstract, Figs. 2–3), 1990–2023 (Fig. 4 title, LMM post-period). Please correct throughout.

Response. Corrected. We thank the referee for pursuing this issue because the inconsistency concealed a more specific problem. Three distinct periods are involved, and the submitted text did not clearly distinguish them: the extent of each dataset (NFI growth indices, 1990 to 2022; tree-ring BAI, through 2023; $\delta^{13}\text{C}$ and iWUE, 1999 to 2022; CRU-derived MJJ SPEI, through 2022); the descriptive figures, which show each series over its full extent; and the analysis window, which is 1999 to 2022 for all segmented regressions (Figure 4, Table S19, and Appendix A) and all mixed models (Eq. 6, Figure 6, and Tables 3 and 4).

The references to 2023 were the source of the inconsistency and have been removed. The pre- and post-2016 periods in Eq. (6) are now given as 1999 to 2015 and 2016 to 2022 for both response variables in Section 2.4.3 (L359) and in the Figure 6 caption. The breakpoint-defined periods in Section 3.2 now

read 2009 to 2022 for Scots pine and 2016 to 2022 for Norway spruce (L459). Section 3.3 now reads 2016 to 2022 (L500), and the title of Figure 4 now refers to the analysis period from 1999 to 2022, with its x-axis redrawn to end in 2022. MJJ SPEI ends in 2022, so 2023 could not have entered the mixed models in any case. Both models are fitted to $n = 144$ plot-years per species (6 plots $\times$ 24 years, from 1999 to 2022; Section 2.4.3, L363 to L367). The period from 1990 to 2022 given in the Abstract and in the captions of Figures 2 and 3 refers to the extent of the growth-index and tree-ring records shown there, which remains unchanged.

Missing references: Girardin et al. 2023 (cited line 39) and Zignol et al. 2025 (cited line 640) do not appear in the reference list. A reference in the bibliography for Anderegg et al. (2018) is listed as “Science, In review”; it was published in Nature according to the linked DOI.

Response. Corrected. The citation at what was line 39 was to (IPCC, 2023) rather than mistaken extra Girardin et al.; also added (Zignol et al., 2025) to the reference list; and the (Anderegg et al., 2018) entry has been corrected to its published form.

“A. Lehtonen, Eyvindson et al., 2023” and “Leppä, Hökkä, et al., 2020; Leppä, Korkiakoski, et al., 2020” (line 683) should probably follow the 2023b / 2020a,b convention used elsewhere; “H. M. Henttonen et al., 2024” and “K. T. Korhonen et al., 2024” carry unnecessary initials. Whatever convention is chosen it is helpful to remain consistent throughout the text.

Response. Corrected. Author initials have been removed from in-text citations, and the a/b year convention is now applied consistently throughout.

I think the site-name spelling in Figure 5 “Vastämäki” may be misspelled (e.g. from the Methods there is “Vastämäki”).

Response. Corrected. Figure 5 now reads Vastämäki, consistent with the Methods and Supplementary info.

Figure 3 y-axis gives iWUE as $\mu\text{mol mmol}^{-1}$; the text uses $\mu\text{mol mol}^{-1}$

Response. Corrected. The axis label now reads $\mu\text{mol mol}^{-1}$, consistent with the text and with the values reported (approximately 60–80 $\mu\text{mol mol}^{-1}$). Figure 3 has been redrawn.

Referee 3

This manuscript provides a timely and compelling investigation into the physiological mechanisms driving forest productivity decline under increasing atmospheric drought (VPD) in boreal forests. By combining large-scale National Forest Inventory (NFI) data with plot-scale dendrochronological and stable isotope measurements, the authors offer robust insights into the contrasting drought adaptation strategies between Norway spruce and Scots pine. The manuscript is well-structured, methodologically sound, and addresses an important knowledge gap in northern forest management.

Response. We thank the referee for this assessment and for the comments below, which have improved the Methods and the Discussion.

Major comments:

1.The climate driver data are extracted at a grid resolution of 0.5 degree. Given that local microtopography (e.g., drained peatlands vs. mineral uplands) strongly modulates microclimate and soil hydrology, the authors should explicitly discuss the potential uncertainties of using coarse grid data to represent plot-scale microclimatic extremes.

Response. We agree that the spatial resolution of the climate products used in this study does not capture plot-scale variation in hydrology and microclimate. We have added a paragraph at the end of Section 4.1 (L648 to L650) stating what the gridded variables do and do not represent.

Local factors such as water-table depth, ditch spacing, microtopography, and canopy structure generate substantial variability in soil moisture within a single climate grid cell. Therefore, the gridded variables should be interpreted as regional climatic forcing rather than as the conditions experienced at individual plots. This distinction also explains a pattern in our results: the local moisture classes captured variation in BAI and iWUE that the gridded drivers did not, which is consistent with the two operating at different spatial scales. Accordingly, SPEI should be interpreted as a regional water-balance index rather than as a direct measure of plot-scale soil water availability. We have added references on microclimatic variability and the limitations of coarse-resolution climate datasets (De Frenne et al., 2021; Lembrechts et al., 2019)

2.The paper identifies 2016 as a sharp breakpoint for spruce growth collapse. Drought-stressed spruce stands in Scandinavia often suffer from secondary biotic disturbances (e.g., European spruce bark beetle). The authors need to clarify whether trees affected by pest/pathogen damage were filtered out or discuss the potential compounding effects of biotic agents post-2016.

Response. We have clarified the sampling practice and its limitations, and we have added a discussion of the potential compounding role of biotic agents.

Trees were selected as living dominant and codominant individuals in apparently healthy condition. At the time of coring, we avoided visibly damaged trees showing crown dieback, stem damage, resin flow, or bark beetle entry holes. This was a field selection practice rather than a recorded criterion. Damage status was not documented for each tree, so we cannot apply a retrospective filter. This is now stated in Section 2.3.4 (L227 to L229).

The implications for interpretation are set out in Section 4.1 (L601 to L607). Drought can increase the susceptibility of Norway spruce to the European spruce bark beetle (*Ips typographus*) and associated fungi, while heat, drought, and windthrow can promote outbreaks (Netherer et al., 2021, 2024). Nevertheless, two lines of evidence indicate that the post-2016 decline is not primarily an artefact of undetected biotic damage. First, the same shift appears in the regional NFI growth index, which is estimated from a sample independent of our plots and is not restricted to surviving dominant trees. Second, the decline coincides with a shift in $\Delta^{13}\text{C}$, a leaf-level physiological signal that stem-boring damage alone would not produce. A compounding role for bark beetles and associated pathogens cannot be excluded. Because tree-ring samples are necessarily restricted to trees that survived until the sampling date, our estimates of stand-level growth loss are conservative.

3.While air temperature (Ta), VPD, and SPEI were selected as the primary climate drivers, the rationale behind choosing these specific three metrics requires a more thorough explanation.

Response. Section 2.1.2 (L151 to L156) now states the rationale explicitly. The three variables were selected because, together, they separate the thermal, atmospheric demand, and water balance

components of water limitation. These components respond differently to warming and affect tree function through different pathways, allowing the growth decline to be attributed to atmospheric demand rather than to the summer water balance. We also state explicitly that T_a was retained for the descriptive and correlation analyses but excluded from the mixed models because of its high collinearity with VPD. Previously, this was stated only in Section 2.4.3.

4.Line 112: better to capitalize first letter for each word in “Standardized precipitation-evapotranspiration index”

Response. Corrected

5.Line 322-324: this part should be given in results section.

Response. Agreed and moved. Section 2.4.2 now describes only the procedure and the harmonisation rule used to define the comparison periods (L329 to L330), while the breakpoint estimates and their detection frequencies are reported in Section 3.3 and Appendix A. The corresponding text now appears in the Results at L484 to L489 and L494 to L503.

6.Line 327: what is “estimated marginal means”?

Response. Section 2.4.2 (L335 to L342) now explains that estimated marginal means are model-adjusted mean response values, in contrast to temporal slopes, which describe annual rates of change. It also explains how they are used in this study. The referee’s question also prompted us to specify the model underlying Tables S13 to S18, which was named but not described in the submitted Methods. For each response and species, we fitted the model $\text{response} \sim \text{group} \times \text{period} \times \text{year}$, where group represents the soil or plot type and year is centred within each period. We then compared slopes and marginal means among groups and periods using the emmeans package (Lenth and Piaskowski, 2017).

7.Line 344: Please clarify whether the reference to "linear mixed-effects models (LMMs)" refers specifically to Equation 6 or to the analytical approach outlined in Appendix A.

Response. The reference is to Eq. (6), and the text now reads “linear mixed-effects models (LMMs, Eq. 6)” (Section 2.4.3, L353).

Figure 2: the line for mineral is not red visually.

Response. Figure 2 caption has been rewritten, and the caption now identifies the colours correctly.