Response to Reviewer 1's Comments

We thank the reviewer for carefully reading our manuscript and for the constructive and detailed comments, and for recognizing the novelty of this approach to NEP data. We address each comment below. The reviewer comments are shown in italic below. Section numbers in our replies refer to the revised manuscript, and supplementary figure numbers in our replies refer to the revised Supplementary Information (see the table at the end).

Reviewer #1

Specific remarks:

Line 14: What do you mean with 'emergent property' here?

> Thank you for this question. By emergent property we mean a pattern of the NEP time series (a result of multiple interactions among ecosystem processes and environmental drivers) that is not observable from any single fluctuation but becomes evident from analyzing the behavior of multiple fluctuations. We have clarified this in the abstract.

Line 18: The authors mention that every year about 14% of the carbon in the atmosphere is exchanged with the terrestrial biosphere. For this claim they refer to Lian et al. (2023). Could you please point me to the precise location in the Lian et al. manuscript where this info can be obtained from?

> Thank you for pointing this out. The 14% actually comes from Zhang et al. (2015), which was accidentally replaced with Lian et al. (2023) during the various revisions before submission. We now use the estimate from Keenan and Williams (2018), who report that plants take up about one fifth of the carbon in the atmosphere each year through photosynthesis (Section 1). We also rewrote the sentences that follow, and we believe there is now a smoother transition from GPP and RECO to NEP.
>
> Zhang, L., Sun, R., Xu, Z., Qiao, C., and Jiang, G.: Diurnal and seasonal variations in carbon dioxide exchange in ecosystems in the Zhangye Oasis area, Northwest China, PLoS ONE, 10, e0120660, https://doi.org/10.1371/journal.pone.0120660, 2015.

Line 19: Personally, I find the sequence here somewhat confusing and perhaps incorrect. At line 18, the exchange flux that is referred to is the gross exchange – the authors mention about 14% of the carbon in the atmosphere is exchanged with the terrestrial biosphere every year. And at line 19 the authors refer to this exchange as Net Ecosystem Productivity…

> Thank you for this comment. We have rewritten the two sentences to separate gross uptake by photosynthesis from net ecosystem productivity, and to state that NEP is the net balance between gross primary productivity and ecosystem respiration.

Line 27: 'the shape of the daily NEP trajectory ... ecophysiological controls that cannot be captured by aggregated totals alone (Dietze et al., 2011; Braswell et al., 2005; Wehr et al., 2016)' This statement is fine, but I do wander whether the cited references are appropriate citations for this statement. If so, could you please give a brief explanation (or indicate specific

sections of the cited manuscripts) showing these are appropriate citations. The same holds for the references on lines 30 and 31.

We thank the reviewer for the opportunity to clarify this.

Dietze et al. (2011) found that model error has a clear peak at the diurnal scale, mostly in the growing season, and that it depends on how photosynthesis and phenology are represented (Abstract). Braswell et al. (2005) showed that eddy flux data constrain model parameters best at daily to seasonal scales but are not useful for annual integrals (Abstract). Wehr et al. (2016) found that respiration is lower during the day than at night, so ignoring this overestimates photosynthesis (Abstract). Together they show that the diurnal and sub-annual shape of the fluxes carries information that annual totals miss.

Lasslop et al. (2010) point out that the daily NEE cycle is often asymmetric on warm, dry days, and that adding a VPD limitation to photosynthesis greatly improves how well it is reproduced (Results and discussion, VPD limitation of the light response curve). Huang et al. (2021) explain that biological and physical factors, including stochastic rainfall, interact in ways that make heterotrophic respiration hard to predict (Abstract and Introduction). Reitz et al. (2021) note that machine learning can capture highly nonlinear relationships (Introduction), and their Table 2 lists predictors such as precipitation, humidity and wind speed (Section 2.3). Together they show that NEP depends on many interacting biological and environmental drivers, and that these relationships can be highly nonlinear.

Schimel et al. (2019) describe how diverse plants and soil microbes are (Section II.5) and how much models simplify this diversity (Section IV). This helps explain why NEP is difficult to interpret and to predict.

Line 34: I suggest to slightly rewrite the sentence 'GPP and RECO respond differently to environmental drivers with different sensitivities and on different timescales throughout the day...', to make it more clear. Currently it is not fully clear whether you mean 'sensitivities of the drivers' or sensitivities of GPP and RECO. Next to that, it is not fully clear to me to what 'Their' refers on line 36.

We thank the reviewer for pointing out the ambiguity. We meant the sensitivities of GPP and RECO to the environmental drivers, so the sentence now reads, "GPP and RECO exhibit distinct sensitivities to environmental drivers, resulting in different diurnal dynamics" (Lines 42–44). At Line 36, we replaced "Their" with "The", so that the temporal asymmetry of NEP is now the subject of the sentence.

Line 35: Similar to one of my previous remarks, may I ask for some information on why you refer to Braswell et al. (2005) here? And to Rastetter et al. (2023) on line 64?.

Braswell et al. (2005) is cited at Line 35 because their model gives photosynthesis and respiration separate parameters, including separate temperature sensitivities (Q10) for vegetation and soil respiration (Methods, Table 1). They also show that how well eddy covariance data constrain these parameters depends on the time scale, with night-time data informing only the respiration parameters (Methods) and the parameters most strongly constrained at daily to seasonal time scales (Abstract). This supports our

statement that GPP and RECO differ in their sensitivities to the drivers and in the time scales on which they respond.

Rastetter et al. (2023) is cited at Line 64 because they show, with a process-based tundra model, that day-to-day weather variability affects ecosystem function independently of extreme events, since many processes respond nonlinearly to environmental drivers (Abstract; Introduction). They also note that studies of climate variability usually focus on extreme events and that few have analyzed day-to-day variability. We have revised the sentence in the Introduction to say that the regime-dependent effects of environmental variability remain less explored.

Line 77: 'Within the context of NEP variability these approaches provide effective measures' ... Do you use these approaches (theories for non-equilibrium fluctuations, fluctuation theorem) in this manuscript? When reading this paragraph I would expect so, but from the methodology section I tend to think you do not? Perhaps this can be made more clear? Please keep in mind that I am not familiar with the fluctuation theorem etc., which might also (partly) explain why this looks somewhat contradictory to me.

Yes, we do use one of these approaches. In particular, we follow the methodology developed by Paneni et al. (2006), where they analyze fluctuations in net energy flux, which for us corresponds to NEE. This is described in Section 2.2, where we state that the approach is directly related to irreversibility in statistical physics (Paneni et al., 2006, 2008; Porporato et al., 2011).

Line 86: Could you perhaps explain what you mean with 'with no reference to either deterministic or stochastic components'? In case you think the explanation is relevant to the reader, you can perhaps include the explanation in the text..

Thank you for this suggestion. We mean that any deviation of NEP from its long-term mean is considered a fluctuation, without any particular decomposition of the time-series (e.g., one could remove average diurnal and seasonal cycles, leaving the remaining fluctuation as "stochastic" variability). We do not separate these components, because the diurnal and seasonal cycles are an important part of the ecosystem response to daily and seasonal environmental forcing, and they are a part of what we aim to study here. We have added this explanation to the end of the Introduction.

Line 87: This is not fully clear to me. When referring to 99th quantile of NEP, to what extent are the values aggregated, i.e. hourly NEP values, minutely values,…? Reading on in the manuscript I found out that the data are half hourly. Do you pool all the half hourly NEP values over the years, and calculate the 99th quantile from this data set? If this is indeed what you did, doesn't this mean that a lot of the summer midday values (even those of relatively normal summer days) end up in the 1% most extreme NEP values? Perhaps this can be made somewhat more clear.

Thank you for raising this point. Yes, the 99th quantile is calculated separately for each site by pooling all half-hourly NEP values over the three-year study period, without separating them by season or time of day.

We agree that this would mean high midday NEP values during the growing season, including those from otherwise typical summer days, can fall within the top 1%. This is consistent with our definition of extremes, which refers to statistically high NEP values relative to the full distribution at each site, rather than unusual climatic events or anomalies relative to seasonal conditions. The threshold is then used to identify extreme fluctuation peaks, with 81 to 378 peaks identified per site (Table S5 of the Supplementary Information), and these peaks are then compared with all peaks.

We have clarified how the threshold is calculated and how these extreme fluctuations are defined in the Introduction and Section 2.2.

Line 87: What do you mean with 'without reference to climatic extremes'?

Thank you for this comment. We clarify that the extremes are defined purely statistically, as the upper 1% of the NEP distribution of each site. We clarified this because in hydrology and earth system science the term "extremes" often refers to hydroclimatic extremes. They are identified from NEP values alone whether or not they occur under extreme climatic conditions such as heatwaves or droughts.

Line 96: 'that met a set of criteria as of June 6th, 2024' … Do these sites meet the criteria throughout the whole period for which data was used?

Thank you for this question. The criteria were used to select sites from the FLUXNET2015 dataset, and our analysis covers the three-year period specified for each site in Table S1. During these periods, we used the reference NEE product (NEE_VUT_REF), which provides a continuous time series through gap-filling. For soil water content, we selected sites whose data have a quality flag ≤ 1, so the averages are based on measured or high-quality gap-filled data. We have also clarified the study periods in Section 2.1.

Line 105 vs 115: Perhaps check whether there is information that is unnecessarily repeated.

We agree. The friction velocity filtering and gap-filling of the ONEFlux pipeline were described twice. We removed them from the sentence on the NEE_VUT_REF product (Line 105), so that the processing is now described only once, in the QA/QC paragraph.

Line 110: You indicate that you downloaded respiration and GPP data. Where in your analysis do you use the GPP and RECO data? Or do you only use the NEE data?

We thank the reviewer for this question. The quantitative asymmetry analysis is based on NEP calculated from NEE_VUT_REF (NEP = −NEE), while GPP and RECO were downloaded and used to help interpret it. We have now made this explicit by analyzing the GPP and RECO average paths around the NEP peak (Fig. S2, SI), which show that GPP carries most of the temporal asymmetry of NEP, that is, the difference between NEP after and before the peak. Because GPP and RECO are model-based partitions of NEE, we use the nighttime partitioning and compare the results with the daytime partitioning (Fig. S2c,d). We state this in Section 2.1.

Line 111 and 102: Is it true that for variables other than SWC and soil temperature, you did not restrict the selection to records with quality flag ≤ 1? If not true, these sentences are written somewhat incorrect or ambiguous in my opinion. For example, line 102 can be read as 'soil temperature and soil water content (...) data that were gap-filled only with high confidence...'. Thus, the 'that were gap-filled' referring to soil temperature and SWC. But the sentence was perhaps rather meant to indicate that the full dataset is restricted to carbon flux records that were gap-filled only with high confidence. A similar issue applies to Line 111. If it are only SWC and soil temperature that are restricted to quality flag ≤ 1, is there a specific reason why you did this only for those variables? It would be good to make sure that the sentences from Line 102 and 111 are clear.

> Thank you for pointing out this ambiguity. We selected sites with soil temperature and soil water content measured at a depth of at least 10 cm, and applied the MDS quality criterion (QC ≤ 1) only to these two variables (TS_F_MDS_1 and SWC_F_MDS_1) during site selection, so that the soil data consist of measured or good-quality gap-filled values. We applied it to these variables because they are gap-filled only by MDS and have no ERA-Interim backup, so we used the quality flag to exclude low-confidence fills (Pastorello et al., 2020). For the other variables, we used the standardized FLUXNET2015 products directly, without an additional quality-flag criterion. NEE_VUT_REF is the reference NEE product of the ONEFlux pipeline, and TA_F, SW_IN_F, VPD_F and P_F are consolidated products that combine measurements or MDS gap-filled data with ERA-derived estimates. We have revised both sentences in Section 2.1 to make this clear.

The heading of Section 2.1 (Data and site selection) might not fully cover the content, i.e. the method of quantifying asymmetry is also discussed here. This can perhaps be made more clear in the heading.

> We agree and have split Section 2.1 into two parts, "Data and site selection" and "Quantification of temporal asymmetry".

Line 127: Does irreversibility here refer to the fact that the fluctuation path cannot be followed in reverse order? I assume it does not refer to the state of the system being permanently altered, as NEP can in principle return to pre-fluctuation values?

> Thank you for pointing this out. We agree that this distinction is important. In Section 2.2 we now state that irreversibility refers to the statistical difference between a fluctuation path and its time-reversed counterpart. It is not a direct thermodynamic interpretation. That is also why we do not emphasize this analogy. It does not imply a permanent change in the ecosystem state, as NEP can return to its pre-fluctuation values, which is reflected in the system's tendency to return to equilibrium described in the same section.

Line 128: 'applies' or 'is applied'? Does the approach really not apply to the partitioned components, or is it just better to apply it to NEE directly because of the potential biases in GPP and RECO?

Thank you for pointing this out. The approach is applied to NEP directly because GPP and RECO are model-based partitions of NEE, and flux-partitioning methods rely on additional assumptions that can introduce method-dependent biases (Lasslop et al., 2010; Reichstein et al., 2005). We have revised the sentence in Section 2.2 to say this. Also, while the statistical approach can in theory be applied to any time-series, it has a more direct physical meaning and connection to irreversibility when applied to net fluxes (Paneni et al., 2008).

Line 130: I very much like the extensive explanation that is given about the fluctuations, and the illustration of the asymmetry calculation procedure in Figure 1.

We appreciate the reviewer's positive feedback on the explanation of fluctuations and on Figure 1.

Line 137: What do you do when two NEP peaks are present (crossing the threshold) temporally close to each other (e.g. half an hour apart) within the same 12h time period? Do you then calculate 2 fluctuation paths and count both paths as two fluctuations (although largely overlapping) for your statistics? Or do you simply take the highest peak within the 12h time period to calculate a single fluctuation path?

Thank you for raising this point. If two peaks above the threshold occur within 6 hours of each other, we keep only the higher one and build a single fluctuation path from it. Each kept peak is counted once, and its path covers 6 hours before and 6 hours after the peak. This is described in Section 2.2.

Caption Figure 1: Is the description given for (b) fully accurate? The authors mention 'Sample half-hourly NEP time-series over one day, showing a large fluctuation (red) that exceeds the NEP threshold and a smaller fluctuation (black) that does not.' The total fluctuation path is shown in red here and the rest in black, isn't it?

Thank you for this comment. In panel (b), the red segment is the fluctuation path, the window of width 2τ centered on the peak, and the remainder of the day is shown in black. We have corrected the caption accordingly.

Caption Figure 1: 'Folded fluctuation path showing the rising and falling limbs overlaid about the peak' … Perhaps this is not fully accurate. I see the rising limb (blue full line), the falling limb (orange line) and the rising limb mirrored around the peak (blue dashed line). Maybe good to make this clear?

Thank you for this helpful observation. We agree and have revised the caption of Figure 1d to describe the three lines explicitly: the rising limb (solid blue), the falling limb (orange), and the rising limb mirrored about the peak (dotted blue).

Caption Figure 1: The step that is not fully clear in the Figure is the normalisation of NEP values applied in Eq. 2. Perhaps you can make it clear in the caption that in subfigures (c) to (e) the NEP values are normalised with peak NEP.

We have clarified in the caption of Figure 1 that the NEP values in panels (c) to (e) are normalized by the peak NEP, as in Eq. 2, and removed the units from their y-axis labels.

Figure 1e: whitespace missing in the y-axis title and check upright vs cursive part of axis title.

Thank you for noticing this. We corrected the y-axis title of Figure 1e by adding the missing spaces between the units and using the same upright font for the whole title.

Line 146: Maybe the flow of the text would be more clear if you would first show Eq. 3 and introduce how you calculate the asymmetry, before you mention the averaging. But this is merely a suggestion, if you think it is more clear the way it is written now then I'm of course fine with it.

We agree. We have reordered Section 2.2 to first define the asymmetry coefficient (Eq. 3) and then describe how it is averaged across fluctuation paths (Eq. 4).

Line 148: I understand that random errors in the timing of the peak will cancel out when averaging. But what if you have a single data point that peaks above the threshold due to a measurement error (e.g. in the evening)? I assume that will than be an erroneous fluctuation that will be included in the statistics? Also at this point in the manuscript I do not understand over what you average? Over all identified fluctuations in the three year periods? Perhaps this can be made somewhat more clear.

Thank you for raising this point. A single erroneous value would have to be very large to exceed the threshold, which for the top 1% is the 99th percentile of NEP at each site. Isolated spikes in the half-hourly NEE data are also limited by the spike detection and quality-control procedures of the FLUXNET2015 processing, as described in Section 2.1. We have also clarified in Section 2.2 that the average in Eq. (4) is calculated over all N fluctuation paths identified at each site over the three-year study period, separately for the top 1% extreme fluctuations and all fluctuations (c = 0).

Line 155: 'for each fluctuation path (or directly for the average fluctuation path) we folded...' When do you calculate it using the average fluctuation path directly and when do you calculate it for each fluctuation path? For the results in Figure 3, did you use the average directly? Can the reader derive this from the context?

Thank you for this comment. We calculate the asymmetry coefficient for each normalized fluctuation path (Eq. 3) and average it over all paths at a site (Eq. 4), which gives the same result as taking the difference between the falling limb and the mirrored rising limb of the mean fluctuation path. Figure 3 shows the mean fluctuation paths and

the resulting asymmetry coefficients. We have removed the unclear wording in Section 2.2 and stated the calculation in the Figure 3 caption.

Line 165: Given that you indicate here that you analyzed max of At for each fluctuation path, should the overbar be removed here? Or should the overbar be kept and the text slightly changed instead? A similar issue occurs for step 4.1 in Table S4, there is no reference to averaging.

Thank you for noticing this. Here A^{max} is the peak asymmetry of an individual fluctuation path, and the overbar is its mean over all fluctuation paths at a site, which is the quantity used in the cross-site analysis. We have defined both in Section 2.2, stated that A^{max} is the value of A_t with the largest magnitude along the path, and added the averaging step to Tables S3 and S4.

Line 176: Why necessarily moderate flux magnitudes?

The MDS algorithm fills a gap with the average of the measured values recorded under similar meteorological conditions (Reichstein et al., 2005; Pastorello et al., 2020). We agree that this does not necessarily give moderate values, so we have removed that statement. We have revised the sentence in Section 2.2 to describe this procedure directly (Lines 195–196), and the paragraph now says that the asymmetry of the top 1% fluctuations is based mainly on directly measured values (87.6% of the extreme peaks).

Line 178: 'attenuate rather than amplify the patterns' ... I understand it will dampen the extreme NEP values, but whether it would attenuate patterns in asymmetry is another step that I do not immediately see.

Dampening extreme NEP values does not necessarily imply weaker temporal asymmetry. We have removed this claim from Section 2.2 and clarified that 87.6% of the identified extreme peaks correspond to directly measured observations.

Line 180: What mean did you use? The mean over all data available for each site?

Thank you for this question. Yes, the mean is calculated over all half-hourly NEP values available for each site. We have clarified this in Section 2.2.

Line 189: Personally, I would like to see a bit more information about the averaging, i.e. you mention you average over three years. But, this can mean multiple things, i.e. you can calculate a single value for the three-year period, or you can create an averaged day with half-hourly values, or you can create an averaged day for each month. I would like to have a bit more details, because at this point in the manuscript it is not clear to me (it becomes more clear when seeing the Figures).

Good point, thank you. The full-record average is the mean of each driver over all half-hourly records of the three-year study period at a site. The matching average is the

mean of each driver over the half-hourly values within the ±6 h windows of all selected fluctuations at that site. We have added these definitions to Section 2.3.

Line 190: Why the word potential here?

The word was redundant, since proxies already convey the meaning, so we have removed it in Section 2.3.

Line 191: What do you mean with 'maximum photosynthetic thresholds' and how does the Duffy et al. (2021) reference support this statement?

Thank you for the question. By maximum photosynthetic thresholds we mean the thermal maximum of photosynthesis, the temperature above which ecosystem photosynthesis declines. Duffy et al. (2021) derive temperature response curves of ecosystem carbon uptake from the FLUXNET network. They find thermal maxima of photosynthesis of 18°C for C3 and 28°C for C4 systems (Fig. 1 of Duffy et al.), which vary by climate region (Methods, Spatial gridding), and show that photosynthesis falls above this maximum while respiration keeps rising. We have replaced the phrase with thermal maxima of photosynthesis and defined it in the sentence in Section 2.3.

Line 195: You refer to Tables S6 and S7 here, but these tables do not show a comparison between matching averages and full-record averages?

Thank you for pointing this out. Tables S6 and S7 list both averages. The rows labeled 3yr give the full-record averages, and the rows labeled Matching give the averages over the windows of the selected fluctuations. These rows were labeled Daily in the preprint, and we have renamed them Matching. Both averages are defined in Section 2.3, and the comparison of the two is given in Section 3.3.

Line 202: Not sure if I'm interpreting this correctly, but in case two models have an equal BIC value, you start looking at AIC, adjusted R2, and RMSE? Is there also priority in the latter three criteria, e.g. looking at AIC first (after BIC)? This is not really clear to me from the text.

Thank you for this question. Yes, that is how the selection works. The model with the lowest BIC is selected, and if BIC values are equal, AIC is considered next, followed by adjusted R^2. RMSE is reported alongside for each model. We have clarified this order in Section 2.3.

Figure 2: Regarding the right (pdf) part of the figure, I assume it shares the y-axis of the left part of the figure (would be logical). However, in the pdf I see non-zero (around 10^-3) pdf values for grassland for NEP values as low as -20, while in the left part of the plot I don't see any fluctuations for grassland that go anywhere near as low as -20. What is the explanation for this?

Thank you for the question. Yes, the PDFs use the same NEP axis as the time series. The distributions were estimated using Gaussian kernel density estimation with a bandwidth factor of 1.5. Since the kernels extend beyond the observed values, the grassland PDF shows a small density at −20, even though the lowest observed NEP is approximately −9.4 μmol CO_2 m^{-2} s^{-1}. We have clarified this in the Figure 2 caption.

Figure 2 (caption): 'Sample time-series of NEP for four selected sites' … I see here data for 4 biomes. Are these 'biomes' a combination of multiple sites having the same biome? In Figure 3 I see that 13 sites and 7 biomes are plotted. I think it would be good to make clear somewhere that these biomes present a combination of some of the used sites (at least, if it is true what I'm assuming).

Each curve in Figure 2 represents an individual site, rather than a combination of sites. The four sites are US-SRM (woody savanna), AU-Rig (grassland), DK-Sor (deciduous broadleaf forest), and US-SRC (open shrubland). These are four of the 13 sites across seven biomes included in Figure 3. We have added the site names to the Figure 2 caption for clarity.

Line 206: sites or biomes?

Thank you for the question. We mean four individual sites, each representing a different biome. We have named these sites in the Figure 2 caption, as described in our response to the comment on that caption, and we now describe the differences in NEP peak magnitude as differences among these sites.

Line 219: 'c = 99% quantile' … I assume you mean that c is chosen such that only values in the 99% quantile are selected? Same for line 307 and caption Figure 6. But ok, it is quite clear what is meant.

Yes, the coefficient (c) is chosen so that the threshold corresponds to the 99th percentile of NEP values at each site, identifying the top 1% of observations. We have clarified this in all places where the expression appears, in Lines 219 and 307 and in the captions of Figures 4 and 6.

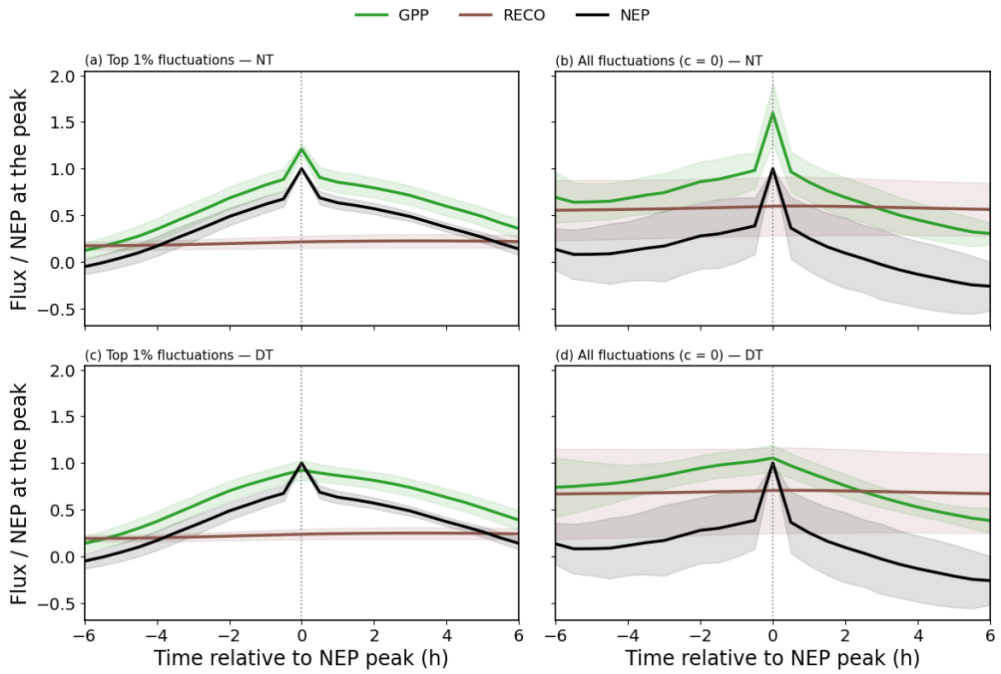
Line 220: Here and throughout the manuscript it seems to be assumed that the peak occurs around midday, with a morning rise and afternoon decline. This is not shown in Figure 3, as relative times are shown. Could you somehow show that a morning rise and afternoon decline is actually what the data shows (and maybe include this in the manuscript)? Does NEP already decline at the end of the morning?

Thank you for this suggestion. We agree that Figure 3 alone does not show this, so we added a new figure to the Supplementary Information (Figure S5). Panel (a) gives the time of day of the selected NEP peaks and panel (b) gives the mean NEP by time of day at each site on days containing a top 1% peak. The top 1% peaks occur around midday,

with a median of 11:30 and about 98% between 08:00 and 16:00. Mean NEP rises through the morning and starts to decline around or before noon, with the mean peaking between 09:30 and 12:30 across the sites. We have added this to Section 3.1.

Line 221: Why at least equally productive instead of more productive? Because you account for the possibility that respiration is higher in the morning? But then the afternoon could theoretically even be less productive.

Thank you for this comment. Positive NEP asymmetry indicates greater net carbon uptake after the peak than at the corresponding time before the peak. We have clarified the interpretation in Section 3.1. The partitioned fluxes (Fig. S2, shown below, and Section 3.3) further show that GPP is higher after the peak and exhibits a larger difference than RECO, suggesting that photosynthesis contributes more strongly to the observed positive asymmetry.



Line 223: 'The same analysis shown separately for all 13 sites is presented as Figure S2 in the Supplementary Information' … In Figure S2 the individual fluctuation paths are shown, perhaps an idea to make this clear when referring to the Figure. A similar suggestion applies to line 256.

Thank you for the suggestion. We have clarified the references to these figures in the manuscript. Figure S4 shows the individual top 1% fluctuation paths and their average for each site. We have also updated the text near Line 256 to clarify that Figure S7 shows the individual fluctuation paths and their averages, while Figure S8 shows the individual asymmetry coefficients and their averages for all fluctuations.

Line 226: 'biome-specific differences under average conditions' … Are they biome specific, i.e. do sites with the same biomes but different locations show large similarity under average conditions?

Thank you for this comment. We compared the mean NEP fluctuation trajectories of Figure 3(c) among sites of the same biome and among sites of different biomes. Some sites of the same biome have very similar trajectories, such as the two evergreen needleleaf forests and the two deciduous broadleaf forests, whereas others differ clearly, particularly the two woody savannas. Overall, the mean correlation between sites of the same biome (0.88) is comparable to that between sites of different biomes (0.91), so the differences do not clearly follow biome type. We therefore changed the wording to site-specific differences.

Line 231: Based on the figure, the max asymmetry looks more like 0.4 than 0.5 (for the site with highest max asymmetry).

Thank you for pointing this out. Indeed, the highest average asymmetry coefficient in Figure 3(b) is about 0.4 (US-SRC), and we have updated the text accordingly.

Line 237: 'This elevated asymmetry may arise from enhanced afternoon respiration during warm winter days' … I struggle with this theory. We are talking about the top 1% NEP fluctuations, i.e. moments when there is a large carbon sink (positive NEP). Now if in the afternoon respiration is enhanced as you suggest, than the sink strength would decrease, right? This would lead to negative asymmetry (or less positive) rather than stronger positive asymmetry, right? If you agree, then please adapt or remove this theory.

Thank you for pointing this out. We agree that increased afternoon respiration would lower NEP after the peak and reduce positive asymmetry, consistent with GPP being 0.20 NEPmax higher 3 h after the peak than 3 h before while RECO is only 0.04 NEPmax higher (Section 3.3, Fig. S2, Table S8). We have therefore removed this explanation and retained the observation that US-SRC has the highest asymmetry coefficient in Figure 3.

Line 238: 'Similar asymmetry is observed when analyzing the top 2% and 5 %, although with a lower magnitude (Figure S6 in Supplementary Information),' Perhaps you should add the word 'generally' here, as in Figure S6a and S6L the asymmetry magnitude for the top 5% is most of the time higher than for the top 1%. Next to that, I suggest to indicate in Figure S6 or in its caption which plot corresponds to which site. Also, I suggest to indicate in the description above the figure that the all fluctuations line (c=0) is also plotted.

Thank you for these suggestions. We added the word "generally" to the sentence, since for AU-Gin (panel A) and AU-Rig (panel L) the top 5% curve is higher than the top 1%. We now list the site shown in each panel of Figure S3 in its caption, and we added a statement that the all-fluctuations line (c = 0) is also plotted.

Line 243: I understand the importance of declining radiation, but 'declining radiation' is listed here as an example of the respiratory and regulatory processes. Could you shortly explain this?

Thank you for this comment. Declining radiation is an environmental driver and not a respiratory process. We have removed it from this sentence, so the respiratory and regulatory processes now include plant and microbial respiration and incipient stomatal limitation.

Line 247: 'For example, lower temperatures limit both photosynthesis and respiration, resulting in relatively balanced rise and fall patterns and lower positive asymmetry' … The step from 'lower temperatures limit both photosynthesis and respiration' to 'lower positive asymmetry' is too big in my opinion. I would argue that it also depends on how temperature behaves before and after the peak. Of course, when (absolute value of) NEP is low, the before vs afternoon differences in NEP will be small as well. However, with your Eq. 2 you normalise with NEPmax, so this does not mean that the asymmetry coefficient cannot be large. Can you explain in more detail in the text why you think lower temperatures lead to lower asymmetry?

Thank you for this comment. Our results show that the asymmetry of the top 1% fluctuations increases with mean air temperature across sites (r = 0.75, Figure 4a). We agree with this point. Because the paths are normalized by NEPmax, the asymmetry depends on the shape of the GPP curve, in particular how quickly GPP rises toward the peak and how it falls afterwards. At higher temperatures GPP rises faster toward the peak, whereas at lower temperatures it rises more smoothly, giving a more symmetric curve. We have rephrased this explanation in Section 3.3 (Figure 4a).

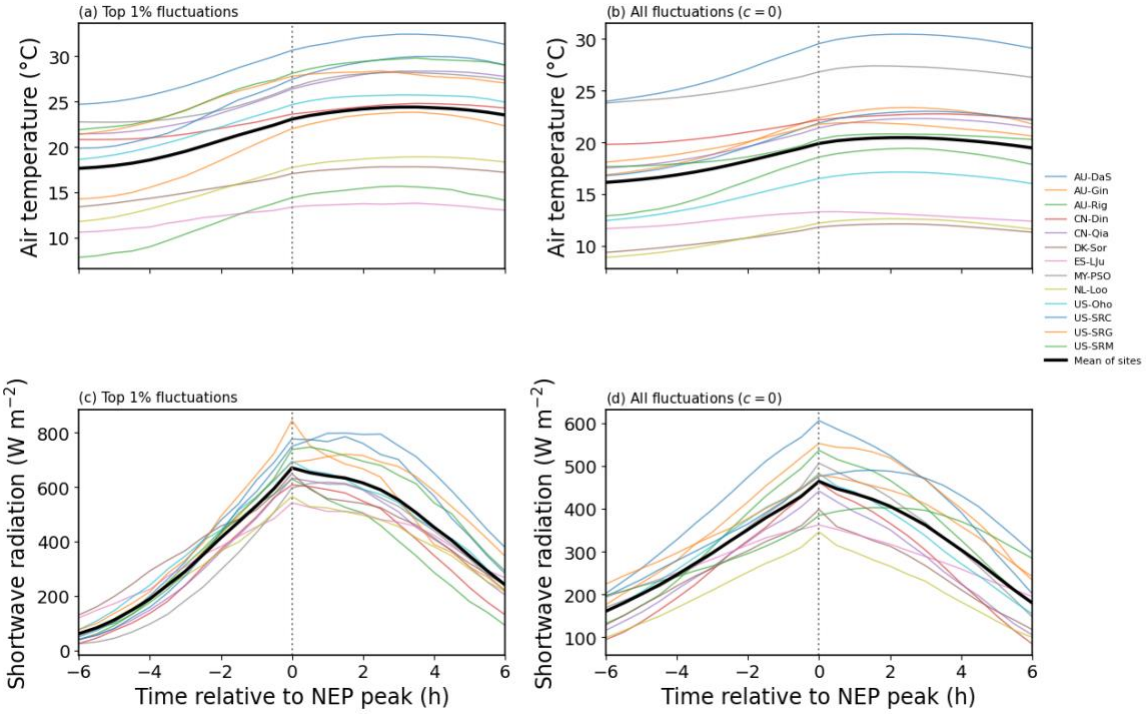
Line 250: Similar to the above, perhaps explain in the text why a GPP-dominated situation leads to positive asymmetry.

We appreciate this suggestion. Please see our reply to Line 247 above, which addresses the same point. Since NEP is GPP minus RECO, the asymmetry depends on the differences in both fluxes between the times after and before the peak, normalized by NEPmax. When changes in GPP values are higher after the peak than before it by more than RECO, the faster morning rise and slower afternoon decline of photosynthesis lead to a positive asymmetry. We have added the GPP and RECO differences and their link to the positive asymmetry to Section 3.3 (Fig. S2 and Table S8).

Line 245: Given that you have all the temperature data etc, have you tried making a plot of temperature on the y-axis and time before and after the peak on the x-axis (separately for the 1% and for all fluctuations)? Do you see a similar asymmetry as in NEP?

Thank you for this suggestion. We had looked at this earlier by splitting the fluctuation paths into cooler and warmer groups and comparing their asymmetry. The groups looked similar, so these plots were not included. We had also extracted the temperature along the fluctuation paths to obtain the matching averages (Tables S6 and S7), but had not included these in the manuscript or Supplementary Information. The dependence on

temperature appears in the long-term average temperature of each site (Figure 4). We have now plotted temperature against time relative to the peak for the top 1% and for all fluctuations (Figure S20). In both cases, temperature is higher after the peak than at the corresponding time before it at all 13 sites. For the top 1% this has the same sign as the positive NEP asymmetry, whereas for all fluctuations the NEP asymmetry is negative (Section 3.2). We added this plot to the Supplementary Information, and in Section 3.3 we state that, given the more complex relationship between driver paths and asymmetry, we relate the asymmetry to integrated measures of the drivers. The new Figure S20 is shown below. Panels (a) and (b) show air temperature and panels (c) and (d) show shortwave radiation, for the top 1% fluctuations and for all fluctuations.



Line 247: Could you provide some information on how the Koyama and Takemoto (2014) reference supports the statements here / why you refer to it?

Thank you for this question. Koyama and Takemoto (2014) measured photosynthesis of Oenothera biennis leaves through the day in the field. Net photosynthesis reached a maximum at mid-morning while sunlight kept increasing toward noon. Photosynthetic rates and stomatal conductance were lowest around early afternoon, when leaf temperature and vapor pressure deficit were high, and part of the reduction was related to stomatal closure (Figs. 1 to 6). Iio et al. (2004) measured Fagus crenata leaves at the canopy top through the day. Photosynthesis peaked in the morning and then fell as light, leaf temperature and vapor pressure deficit rose. In their tests the decline was driven mainly by vapor pressure deficit through stomatal closure. We cite both papers for how light, temperature and vapor pressure deficit relate to photosynthesis through the day. The faster morning rise and more gradual decline of NEP is the pattern we observed in our results. We have split the sentence in Section 3.1 so that the citations follow only the first part.

Line 258: 'while respiration remains more steady' ... Or is it more likely that it even increases (higher temp in afternoon)? Is this something you deduced from the Fluxnet respiration data? Might be good to make this clear.

To clarify this, we examined the FLUXNET ecosystem respiration estimates (RECO_NT_VUT_REF, the nighttime partitioning product, which is estimated using air temperature) around the NEP peaks at all 13 sites (Figure S2). RECO varies much less than GPP, as intended by our original description of respiration as 'more steady.' However, RECO also increases slightly after the peak, being higher 6 h after the peak than 6 h before it at all 13 sites, consistent with the higher afternoon air temperature (Figure S20). When considering all fluctuations, this increase in RECO, together with the decline in GPP, contributes to the negative asymmetry. We have revised the sentence in Section 3.2 and the corresponding statement in Section 3.4 to clarify this.

Line 262: 'The Savanna ecosystem (AU-DaS) has the highest negative asymmetry coefficient' … I suggest to add the word 'generally' or 'for most of the time' here, as this does not hold at all times based on Fig 3d.

Thank you for pointing this out. AU-DaS has the most negative asymmetry coefficient over most of the time range shown in Fig. 3d, and we have added the word "generally" to the sentence in Section 3.2 to reflect this.

Figure S17: I don't understand why there is a difference between Figure S12b and Figure S17a. Aren't these supposed to be the same variables? And why are the asymmetries positive in Figure S17a? And why a Hill function in S17a and a linear function in S12b?

Thank you for pointing out this ambiguity. Both figures use matching average temperature, but they represent different fluctuation regimes. Figure S14b (S12b in the original numbering) shows all fluctuations (c = 0), where NEP asymmetry is generally negative and the y-axis displays its negative ($-\bar{A}^{max}$). Figure S19a (S17a in the original numbering) shows the top 1% fluctuations (c = 99%), where asymmetry is positive and the y-axis displays $\bar{A}^{max}$. The matching average temperatures also differ because they are calculated separately for each regime. The best-fitting functions were selected separately, resulting in a linear fit for all fluctuations and a Hill fit for the top 1%, as reported in Tables S7 and S6, respectively. We have clarified these distinctions in the captions of Figures S14 and S19 in the Supplementary Information.

Figure S17: Daily VPD average is not plotted?

The daily average VPD for all fluctuations is shown in Figure 4b of the main text, which is also the source of the fit statistics quoted in Section 3.3. Figure S19b shows the 3-year average VPD for the same regime. We have corrected the figure reference in Section 3.3 and clarified the caption of Figure S19 in the Supplementary Information.

Figure 3: The legend indicates that 'CN-Qia' and 'Evergreen needle Forest' are plotted as full lines, but for one of them this is not true (see the actual lines in the figure).

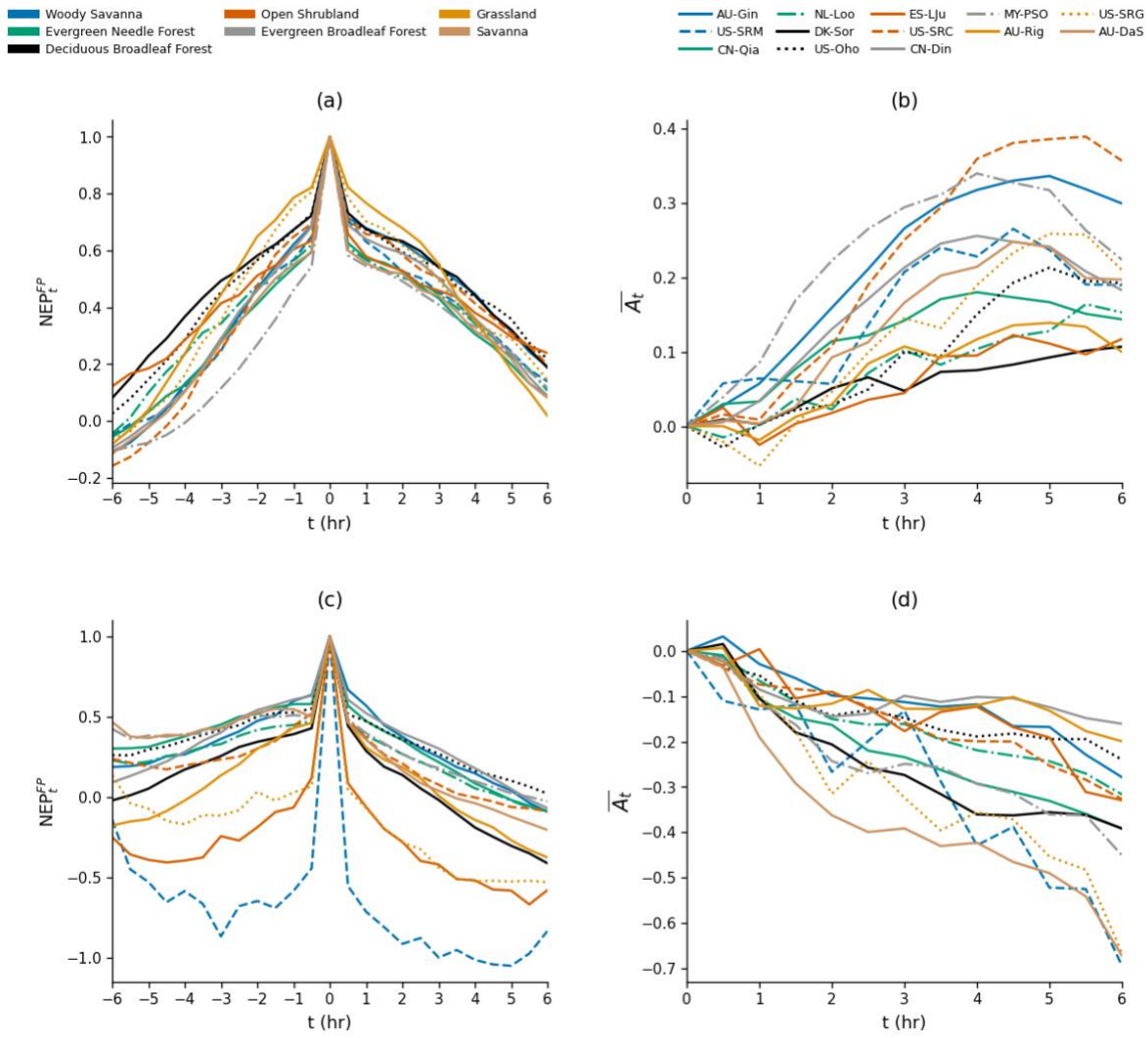
Thank you for this comment. We have revised the legend of Figure 3 so that it matches the plotted lines. Biomes are now shown by color, and each site is shown by its line style, using the same style as in the panels. The two Evergreen Needle Forest sites are now distinguishable, with CN-Qia solid and NL-Loo dash-dotted. We also added a sentence to the figure caption explaining that colors denote biomes and line styles denote sites. For consistency, the legends of Figures 4 and 5 now follow the same convention.

Figure 3: I see that you put NEE_FP instead of NEP_FP on the y-axis? I assume it does not matter for the signs as they will cancel out due to the normalisation.

Thank you for pointing this out. Figure 3 shows NEP, which is NEE multiplied by -1 (NEP = -NEE). Because each path is divided by its peak value, this sign change does not affect the normalized paths. We have corrected the y-axis label to NEP and added this explanation to the manuscript.

Figure 3: Please check the legend: According to the legend there should be 14 full lines in the figure panels. However, I only count 7, so apparently the legend symbols are not an accurate view of the actual lines. This makes it more difficult to identify the correct lines.

Thank you for pointing this out. As described in our response to the previous comment on the Figure 3 legend, biome names are now shown by color in a separate key, while line styles correspond only to the individual sites. The legend now accurately reflects the plotted curves, with 7 solid lines matching the 7 solid lines in the panels, and makes the sites easier to distinguish. The change is made in Figure 3 and its caption. The corrected Figure 3 is shown below.



Line 273: Is there a reason why you include the overbar on Amax in Figure 4 but not in the text here?

Thank you for pointing this out. We have updated the notation in the text to $\overline{A^{max}}$ for consistency with Figure 4, in the methods and results sections that relate asymmetry to the environmental drivers.

Line 274: I suggest to refer to Table S6 here.

Good point. We have added a reference to Table S6, which lists the fit statistics for the top 1% fluctuations, to this sentence in Sect. 3.3.

Line 276: At 22°, the curve seems closer to 0.35 than to 0.45… A typo? Or do you look at the data points instead of the curve (text mentions curve)?

Thank you for pointing this out. The value of approximately 0.45 refers to the observed data points rather than the fitted curve. We have revised the text in Section 3.3 to distinguish between the observed values and the fitted trend, which reaches approximately 0.35 at 22°C.

Line 279: Similar to one of my previous remarks, I'm not sure whether I find the explanation sufficient here. I suggest that you provide a clear and elaborate explanation on why higher GPP produces higher asymmetry. Not necessarily at this point in the manuscript, it could also be earlier. Of course, larger GPP allows for larger differences in GPP before and after the peak. But given that the NEP values are normalised with NEPmax, this is on itself not a reason I assume. I also suggest you elaborate on 'the differential response between GPP and RECO saturates'.

Thank you for this comment. Since each path is divided by NEPmax, asymmetry is the NEP difference after versus before the peak, expressed as a fraction of NEPmax. Because NEP = GPP − RECO, this equals the post-peak minus pre-peak difference of GPP minus that of RECO, both divided by NEPmax. The average paths in Table S8 show GPP 0.20 NEPmax higher 3 h after the peak than 3 h before, against 0.04 for RECO, so GPP accounts for most of the NEP asymmetry of 0.16. Asymmetry is greater when GPP stays closer to its peak after the NEP peak. As temperature rises, GPP increases more slowly as it approaches the optimal temperature for photosynthesis (Duffy et al., 2021; Yang et al., 2021). As a result, the increase in asymmetry also slows down, while the difference in RECO remains small. We have added these values and clarified the explanation in Section 3.3.

Caption Figure 4: 'Statistical metrics for each fit are reported in the legend boxes' … They seem to be missing?

Thank you for pointing this out. The statistical metrics for each fit are reported in Table S6 for the top 1% fluctuations (panel a) and Table S7 for all fluctuations (panel b). We have revised the Figure 4 caption accordingly.

Line 284: 'reflecting its role in steep morning NEP rises' … yes, but to directly contribute to asymmetry, the afternoon descent should be less steep, right? And radiation is rather symmetric I assume.

Thank you for pointing this out. Positive asymmetry depends on both the morning rise and the afternoon decline of NEP. Incoming shortwave radiation is approximately symmetric around solar noon, but NEP peaks usually occur before the daily radiation maximum, so radiation around the NEP peak is higher after the peak than before it. For the top 1% fluctuations, radiation at +3 h exceeds radiation at −3 h at all 13 sites (553 versus 291 W m^{-2} on average) (Fig. S20c), consistent with the more gradual afternoon decline of NEP relative to its morning rise in this regime (Fig. S2). We have revised the sentence in Section 3.3 and added shortwave radiation to Fig. S20.

Line 285 and 286: I don't really understand how VPD could reinforce positive asymmetry and show a positive correlation with Amax. My reasoning would be that under high VPD, photosynthesis would be depressed. High VPD values mostly occur in the afternoon (after the peak), when temperatures are high. Thus, under high VPD conditions I would expect a strong afternoon NEP decline, and thus a less positive or even negative asymmetry... Am I making a mistake in my reasoning?-

Thank you for this comment. Our results show that VPD acts differently in the two regimes. For all fluctuations, higher VPD reduces GPP in the afternoon and is associated with more negative asymmetry (r = −0.59, Fig. 4b). The top 1% fluctuations occur on days that are generally favorable for photosynthesis. Radiation remains higher after the NEP peak than before it at all 13 sites (Fig. S20c), temperatures are warm (Fig. S20a), and GPP remains greater than RECO after the peak (Fig. S2a, Table S8). VPD also increases after the peak on these days and can reduce GPP, but GPP is still 0.20 NEPmax higher 3 h after the peak than 3 h before, whereas for all fluctuations it is 0.18 NEPmax lower (Fig. S2). The asymmetry therefore stays positive despite the increase in VPD. The positive correlation between VPD and asymmetry for the top 1% (Table S6) describes a relationship across sites, not a within-day effect, since sites with higher mean VPD also have higher radiation and greater positive asymmetry (Figs. S9 and S12). The relationship across sites is described in Section 3.3, and the regime dependence in Section 3.4. Stomatal opening in fact does not depend only on VPD, but also on conditions that favor photosynthesis, such as plant water status. VPD matters most when plants are already water stressed, and less when moisture is plentiful. We have added this to Section 3.3.

Line 286: 'when coupled with high radiation' ... Is this based on reasoning, or is this based on some model test or so?

Thank you for this question. The statement was based on our interpretation of the individual correlations between asymmetry and VPD and shortwave radiation (Table S6), rather than a model testing their combined effects. We did not perform a VPD–radiation interaction analysis. As described in our response to the previous comment, we have revised the sentence in Section 3.3 to clarify the interpretation of the VPD correlation.

Line 293: You could also refer to Table S7

Thank you for the suggestion. We have added a reference to Table S7, which lists the fit statistics for all fluctuations, to this sentence.

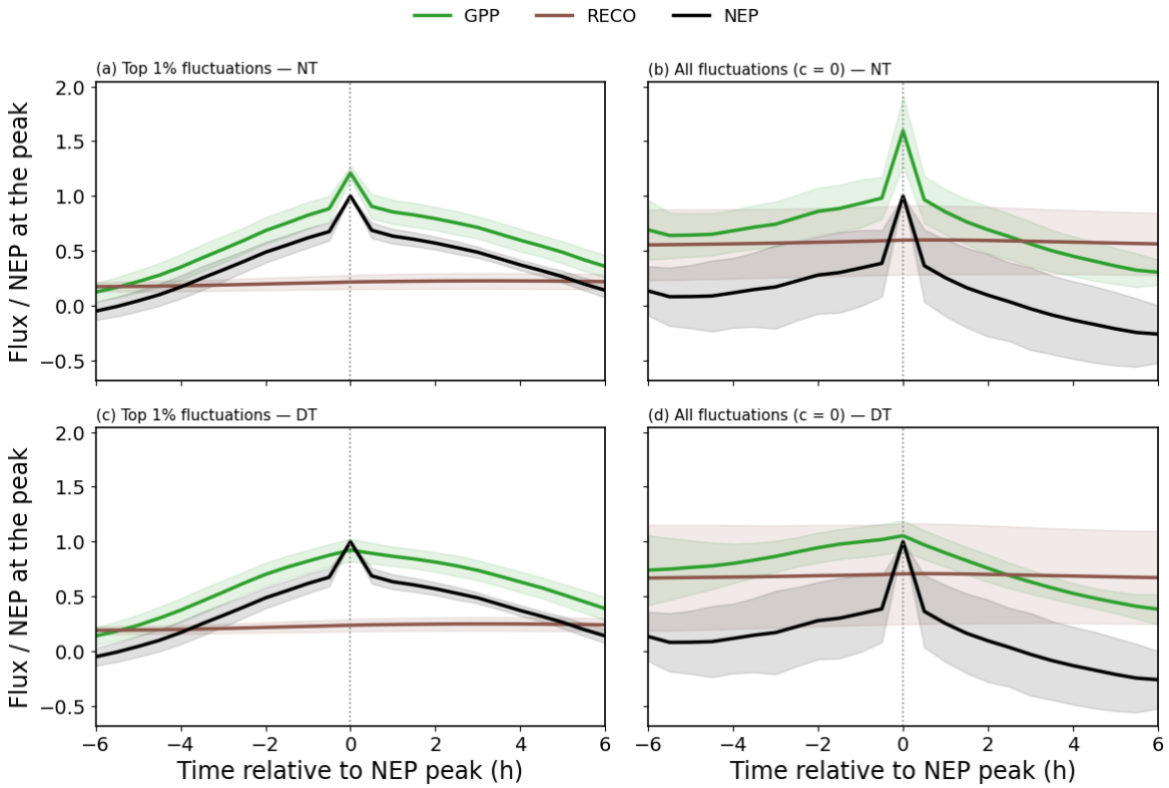
Line 303: the text mentions r=-0.2 for the T:P ratio. Figure S16 shows r=-0.2624 and r=-0.3825 for 3yr and daily data, respectively. Table S7 mentions r=-0.2624 and r=-0.2111 for 3yr and daily data,respectively. This is not fully consistent.

Thank you for pointing out this inconsistency. The correct correlation values are reported in Table S7: r = −0.26 for the 3-year and r = −0.21 for the matching MAT:MAP ratio. Figure S18b previously showed the temperature relationship instead of the matching

MAT:MAP ratio, which we have now corrected. Figure S18 and Table S7 are now consistent, and we have updated the manuscript text to report both values and revised the figure caption accordingly.

Line 310: Can't Reco be a part of the explanation as well? Have you analysed GPP and Reco separately? If not, it might be better to make clear that the text here is an hypothesis and does not directly follow from the data.

Yes, we have examined GPP and RECO separately using the FLUXNET-partitioned fluxes (Fig. S2 in the SI). Each path is divided by the NEP at its peak, the same as for the asymmetry coefficient, so that the GPP and RECO changes can be compared directly with the NEP asymmetry. For the top 1% of fluctuations, GPP is higher 3 hours after the peak than 3 hours before (+0.20), compared with a much smaller difference in RECO (+0.04), which slightly lowers the asymmetry. This shows that GPP is the main contributor to positive NEP asymmetry. For all fluctuations, GPP is lower after the peak (−0.18), while RECO changes only slightly (+0.02). The decline in GPP therefore explains much of the negative NEP asymmetry, with a smaller contribution from RECO. We found similar results using the daytime partitioning method (Fig. S2c,d) and state in Section 2.1 that both partitioning methods are used. The new Figure S2 is shown below. Panels (a) and (b) use the nighttime partitioning and panels (c) and (d) the daytime partitioning, for the top 1% fluctuations and for all fluctuations.



Line 311: 'These differing responses between extremes and typical daily peaks are due to differences in the timing of how photosynthesis and respiration respond to daily changes in the environmental drivers' ... Aren't it at the end the differences in the drivers between the extreme and daily situations that, via the GPP and RECO responses, lead to the observed difference in asymmetry between daily and extreme fluctuations?

Yes, differences in environmental conditions between extreme and typical fluctuations influence the responses of GPP and RECO, which together shape the observed differences in NEP asymmetry. Figure S20 shows that temperature and shortwave radiation around the NEP peak are higher for the top 1% than for all fluctuations and change more across the peak, while Figure S2 shows the corresponding GPP and RECO responses. We have revised the sentence to include the role of environmental conditions and the timing and magnitude of the GPP and RECO responses.

Figure 5: Is there a reason why, for US-SRM, there are horizontal black bars below and above the box for top 1% fluctuations but not for the box of all fluctuations?

Thank you for noticing this. The black bars are the ends of the whiskers. We set the whiskers to the 25th and 75th percentiles, so the boxes show the interquartile range and the whiskers end at the box edges. For US-SRM with all fluctuations, the interquartile range extends beyond the plotted y-axis range of −2 to 1, so its box edges are outside the plot and no bars appear. For the top 1%, the box lies within the range, so the bars are visible. We have added a description of the boxes to the caption.

Figure 5: I don't understand why there is an overbar on Amax here. You show the distribution for each site with boxes, so where does the averaging happen?

Thank you for pointing this out. The boxes show A^{max} for the individual paths at each site, so no averaging is involved. Only the red lines are averages, the site means $\overline{A^{max}}$. We removed the overbar from the y-axis label and clarified this in the caption.

Line 316: To what extent is this an hypothesis, or actual results? Do the data show an accelerated morning GPP rise when temperature is high, or is this an assumption? And similar, do the actual data show a steep afternoon GPP decline while RECO remains more steady in the 'average' (all fluctuations) situation?-

Thank you for this comment. It is a hypothesis with good support from the data. The temperature and VPD associations are shown in Fig. 4, and the average GPP and RECO fluctuation paths in Table S8 and Fig. S2 show the patterns described. For the top 1%, GPP is at 0.42 of its peak value 3 h before the NEP peak and 0.59 of its peak value 3 h after, so GPP rises faster toward the peak than it declines after it, while RECO stays nearly flat. For all fluctuations, GPP drops after the peak (−0.18 between −3 h and +3 h) while RECO changes little (+0.02), with the same result for the daytime partitioning (Fig. S2c,d). Our interpretation is consistent with well-known physiological mechanisms (Duffy et al., 2021; Yang et al., 2021; Chaves et al., 2016; Novick et al., 2016). We have added the figure and table references to Section 3.4.

Line 322: 'the amplitude of NEP fluctuations' ... there is a sign shift, so is the amplitude the most appropriate word here?

Good observation. Here "fluctuation" refers to any deviation from the mean, so by "amplitude" we refer to how far the NEP peak is from the site mean, in line with how it is defined in the original literature where the method was developed.

Line 329: 'sustains GPP under near-optimal conditions'?

Thank you for this comment. We meant that photosynthesis depends on temperature and is highest near an optimum temperature (Duffy et al., 2021; Yang et al., 2021). We have reworded the phrase to make that clear.

Line 324: '...one in which NEP is governed primarily by changes in GPP (linked to temperature) and another in which RECO plays a larger role' ... This sounds very plausible, but personally I don't see data-based evidence whether this is really what is happening. Would it be possible to do an analysis (or give an explanation) that more explicitly proves that temperature-induced GPP changes are indeed the main reason for the positive asymmetry and Reco plays a larger role in the other case? If not, I suggest that you try to make it somewhat more clear that this a theory, and does not directly follow from the data. Or perhaps your statements follow from reasoning, but then I suggest you give a more full explanation.

Thank you for this comment. Yes, the data support this interpretation. We analyzed GPP and RECO separately using the FLUXNET-partitioned fluxes (Fig. S2 and Table S8 in the SI). For the top 1% fluctuations, GPP is 0.20 NEPmax higher 3 h after the peak than 3 h before, against 0.04 for RECO, so the positive NEP asymmetry follows GPP, and the asymmetry increases with mean air temperature across the 13 sites ($r = 0.75$, Fig. 4a). For all fluctuations, RECO is a larger share of the NEP peak than for the top 1% (about 0.6 of NEPmax in Fig. S2b, compared with 0.22 in Table S8), while GPP still drops after the peak (−0.18, against +0.02 for RECO). The daytime partitioning gives the same pattern (Fig. S2c,d). We have added these references in Section 3.4.

Line 333: Regarding my point above, in this paragraph there is an additional argument supporting your statement that NEP is governed primarily by changes in GPP for the top 1% fluctuations, i.e. the logical but important reasoning that GPP should strongly exceed Reco in order to end up in the top 1% fluctuations. You could consider slightly changing the order of some of the information to make this clear earlier in the section.

Thank you for this suggestion. For the top 1% fluctuations, GPP substantially exceeds RECO at the NEP peak, as shown in Fig. S2a and Table S8. On average across the 13 sites, GPP is approximately 1.2 times the peak NEP, compared with 0.2 for RECO. This supports our interpretation that GPP plays the dominant role in these fluctuations. We have included this explanation earlier in the section to make the reasoning clearer.

Line 333-339: 'Under such conditions, afternoon stomatal limitation is reduced because VPD is less likely to reach levels that trigger strong stomatal closure' ... I understand that at the time of the peak, conditions should be very favourable for a high photosynthetic rate. However, this does not necessarily mean that a few hours later – after the peak – this is still the case, as e.g. VPD can be strongly increased. Are there convincing arguments for assuming favourable photosynthetic conditions over the whole fluctuation period instead of only around the peak?-

Thank you for this question. We looked at the conditions around the NEP peak for the top 1% fluctuations, and the data show that photosynthetic uptake remains high well after the peak (Fig. S2a and Table S8 in SI). GPP stays above RECO at all 13 sites from 3.5 hours before the peak to 6 hours after it, and 3 hours after the peak GPP is still about 59% of its peak value, compared with 42% three hours before. This happens even though VPD rises from 7 to 16 hPa between 3 hours before and 3 hours after the peak, a larger rise than for all fluctuations (9 to 14 hPa; Table S8), so carbon uptake continues despite the higher evaporative demand. Soil water content at the NEP peak is also higher for the top 1% fluctuations than for all fluctuations at nine of the 13 sites, by more than one percentage point at five of them (Table S9 in SI). With more water available, plants can be less dependent on VPD, so photosynthesis can stay high after the peak. We have revised the paragraph to describe this.

Line 343: What is meant with cancellation artifacts here? And with 'cancellation effects' on line 346?

Thank you for pointing this out. By cancellation we meant that positive and negative values of A^{max} from individual fluctuation paths could offset each other in the site mean. Figure 5 shows the A^{max} distributions at each site, and the site means are positive for the top 1% fluctuations and negative for all fluctuations at all 13 sites. Since this is already described at the start of Section 3.4, we have removed these sentences.

Line 350: Soil moisture is expressed here in mm instead of the more traditional m3/m3or %. Can you explain these units, perhaps also in the manuscript? Or is it a typo?

Thank you for pointing this out. We apologize for the confusion regarding the units. The soil water content (SWC_F_MDS_1) used in our analysis is reported in % in FLUXNET2015. We have corrected the unit in Section 3.5 and clarified it in Section 2 and Table S2.

Line 351: Does a low MAT:MAP ratio contributes to water limitation? Isn't it the other way round?

Yes, a higher MAT:MAP ratio generally indicates greater water limitation. For AU-DaS, the water limitation is better explained by its very low soil moisture (5.8%) and high VPD (~15 hPa). Its MAT:MAP ratio (~300) is closer to the middle of the range across the 13 sites (Fig. S18a). We have revised Section 3.5 accordingly.

Line 356: Is there a reason why you sometimes mention 'MAT:MAP ratio' and sometimes 'T:P ratio'? Is the latter not averaged?

Thank you for noticing this. Both names refer to the same quantity, the ratio of the mean air temperature to the mean precipitation, calculated from the averages. We now use MAT:MAP consistently in Sections 3.3 and 3.5.

Line 356: Is the 'very high shortwave radiation' at US-SRC beneficial for sustained afternoon assimilation?

Thank you for this comment. US-SRC has the highest mean shortwave radiation among the 13 sites (~245 W m^{-2}, Fig. S15), along with warm temperatures (~20 °C). These conditions are consistent with more sustained afternoon photosynthesis and the weaker negative asymmetry observed at this site, despite its high VPD. This is described in Section 3.5.

Supplementary Figure S16: Why is Au-Das only shown in panel (a)? And I assume the difference between panels (a) and (b) is that in (b) you calculate for every day the ratio and average afterwards over the days, while in (a) you calculate the mean T and mean P, after which you calculate the ratio? I suggest to explain this more.

Thank you for your comment. Panel (b) in the earlier version of Figure S16 (now Figure S18) showed a temperature plot (x-axis in °C) instead of the daily MAT:MAP plot, and AU-DaS was not visible. We have corrected the figure, and both panels now include all 13 sites. Regarding the calculation, your understanding of panel (a) is correct. We use the mean temperature and mean precipitation over the full 3-year record to calculate the ratio. For panel (b), we use the same approach but calculate the means over the windows of the selected fluctuations. We have clarified this distinction in the captions of Figures S13 and S18.

Line 357: 'US-SRC again deviates from expectation' ... this seems to depend a lot on whether you take the daily ratio or 3yr ratio, for the daily ratios the point is relatively close to the line.

Thank you for your comment. Figure S18 shows that US-SRC lies at a similar distance from the fitted line in both panels, while AU-DaS shows the largest departure from the fitted relationships. We have revised Section 3.5 to focus on AU-DaS and include its MAT:MAP values for both averaging periods.

Line 358: 'despite a relatively high T:P ratio (~ 600)' ... Are you sure you did not mistakenly look at the AU-Rig datapoint instead of the US-SRC datapoint? Because, if I read the legend well, Us-SRC has a (3yr) MAT:MAP ratio above 1000.

Yes, thank you for catching it! The 3-year MAT:MAP ratio for US-SRC is above 1000 (Fig. S18a), while the value of approximately 600 corresponds to AU-Rig. We have

updated Section 3.5 to clarify the interpretation, focusing on AU-DaS and its MAT:MAP values.

Line 359: 'reflecting the overriding influence of extremely high radiation' ... Why would extremely high radiation lead to very negative asymmetry? Afternoon radiation will generally not be higher than morning radiation, so why would radiation depress GPP or increase Reco in the afternoon? Is the effect indirect via increasing VPD? Perhaps good to make this clear

Thank you for the comment. Yes, radiation affects asymmetry at AU-DaS indirectly. High radiation at this site occurs together with high VPD and low soil moisture. Such conditions affect stomatal conductance and contribute to a steeper afternoon decline in GPP, as described in Section 3.2. In addition, high radiation can increase atmospheric evaporative demand, thus raising VPD and reducing stomatal conductance and GPP in the afternoon (Novick et al., 2016; Grossiord et al., 2020). We have described this in Section 3.5.

Line 363: '...gap filled hourly values of NEE using' ... You mentioned earlier that you use half-hourly NEE data. How do you go from the hourly to the half-hourly data?

This was a typo. All 13 sites use half-hourly NEE. We have corrected the sentence.

Line 366: What do you mean with 'over a few extreme fluxes'? These lines are not really clear to me.

Thank you for pointing this out. We intended to explain that the MDS method selects observations with similar meteorological conditions. The search window is widened when suitable observations are unavailable to find additional data (Reichstein et al., 2005). So, the resulting estimates can capture extreme fluxes, although averaging may smooth individual peaks. We have revised the paragraph in Section 3.5 to clarify this explanation.

Line 366: 'If any of these values exceed the top 1%' ... I assume 'these values' refers to the values of the missing data. And I assume that with 'they would most likely reduce' you mean the gap-filled values are most likely reduced compared to the original (unknown) fluxes. If you agree that these lines are currently not fully clear, perhaps they can be written slightly differently... But I probably don't fully understand what you want to say here, especially the part about the bias.

Yes, by these values we mean the gap-filled NEE values. The point was that a gap-filled value is an average over similar records, so it can be lower than the true, unknown flux. We agree the sentence was not clear and have removed the statement about bias. The revised paragraph in Section 3.5 says that averaging may smooth individual peaks, and refers to Section 2.2, where 87.6% of the 2,522 top 1% peaks are directly measured and 12.4% are gap-filled (Table S5).

Line 367: I understand the extreme values might be 'dampened' because of the averaging, but why a positive bias? When NEP is underestimated, isn't it a negative bias? Or am I misunderstanding the reasoning here?

Thank you for pointing this out. MDS gap-filling may smooth unusually large flux values, potentially reducing the magnitude of extreme NEP peaks. We have revised Section 3.5 to explain this effect more clearly and removed the reference to positive bias. The revised paragraph also states that only 12.4% of the identified extreme peaks are gap-filled, as shown in Section 2.2 (Table S5).

Line 366: If extreme meteorological conditions would be causing the extreme fluxes, than averaging over similarly extreme conditions would not decrease the fluxes? Are generally mostly less extreme conditions selected for calculating the average in such a case?

Not necessarily. MDS matches the missing value with records under similar meteorological conditions, so extreme conditions are matched with similar conditions and not with less extreme ones (Reichstein et al., 2005). The matching is relaxed only when similar records are not found (Pastorello et al., 2020), and only then can the average end up smoother than the true actual flux. But for the top 1% fluctuations, the similar records used for gap-filling can only be less extreme than the selected peaks, so averaging them can only bias the flux down, not up. We have clarified this explanation in Section 3.5.

Line 371: But do you use also the RECO and GPP data (as they are not shown)?

Thank you for this question. We use NEP (−NEE) to calculate asymmetry, and GPP and RECO to interpret the observed patterns. The GPP and RECO results are presented in Figure S2 and Table S8 in the SI. We have clarified their use in Section 3.5.

Figure 6: I suggest to put arrow heads under the lines pointing to the word 'temperature', as you do for VPD

Thank you for the suggestion. We have added arrowheads below the temperature lines in Figure 6, pointing to the word Temperature, as for VPD.

Figure 6 caption: 'to temperature under extremes' … you could consider to write 'to temperature under positive extremes' or something similar, as you did not look to negative extremes

We have revised the Figure 6 caption to specify that the shift toward temperature occurs under positive extremes, consistent with the top 1% NEP fluctuations examined in our analysis.

Figure 6 and e.g. line 374 in the text: I doubt somewhat whether we can firmly talk about a temperature driven regime in extreme conditions and a VPD-driven regime under normal conditions. Figure 4b indeed shows a pattern for VPD, but the curve fit is not super, and r=-0.59 is not overwhelmingly convincing either I would think. I know VPD gave the best fit of all variables for the typical conditions, but, based on the findings in this manuscript, can we really firmly talk about a VPD-driven regime and a temperature controlled regime? What is your view on this?

Thank you for this question. Yes, the data support a different leading variable in each regime. The main finding of our study is the reversal in temporal asymmetry across all 13 sites, which holds independently of the driver relationships. We also find that the variable most strongly related to asymmetry differs between the two regimes: temperature for the top 1% of NEP fluctuations (Figure 4a: $r = 0.75$, $p = 0.003$; Table S6) and VPD for all fluctuations (Figure 4b: $r = -0.59$, $p \approx 0.035$; Table S7), and in each regime it is the variable with the best fit and the only significant correlation. These findings are also consistent with the known effects of high VPD on stomatal closure and afternoon GPP (Novick et al., 2016; Grossiord et al., 2020), and the role of temperature in photosynthetic capacity (Duffy et al., 2021; Yang et al., 2021), as discussed in Section 3.4. Figure 6 summarizes these two relationships conceptually, and its caption now refers to the dominant association in each regime and to positive extremes (the top 1% of NEP).

Line 392: Are you talking here about formulating a nonlinear response function of NEP to environmental drivers, e.g. temperature? How would you translate from asymmetry to actual NEP? I am not familiar with these types of analyses

Thank you for this question. Yes, we are referring to the nonlinear response of ecosystems to environmental drivers such as temperature, as represented in biosphere models. We do not formulate these response functions in this study, since asymmetry is calculated directly from the observed NEP trajectories and describes the differences between the rising and falling limbs around the daily peak. In future work, the same asymmetry metric could be calculated from NEP simulated by biosphere models and compared with observations. This would help examine whether the models capture the asymmetry and its switch across regimes. We have clarified this in Section 4 (Implications).

Line 398: ‘Terrestrial biosphere models and machine learning techniques are common approaches to predict NEP’ … you could consider giving an example reference

Thank you for this suggestion. We have added a reference for each approach. Fisher et al. (2014) is a review of terrestrial biosphere models and the processes they represent, including NEE and NEP as model outputs. Jung et al. (2020) shows how machine learning, using meteorological and remote sensing data, is used to estimate NEE from eddy covariance measurements.

Line 403: ‘the tails of the NEP distribution that matter most under disturbance or climate extremes’ … depends somewhat on the kind of climate extremes whether you end up in the NEP tails I would think.

> Thank you for this comment. Our analysis focuses on statistical extremes in NEP, defined as the upper tail of the NEP distribution, rather than on climate extremes. We have updated the sentence to clarify this point.

Line 407: ‘typically associated to stronger environmental forcings’ … or just more ideal conditions for sustaining high afternoon photosynthesis… is a stronger forcing typically required for that?

> Thank you for this comment. Reaching the top 1% of NEP requires GPP to exceed RECO substantially (Section 3.4), and the positive asymmetry we observe is associated with conditions that favor sustained afternoon photosynthesis. In the top 1%, GPP remains elevated after the NEP peak (normalized +3 h minus −3 h difference of +0.20, compared with +0.04 for RECO), whereas for all fluctuations it is negative (−0.18; Table S8 and Fig. S2 in SI). We have added this description to Section 3.3.

Line 412: What do you mean with strategies for sustaining carbon balance in this context? Is this desirable (e.g. net uptake might be better)? Perhaps check whether this sentence really states what you want it to.

> Thank you for this comment. We meant keeping the ability of ecosystems to take up carbon, in other words net uptake. That is generally the desirable outcome, since a net sink helps stabilize the Earth’s climate (see the Introduction). Knowing how NEP behaves during large fluctuations may help evaluate how well ecosystems maintain net uptake as variability in environmental conditions increases. We have rewritten the sentence in Section 4 to say this.

Line 413: ‘Finally, while here we focused on ecosystems in their stationary state’ … Are they indeed all in stationary state? Are do you rather mean relatively undisturbed?

> Thank you for this comment. We meant relatively undisturbed ecosystems, and we have revised the wording accordingly.

Supplementary Table S3: In the definition of Amax (with overbar), there is no reference to ‘average’ given, while this is what the overbar indicates I assume.

> Thank you for this comment. The overbar indicates the average over all fluctuation paths of a site. We have revised the definition of $\overline{A^{max}}$ in Table S3 to state this and added a separate entry for A^{max} (single path).

Figure S12: I suggest to also refer to daily temperature average in the caption (you also refer to daily average variables in the captions of the other figures), instead of just to the 3-yr average

Thank you for this suggestion. We have revised the caption of Figure S14 (S12 in the original numbering) to mention both the 3-year average temperature and the matching average temperature, which is the mean over the windows of the selected fluctuations, as in the captions of the other figures.

Technical remarks:

Line 71: 'which have been proving useful' … 'which have been proven useful' instead?

Thank you, we have corrected this to "have been proven useful".

Line 148: 'The asymmetry metric' instead of 'The asymmetric metric'?

Thank you, we have corrected this to "The asymmetry metric".

Line 215: 'focus' instead of 'focuses'? Or alternatively, 'analysis' instead of 'analyses'?)

Thank you, we have corrected this to "the analysis here focuses on".

Line 223: full stop instead of comma after 'biomes'?

Thank you, we have replaced the comma with a full stop.

Line 231: 'reaching' instead of 'reaches'?

Thank you, we have corrected this to "ultimately reaching".

Line 349: 'falls' instead of 'falling' to make the sentence correct?

Thank you, we have corrected this to "falls below the fitted curve".

Line 369: 'artifacts' instead of 'artifact'?

Thank you, we have corrected this to gap-filling artifacts.

Line 387: Bracket should be moved in front of 'Niu'. A similar issue on line 389.

Thank you, we have corrected the citation format at both places so that the reference appears in brackets.

Line 390: systems' instead of system's (and perhaps responses)? Depends on what you want to say, but the current sentence seems grammatically incorrect to me.

Thank you for this comment. We have corrected the grammar and now write "a system's nonlinear response", since the statement refers to a system in general.

Figure 2: It seems there is a whitespace between C and O in the y-axis title that should not be there. Also, the 'C' seems to be upright text while the 'O' is cursive.

Thank you for this comment. We have corrected the y-axis title of Figure 2.

Caption Figure 6: Please check the font of QNEP.

Thank you for this comment. We have corrected the font of Q_{NEP} in Figure 6 and its caption.

Supplementary Table S1: 'MEAN NEP FOR STUDIED PERIOD (µmolCO2m-2s-1)' … Please check subscript vs normal font.

Thank you for this comment. We have corrected the subscript and superscript formatting of the units in the header of Table S1.

In supplementary Table S3, it seems the overbar is missing for 'The average asymmetry coefficient across all fluctuation paths'. At least, I don't see it in the pdf on my computer.

Thank you for this comment. As in our response to the earlier comment on Supplementary Table S3 (on the definition of Amax with overbar), we have added the overbar to $\overline{A_t}$.

Whitespace missing in heading Table S5: 'top1%', while the whitespace is included in the section title on the same page.

Thank you for this comment. We have added the missing space in the Table S5 heading.

Table S4: Please check the placement of the formulas in the Table, part of the formula symbols are obscured

Thank you for this comment. We have corrected the placement of the formulas in Table S4 so that all symbols are fully visible.

Table S4: Please check the formatting and missing overbar in the formula of step 3.2

Thank you for this comment. As in our response to the previous comment on Table S4, we have corrected the formatting of the formula in step 3.2 and added the overbar on $\overline{A_t}$.

Revised Supplementary figure numbers

The Supplementary Figures were renumbered in the order in which they are first cited in the revised manuscript. The reviewer comments above use the numbers of the preprint, and our replies use the numbers of the revised Supplementary Information. Tables S1 to S9 keep their numbers.

Number in the reviewer's comments	**Number in our replies and the revised SI**	**What it shows**
S1	S1	Example NEP time series
S2	S4	Top 1% fluctuation paths, all sites
S3	S7	All fluctuation paths, all sites
S4	S6	Asymmetry coefficients, top 1%
S5	S8	Asymmetry coefficients, all fluctuations
S6	S3	Threshold sensitivity (top 1%, 2%, 5%)
S7 – S11	S9 – S13	Drivers, top 1% (shortwave radiation, soil moisture,

		precipitation, VPD, MAT:MAP)
S12 – S16	S14 – S18	Drivers, all fluctuations (temperature, shortwave radiation, soil moisture, precipitation, MAT:MAP)
S17	S19	Matching temperature (top 1%) and 3-year VPD
New	S5	Timing of NEP peaks
New	S20	Temperature and shortwave radiation around the peak
New	S2	GPP, RECO and NEP around the peak

Figures S2, S5 and S20 are new and were added in response to the reviewer's comments.