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

Response to RC2

Reviewer 2

General comments

This paper reports a modelling experiments that aims at disentangling the effects of warming to that of rainfall changes, that are occurring under climate change. It also incorporates different times scales at which warming and rainfall changes occur. I found accurate and novel an approach that investigates warming and drought patterns across different time scales, from daily events to multi-year climate oscillations. I also found sound using a process-based model that incorporates both physiology and demography (BiomeE), as it has been showed previously that both are important in tropical wet forests. The choice of well documented sites is sound as well.

However I found too many major shortcomings to this study. May be they are mainly a matter of presentation, or the work itself is questionable. From the current state of the article I cannot say, as too much information is lacking.

- We thank the reviewer for the positive assessment of the study objective, the multi-timescale framing, the use of the BiomeE model, and the choice of well-documented study sites. We also appreciate the reviewer's concerns regarding the presentation of the study and the level of detail currently provided in the manuscript. As discussed in detail below, these comments identify important areas where the manuscript needs to be clarified and strengthened, particularly in the description of model setup, simulation design, and interpretation of the results. Where clarification alone is not sufficient, we will address the reviewer's concerns through additional targeted analyses and simulations. Overall, these comments are helping us improve the transparency, reproducibility, and presentation of the work, and to place the conclusions on a clearer and better supported basis.

Major comments

- There is a massive lack of information on model parameters. All parameters used need to be clearly stated, along with their sources. From what I could gather, I suspect that maximum leaf area index (LAI), and soil water holding capacity were unrealistic. Too low for LAI, and too high for soil water. Both are paramount for correctly simulating the water cycle.
  - We thank the reviewer for raising this important point. We agree that the original manuscript does not provide sufficient documentation of the parameterization, particularly for the hydrologic setup, and that this limits the reader's ability to evaluate the realism of the simulations. We also agree that the treatment of soil water storage requires closer scrutiny. Published site studies indicate strong spatial heterogeneity in soil properties at Paracou (Bonal et al., 2008; Epron et al., 2006; Van Langenhove et al., 2021; Sabatier et al., 1997)

and deep, well-drained clay-rich soils at Tapajos (Silver et al., 2000; Nepstad et al., 2002; Quesada et al., 2010). In addition, Longo et al. (2018) and its Supporting Information show that deep-soil assumptions can lead to overestimation of plant-available water, especially at GYF. We therefore agree that the current hydrologic treatment needs to be reassessed. Specifically, we will provide a complete parameter table with sources, clarify the distinction between the model trait LAImax and simulated stand-scale LAI, and perform targeted additional simulations using the updated model configuration with improved drainage and soil-water treatment to evaluate the sensitivity of the results to these assumptions.

- Some modelling choices were too simplistic. Using one plant functional type seems oversimplifying, given the predicted shifts in plant functional types under climate change in other studies (e.g. Levine et al. 2016 doi: 10.1073/pnas.1511344112; Longo et al. 2018 doi: 10.1111/nph.15185). Plus it seems contradictory with a model that is designed to handle multiple plant types. Another, probably more concerning choice, was to maintain relative humidity equal to observed climate under all scenarios. Variations in rainfall and temperature impact on relative humidity, which impacts on water vapor pressure deficit, which is one of the main drivers of forest response in this study.
  - We thank the reviewer for this important comment and agree that both the single-PFT setup and the treatment of relative humidity are simplifying assumptions. The original purpose of this study was to use BiomeE as an idealized sensitivity framework to isolate the effects of climate extremes on a representative tropical forest system, rather than to produce fully site-specific forecasts of compositional change. Nevertheless, we agree that the single-PFT setup limits the ecological realism of the analysis. To address this, we will extend the analysis to include simulations with two plant functional types, representing evergreen and deciduous strategies, in order to better capture functional heterogeneity and improve comparison with site characteristics and observations.
  - We also agree that prescribing relative humidity under all perturbation scenarios is an idealization that should be stated more clearly. In an offline model framework, the humidity response to imposed temperature and rainfall perturbations is not uniquely determined, because the model does not include two-way land-atmosphere coupling. One alternative would be to prescribe humidity changes from observations, reanalysis, or external climate-model output, but this would move the study away from its intended idealized and controllable sensitivity framework toward a more case-specific scenario design. Another possible assumption would be to hold actual vapor pressure fixed, but this is likewise an imposed idealization; it would generally intensify VPD increases under warming and would not resolve how atmospheric humidity should respond to rainfall perturbations. In the present study, relative humidity was held fixed so that temperature perturbations modify atmospheric demand through the Clausius-Clapeyron relationship while preserving ambient fractional saturation. For rainfall perturbations, however, holding relative humidity fixed also implies that near-surface atmospheric moisture is maintained independently of the local precipitation anomaly, rather than co-varying with it. This choice was made to isolate the effect of reduced rainfall on water supply to the soil-plant system without simultaneously imposing an additional humidity response. We therefore view this as a first-order idealized approximation, rather than a fully realistic prediction of atmospheric adjustment.
  - To make this assumption explicit in the manuscript, we will add the following paragraph to the Methods section:

    > Other meteorological variables, such as relative humidity and solar radiation, were kept constant to isolate the direct effects of rainfall perturbations on water supply and temperature perturbations on atmospheric demand. In this offline framework, the humidity response to imposed temperature and rainfall perturbations is not uniquely determined because two-way land-atmosphere coupling is not represented. Holding relative humidity fixed therefore constitutes an idealized forcing assump-

tion: it preserves ambient fractional saturation during temperature perturbations and, during rainfall perturbations, implies that near-surface atmospheric moisture is maintained independently of the local precipitation anomaly rather than co-varying with it. This choice was made to retain a simple and controllable sensitivity framework that separates first-order supply-side and demand-side effects, rather than to reproduce the full atmospheric adjustment to climate extremes.

- The paper does not present any convincing model test, nor does it refer to previous tests. The only one, that is presented in fig A1, is against gross primary production obtained from eddy covariance measurements, and predictions do not match observations over multiple years. This does not lend confidence in the model's ability to handle forest response to inter-annual climate variability. Eddy covariance data also include net ecosystem exchange, and evapotranspiration, which could also have been used for model testing. Why not? The chosen sites are well documented and present other datasets that could be used for further testing. Why not?
  - We thank the reviewer for this valuable comment and agree that the current manuscript does not present a sufficiently strong model evaluation. The comparison currently shown in Figure A1 was intended as a qualitative consistency check, not as a formal validation, and we agree that this was not communicated clearly enough. The purpose of this study is not to reproduce all observed interannual variability at the study sites, but rather to use a mechanistic vegetation model to examine the first-order effects of imposed climate extremes on tropical forest function. Achieving accurate prediction of year-to-year observed flux variability would likely require additional site-specific parameter constraint and data assimilation, which is beyond the scope of the present analysis.
  - Nevertheless, we agree that the baseline evaluation should be strengthened. In response to this comment, we will expand the comparison with available observations and site benchmarks, including evapotranspiration and net ecosystem exchange where observational coverage and model-output definitions are comparable, to better assess whether the model reproduces the mean state and broad flux behavior under ambient conditions, rather than relying on a single qualitative GPP comparison. We will also revise the manuscript text to make clearer that the current evaluation is intended as a consistency check for the ambient baseline, not as a claim of exact predictive skill for interannual variability.
  - We also agree that the presentation of Figure A1 and the initialization sequence required clearer explanation. The comparison shown in Figure A1 was not intended to display the initial adjustment of the model from year 1. Instead, it shows one equilibrated forcing cycle from the repeated ambient simulations, which is the same part of the integration used for the subsequent perturbation analysis. The different durations reflect the different forcing records used at each site: 30 years for BCI and 50 years for GYF and TNF. For this reason, the comparison is shown over years 270–300 for BCI and 450–500 for GYF and TNF within the repeated integrations, rather than from year 1. We also agree that the initialization sequence should be described in the Methods section rather than in the model-structure section, and that the original wording about changed "parameters" was misleading in this context. The perturbation scenarios were generated by modifying the meteorological climate forcing, not by changing internal model parameters. The revised manuscript will state: "Before any perturbations were applied, BiomeE was integrated under repeated ambient climate forcing until equilibrium was reached. Equilibrium was defined as a state in which interannual variations in carbon and water fluxes, such as gross primary production and transpiration, and forest structure, such as leaf area index, fluctuated around a stable long-term mean with no consistent trend. The perturbation experiments were then initialized from these equilibrated ambient states, with the imposed perturbations introduced after nine full forcing cycles (year 270 for BCI and year 450 for GYF and TNF)." The revised manuscript will also state: "These modified datasets were used as model forcing for post-equilibrium simulations to assess long-term impacts on forest structure, carbon and water fluxes, and ecosystem resilience."

- Then I found the simulation design was not well presented. It should state all scenarios, the number of simulations. Particularly I found the setup of the warming scenarios confusing. It apparently mixes periods of warming and cooling, without being very explicit about the latter. Obviously, introducing cooling severely hampers the realism of the predictions. Also, it is important to be able to compare those warming and drought treatments to the standard future climate projections under varying emission scenarios. For instance, the figures don't show us if warming is close to 2, 4, or more °C.
  - We thank the reviewer for pointing out this source of confusion and agree that the presentation of the temperature experiments needed to be clarified. In particular, the manuscript previously used the term "warming" too broadly. Among the temperature experiments, only the mean-temperature offset case corresponds to uniform warming in the usual sense and can be related directly to an absolute increase in degrees Celsius, as shown in the mean-temperature panel of the temperature-perturbation figure. The remaining temperature experiments are better described as idealized temperature-variability perturbations, because the Fourier-based amplitude modifications amplify variability at daily, seasonal, or multi-year timescales and therefore include both warmer and cooler phases. Their purpose is not to represent monotonic future warming trajectories, but to test how ecosystem response depends on the timescale of imposed temperature extremes.
  - We also agree that the simulation design should be described more explicitly. To address this, we will revise the Methods section to state the scenario families and the simulation structure more clearly, including the distinction between deterministic temperature perturbations and stochastic drought ensembles, and the number of realizations used for the daily and yearly drought experiments. We will also revise the terminology throughout the manuscript so that the study is framed consistently in terms of climate extremes and temperature perturbations, while reserving the term "warming" for the mean-temperature offset experiment.
  - The revised Methods text will state:

    > Four simplified temperature-perturbation scenario families were implemented to capture a range of thermal stresses across timescales: (i) daily temperature variability increase, (ii) seasonal temperature amplification, (iii) multi-year temperature oscillation increase, and (iv) mean temperature offset increase. Only the mean-temperature offset case corresponds to uniform warming in the usual sense. The variability scenarios (i)–(iii) instead amplify selected Fourier components of the temperature signal and therefore include both warmer and cooler phases while increasing the temperature range at the targeted timescale. Their purpose is to provide controlled temperature-extreme perturbations for testing timescale sensitivity, rather than to reproduce monotonic future warming trajectories.
    >
    > All simulations were initialized from equilibrated ambient states at each site and then forced with modified precipitation, temperature, or combined temperature–drought scenarios. Per site, the experiment set consisted of one ambient control together with drought-only, temperature-only, and compound temperature–drought perturbation families. The drought-only family comprised 16 daily dry-spell extension bundles with 20 stochastic realizations each, one seasonal dry-season extension family with 9 deterministic cases, and 12 yearly drought bundles with 20 stochastic realizations each. The temperature-only family comprised one mean-temperature offset sweep and variability-amplification sweeps at the daily, seasonal, 3-year, 5-year, and 7-year timescales; these temperature perturbations were deterministic, with the number of amplitude levels varying by site and timescale to span comparable stress ranges. The compound family comprised seven yearly compound bundles and seven daily compound bundles. This distinction between stochastic drought ensembles and deterministic temperature perturbations is important for

interpreting the number of simulations and the averaging procedures described below.

- The result section only focuses on forest scale fluxes such as GPP and transpiration. It does not show any prediction of some key variables that influence the forest response to warming and drought, such as LAI, or plant water potentials. This prevents a thorough understanding of the underlying mechanisms controlling the flux patterns. Has there been mortality, has there been shifts in LAI? Without going into further details, the analysis remains superficial.
  - We thank the reviewer for this important comment and agree that the current Results section focuses too strongly on ecosystem-scale fluxes and therefore does not yet fully expose the vegetation and hydraulic mechanisms underlying those flux responses. BiomeE does predict structural and demographic variables such as leaf area index (LAI) and mortality, and we agree that these should be shown more explicitly.
  - In the revised analysis, we will expand the Results to include additional diagnostics of canopy and vegetation state, prioritizing LAI and mortality and, where robust, hydraulic stress indicators such as plant water potential. Because we are also reassessing the hydrologic treatment and plant functional representation, we prefer not to overinterpret the current state-variable outputs before those additional simulations are completed. Our plan is therefore to retain the current drought and temperature figures as a first-order overview of flux sensitivity, while adding a more mechanistic figure set that links those flux responses to vegetation-state changes under drought-only, temperature-only, and compound forcing, with additional diagnostics included in the appendices as needed. We agree that this will make the manuscript substantially more informative and will better explain the vegetation-state and hydraulic mechanisms that shape the simulated forest responses.
- From what is presented in the current paper, I don't know if the outcomes would have been different with a big leaf model using prescribed LAI, than with a more refined model such as BiomeE with more physiology and demographics.
  - We thank the reviewer for this comment and agree that, in the current manuscript, the added value of using BiomeE over a simpler big-leaf framework is not yet demonstrated clearly enough. This is closely related to the previous point: by focusing mainly on fluxes, we have underused the model's explicit representation of demography, functional diversity, individual-based competition, size structure, and compositional dynamics.
  - In the revised analysis, we will address this more directly by using simulations with multiple PFTs and by showing how variables such as LAI and mortality evolve under the imposed stress scenarios. This will make clearer how the model's structural and demographic capabilities shape the ecosystem response to drought, temperature perturbations, and compound stress, and why these responses cannot be reduced to a framework with prescribed LAI and no explicit vegetation dynamics. We agree that making this distinction more visible is important for justifying the use of BiomeE in the study.

Minor comments

- **Line 43**

  please explain what ENSO means.
  - We thank the reviewer for noting this omission. We will define ENSO at its first occurrence in the Introduction as "El Nino–Southern Oscillation (ENSO)".
- **Lines 120-125**

  spin up runs need more details, and ideally should be in a separate section than the model description. Were they followed by warm up runs?

- We agree that the initialization procedure required clearer presentation. To address this, we will revise the opening of the "Experimental Design and Output Analysis" subsection in the Methods so that the spin-up and perturbation timing are described directly within the simulation-design description.
- The revised manuscript will state: "Before any perturbations were applied, BiomeE was integrated under repeated ambient climate forcing until equilibrium was reached. Equilibrium was defined as a state in which interannual variations in carbon and water fluxes, such as gross primary production and transpiration, and forest structure, such as leaf area index, fluctuated around a stable long-term mean with no consistent trend. The perturbation experiments were then initialized from these equilibrated ambient states, with the imposed perturbations introduced after nine full forcing cycles (year 270 for BCI and year 450 for GYF and TNF)." This change will make the initialization sequence explicit in the manuscript.

- **Line 123**

  the use of the term "parameter" can lead to confusion, here "climate forcing" seems more appropriate.

  - We agree and will revise this wording in the manuscript. Where the original phrasing could be interpreted as changing internal model parameters, the revised text will refer instead to modifications of the meteorological climate forcing used to generate the perturbation scenarios.

- **Line 127**

  using only one PFT seems contradictory with the model choice. What is the justification?

  - We agree with the reviewer. In the revised manuscript, we will state more explicitly that "The use of a single PFT was intended to isolate first-order ecosystem sensitivity to imposed climate extremes while limiting additional compositional degrees of freedom. However, it does not capture shifts in functional composition..." As discussed in our response to the related major comment above, we will address this limitation by performing additional simulations with more than one plant functional type.

- **Line 127-129**

  please provide all details of parameters. As it is, how the reader knows what value comes from what study?

  - We agree with the reviewer that the current manuscript does not document the parameterization with sufficient clarity. In the revised manuscript, we will add a complete parameter table with sources, clearly identifying which parameters are inherited from the default BiomeE tropical configuration, which were varied by site, and which were informed by site-based observations or literature values. This will allow the reader to trace the origin of the parameter choices directly, rather than inferring them from the narrative description alone.

- **Line 131**

  is it realistic to use the same hydraulic parameters for all sites? It seems contradictory for three sites covering a rainfall gradient. What data is there to back up this choice? A separate section on model parameter values and sources is needed.

  - We thank the reviewer for raising this point. We agree that applying the same hydraulic assumptions across sites spanning a rainfall gradient is too restrictive unless it is very clearly justified, and the current manuscript does not do that sufficiently. Our original intent was to retain a common hydraulic framework while varying a limited set of site-specific structural and physiological traits, so that the simulations could isolate first-order differences in climate response without introducing a full site-by-site reparameterization of all hydraulic traits. However, we agree that this choice needs stronger observational grounding because it directly affects water supply to plants.

  - In the revised manuscript, we will expand the site description and parameter presentation to document a site-dependent hydraulic gradient designed to better reflect the observational constraints at each site. We will present these assumptions, together with their data basis and the full parameter table, within the site-description / parameter-description part of the Methods rather than in a separate standalone section. We will also use this revised hydraulic treatment in the updated simulations.

- **Lines 132-135**

  Exactly how these parameters were "tuned" to match observational data?

  - We agree that the original wording was imprecise. No formal parameter tuning, optimization, or data-assimilation procedure was carried out. Instead, the representative tropical PFT was initialized from the default BiomeE tropical configuration, and a limited set of site-varying traits was adjusted using available site information and literature-based ranges so that the resulting parameterization remained broadly consistent with the observed structural and physiological differences among sites.
  - The revised manuscript will state: "For this study, BiomeE was configured with a single plant functional type (PFT), representing a representative evergreen broadleaf tropical tree, initialized from the model's default tropical PFT parameterization. No formal parameter tuning, optimization, or data-assimilation procedure was carried out. Instead, a limited set of site-varying traits was adjusted using available site information and literature-based ranges so that the resulting parameterization remained broadly consistent with the observed structural and physiological differences among sites." We will revise the wording to avoid implying a formal calibration procedure, while also noting that we will rerun the model with greater functional diversity at these sites by including more PFTs.

- **Lines 137-139**

  previous studies, though, have emphasized the importance of shifts in PFT composition under climate shifts.

  - We agree with the reviewer. As noted above, the lack of PFT-composition shifts is a limitation of the current single-PFT setup, and for this reason we will perform additional simulations with more than one plant functional type.

- **Line 146-148**

  source of dataset?

  - We thank the reviewer for pointing this out. We will revise the BCI site description in the manuscript to identify the observational source more clearly. The revised manuscript will state: "The BCI climate dataset used to drive BiomeE spans from 01 January 1991 to 31 December 2020 and combines long-term records of rainfall and temperature with additional meteorological variables (e.g., radiation, humidity); the associated BCI flux observations used in the ambient evaluation follow Detto and Pacala (2022)." We will also correct the BCI observational citation in the Figure A1 discussion so that the flux comparison cites Detto and Pacala (2022).

- **Lines 148-151**

  source(s) of data? Max LAI seems low for a tropical wet forest.

  - We agree that the manuscript currently does not distinguish clearly enough between observed ecosystem LAI and the model parameter maximum LAI. The value reported here is a PFT-level model parameter used in the current single-PFT configuration, not a direct stand-scale observational LAI estimate. In this configuration, the crown representation does not resolve vertical variation in leaf mass per area or leaf nitrogen within the crown, so a lower crown-level LAI was used to avoid unrealistically carbon-negative lower canopy leaves; this is a compromise between model complexity and performance. We will revise the manuscript to make this distinction explicit, provide the source of the associated BCI site traits in the parameter table, and reassess the BCI LAI parameterization in the updated simulations.

- **Lines 154-158**

  source(s) of data? Max LAI is low compared to measured LAI at Guyaflux (Bonal et al. 2008 report values well above 6).

  – We agree with the reviewer. As in the previous response, we will clarify in the revised manuscript that the listed maximum LAI is a model PFT parameter rather than a direct observed stand LAI, and we will document the source attribution more clearly in the site-description / parameter-description part of the Methods and in the full parameter table. We will also reassess the GYF LAI parameterization in the updated simulations.

- **Lines 163-166**

  source(s) of data? Max LAI seems low.

  – We agree with the reviewer. As explained above, the listed maximum LAI is a model PFT parameter rather than a direct observed stand-scale LAI estimate, and the lower crown-level LAI reflects the simplified crown representation used in the current configuration. We will document the source attribution more clearly in the site-description / parameter-description part of the Methods and in the full parameter table, and we will revisit the TNF LAI parameterization in the updated simulations.

- **Lines 167-172**

  what is the point of providing a test, if it then discarded as not fully relevant? I would argue that it is relevant, not sufficient, and rather strongly suggests that the model is not suitable to "examine the relative ecosystem responses to imposed climate stressors".

  – We thank the reviewer for raising this point. We agree that the ambient model evaluation is relevant and should be interpreted as an important baseline consistency check. However, the goal of this study is not to reproduce the full observed interannual variability in GPP, but to use an idealized process-based framework to examine first-order and relative ecosystem responses to imposed climate extremes. In that context, imperfect agreement with observed year-to-year variability does not by itself invalidate the sensitivity analysis, although it does indicate that the baseline evaluation should be strengthened. We will therefore expand the comparison with available observations and site benchmarks, rerun the relevant simulations with the updated model configuration, and add the vegetation-state and hydraulic analyses needed to provide a clearer assessment of the ambient simulations and of BiomeE's physiological and demographic suitability for this study.

- **Fig. 1**

  figure seems identical to a previous publication, please ensure proper reference. There is no mention of senescence/turnover of plant tissues in this figure, nor in the text. Surely mortality is not the only way tree C and N are transferred to the soil organic matter? Also full names of all acronyms and symbols should be provided in the legend. Not clear what the spatial representation is. I assume it is 1D (only vertical layers) but the figure suggests otherwise, please be more precise.

  – We thank the reviewer for this comment. The schematic used in the manuscript is an adapted version of the published BiomeE model schematic, and we agree that the source should be cited explicitly in the figure caption. We also agree that the model description and caption should be more self-contained and should clarify the role of tissue turnover, the meaning of the acronyms, and the spatial interpretation of the schematic.

- In the revised manuscript text, we will state: "BiomeE also integrates biogeochemical processes driven by both plants and soil microbes, with carbon and nitrogen transferred from vegetation to soil through both mortality and tissue turnover, especially from leaves and fine roots. The canopy representation is vertically resolved but horizontally implicit under the perfect plasticity approximation framework: cohorts are organized by height and crown area rather than placed in an explicit three-dimensional canopy."
- The caption will state: "Adapted from the BiomeE model schematic of Weng et al. (2019). (A) Vegetation dynamics: plant cohorts are organized into vertical canopy layers and linked to physiological processes including photosynthesis, respiration, transpiration, and allocation, together with demographic processes including reproduction, growth, and mortality. (B) Biogeochemical and hydraulic structure: the schematic shows plant carbon (C), nitrogen (N), and water pools, including leaves, sapwood, heartwood, fine roots, non-structural carbon (NSC), and non-structural nitrogen (NSN), together with soil biogeochemistry (BGC) pools. Fine litter (x_1), coarse litter (x_2), microbial biomass (x_3), fast soil organic matter (SOM; x_4), slow SOM (x_5), and mineral N represent belowground transfers associated with mortality, tissue turnover, decomposition, and nitrogen cycling. Blue arrows indicate plant-hydraulic and water fluxes, and the dashed line separates aboveground and belowground processes. The schematic represents a one-dimensional vertically resolved canopy-soil system rather than a fully spatial three-dimensional forest."

- **Line 175**

  in this section about the climate scenarios, it is not obvious that temperature and rainfall treatments are mixed and how. It should state how many scenarios there are, and also how they compare with current climate model prediction scenarios.

  - We thank the reviewer for pointing out this lack of clarity. We agree that the original presentation of the climate-scenario design did not describe the relationship among drought-only, temperature-only, and compound perturbations clearly enough. To address this, we will revise the Methods section to state explicitly that each site includes one ambient control together with drought-only, temperature-only, and compound temperature–drought perturbation families, and we will describe the ensemble structure and number of realizations for each family directly in the experimental-design subsection.
  - The revised manuscript will state: "All simulations were initialized from equilibrated ambient states at each site and then forced with modified precipitation, temperature, or combined temperature–drought scenarios. Per site, the experiment set consisted of one ambient control together with drought-only, temperature-only, and compound temperature–drought perturbation families. The drought-only family comprised 16 daily dry-spell extension bundles with 20 stochastic realizations each, one seasonal dry-season extension family with 9 deterministic cases, and 12 yearly drought bundles with 20 stochastic realizations each. The temperature-only family comprised one mean-temperature offset sweep and variability-amplification sweeps at the daily, seasonal, 3-year, 5-year, and 7-year timescales; these temperature perturbations were deterministic, with the number of amplitude levels varying by site and timescale to span comparable stress ranges. The compound family comprised seven yearly compound bundles and seven daily compound bundles."
  - We will also clarify that these scenarios are intended as idealized, controlled perturbations rather than direct reproductions of specific future climate-model trajectories. In this framework, only the mean-temperature offset experiment maps directly onto conventional warming magnitudes, whereas the variability-amplification experiments are designed to isolate the role of timescale-specific temperature extremes. We will therefore revise the terminology throughout the manuscript so that the study is framed consistently in terms

of temperature perturbations and climate extremes, while reserving "warming" for the mean-temperature offset case.

- **Line 200**

  I consider using a constant relative humidity an issue in this study, and in contradiction with the level of details in the changes in temperature and precipitations in the different climate scenarios.

  - We agree with the reviewer and refer here to our response to the related major comment above regarding the treatment of relative humidity. As described there, we will clarify in the manuscript that holding relative humidity constant is an idealized forcing assumption in the offline framework, rather than a fully realistic prediction of atmospheric adjustment, and we will explicitly discuss its implications for both temperature and rainfall perturbations.

- **Lines 208-210**

  I do not figure out how drought can be lengthened at the daily scale. Is it only by reducing the amount of rain on rainy days?

  - We thank the reviewer for raising this point. In the daily drought scenario, drought is lengthened by extending sequences of dry days in the daily rainfall series rather than by reducing rainfall intensity on days that remain wet. We agree that this needed to be stated more explicitly in the manuscript.
  - The revised manuscript will state: "In this daily extension, original dry days remain rain-free and any newly added days in the extended dry spell are also assigned zero rainfall for all 24 hours, while days that remain wet retain their original hourly rainfall values." This wording is intended to make clear that the daily drought treatment acts by adding zero-rain days to lengthen dry spells, not by weakening rainfall on the remaining wet days.

- **Lines 243-261**

  Water stress arises when the water resource does not meet the water demand. Trees can to a certain extent adapt their water demand by modifying their leaf area. Mortality can also reduce the water demand. Hence rainfall, soil water content, and transpiration are not direct indicators of tree water stress. I find strange and sorely lacking that biological water stress indicators, such as plant water potentials, which the model computes, are not provided in the results. The differences between sites largely depend on model parameters, which were not provided: stomatal conductance, soil water holding capacity, etc. An isohydric or anysohydric behavior is defined by plant hydraulic parameters. Therefore the sentence "suggesting a more anisohydric strategy" (l 252), is strange, especially since it is stated that the same hydraulic parameters were used at all sites (l 131).

  - We thank the reviewer for this important comment and agree that the current analysis does not provide sufficient hydraulic diagnostics to support terms such as "anisohydric strategy." The intent was not to make a formal classification of site behavior as isohydric or anisohydric, but rather to suggest a possible interpretation of the WUE and transpiration responses. However, we agree that even this interpretation is too strong without direct analysis of modeled plant water-status variables and hydraulic behavior, and especially without fuller documentation of the hydraulic parameterization. We will therefore revise the manuscript to remove this wording and interpret these responses more conservatively.
  - The revised manuscript will state: "In contrast, TNF exhibited increasing WUE as rainfall declined, with GPP maintained longer than transpiration, indicating that carbon uptake and water loss became increasingly decoupled under drying conditions in the current model configuration."
  - We will also revise the later discussion accordingly. The revised manuscript will state: "This behavior increases short-term carbon gain in the simulations, but may also elevate

hydraulic risk once soil moisture falls below a critical level."

  - We also agree that rainfall, soil water content, and transpiration alone are not sufficient indicators of biological water stress. BiomeE computes internal hydraulic variables and includes a plant-hydraulic representation that can retain hydraulic redundancy under stress. In the updated simulations, we will use this capacity and expand the presentation of hydraulic and vegetation-state diagnostics, including cohort-scale water relations, so that the mechanisms underlying the ecosystem responses can be interpreted more directly from internal plant behavior rather than from fluxes alone.

- **Fig. 2**

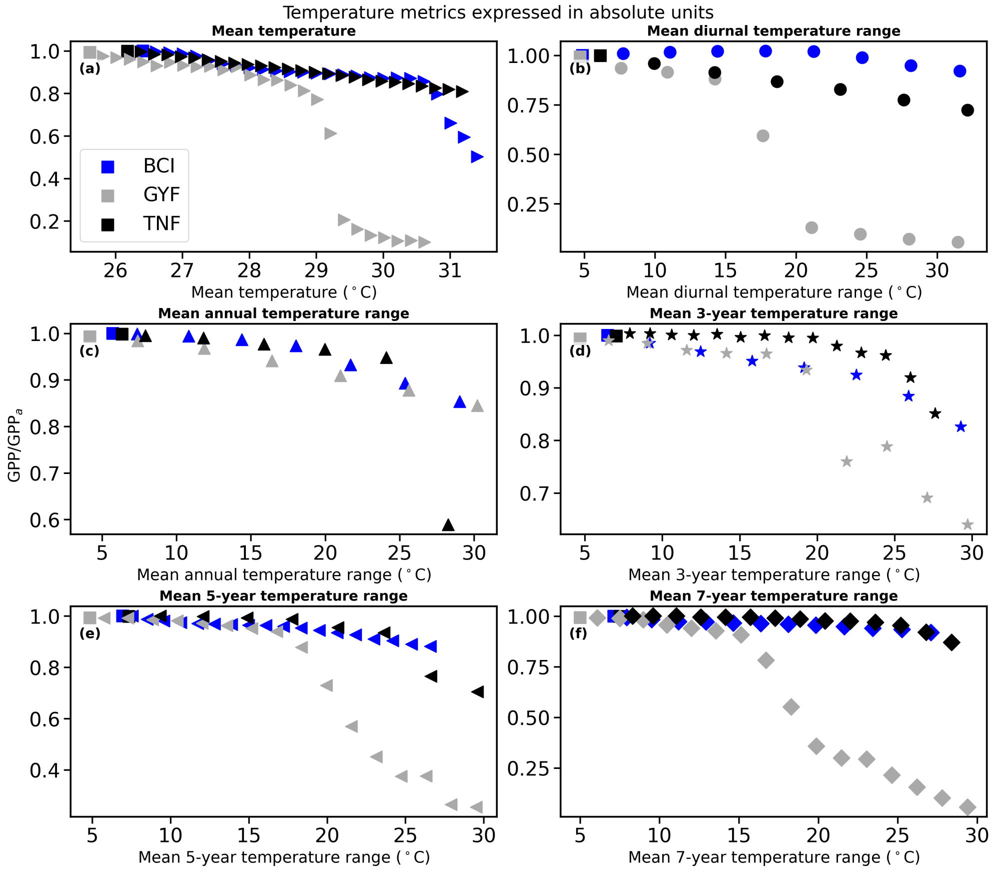
  Soil water and transpiration should have the same unit (e.g. mm) for better readability. Why displaying GPP in one case and NPP in another? What is the unit of WUE: g C of GPP or of NPP? Please make units more consistent. Soil water contents are very high, especially if they are multi-decadal averages. Can those soils actually hold that much water? Bonal et al. 2008 report max soil humidity of 840mm at Guyaflux.

  - We thank the reviewer for this comment and agree that the original figure was confusing because it mixed different variables and different units in the same place. In the revised manuscript, we will separate these relationships so that each figure shows one clearer part of the drought response. The main drought figure will show normalized GPP versus normalized rainfall, the soil-water response will be shown separately as normalized soil water content versus normalized rainfall, and the transpiration relationship will be shown using GPP rather than NPP.
  - The revised manuscript will state in the planned WUE/SWC figure caption (revised Figure 3): "Average water use efficiency (WUE, defined here as GPP divided by transpiration; gC kgH_2O^{-1}) as a function of normalized soil water content (SWC) ..."
  - We also agree that the large soil-water values point to a more important issue with the hydrologic setup. As discussed in our earlier response on parameterization, we are reassessing the soil-water storage treatment in the updated simulations, and we will revisit the soil-water diagnostics using that revised setup.

- **Fig. 3**

  I do not understand what "temperature range" means. Surely we are not talking about ambient temperatures increased 7 fold! But then, what T/Ta means? How do we compare those predictions to future climate changes of +2, +4, or +6°C? A more tractable index (e.g. increase in °C) would be easier to interpret.

  - We thank the reviewer for this comment and agree that the original temperature axes were not easy to interpret. We will therefore revise the manuscript caption to define the temperature metrics more clearly, and for additional clarity we include below a companion version of the figure in physical units (°C).
  - The revised manuscript will state in the temperature-perturbation figure caption: "Here, ΔT_d is computed as daily maximum minus daily minimum temperature averaged over the analysis cycle, whereas ΔT_y, ΔT_{3y}, ΔT_{5y}, and ΔT_{7y} are computed from the daily-mean temperature series as the difference between the maximum and minimum daily-mean temperature within the corresponding time window, and then normalized by the ambient value."



Response Figure R1. Companion version of the temperature-perturbation figure in physical units. Panel (a) shows mean temperature in °C. Panel (b) shows mean diurnal temperature range in °C. Panels (c)–(f) show annual and multi-year temperature-range metrics in °C, computed from the daily-mean temperature series over the corresponding time window.

- As shown in Response Figure R1, only panel (a) can be compared directly with familiar mean-warming values such as +2 or +4°C. The remaining panels do not represent a several-fold increase in absolute air temperature. Instead, they show idealized changes in temperature range at daily, annual, and multi-year timescales. We will keep this distinction explicit in the updated simulations and figure presentation.

- **Fig. 4**

  what is the maximum VPD? Daily, yearly, over the whole simulation?

  - We thank the reviewer for pointing out that this definition was not stated clearly enough. In the VPD/transpiration figure, the VPD metric is the maximum VPD reached within one full forcing cycle for a given temperature-perturbation case, expressed as the relative change from the corresponding ambient maximum at the same site, i.e. $(\text{VPD}_{\text{max}} - \text{VPD}_{\text{max},a})/\text{VPD}_{\text{max},a}$. For BCI this forcing cycle is 30 years, and for GYF and TNF it is 50 years. It is therefore not a yearly mean or yearly maximum diagnostic.

  - The revised manuscript will state in the VPD/transpiration figure caption: "Yearly-averaged normalized transpiration, computed over the first full stressed cycle after perturbation onset (30 years for BCI; 50 years for GYF and TNF), as a function of the relative change in maximum vapor pressure deficit, defined as $(\text{VPD}_{\text{max}} - \text{VPD}_{\text{max},a})/\text{VPD}_{\text{max},a}$, where VPD_{max} is the maximum VPD reached within one full forcing cycle for a given temperature-perturbation case and $\text{VPD}_{\text{max},a}$ is the corresponding ambient value at the same site." We will also correct the caption so that the ambient case is identified as the reference point at $(0, 1)$ in each panel.

- **Lines 284-285**

details on simulation design should all be presented in the methods section.

  - We agree with the reviewer and will move these details into the Methods section. In the revised manuscript, the *Experimental Design and Output Analysis* subsection will state the full simulation structure, including the separation between drought-only, temperature-only, and compound experiments, the ensemble sizes for stochastic drought runs, and the definition of the averaging cycles used in the figures.
  - The revised manuscript will state: "Per site, the experiment set consisted of one ambient control together with drought-only, temperature-only, and compound temperature–drought perturbation families..." It will also state: "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."

- **Fig. 5**

this figure shows simulation results at higher VPD than in fig 4 for BCI and GYF, why? I do not understand what the cycles mean, hence I do not understand the right panel at all. What was the rationale for selecting the simulations to be displayed in the right panel?

  - We thank the reviewer for this comment and agree that the original presentation of the compound figures was confusing. The VPD/transpiration figure was plotted against the relative change in maximum VPD, whereas the yearly and daily compound figures were plotted against normalized maximum VPD, and the yearly and daily compound figures were not shown over the same displayed VPD range. We also agree that the meaning of the cycle axis and the rationale for the selected scenarios in the right panels were not explained clearly enough.
  - To address this, we will revise the compound-figure presentation in two ways. First, we retain Figure C5 as the direct yearly-versus-daily comparison at matched VPD levels. Second, we will revise the main compound figures so that the yearly and daily experiment families are displayed over the same VPD range, making the left panels easier to compare visually.
  - We will also clarify the caption and the main text. The yearly compound figure caption will state: "Symbols (star, circle, and square) mark three representative climate scenarios in VPD-rainfall space (A, B, and C), selected to illustrate mild, intermediate, and strong compound stress. Right panels show the corresponding sequence of cycle-averaged GPP (kgC m^{-2} yr^{-1}) as a function of cycle number for the same three scenarios, using the same marker shapes as in the left panels to link forcing conditions to ecosystem responses through time." The Results text will state: "The left panels summarize the first averaged cycle after perturbation onset, showing how normalized GPP varies across the normalized VPD-rainfall forcing space. The right panels do not show additional forcing combinations; instead, they track the sequence of cycle-mean GPP values for the three representative scenarios marked in the left panels. One averaging cycle corresponds to one full pass through the forcing record, that is 30 years for BCI and 50 years for GYF and TNF."

- **Fig. 6**

same comments as fig 5.

  - We agree and will make the same clarifications for the daily compound figure. In particular, the revised caption will explain that the symbols in the left panels identify representative mild, intermediate, and strong compound-stress cases, and that the right panels show the sequence of cycle-averaged GPP values for those same scenarios through time. As noted above, the direct yearly-versus-daily comparison will be provided separately in Figure C5, while the main yearly and daily compound figures will be displayed over the same VPD range for easier visual comparison.

- **Lines 313-318**

  what mechanisms, within the modelling framework, explain those patterns?

  - We thank the reviewer for this comment and agree that the original discussion was too descriptive at this point. Within BiomeE, this transition in ecosystem response emerges because reduced rainfall progressively depletes the modeled rooting-zone water store. As soil water declines, plant water supply becomes more limited, hydraulic stress increases, and GPP begins to fall once the remaining water store is no longer sufficient to buffer drought. The site differences arise because the same fractional rainfall reduction maps onto different absolute water availability and different modeled water-use behavior across sites. We will revise the manuscript to make this mechanism more explicit. We also agree that this interpretation should be supported more directly with internal ecosystem diagnostics, and in the updated simulations we will show how hydraulic and vegetation-state variables change within the system so that these mechanisms are demonstrated more explicitly rather than inferred from fluxes alone.
  - The revised manuscript will state: "Within BiomeE, this transition in ecosystem response emerges because reduced rainfall progressively depletes the modeled rooting-zone water store; as soil water declines, plant water supply becomes more limited, hydraulic stress increases, and GPP begins to fall once the remaining water store is no longer sufficient to buffer drought."

- **Line 325**

  what means "still retains more water"?

  - We agree that this wording was too vague. Here we meant that, under the same fractional rainfall loss, GYF retains a larger modeled rooting-zone water store than TNF because it starts from a much wetter baseline climate. We will revise the manuscript to state this more explicitly. As noted above, we will also use the updated simulations to show how the internal hydraulic and vegetation-state variables evolve, so that statements about site differences are supported more directly by the modeled ecosystem behavior.
  - The revised manuscript will state: "GYF, even after losing 60% of its rainfall, retains a larger modeled rooting-zone water store than TNF because it starts from a much wetter baseline climate, allowing the ecosystem to persist longer in the simulations."

- **Line 327**

  what about variations in LAI? or mortality?

  - We agree and refer the reviewer to our response to the related major comment above regarding vegetation-state diagnostics. As noted there, we will expand the revised analysis to show how variables such as LAI and mortality evolve under the imposed stress scenarios. This will allow the manuscript to link the flux responses more directly to changes in canopy structure and demographic state in the updated simulations.

- **Lines 322-344**

  predictions of hydraulic failure occurrence, LAI variations, and mortality, are necessary to discuss the differences between sites.

  - We agree with the reviewer. As noted in our responses above, discussing the site differences only in terms of fluxes is not sufficient. In the updated simulations, we will expand the analysis to show how hydraulic failure-related diagnostics, LAI, and mortality change across sites and forcing scenarios, so that the differences among sites are supported more directly by the modeled ecosystem responses rather than inferred from GPP and transpiration alone.

- **Line 334**

  in simulations in which only rainfall was manipulated, then of course, soil moisture – fed by rainfall- will be the limiting factor...

  - We agree with the reviewer. In a drought-only setup where rainfall is the perturbed driver, it is expected that declining soil moisture will be the primary proximate limitation. Our intent was not to present this as an unexpected result, but rather to clarify that, within this experimental setup, the site differences arise from how comparable rainfall reductions translate into rooting-zone water availability and hydraulic limitation. We will revise the manuscript to make this point more clearly.
  - The revised manuscript will state: "Overall, as expected in the drought-only experiments, ecosystem instability in BiomeE is governed primarily by declining soil moisture and the resulting reduction in water supply to plants. Under this setup, the key site differences arise from how rainfall reductions propagate into rooting-zone water availability and hydraulic limitation."

- **Lines 334-335**

  plants hydraulic strategies and community structure actually define how soil water decreases.

  - We agree with the reviewer. As noted in our responses above, differences in hydraulic behavior and vegetation structure are important for interpreting how soil water declines and how ecosystem responses diverge among sites. In the updated simulations, we will therefore show the evolution of hydraulic and vegetation-state variables more explicitly so that these mechanisms are demonstrated directly.

- **Line 341**

  again, what about LAI variations and mortality?

  - We agree and refer the reviewer to our responses above regarding LAI, mortality, and hydraulic diagnostics. In the updated simulations, we will show how these vegetation-state variables evolve so that the differences among sites are supported more directly by the modeled ecosystem response.

- **Line 344**

  in the figures, I do not know where the "4°C above ambient" is.

  - We agree that this was not clear enough in the original presentation. The "4°C above ambient" statement refers specifically to the mean-temperature offset experiment. Because the main figure was plotted using normalized mean temperature, this level was not as easy to identify as it should have been. For clarity, we include Response Figure R1 above, where panel (a) shows the same experiment in °C, so the approximate +4°C critical level can be read directly. We will revise the main figure presentation accordingly, and we will refer explicitly to the mean-temperature offset panel in the manuscript discussion.
  - The revised manuscript will state: "Under the mean-temperature offset experiment, GYF was the most warming-sensitive: GPP declined rapidly once temperatures rose beyond roughly 4°C above ambient, leading to ecosystem collapse."

- **Line 355**

  a figure to illustrate "partial hydraulic and carbon recovery" would be useful.

  - We agree with the reviewer. As discussed above, the current manuscript does not yet show the internal state variables needed to demonstrate this recovery process directly. In the updated simulations, we will include additional diagnostics of both carbon status and hydraulic state, including non-structural carbon (NSC) and leaf water potential (Ψ_L), so that the partial recovery between successive hot–dry periods can be shown explicitly rather than inferred only from GPP and transpiration.

- **Lines 394-395**

  so LAI was fixed? Actually reduction in LAI also reduces plant water use, relieves water stress, and represents a major management mean to reduce water stress, so this statement about additional water loss when LAI decreases is more than questionable.

  - We thank the reviewer for this comment and agree that the original wording was incomplete. LAI was not fixed in BiomeE. Our point was not that a reduction in LAI necessarily increases total water loss, but rather that canopy loss can introduce competing feedbacks on the water balance that were not analyzed explicitly in the present discussion. On one hand, lower LAI can reduce transpiration demand and partly relieve plant water stress. On the other hand, it can increase radiation reaching the soil and potentially enhance soil evaporation and local heating. We will revise the manuscript to clarify this point. As noted above, we will also show LAI changes more explicitly in the updated simulations.
  - The revised manuscript will state: "Fourth, the simulations omit secondary feedbacks associated with canopy loss. Reductions in LAI can decrease transpiration demand, but they can also increase radiation reaching the soil and potentially enhance soil evaporation and local heating."

- **Lines 399-405**

  this positioning is hardly justifiable, given than those variables are key to understand the results presented in this paper.

  - We agree with the reviewer. These variables are important for interpreting the results and making the study more defensible, and we should not leave them only as a future direction. As noted in our responses above, we will therefore include these vegetation-state and hydraulic diagnostics in the revised analysis so that the mechanisms behind the reported flux responses are shown more directly.
  - We will also remove the previous wording that placed these variables only beyond the scope of the present study and left them for future work, so that the manuscript framing is consistent with this revised analysis plan.

- **Lines 407-417**

  a lot of factors can have an influence in the model response to climate change (fire, pest outbreak, species composition shifts, etc). One that is known to directly influence plant physiology, and thus possibly more relevant here, is air CO2, and the ongoing debate on the compensating effect of increasing CO2 in a warmer and drier climate. What about that?

  - We agree with the reviewer that atmospheric CO_2 is important because it can directly affect plant physiology and potentially modify ecosystem responses to warming and drought. In the main experiment set, CO_2 was held fixed so that the imposed rainfall and temperature perturbations could be interpreted in isolation. To address this concern while preserving the controlled design of the main experiments, we will add an additional elevated-CO_2 sensitivity simulation for selected warm–dry scenarios using the model's existing elevated-CO_2 forcing option. This additional comparison will evaluate whether elevated CO_2 modifies photosynthesis, transpiration water cost, and plant hydraulic stress relative to the fixed-CO_2 cases.
  - The revised manuscript will state: "Fourth, atmospheric CO_2 was held fixed, so the simulations do not capture the full range of potential CO_2–warming–drought interactions or their effects on photosynthesis, transpiration water cost, and plant hydraulic stress."

- **Line 437**

  Please be more accurate when citing sources, e.g. Bonal et al. 2008 does not provide data for the 2004-2014 period!

  – We thank the reviewer for this correction. We agree that Bonal et al. (2008) is a site-description paper and not the correct citation for the full 2004–2014 GYF eddy-covariance dataset used in Figure A1. We will therefore replace that citation with the corresponding FLUXNET2015 GF-Guy dataset reference.

  – The revised manuscript will state: "Figure A1 presents the annual Gross Primary Productivity (GPP) at the Paracou Field Station in French Guiana (GYF) and Barro Colorado Island (BCI), based on both observed eddy covariance data (Bonal and Burban, 2015; Detto and Pacala, 2022) and BiomeE simulations."

- **Lines 438-443**

  I do not agree, see comment for Lines 167-172 above. Model testing also need statistics (R2, RMSE, bias...) to be interpreted.

  – We agree and refer the reviewer to our response above to the related comment on Lines 167–172. As stated there, the ambient comparison is relevant and should be strengthened, even though the main goal of the present study is to examine relative ecosystem responses rather than to provide a full predictive validation. In the updated simulations, we will therefore expand the comparison with available observations and include summary statistics, such as R^2, RMSE, and bias, so that the realism of the ambient baseline is documented more clearly.

- **Lines 495-499**

  If I understand correctly (although I do not understand the concept of the replicates described here), simulations for GYF, TNF, and BCI covered 900, 900, and 540 years, respectively. From those years, the 450-500, 450-500, and 270-300 years were used for presenting the results, respectively. So my question is: what was the point of simulating the remaining, 400, 400, and 240 years?

  – We thank the reviewer for this question and agree that the rationale was not explained clearly enough. The first stressed cycle was used for the main left-panel summaries because it isolates the short-term ecosystem response to the imposed forcing. The simulations were continued for additional cycles because the forcing pattern repeats, but the ecosystem state does not: mortality, biomass, and canopy structure can continue to change from one cycle to the next. This longer integration is therefore needed to assess whether the system stabilizes at a lower-productivity state or continues toward collapse. This long-term behavior is already illustrated in the right panels of the compound figures, where some cases remain in a reduced but persistent state while others decline further under repeated stress.

  – We will revise the manuscript to make this distinction more explicit. The revised manuscript will state: "The first stressed cycle was used for the main cross-scenario comparisons because it isolates the initial ecosystem response, while the later cycles were retained to assess longer-term outcomes, including delayed mortality, structural reorganization, and whether the system approached a lower-productivity equilibrium or continued toward collapse." It will also state: "These multi-cycle trajectories show that some cases settle into a reduced but persistent productivity state, whereas others continue toward collapse under repeated stress."

  – As discussed above, in the updated simulations we will also expand the long-term diagnostics so that these later structural and hydraulic changes are documented more explicitly.

- **Fig C2**

  From the text I found hard to understand what the warming scenarios actually implied. From this figure it seems that warming is mixed with cooling. I beg to understand what the cooling is doing in a study on drought and warming, and I can only think it compromises the realism of the predictions.

  - We agree that this needed to be explained more clearly, and we refer the reviewer to our response above regarding Figure 3 and the temperature-scenario design. As clarified there, only panel (a) represents a mean-warming experiment, whereas the other temperature experiments were constructed to amplify timescale-specific temperature variability and therefore include both warmer and cooler departures from ambient by design. We will revise the manuscript text and caption accordingly so that these cases are described as temperature-variability or temperature-extreme experiments rather than as uniform warming scenarios.