

Tropical forest responses to climate extremes: an analysis using an individual-based demographic vegetation model

Response to RC1

Reviewer 1

General comments

This study investigates how tropical forests respond to drought, warming, and their compound effects across multiple timescales using a demographic vegetation model (BiomeE). By applying controlled perturbations to precipitation and temperature, the authors aim to disentangle the relative roles of soil water supply and atmospheric demand in shaping ecosystem stability. The results highlight the importance of vapor pressure deficit (VPD), baseline hydroclimate, and stress persistence in determining forest resilience and collapse transitions across contrasting tropical sites.

I find that the topic is relevant and timely, and the modeling framework provides useful mechanistic insights into the interactions between drought and warming in tropical ecosystems. The manuscript is well-structured and addresses an important question in the context of future climate extremes. However, I have several concerns about the consistency between the figures and their interpretations, as well as the clarity of some key definitions and the conceptual framing. These issues should be addressed to improve the robustness and readability of the manuscript.

- We thank the reviewer for the positive assessment of this study. We also thank the reviewer for raising concerns about the consistency between the figures and their interpretation. In response to these comments, the revised manuscript will reorganize the figures and revise the corresponding discussion to improve their consistency and clarity, as detailed under each comment below. Where these issues affect more than one section, we will propagate the corresponding revisions throughout the manuscript, including the Results, Methods, figure captions, Abstract, and Discussion.

Major comments

- **Figure 2**

I find this figure somewhat difficult to interpret due to inconsistent axes and response variables across panels (a-c). While each panel individually provides useful information, their combined presentation obscures the underlying logic and makes cross-panel comparison challenging.

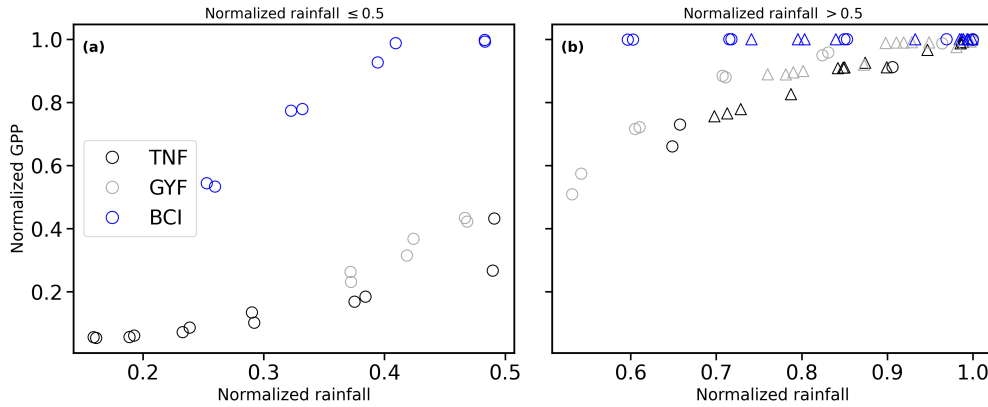
Specifically, panel (a) uses rainfall as the x-axis and GPP as the response, panel (b) uses soil water content and WUE, and panel (c) uses transpiration and NPP. These shifts in both explanatory and response variables make it difficult for the reader to follow the causal chain from water availability to ecosystem response.

I would suggest the authors consider the following improvements:

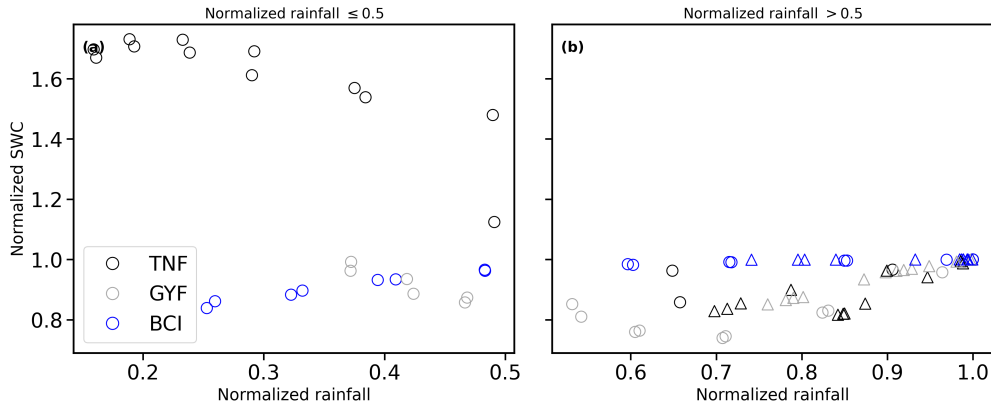
- A. Clarify the conceptual role of each panel in the main text, explicitly stating that panels (a-c) represent different stages of the water-carbon pathway (e.g., external forcing \rightarrow internal state \rightarrow flux response).
- B. Unify response variables where possible (e.g., use GPP consistently across panels, or explicitly justify the use of NPP in panel (c)).
- C. Consider reorganizing the figure, for example:
 - Using a consistent x-axis (e.g., soil moisture or rainfall) across panels, or
 - Splitting the figure into multiple figures with clearer thematic grouping (e.g., transition points vs. water-use strategies vs. flux relationships).
- D. In panel (c), the linear relationship appears to mask important nonlinear behavior (as noted for TNF). It would be helpful to explicitly highlight this in the figure (e.g., by showing the raw points more clearly, adding a breakpoint, or avoiding a simple linear fit).

Overall, improving the structural consistency of Figure 2 would significantly enhance readability and clarify the key model behavior.

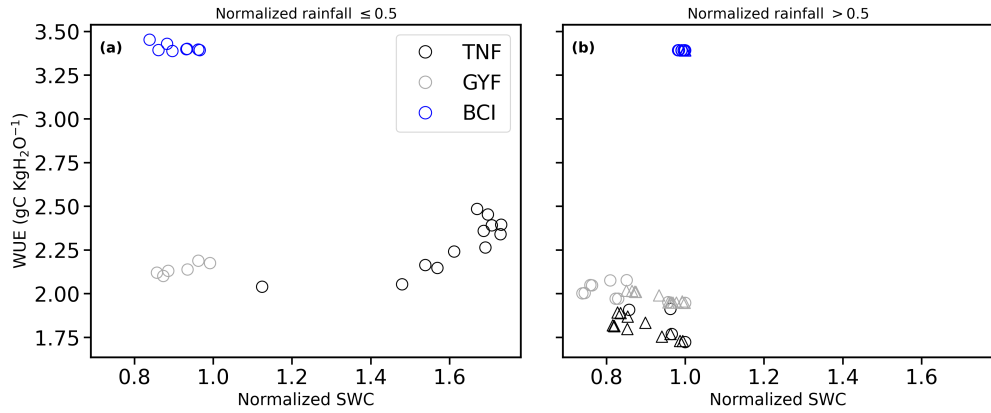
- We thank the reviewer for this helpful comment. We agree that the original drought figure combined different explanatory and response variables in a way that obscured the underlying logic. In the revised manuscript, we will separate this material into a clearer sequence from external forcing to ecosystem response. Figure 2 will show normalized GPP as a function of normalized rainfall, linking the imposed forcing to the ecosystem flux response. Figure C3 will show the relationship between normalized rainfall and normalized soil water content, clarifying how the modified forcing translates into water available in the rooting zone. Figure 3 will show water use efficiency as a function of normalized soil water content, emphasizing the internal physiological response. The transpiration-related relationship will be retained in Figure C4 as supporting evidence for the nonlinear behavior, and the main drought presentation will consistently emphasize GPP rather than a mixed GPP/NPP framing. We will also revise the corresponding text so that Figure 2 is described as the external forcing–response relation, Figure C3 as the link to internal water supply, and Figure 3 as the internal physiological response. Because the revised manuscript is not being submitted at this author-comment stage, we include the planned figure sequence below as response figures.



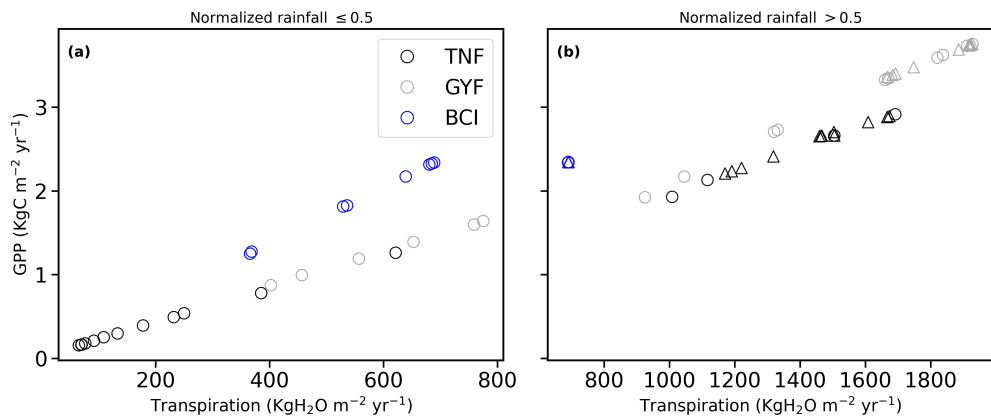
Response Figure R1. Planned revised main drought figure, to appear as Figure 2 in the revised manuscript. The figure shows normalized GPP as a function of normalized total rainfall for the yearly and daily drought scenarios, separating cases with normalized rainfall ≤ 0.5 and > 0.5 for clarity.



Response Figure R2. Planned supporting soil-water figure, to appear as Figure C3 in the revised manuscript. The figure shows how normalized rainfall reductions translate into normalized soil water content across sites and drought-scenario families.



Response Figure R3. Planned revised physiological-response figure, to appear as Figure 3 in the revised manuscript. The figure shows water use efficiency (WUE, defined as GPP divided by transpiration) as a function of normalized soil water content.



Response Figure R4. Planned supporting transpiration relationship, to appear as Figure C4 in the revised manuscript. The figure shows GPP as a function of transpiration and is retained as supporting evidence for the nonlinear behavior without mixing GPP and NPP in the main drought figure.

To make this new sequence explicit in the manuscript, we will add the following text to the Results section:

“Figure 2 shows the external forcing–response relationship for the yearly (triangles) and daily (circles) drought scenarios by plotting normalized GPP against normalized rainfall. Each point represents the mean across replicate simulations and over the first full stressed cycle after drought onset (30 years for BCI and 50 years for GYF and TNF; see Appendix B). For clarity, the results are shown separately for normalized rainfall ≤ 0.5 (Figure 2a) and normalized rainfall > 0.5 (Figure 2b); this division is intended only to improve visualization and does not imply a fixed ecological cutoff. Across all sites, GPP remained close to ambient under relatively high rainfall and then declined rapidly once rainfall fell below a site-dependent transition point.”

“Figure C3 provides the corresponding link between the external forcing and internal water supply by showing how rainfall reductions translate into water available in the rooting zone.”

“Figure 3 then shows the internal physiological response to that reduced water supply through the relationship between WUE and normalized soil water content.”

- **Figure 3**

The definition of the x-axis variable “T/Ta” in Figure 3 is unclear. The figure caption does not explicitly define this quantity, and it appears that different panels represent different temperature metrics (e.g., mean temperature vs. temperature range at various timescales).

This lack of clarity makes it difficult to interpret the magnitude of warming and to compare across panels. I suggest that the authors:

- A. Explicitly define “T/Ta” in the figure caption, including whether it refers to mean temperature, temperature range, or another metric.
- B. Use distinct labels for different temperature metrics if they are not directly comparable (e.g., mean temperature vs. diurnal or interannual range).

– We thank the reviewer for raising this point of confusion. We agree that the original shared temperature notation was too generic for panels representing different temperature metrics. In the revised manuscript, the temperature-perturbation figure will carry metric-specific labels: panel (a) will use \bar{T}/\bar{T}_a for mean temperature relative to ambient, whereas panels (b)–(f) will use $\Delta T_x/\Delta T_{x,a}$ for temperature-range metrics relative to ambient. Here, the subscript x identifies the timescale of the metric: daily (d), annual (y), 3-year ($3y$), 5-year ($5y$), and 7-year ($7y$). We will also revise the caption to define these quantities explicitly. The caption will read: *Impact of temperature-perturbation experiments on the three studied sites. Results are shown for BCI in blue, GYF in gray, and TNF in black. Squares denote the ambient case, the reference point at (1,1) in each panel. Each point is the yearly average over the first full stressed cycle after perturbation onset (30 years for BCI; 50 years for GYF and TNF) and is normalized to the corresponding ambient simulation at that site. Panels show normalized GPP as a function of panel-specific temperature metrics: (a) \bar{T}/\bar{T}_a , where \bar{T} is the mean temperature after a uniform temperature offset and \bar{T}_a is the ambient mean temperature; (b) $\Delta T_d/\Delta T_{d,a}$, where ΔT_d is the mean daily temperature range; (c) $\Delta T_y/\Delta T_{y,a}$, where ΔT_y is the mean annual temperature range; (d) $\Delta T_{3y}/\Delta T_{3y,a}$, where ΔT_{3y} is the mean 3-year temperature range; (e) $\Delta T_{5y}/\Delta T_{5y,a}$, where ΔT_{5y} is the mean 5-year temperature range; and (f) $\Delta T_{7y}/\Delta T_{7y,a}$, where ΔT_{7y} is the mean 7-year temperature range. For panels (b)–(f), each ΔT metric is computed as the difference between the high and low temperatures over the corresponding timescale and then normalized by its ambient value.*

- **Figure 5 & 6**

The patterns shown in Figures 5 and 6 appear to be inconsistent with the interpretation provided in the text, particularly for GYF.

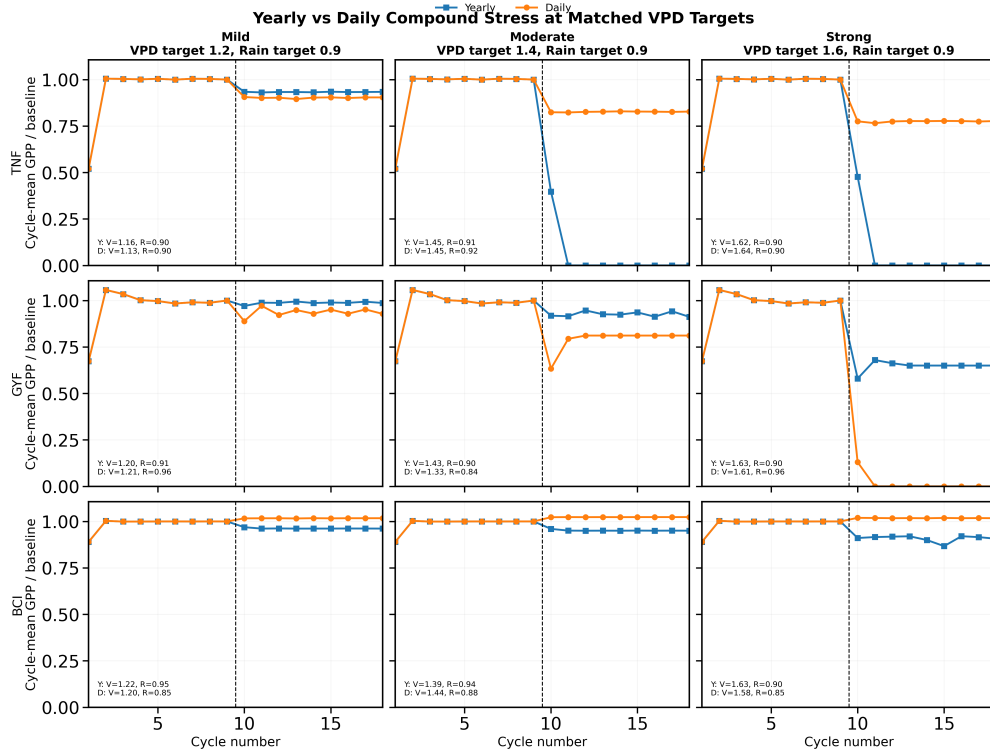
In Figure 5 (yearly-scale perturbations), GPP declines relatively gradually with increasing VPD, whereas in Figure 6 (daily-scale perturbations), the decline is more abrupt. This seems counter to the general expectation stated in the manuscript that longer-term stress should lead to stronger or more abrupt ecosystem responses.

Would the authors clarify this apparent discrepancy? Specifically:

A. Is the difference due to the way VPD or GPP is normalized or aggregated across timescales? If so, the normalization procedure should be clearly defined and justified, as it may affect the apparent sensitivity across figures.

B. Are the VPD ranges directly comparable between Figures 5 and 6?

– We thank the reviewer for raising this important point. We agree that the original presentation invited a direct comparison between the original compound figures that was not sufficiently justified. The two panels were constructed from different simulation families, and the sampled VPD ranges were not directly comparable. To address this, we generated a direct comparison between yearly and daily compound stress at comparable VPD targets. We will add this matched comparison to the revised manuscript as Figure C5, and for transparency we also include it below as Response Figure R5.



Response Figure R5. Direct comparison between yearly and daily compound stress at matched normalized maximum VPD targets. Columns correspond to mild ($\max(\text{VPD})/\max(\text{VPD}_a) \approx 1.2$), moderate (≈ 1.4), and strong (≈ 1.6) compound stress. Rows correspond to TNF, GYF, and BCI. In each panel, the plotted trajectories are the cycle-mean normalized GPP for the nearest yearly and daily compound cases at the target VPD level. This figure will appear as Figure C5 in the revised manuscript.

As shown in Response Figure R5, the timescale effect is site dependent. At matched VPD targets, TNF and BCI are more strongly affected by yearly compound stress, whereas GYF shows stronger reductions under the daily compound case. We therefore agree that our original wording was too broad. To avoid overgeneralizing from figures that sample

different forcing domains, the revised manuscript will use the more conservative statement that the effect of stress timescale is site specific rather than universal, and it will explicitly note that the revised compound figures are not directly comparable without the matched comparison in Figure C5.

We will also add an explicit Methods clarification to define the averaging-cycle procedure used in the drought, temperature, and compound figures:

“Ecosystem outputs were organized into complete forcing cycles, where one averaging cycle denotes one full pass through the site forcing record: 30 years for BCI and 50 years for GYF and TNF. All reported values were first averaged within these cycles. For the drought and temperature-perturbation figures, the plotted responses correspond to the first full stressed cycle after perturbation onset. In the compound figures, the left panels show the cycle-mean response for that same first stressed cycle, whereas the right panels show the sequence of cycle means across the stressed period. At BCI, this first analysis cycle corresponds to years 270–300; at GYF and TNF, it corresponds to years 450–500 within the repeated integrations.”

The revised compound-results text will state:

“Because the yearly and daily compound figures were generated from different compound experiment families, the sampled VPD–rainfall spaces are not directly comparable between the two figures. Figure C5 therefore provides a matched comparison between yearly and daily compound cases at similar normalized maximum VPD.”

“The compound experiments showed that the effect of stress timescale was site specific rather than universal.”

We will also revise the captions of the yearly and daily compound figures to state explicitly that the left panels show the first averaged cycle after perturbation onset, while the right panels show the sequence of cycle-averaged responses. The main text will explain that the forcing domains in the two figures are not intended for direct one-to-one comparison, and that clarification will be propagated throughout the manuscript wherever the compound-response interpretation is discussed.

- **Conceptual inconsistency between soil moisture limitation and VPD dominance**

The manuscript presents soil moisture limitation (Section 4.1) and atmospheric demand (VPD; Sections 4.2-4.3) as primary drivers of ecosystem instability, but the relationship between these two mechanisms is not clearly articulated.

In particular, Section 4.1 emphasizes cumulative soil water supply as the determinant of ecosystem decline, whereas Section 4.3 suggests that VPD dominates ecosystem response under compound stress. It remains unclear whether these are complementary processes or competing explanations.

I recommend that the authors clarify the conceptual framework linking water supply (soil moisture) and atmospheric demand (VPD), and explicitly state the conditions under which each mechanism becomes dominant.

- We thank the reviewer for raising this important inconsistency. We agree that the original manuscript did not clearly distinguish the mechanisms operating in the drought-only, temperature-only, and compound experiments. In the revised manuscript, we will separate these cases more explicitly and remove the earlier overstatement that the compound response could be described by a single VPD-dominated pathway.

In Section 4.1, the previous sentence is:

“Overall, these patterns show that soil moisture limitation is the primary driver of ecosystem instability in BiomeE.”

It will read:

“Overall, these patterns show that, in the drought-only experiments, ecosystem instability is governed primarily by declining soil moisture and the resulting reduction in water supply to plants. Because temperature and relative humidity are not perturbed in these simulations, the demand-side forcing through VPD is held approximately fixed by construction.”

In Section 4.3, the previous sentence is:

“Notably, in the yearly compound case, ecosystem response was dominated by VPD, with rainfall playing only a secondary role in determining GPP declines.”

It will read:

“These patterns indicate that compound responses cannot be reduced to a single VPD-dominated pathway; instead, they emerge from the interaction between reduced soil-water supply and increased atmospheric demand, with the balance between the two varying across sites and timescales.”

We will also revise the compound-results subsection, the Methods sentence describing constant relative humidity, the abstract, and the conclusions so that the same framework is used consistently throughout the manuscript. These additional changes will make clear that the compound response is not a single direct effect of only one driver, but instead reflects the simultaneous action of reduced water supply and increased atmospheric demand.

Minor comments

- **Definition of normalized variables**

The manuscript frequently refers to normalized variables (e.g., normalized GPP and normalized VPD), but the normalization procedures are not clearly defined. It would be helpful to explicitly state how these variables are normalized (e.g., relative to mean ambient conditions, baseline simulations, or maximum values), as this affects the interpretation of sensitivity across scenarios.

- We thank the reviewer for pointing out this confusion. As suggested, the revised manuscript will include a new subsection called “Experimental Design and Output Analysis”, which describes the number of simulations, the different experiment families, and the averaging and normalization procedures used to generate the plotted responses. This text will read as follows:

“Ecosystem outputs were organized into complete forcing cycles, where one averaging cycle denotes one full pass through the site forcing record: 30 years for BCI and 50 years for GYF and TNF. All reported values were first averaged within these cycles. For the drought and temperature-perturbation figures, the plotted responses correspond to the first full stressed cycle after perturbation onset. In the compound figures, the left panels show the cycle-mean response for that same first stressed cycle, whereas the right panels show the sequence of cycle means across the stressed period. At BCI, this first analysis cycle corresponds to years 270–300; at GYF and TNF, it corresponds to years 450–500 within the repeated integrations. For the stochastic drought ensembles, outputs were first averaged across the 20 realizations and then normalized against the equivalently averaged ambient case.”

“Unless otherwise stated, normalized quantities are expressed as X/X_a , where X denotes the value from the perturbed simulation and X_a the corresponding ambient value at the same site over the same analysis cycle. In the drought figures, normalized rainfall denotes total rainfall over the analysis cycle relative to the ambient total, whereas normalized GPP and normalized soil water content denote cycle-mean values relative to their ambient means. In the temperature-perturbation figures, normalized GPP is plotted against either mean temperature or the corresponding timescale-specific temperature range, each normalized by

its ambient value. In the compound figures, normalized rainfall and normalized maximum VPD are both expressed relative to their ambient values over the same analysis cycle.”

- **Definition of ”threshold.”**

The term ”threshold” is used throughout the manuscript, but it is not formally defined. Please clarify how thresholds are identified (e.g., visually, statistically, or based on a predefined criterion), as this would improve reproducibility and interpretation.

- We thank the reviewer for raising this concern. We agree that the word “threshold” was not well used in these parts of the manuscript, because no formal method was applied to identify such values and the term was only being used descriptively to indicate that a transition in the response was occurring. For this reason, we will revise the Results and Discussion sections to replace this wording with more appropriate terms such as “transition”, “onset of decline”, or “critical level”, depending on context. Some examples include:

Before:

“Across all sites, GPP remained close to ambient under relatively high rainfall and then declined rapidly once rainfall fell below a site-dependent threshold.”

After:

“Across all sites, GPP remained close to ambient under relatively high rainfall and then declined rapidly once rainfall fell below a site-dependent transition point.”

Before:

“At longer periods (seasonal to multi-year), warming persisted for months and produced similar thresholds across sites, aligning with the fourfold increase in temperature range identified in the Results.”

After:

“At longer periods (seasonal to multi-year), warming persisted for months and produced similar transition points across sites, aligning with the fourfold increase in temperature range identified in the Results.”

Before:

“Across sites, drought impacts emerged only once rainfall fell below specific thresholds, confirming that cumulative water supply determines when ecosystems begin to decline.”

After:

“Across sites, drought impacts emerged only once rainfall fell below site-specific critical rainfall levels, confirming that cumulative water supply determines when ecosystems begin to decline.”

We will retain the term “threshold” only where it refers to an explicit methodological cutoff, such as the 25% monthly-rainfall criterion used to classify dry months and the 20th-percentile criterion used to classify dry years.

- **Use of GPP versus NPP**

The manuscript switches between GPP and NPP across figures (e.g., Figure 2), but the rationale for this distinction is not clearly explained. A brief justification of when and why each metric is used would improve clarity.

- We thank the reviewer for raising this point. As noted in our response to the Figure 2 comment, we will revise the main drought figure set so that the manuscript consistently focuses on GPP rather than mixing GPP and NPP across panels. We chose GPP because it is the more direct ecosystem-scale carbon flux response to the imposed drought and

temperature perturbations, and it is also the quantity used throughout the main Results and Discussion to interpret changes in ecosystem function. In contrast, NPP additionally reflects respiration and allocation effects, which makes the comparison across panels less direct in this context. For this reason, the transpiration-related relationship will be retained only as supporting material, while the main text figures will use GPP consistently.

- **Figure captions**

Some figures (e.g., Figure 3) lack sufficient information in the captions to be interpreted independently of the main text. Improving figure captions to be more self-contained (e.g., defining variables and explaining axes clearly) would enhance readability.

- We thank the reviewer for this suggestion. We will revise the figure captions to make sure that they contain the key information needed for interpretation without requiring the reader to refer back to the main text. For example, the caption of the temperature-perturbation figure will read as follows: *Impact of temperature-perturbation experiments on the three studied sites. Results are shown for BCI in blue, GYF in gray, and TNF in black. Squares denote the ambient case, the reference point at (1, 1) in each panel. Each point is the yearly average over the first full stressed cycle after perturbation onset (30 years for BCI; 50 years for GYF and TNF) and is normalized to the corresponding ambient simulation at that site. Panels show normalized GPP as a function of panel-specific temperature metrics: (a) \bar{T}/\bar{T}_a , where \bar{T} is the mean temperature after a uniform temperature offset and \bar{T}_a is the ambient mean temperature; (b) $\Delta T_d/\Delta T_{d,a}$, where ΔT_d is the mean daily temperature range; (c) $\Delta T_y/\Delta T_{y,a}$, where ΔT_y is the mean annual temperature range; (d) $\Delta T_{3y}/\Delta T_{3y,a}$, where ΔT_{3y} is the mean 3-year temperature range; (e) $\Delta T_{5y}/\Delta T_{5y,a}$, where ΔT_{5y} is the mean 5-year temperature range; and (f) $\Delta T_{7y}/\Delta T_{7y,a}$, where ΔT_{7y} is the mean 7-year temperature range. For panels (b)–(f), each ΔT metric is computed as the difference between the high and low temperatures over the corresponding timescale and then normalized by its ambient value.*