Author comment: response to referee n°1

How beech ecophysiology shapes temperate forest gross primary productivity – Part 2: Identifying critical timeframes across phenological stages

Jonathan Bitton[1], Bernard Heinesch[1], Catherine Charles[1] and Bernard Longdoz[1]

[1]Biosystems Dynamics and Exchanges, TERRA Teaching and Research Center, Gembloux Agro-Bio Tech, University of Liege, Gembloux, Belgium

Correspondence to: Jonathan Bitton (Jonathan.bitton@uliege.be)

We sincerely thank the referees for their thorough and constructive evaluation of our manuscript. Their comments have been extremely helpful in improving the clarity, structure, and overall quality of the work.

This document provides a detailed, point-by-point response to all comments. We have addressed each referee separately, while ensuring consistency in the revisions throughout the manuscript. For transparency, responses are organized as follows: an overall assessment, followed by general comments and, where applicable, specific line-by-line comments as well as comments related to figures and tables.

Although some comments addressed similar aspects of the study (e.g., indicator definitions and statistical approach), we retained individual responses for each referee, as their remarks were complementary and led to specific revisions in different parts of the manuscript.

All modifications made in the manuscript are indicated at the end of each response and introduced by the symbol "⇨". References to sections, figures, and tables in the revised manuscript are highlighted in blue.

Overall assessment

"In this study, the authors analyzed 24 years of carbon flux data at a European beech forest in northeastern France. They applied a Wavelet Area Interpretation approach to define GPP metrics characterizing (a) the rate of early season GPP increase, (b) the magnitude of peak GPP, and (c) declining late summer GPP. They then analyzed the effect of environmental variables on each metric by correlating various versions of the variables with the metrics, with versions defined by window length (e.g. VPD averaged over 1 week, 2 weeks, 3 weeks, etc.) and start date (referenced to the start of spring; 1 day increments). The authors provide an extensive and well-cited discussion of the correlation results in the context of beech's physiological and ecophysiological seasonal patterns. They conclude, among many specifics about the correlations and their likely (eco)physiological implications, that "forest carbon dynamics must be understood through the lens of seasonal physiology, not annual summaries" (lines 831–832). This is a useful study, carefully executed and thorough. I appreciate the authors' deep thinking about their study site, and I hope, after some clarification and streamlining of the narrative, that this will be published in Biogeosciences."

We thank the referee for the thorough and insightful evaluation of our manuscript, and for the positive assessment of the study. The comments provided were extremely helpful in improving both the clarity and structure of the paper. In particular, the referee's suggestions led us to (i) clarify the definition and interpretation of the GPP indicators, (ii) substantially revise the description of the statistical framework, and (iii) strengthen the connection between results and their ecological interpretation throughout the manuscript. We believe that these revisions have significantly improved the readability, consistency and overall robustness of the manuscript.

General comments

<u>**Comment n°1**</u>**: "*Regarding your three GPP metrics: why these particular metrics? (Why not, for example, the rate of decline rather than the offset between two peaks)? Do you really need to name them "IRise," "IPeak," and "IDrop," rather than using short phrases to call them what they are? At the least, please include a detailed schematic labeling these metrics when they are introduced; I didn't understand what IDrop was until I made it to Figure 4, and even then, I still had some questions (is there always a second peak? What is the usefulness of this particular metric?).*"**

We thank the referee for this comment, which helped us improve the clarity of the indicators.

Our choice of IRise, IPeak and IDrop was guided by the objective of capturing the three main, well-documented phases of seasonal GPP dynamics in temperate deciduous forests:

1. Early-season development (IRise): this phase is strongly linked to leaf ontogeny and reserve mobilization, both of which have been extensively studied in literature. The rate of increase in GPP provides a direct proxy for canopy development dynamics.
2. Peak productivity (IPeak): peak GPP reflects the period of maximum photosynthetic capacity, when environmental constraints are minimal and canopy structure is fully developed.
3. Mid-season decline (IDrop): this phase captures the reduction in GPP during summer, which is typically associated with increasing water and atmospheric stress. While we agree that this indicator is less intuitive, it was specifically designed to quantify the intensity of the decline relative to seasonal dynamics, rather than simply the senescence dynamics in autumn. This makes it particularly suited to identifying drought-related responses.

We acknowledge that IDrop was not sufficiently clear in the original version. The indicator does not rely on the systematic presence of a second peak in the GPP curve. Instead, it is derived from the wavelet representation of the signal, where the right-hand side of the wavelet centered on the GPP maximum captures the decline phase. The presence of a plateau or secondary peak only modulates the magnitude of the signal, without affecting the definition of the decline itself.

Regarding the number of indicators, we deliberately limited the analysis to three metrics to maintain clarity and avoid overcomplicating interpretations. Additional indicators, such as decline rate, were considered but not retained due to (1) weaker signal detectability, (2) higher uncertainty and (3) limited support in the literature.

Concerning terminology, we retained the IRise/IPeak/IDrop notation for consistency with the companion paper and for conciseness throughout the manuscript, as these indicators are repeatedly used.

⇨ Changes in Manuscript: Following the referee's suggestions, we have:

i. added a reference to Fig. 4 in Sect. 2.4 providing a schematic representation of the indicators;
ii. clarified definition of all indicators in Sect. 2.4 to improve accessibility;
iii. added a supplementary figure (Fig. S1, see below) illustrating the wavelet-based computation of the indicators.

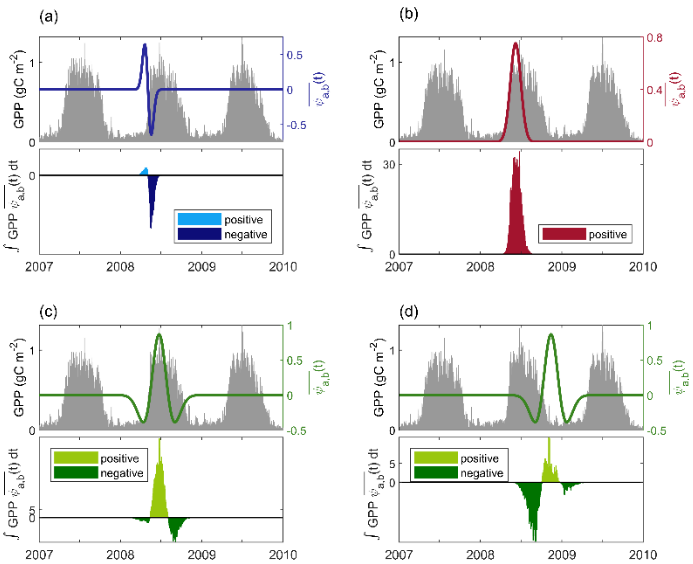

Figure S1: Computation steps for deriving (a) IRise, expressed as the DOG1 April peak at a 3-month period on a normalized GPP signal, (b) IPeak, computed using the opposite of the right lobe of the DOG1 April peak, which is equivalent to a Gaussian function, at a 7-month period, and (c)/(d) IDrop, defined as the ratio of the DOG2 (c) June to (d) November peaks at 6-months periods. For each panel, the top graph displays the wavelet at selected scale and time with the GPP signal in the background (restricted to positive values for readability). The graph below illustrates the area under the product curve of the wavelet and the signal. The algebraic sum of the contributing areas, each linked to wavelet lobes, corresponds to the indicator value yielded by the WT (after taking the ratio c/d for IDrop). Computations are presented for the year 2008.

__Comment n°2__: ***"Statistical methods require some clarification. I don't understand: (i) how longer correlation windows were "decomposed" into shorter windows (L240, L251–254); (ii) how "overlapping correlations" (L243) were defined (for example: what if the window for A overlaps with B and the B window overlaps with C, but C and A don't overlap – are these all in the same group?); or (iii) how the partial correlation analyses were done to identify the most influential variables. What if VPD and Tair both, individually, "rendered remaining [variables] non-significant" (L247) – how would you decide which was more influential?"***

We agree that the original description of the statistical approach was not sufficiently explicit and have substantially revised Sect. 2.5 to improve clarity. The revised description now presents the method as a structured, stepwise procedure, including four main steps:

1. the identification of coherent regions of influence in correlograms across the 2D space of window start dates and lengths, reflecting temporally coherent ecological processes;
2. the decomposition of broad-scale correlations into shorter characteristic windows through the extraction of local maxima;
3. the definition and grouping of overlapping windows, including cases of indirect overlap (e.g. if A overlaps with B and B with C, all belong to the same group);
4. the selection of dominant variables using partial correlations. A variable was considered dominant if it rendered the others non-significant when controlled for. In situations where multiple variables remained significant, they were interpreted as co-dominant, reflecting shared physiological control.

We chose this data-driven, pattern-based approach to prioritize ecological interpretability and parsimony rather than strict predictive optimization.

⇨ Changes in Manuscript: We have fully rewritten Sect. 2.5 to provide a structured description of the method, including the four steps above.

__Comment n°3__: ***"I appreciate your well-researched introduction and discussion, but they often stray quite far from the actual study. Wherever possible, it would be helpful if you can directly connect to your study topic and results. (See line-by-line notes for, e.g., L43, 66, 621–636). In the discussion, I suggest leading with the main result and then following with the (possibly) explanatory details, continuing to reference results as useful."***

We agree that some sections of the Introduction and particularly the Discussion could be more explicitly connected to the core results of the study.

Our initial intention was to provide short contextual elements to each subsection in order to place the analyses in a broader ecophysiological context, ensuring accessibility to readers from different communities (e.g. flux science, physiology, ecosystem ecology). However, we acknowledge that this occasionally led to sections that were insufficiently anchored in the results.

⇨ Changes in Manuscript: To improve the clarity of the manuscript while preserving its overall readability and integrative perspective, we have:

i. revised the manuscript so that each Discussion subsection begins with a clearer reference to the main findings, followed by the corresponding interpretation and contextualization;
ii. streamlined passages in the Introduction to improve their direct relevance to the study (see line-by-line comments).

__Comment n°4__: "***Your discussion is thorough in contextualizing results, but it is missing some key points related to the methods. How do you expect that the fact that beech is only ~75 % of the basal area to influence results, if at all? What is the implication, if any, of using the nighttime partitioning method? Why did you choose to test just one year of legacy effect? Could you discuss the variability of timing in the GPP indicators across years and how that might influence the correlations with environmental variables quantified in particular windows?***"

We thank the referee for this insightful comment. We agree that several methodological aspects required further clarification and have addressed them explicitly:

1. Species composition

Beech represents more than 75 % of basal area (> 80 % stems/ha) within the footprint and therefore dominates the GPP signal. Contributions from other species are expected to be limited. The seasonal patterns identified in our analysis are consistent with known beech phenology, supporting this interpretation. Minor influences from co-occurring species (mostly oak and hornbeam) cannot be entirely excluded, particularly for early-season dynamics, but are unlikely to explain the observed magnitude of variations in GPP indicators. In addition, no clear shifts in species-driven canopy dynamics were observed following thinning events, which represent the main driver of structural and compositional changes in the stand, further supporting the limited influence of secondary species on the GPP signal.

2. Partitioning approach

We used the nighttime partitioning method, which is widely applied in eddy covariance studies. We favored it over daytime methods because it minimizes assumptions about GPP response functions and avoids introducing additional model dependencies. While this approach relies on modelling ecosystem respiration and may introduce some uncertainty in absolute flux estimates, its influence on the temporal patterns analyzed here is expected to be limited. This is partly because the indicators used integrate GPP over multi-day to seasonal timescales, which reduces the impact of short-term partitioning errors. Consistently, additional tests using alternative processing approaches (e.g. daytime partitioning or outlier filtering) resulted only in minor shifts in correlation windows and strengths, without altering the overall interpretation.

3. Legacy effects

We tested legacy effects over one year only, assuming that the strongest impacts occur closest in time to the stress event. Extending the analysis to multi-year lags would reduce statistical power given the limited sample size. This choice is now clarified in the manuscript.

4. Timing variability

GPP indicators are derived using a wavelet-based approach that aligns them with intrinsic discontinuities in the GPP signal, allowing adaptation to interannual phenological shifts. Environmental drivers are also computed relative to the start of season (SOS), ensuring temporal alignment between drivers and ecosystem response. This approach implicitly assumes that phenological timing relative to SOS is broadly consistent across years. This hypothesis is now explicitly stated in the manuscript.

⇨ Changes in Manuscript: We have added a dedicated limitations section (Sect. 4.6) addressing these four aspects.

Line-by-line comments

<u>L18</u>: "*At this stage in the manuscript, it's not clear why $VPD_{92–106}$ and $VPD_{107–114}$ would be considered separate windows.*"

⇨ Changes in Manuscript: We have modified the sentence in the Abstract to distinguish the two windows as reflecting pre-drop and post-drop onset phases of stress.

<u>L43</u>: "*Short-term climatic drivers of what?*"

⇨ Changes in Manuscript: We have clarified the sentence by specifying that these are short-term climatic drivers of forest carbon exchanges.

<u>L66</u>: "*Re. "how beech responds" – in terms of what? Carbon? Mortality? Something else? This is still very broad.*"

⇨ Changes in Manuscript: We have clarified this statement to explicitly refer to beech responses in terms of carbon uptake dynamics and physiological functioning.

<u>L110</u>: "*What are "Warm Winter records"?*"

"Warm Winter records" refers to a specific dataset produced for the ICOS processing chain.

⇨ Changes in Manuscript: We simplified this sentence to avoid unnecessary terminology and now refer only to "ICOS-processed data" which keeps the description concise.

L128–133: “***Provide equations please (and number all equations, to be referenced in the text).***”

⇨ Changes in Manuscript: We have added Eqs. (1)-(4), illustrating index computation for a non-bissextile year y, and numbered all equations consistently throughout the manuscript.

$$\mathrm{IREW(y)} = \frac{\sum_{i=1}^{365} \max(0, 0.4 - \mathrm{REW_i})}{0.4} \tag{1}$$

$$\mathrm{IVPD(y)} = \frac{\sum_{i=1}^{365} \max(0, \mathrm{VPD_i} - 1.5)}{\max(\mathrm{VPD} \; - 1.5)} \tag{2}$$

$$\mathrm{ITair(y)} = \frac{\sum_{i=1}^{365} \max(0, \mathrm{Tair_i} - 25)}{\max(\mathrm{Tair} \; - 25)} \tag{3}$$

$$\mathrm{nDI(y)} = \frac{\mathrm{IREW(y)} - \langle \mathrm{IREW} \rangle}{\sigma_{\mathrm{REW}}} + \frac{\mathrm{IVPD(y)} - \langle \mathrm{IVPD} \rangle}{\sigma_{\mathrm{VPD}}} + \frac{\mathrm{ITair(y)} - \langle \mathrm{ITair} \rangle}{\sigma_{\mathrm{Tair}}} \tag{4}$$

where ⟨ ⟩ denotes the interannual mean and σ the standard deviation.

L142: “***What is the relationship between "trunk" and "shoot"? What data support the inclusion of root biomass in "B"?***”

We agree that the terminology used in the original version could be ambiguous. We have therefore adopted the more standard terms “aboveground” and “belowground” biomass. Aboveground biomass includes all aerial compartments (stem, branches and twigs), while belowground biomass refers to root biomass.

Total tree biomass (B) was derived from a site-specific allometric relationship, which directly estimates biomass as a function of tree circumference and explicitly includes both aboveground and belowground components.

Because the different biomass compartments follow similar power-law relationships with circumference (see below), we retained a single total-biomass formulation rather than summing compartment-specific equations.

$$\mathrm{B\ [kg]} = 0.0165 \cdot \mathrm{circ\ [cm]}^{2.32}$$
$$\mathrm{B_{above}\ [kg]} = 0.156 \cdot \mathrm{circ\ [cm]}^{2.29}$$
$$\mathrm{B_{below}\ [kg]} = \mathrm{B} - \mathrm{B_{above}}$$

These relationships have been previously established for the Hesse site and validated (see Fig. S2 below). Their applicability is further supported here by comparison with independent allometric equations available for beech in the GlobAllomeTree database (Henry et al., 2013[1]).

⇨ Changes in Manuscript: We have (i) replaced the previous terminology describing tree compartments with the terms “aboveground” and “belowground” biomass and (ii) added a dedicated Supplementary section (Sect. S2) and figure (Fig. S2, see below) describing the allometric relationships in detail and supporting their applicability at the study site.

---

[1] Henry, M., Bombelli, A., Trotta, C., Alessandrini, A., Birigazzi, L., Sola, G., Vieilledent, G., Santenoise, P., Longuetaud, F., Valentini, R., Picard, N., and Saint-André, L.: GlobAllomeTree: international platform for tree allometric equations to support volume, biomass and carbon assessment, iForest - Biogeosciences and Forestry, 6, 326, https://doi.org/10.3832/ifor0901-006, 2013.

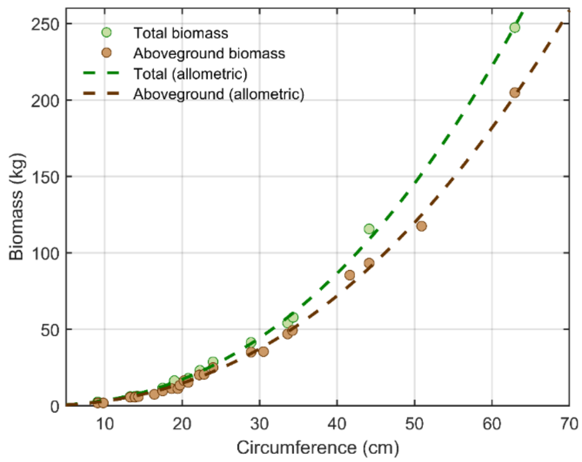

Figure S2: Allometric relationships used to estimate tree biomass at FR-Hes. Total biomass (green dashed line) and aboveground biomass (brown dashed line) are shown as a function of tree circumference, together with observed data points. The relationships follow power-law formulations calibrated from destructive sampling (Ottorini, 2012, unpublished dataset). The figure is adapted from original material provided by J. M. Ottorini. Model parameters are highly significant for both relationships ($p < 0.001$), supporting their use for biomass estimation across the study period.

L151–154: "***Provide equations please***"

⇨ Changes in Manuscript: We have added the corresponding equations and clarified the definitions of all derived variables as follows:

- Median daily growth proportion (g_{pct}), computed by dividing daily growth by the annual growth range according to Eq. (6) and taking the median across trees. The annual growth range (AGR) was defined as the difference between median circumference values after (Oct–Dec) and before (Jan–Mar) the rapid growth phase.

$$g_{pct}(d) = \frac{circ(d) - circ(d-1)}{AGR} = \frac{circ(d) - circ(d-1)}{circ_{Oct-Dec} - circ_{Jan-Mar}} \quad (6)$$

- Daily carbon allocation to biomass (S_b), calculated as the product of the above proportion and S_g, divided by daily GPP:

$$S_b = \frac{g_{pct} \cdot S_g}{GPP_d} \quad (7)$$

<u>L175–177</u>: “*What sensors were used to measure albedo and NDVI?*”

Albedo was evaluated as the ratio of reflected to incoming short-wave radiation, both measured using a pyranometer. NDVI was computed as the difference between near-infrared and red reflectance divided by their sum. Reflectance at each wavelength was measured using sensors recording incoming and reflected radiation in the red (approx. 650 nm) and near-infrared (approx. 810 nm) spectral bands.

⇨ Changes in Manuscript: We have clarified the description of albedo and NDVI by explicitly specifying how the relevant radiation components were measured and derived at the canopy level.

<u>L182–185</u>: “*Why were there times when the GPP definition wasn’t available? Did these definitions, when multiple were available, typically agree (this could be in a supplement)?*”

We clarify that GPP itself was always available from flux partitioning; the reported percentage refers to the availability of independent SOS detection metrics (RLT, albedo, NDVI).

These metrics were not used as independent SOS estimates but as sequential consistency filters to identify a single date combining canopy structural change (RLT) and spectral response (NDVI/albedo).

Because the objective was not to compare methods but to ensure a robust and ecophysiologically consistent SOS detection, we did not perform a systematic comparison between metrics.

⇨ Changes in Manuscript: We have clarified these points in Sect. 2.3.3 to avoid ambiguity.

<u>L220</u>: “*Are all the correlations linear?*”

Yes, all correlations are linear (Pearson). We chose Pearson correlations because they provide direct interpretability, are consistent with the use of partial correlations, and yielded results very similar to rank-based approaches (Spearman) in this dataset.

⇨ Changes in Manuscript: We have clarified the definition of both methods in Sect. 2.5.

<u>L229</u>: “*I suggest a "for example" parenthetical here to clarify: "This approach produced one average value per year for each window configuration (for example, one VPD value for a week-long window beginning at SOS+0, one for a week long window beginning SOS+1, etc.)."*”

Thank you for this suggestion.

⇨ Changes in Manuscript: We have added the suggested example.

L256: "***Residuals of what model?***"

The residuals are derived from simple linear regressions between each indicator and its associated environmental driver (i.e., the same relationships used in the Pearson correlation analysis). Residuals therefore represent deviations from the expected linear response of the ecosystem to environmental conditions.

⇨ Changes in Manuscript: We have clarified this point in Sect. 2.5 to avoid confusion.

L306: "***Again, I'm confused by where these residuals are coming from.***"

See response to comment on L256.

L342: "***$Rg_{15–29}$ is also in Fig. 4 (in addition to Rg10–31), and it has a larger correlation than $Rg_{10–31}$ – why isn't is considered here?***"

The $Rg_{15–29}$ window is indeed shown in Fig. 4, although it has a lower correlation with IRise (0.741) than $Rg_{10–31}$ (0.782). Both windows were considered; however, $Rg_{15–29}$ represents a shorter sub-period nested within the broader $Rg_{10–31}$ window and was therefore only discussed in relation to its second phase (SOS+15 to SOS+30). The broader $Rg_{10–31}$ window, which integrates both phases, is discussed in L334–340.

L379: "***The Rg significant window was actually 9 days later than maximum incident radiation; does this really correspond?***"

We agree that the term "correspond to" was not appropriate.

⇨ Changes in Manuscript: We replaced it with "are linked to," to better reflect the temporal offset while maintaining the ecological interpretation.

<u>L385–386</u>: "*I don't see any thresholds in these figures…*"

Threshold values were not reported in the initial version because they were not explicitly quantified through a formal threshold analysis. However, as they help interpret the observed relationships, we now provide approximate values (250 W m⁻² for Rg, 16 °C for Tair, and 10 hPa for VPD) and illustrate both full and filtered relationships (i.e. excluding summer-extreme years) within the same figure.

⇨ Changes in Manuscript: We have (i) added threshold values directly to the figure caption and corresponding section of the manuscript (Sect. 3.3.1) and (ii) updated Fig. 6 accordingly (see below).

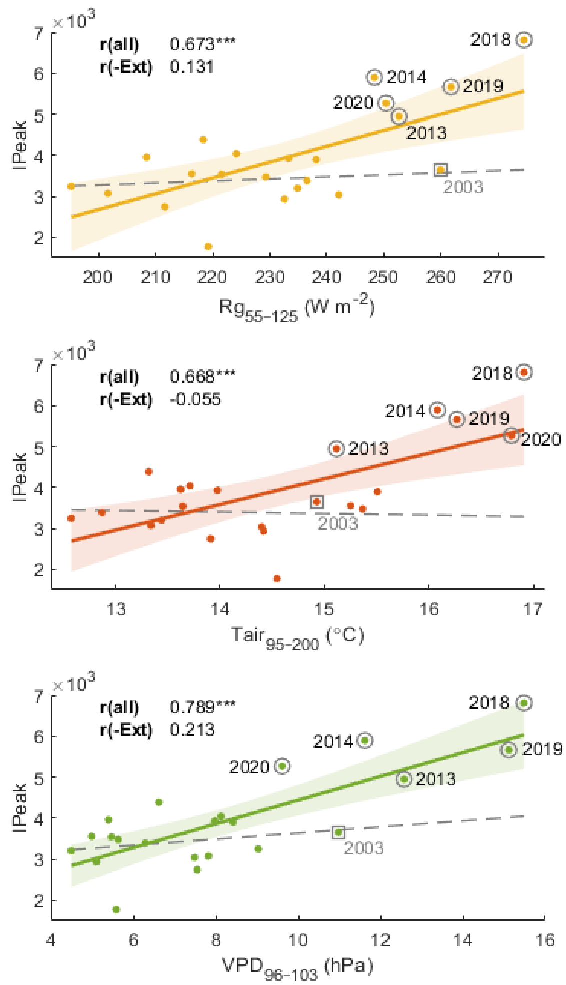

Figure 6: Scatter plots illustrating the strongest correlations between IPeak and mean environmental drivers. Each panel displays the fitted linear regression (solid line) with its 95 % confidence interval (shaded area). The dashed grey line represents the regression excluding extreme-summer years, which are indicated by grey circles. The outlier year 2003 is additionally highlighted with a grey square. Correlation coefficients are reported both for the full time series and for the subset excluding extreme-summer years. Statistical significance is indicated by one, two or three stars for $p \leq 0.05$, 0.01 and 0.001, respectively. Approximate threshold values (250 W m⁻² for Rg, 16 °C for Tair, and 10 hPa for VPD) mark the conditions above which IPeak responds to environmental drivers. Although IPeak retains the units of GPP, it represents a wavelet-derived metric and should be interpreted comparatively rather than as an absolute flux.

L387: "***Consider labeling 2003 on Fig. 6***"

Thank you for this suggestion.

⇨ Changes in Manuscript: We have added a specific labelling for 2003 (see updated Fig. 6 in response to comment on L385–386).

L413: "***What about prior year $P_{56–63}$?***"

This effect is discussed in the manuscript but embedded within the broader $P_{57–120}$ window (see L414).

L424: "***But none of the extreme years were next to each other except 2018/19***"

We agree with this observation.

⇨ Changes in Manuscript: We removed the sentence for clarity as it did not add substantial value to the interpretation.

L426: "***I don't see "Rg began correlating with IDrop from SOS+8 onward" in either Fig. 4. Or Fig. S3 – what am I missing?***"

Thank you for pointing this out.

⇨ Changes in Manuscript: The reference has been corrected to the appropriate figure (Fig. S7 – Fig. S9 in the revised manuscript) and table (Table S5).

<u>L433</u>: "*But VPD and REW are correlated (and causally related) – how did you decide which of these was more important (see broad note about statistical method clarification)*"

The distinction between VPD and REW was based on a combination of (1) correlation strength and (2) partial correlation analysis, which allowed us to assess whether one variable explains the signal of the others (see response to broad note).

In this case, VPD consistently showed stronger correlations and remained significant when controlling for REW, whereas the reverse was not true (see table below). This indicates that VPD is the dominant driver of IDrop, while REW plays a secondary role.

Control: VPD				**Control: REW**			
Variable	Partial correlation	p-value	Significance level	**Variable**	Partial correlation	p-value	Significance level
Rg	−0.06	0.80	/	**Rg**	0.54	0.016	*
P	−0.44	0.06	/	**P**	−0.37	0.12	/
Tair	0.11	0.66	/	**Tair**	0.61	0.006	**
REW	−0.27	0.26	/	**VPD**	0.76	0.0002	***

<u>L446–449</u>: "*Why should stress AFTER the drop influence the drop?*"

We clarify that this second window refers to conditions occurring after the onset of the drop, not after its completion. These post-onset conditions influence the intensity and duration of the decline, rather than triggering it. We further discuss the relevance of this window in L662–663.

⇨ Changes in Manuscript: We have revised the text accordingly.

<u>L455</u>: "*This seems to conflict with L438.*"

We acknowledge that the wording was unclear. VPD is the dominant driver over most of the period (see correlation values, number of selected VPD windows in Fig. 4 / Table S5 and response to L433), while REW plays a secondary and more localized role, particularly late in the season.

⇨ Changes in Manuscript: We have revised this paragraph to clarify this distinction.

L513–514: **"*I don't really understand this.*"**

⇨ Changes in Manuscript: We have reformulated this statement to explain that environmental stress can slow leaf expansion rates while simultaneously extending the duration of expansion (as described in Granier and Tardieu, 2009). As a result, final leaf area may remain unchanged despite reduced instantaneous growth rates.

L541: **"*It looks like pretty strong coupling in Figure 6 to me.*"**

We clarify that this statement refers to typical years, when excluding extreme-summer years.

⇨ Changes in Manuscript: We now explicitly refer to the relationships observed when excluding extreme years to avoid confusion.

L542: **"*2013, 2014, 2018, 2019, 2020 are not notably less scattered around the fit line than the other years.*"**

⇨ Changes in Manuscript: We now rely on the updated figure (excluding summer-extreme years, see response to comment on L385–386) to better illustrate the relationship.

L621–636: **"*How is this section connected to your analyses?*"**

We agree that the connection between this section and our results was not sufficiently explicit in the original version.

⇨ Changes in Manuscript: To address this, we have revised the beginning of Sect. 4.3 to clearly anchor the discussion in our findings. In particular, we now explicitly state that mid-season dynamics are characterized by a shift in GPP control, with drought-related variables (VPD/REW) emerging as dominant drivers of IDrop (Fig. 4). We also added a transition sentence clarifying that the subsequent paragraph is intended to provide the physiological context needed to interpret these observed patterns. This ensures continuity between our results and the broader eco-physiological framework discussed in this section.

L644: **"*Higher sensitivity to atmospheric demand than to what?*"**

⇨ Changes in Manuscript: We have clarified the sentence by specifying that beech at mesic sites shows higher sensitivity to atmospheric demand than to soil water limitation (which is consistent with recent literature, e.g. Green, 2024; Novick et al., 2016).

L645–646: "***This is a convoluted sentence, and I don't think it's supported by the fact that VPD correlates well with IDrop.***"

We agree that the original formulation was too categorical.

⇨ Changes in Manuscript: We have revised the sentence to better reflect that the relative importance of soil versus plant hydraulic limitation is context-dependent. The revised wording emphasizes that our interpretation applies specifically to site conditions and is consistent with theoretical frameworks describing how the limiting hydraulic element governs sensitivity to VPD versus soil drying (Wankmüller et al., 2024).

L779–784: "***I don't think this all follows. Severe thinning reduces IPeak whereas minor thinning doesn't; how does this suggest that loss of photosynthetic surface outweighs effects of increased light availability and competition release? – this would only be the case for severe thinning, presumably.***"

We agree that the original formulation was unclear and could be misinterpreted. We have revised the paragraph to clarify that the observed pattern reflects a contrast between thinning intensities. Severe thinning (e.g., 2005) leads to a strong reduction in photosynthetic surface that outweighs potential gains in light availability and reduced competition. In contrast, lighter thinning events induce only minor structural changes, which are insufficient to significantly affect stand-level productivity.

⇨ Changes in Manuscript: This distinction is now made explicit in Sect. 4.5.1.

L63: "***influences → influence***"

L127: "***aerial drought → atmospheric drought***"

L244: "***influent → influential***"

L303: "***indicating → suggesting***"

L461: "***show with → have***"

L660: "***shot → short***"

L664: "***aerial → atmospheric***"

L671: "***drought → droughts***"

⇨ Changes in Manuscript: Corrected.

Figures and Tables

<u>Figure 1</u>: "*Nice figure! Please connect the references with the processes using, for example, superscripts.*"

Thank you for your positive comment and for this suggestion.

⇨ Changes in Manuscript: We have added connections between references and corresponding processes in Fig. 1.

<u>Figure 3</u>: "*What are the units for IRise, IPeak, and IDrop? The second column of figures is NOT residuals from the models in the first column, which is confusing… where do these residuals come from? Consider labeling all the years, not just extreme years. Re. "Years affected by thinning" – what makes you think that the thinning is only influential in the year it is done?*"

Thank you for these helpful remarks.

- IRise, IPeak and IDrop are derived from wavelet coefficients. Because the wavelet transform preserves the dimensionality of the original signal (using a L1-normalisation), IPeak is expressed in the same units as GPP (gC m^{-2}), while IRise (computed from a normalized signal) and IDrop (defined as a ratio of wavelet coefficients) are dimensionless.
- We clarified that the residuals are derived from separate linear relationships, not from the panels shown alongside them.
- While labeling all years would improve completeness, it resulted in substantial visual overcrowding and reduced readability of the figure. We therefore retained labels for extreme years only, as they are central to the interpretation of the results.
- We also rephrased "years affected by thinning" to "years during which thinning operations were conducted," to avoid implying that effects are limited to that year.

⇨ Changes in Manuscript: We have:

i. improved the caption accordingly;
ii. added units in Fig. 3, Fig. 6 and Fig. S2 (Fig. S4 in the revised manuscript), while stating in the captions and in Sect 2.4 that indicators should be interpreted in a relative sense (e.g., across years or events).

Figure 4: "***What a great figure! It really gets a lot of information across in an intuitive way. Would it be useful to scale the thickness of the bars to the correlation coefficient (maybe not?). How did you decide how to order the correlation bars from top to bottom within each section? Consider ordering by correlation magnitude or by beginning date of the correlation window. **Which correlations are shown? Sometimes, in the text, windows are referenced as though they're significant but they're not in this figure – e.g. L355, L382, L427–430, L518. Or, windows are in the figure but are omitted from the text as though they're not significant – e.g. L413.***"

Thank you for this comment and for the positive feedback on the figure.

Figure 4 displays only the dominant correlations selected through our variable-selection procedure (see Methods), which explains why some significant windows mentioned in the text are not shown.

Regarding the specific examples:

- L355, L413, L518: the windows $P_{95–137}$, $P_{57–120}$, $Tair_{22–43}$ and $Tair_{9–16}$ are included in Fig. 4.
- L427–430: only $VPD_{53–186}$ was retained after variable selection; we also corrected an inversion in the starting dates for P (SOS+58) and VPD (SOS+53) in L429–430.
- L382: this window was not retained because it could be explained by other variables; we have clarified this explicitly in the text.

In the figure, long-scale windows are systematically displayed above short-scale ones, which are displayed chronologically (except a minor ordering error regarding the $VPD_{9–16}$ window for IRise). We will try implementing suggested adjustments to improve visual readability (in particular, correlation strength).

⇨ Changes in Manuscript: We have:

i. clarified the ordering of correlations in the caption;
ii. corrected the ordering of short-scale windows for IRise;
iii. updated L429–430 (Sect. 3.4.1) with the correct starting dates for P and VPD;
iv. added a clarification in L382 (Sect. 3.3.1) regarding the rejection of the $Rg_{151–158}$ window.

Figure 7: "***This is a particularly information-light figure; add all year labels; consider scatter plots with Reco, g_c on the y axis and extremity on the x axis.***"

Thank you for this suggestion. We updated Figure 7 to improve its informativeness and readability, including labeling all years directly and clarifying the ranking of variables. During this revision, we identified an inconsistency in the original version: bars associated with g_c were inadvertently derived from latent heat data rather than g_c. This has been corrected. The overall interpretation of the results remains unchanged.

⇨ Changes in Manuscript: We have:

i. updated Figure 7 (see below) accordingly;
ii. revised the corresponding statements in the manuscript (L405–406 and L544) to reflect the updated pattern, emphasizing that canopy conductance is generally high across years but particularly elevated in 2013, 2014 and 2018, rather than consistently across all extreme years.

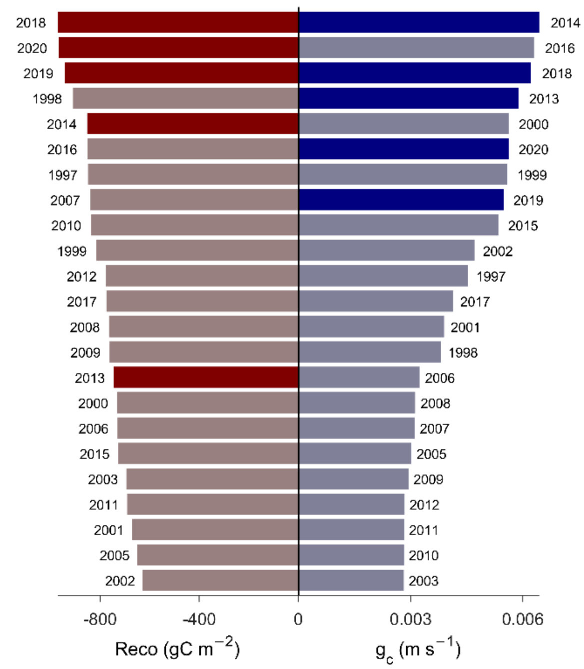

Figure 7: Cumulative ecosystem respiration (Reco, left side) and canopy conductance (g_c, right side) integrated over the IPeak integration period (SOS to SOS+120). Years are ranked independently for each variable in decreasing order of magnitude. Each horizontal bar represents a single year, with labels positioned on the corresponding side of the axis. A vertical line separates the two variables for clarity.

Table S1: ***"What does bold mean? It's not clear which effects are dominant."***

⇨ Changes in Manuscript: We have:

i. replaced the use of "bold" formatting with "blue" to improve clarity in Tables S1–S5;
ii. updated the caption of Table S1 accordingly;
iii. corrected minor inconsistencies in Table S5 (grouping, correlation values and windows), which do not affect the interpretation of the results.
