***1st Reviewer***

*1. I have only one remaining minor concern. Because the authors clarify that the sensitivity metric is not intended to isolate a direct $CO_2$ effect, the wording throughout the manuscript should be checked carefully to avoid implying a causal $CO_2$-only response. The sensitivity should be consistently described as an ACGR-associated integrated ecosystem response under coupled $CO_2$, climate variability, and land-use change.*

→ Helpful comments! Following the reviewer's comment, we revised the related sentences to specify that the sensitivity indicators reflect the combined effects of ACGR-related integrated ecosystem responses throughout the manuscript.

*2. In addition, the model-selection criterion based on the most recent 20-year window and the monsoon/non-monsoon partition should be clearly acknowledged as simplified methodological choices.*

→ Valuable comments! Responding to the reviewer's comments, the earlier version has been revised as follows:
"The most recent 20-year period (2004-2023) was selected as the base window to reflect current climate conditions and ecosystem responses; this represented a methodological simplification for model selection."
"Note that, because p-values cannot be meaningfully averaged into a single value, the statistical significance of NEP sensitivity was assessed using a simplified method involving the last 20-year (2004-2023) moving window to reflect ecosystem changes in the recent decades."

***2nd Reviewer***

*Following the first round of reviews of the paper by Na. et al., the authors responded to most of the minor comments and performed an additional statistical analysis in response to some of the major comments. While the tasks of the paper and the results achieved are now clearer, there are still important issues in the methods section to consider.*

→ We sincerely thank the reviewer for the helpful and insightful comments, which have enabled us to improve the manuscript again. Following the reviewer's comments, we further revised the earlier version of the manuscript accordingly.

***Major comments***

*1. The characteristics of the LOWESS method in the process of removing long-term trends are highly dependent on the specific task. In the cited paper by Li et al. (2024), the authors additionally investigate the cyclical dependencies of trends and their relationships, which led to the choice of smoothing span (f) = 0.25. However, the task addressed by Na et al. does not involve determining the periodicity. The goal is simply to remove the overall trend in data changes within a time window. To achieve this, the window parameters are expected to be different from Li et al. This requires justification and explanation of the choice, which is currently absent. Therefore, it remains unclear why the authors used (f) = 0.25, robustness iterations = 3 for the detrending procedure.*

→ Constructive comments! We initially adopted f = 0.25, following the study by Li et al. (2024) although the number of robustness iterations was not explicitly reported in their study. In the present study, the number of robustness iterations was set to 3, which was the base value in the implementation. To verify whether the results depend on this specific parameters, we fixed the number of robustness iterations at 3 and repeated the analysis using smoothing span (f) = 0.2 and 0.3 (Fig. R1). We also fixed f at 0.25 and performed tests with the number of robustness iterations set to 2 and 4 (Fig. R2). Indeed, the temporal variations in sensitivities were generally consistent with the results from the primary setting (i.e., f = 0.25 and robustness iterations = 3). Although some minor differences in scale and statistical significance were observed in certain individual moving windows, these differences did not affect the main results of the study.

Responding to the reviewer's comments, we add the following sentences in the revised manuscript: "Sensitivity analysis using f=0.2 and 0.3 with robustness iterations fixed at 3, and using 2 and 4 robustness iterations with f fixed at 0.25, produced consistent temporal patterns, indicating that the main results are not sensitive to variations in the LOWESS parameters."

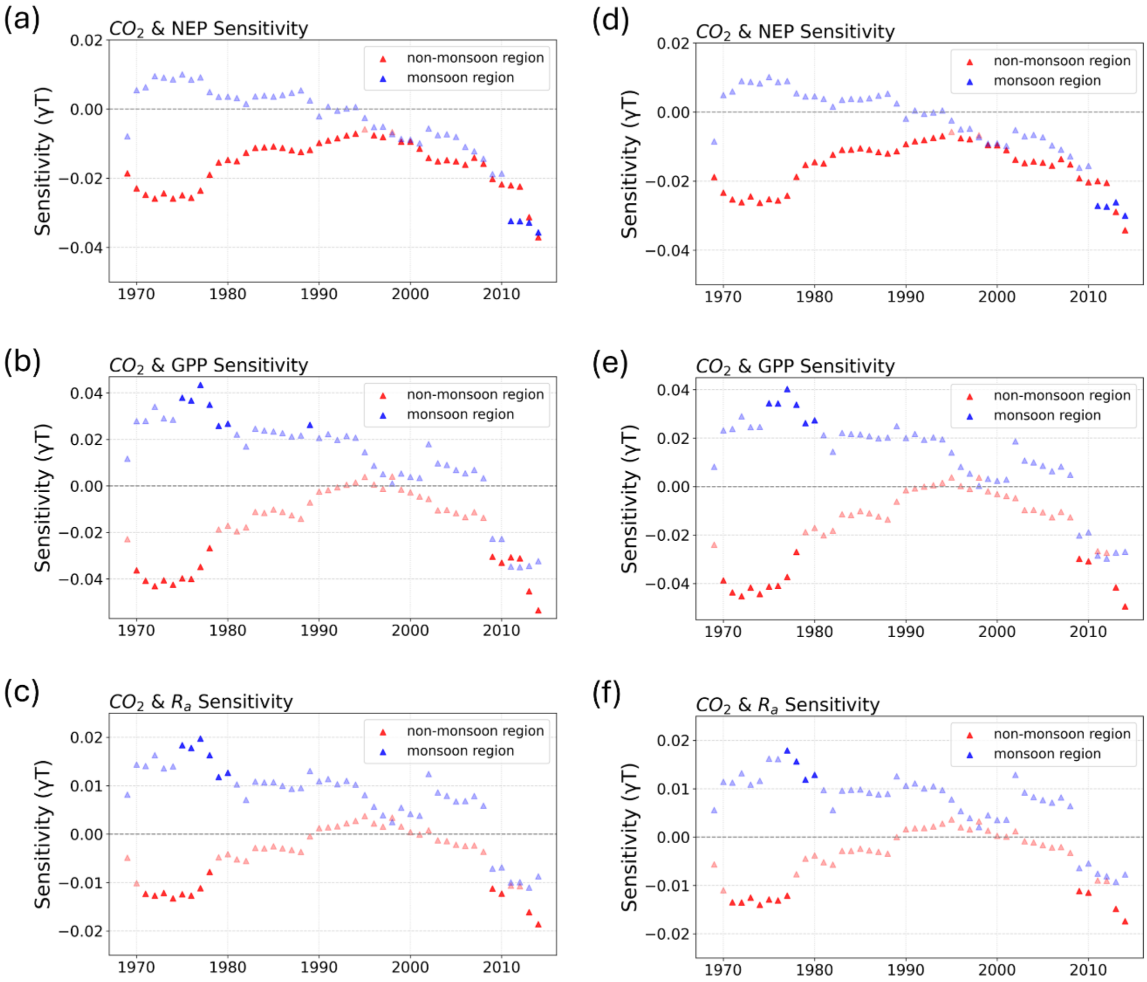


Figure R1. Time series of the sensitivity between atmospheric $CO_2$ growth and (a, d) NEP, (b, e) GPP and (c, f) $R_a$ using 20-year moving windows over the entire period from 1959 to 2023. Panels (a-c) show the results with a LOWESS smoothing span of f = 0.2, whereas panels (d-f) show the results with f = 0.3; the number of robustness iterations was fixed at 3 in both cases. Markers show the ensemble mean of nine TRENDY models for the non-monsoon (red) and monsoon (blue) regions. Dark markers indicate statistically significant regression slope within each moving window, whereas faded markers indicate non-significant slope; $p>0.1$ was considered statistically non-significant.

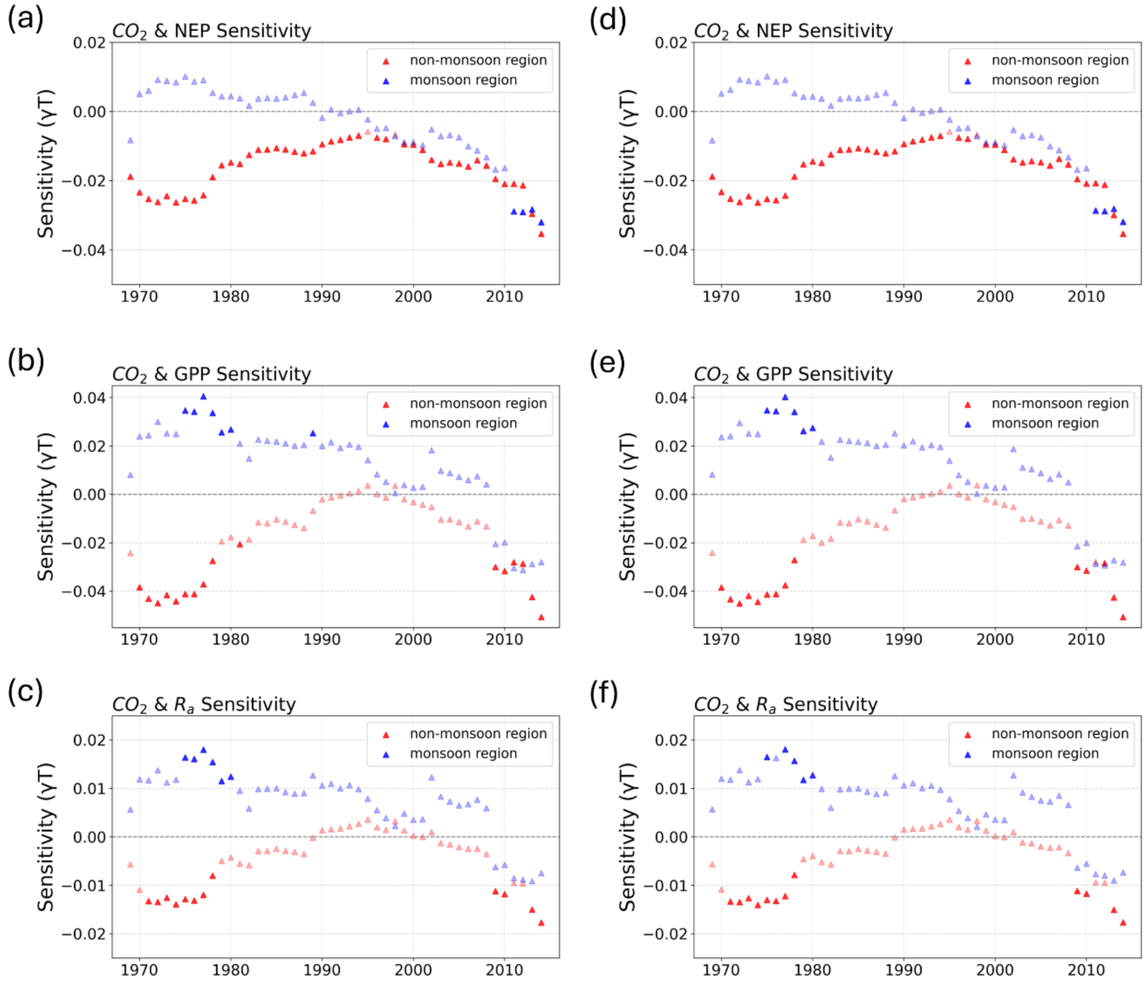


Figure R2. Time series of the sensitivity between atmospheric $CO_2$ growth and (a, d) NEP, (b, e) GPP and (c, f) $R_a$ using 20-year moving windows over the entire period from 1959 to 2023. Panels (a-c) show the results with 2 robustness iterations, whereas panels (d-f) show the results with 4 robustness iterations; the LOWESS smoothing span was fixed at $f = 0.25$ in both cases. Markers show the ensemble mean of nine TRENDY models for the non-monsoon (red) and monsoon (blue) regions. Dark markers indicate statistically significant regression slope within each moving window, whereas faded markers indicate non-significant slope; $p>0.1$ was considered statistically non-significant.

*2. In their response to the comment, the authors assert that the results of the sensitivity analysis do not depend on the window length. This could be stated if there are more than two sets of results with different window lengths - very short, medium, and long (it is a required minimum). Furthermore, we cannot ignore the specific characteristics of the LOWESS algorithm, which also uses a sliding window algorithm that could potentially influence the use of similar algorithms in subsequent stages. Therefore, the window length parameters used to determine sensitivity and those used for detrending must be consistent and synchronised. This is currently lacking in the paper, and the authors' response cannot be used as a basis for selecting the time window length.*

→ Practical comments! To address these concerns, moving windows of 10, 15, 20, 25 and 30 years were applied to cover very short, medium and long term scales (Figs. R3-R5). The sensitivity analysis showed that short term variability was greater when a 10-year period was used, whereas the temporal variations gradually became more moderate as longer periods were used. In particular, certain regional peaks and occasional changes in sign were more distinct with 10-year window.

Responding to the reviewer's comments, the following sentence was revised in the revised manuscript: "In addition, to examine whether the sensitivity pattern is influenced by the length of the moving window, 10, 15, 25 and 30-year moving windows were applied (figure not shown). The shorter the moving window, the greater the short-term variability, while the longer the window, the more gradual the temporal variation; however, the overall temporal patterns and trends were generally consistent with the results observed using a 20-year moving window."

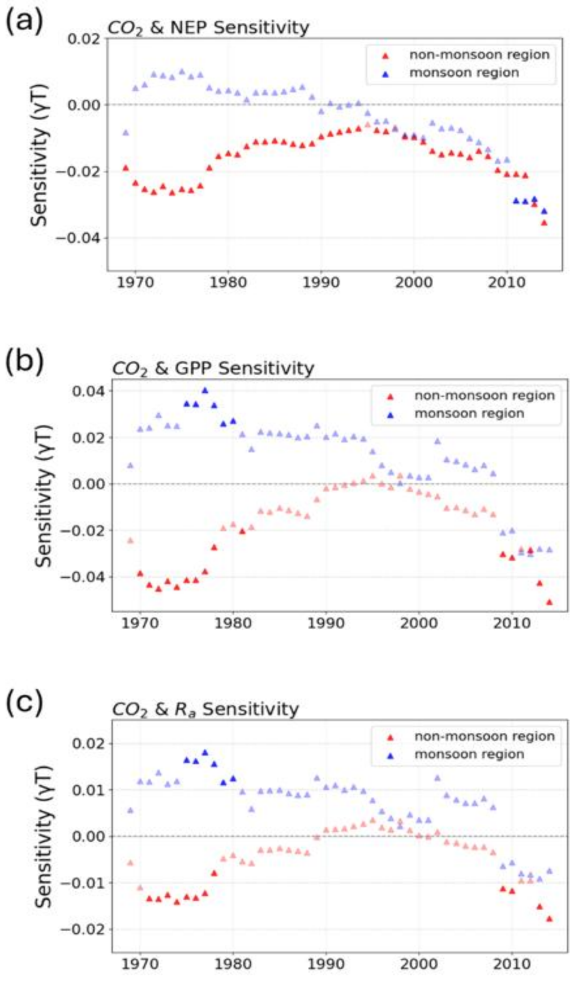


Figure R3. Figure 3 in the manuscript. Time series of the sensitivity between atmospheric $CO_2$ growth and (a) NEP, (b) GPP and (c) $R_a$ using 20-year moving windows over the entire period from 1959 to 2023. Markers show the ensemble mean of nine TRENDY models for the non-monsoon (red) and monsoon (blue) regions. Dark markers indicate statistically significant regression slope within each moving window, whereas faded markers indicate non-significant slope; $p>0.1$ was considered statistically non-significant.

Furthermore, regarding the reviewer's concern about the synchronization of the LOWESS 'f' parameter and the window length used in the sensitivity analysis, we would like to clarify that these two variables serve different purposes. In LOWESS, the parameter f determines the smoothing level used to estimate and remove the long-term trend from the entire time series, while the moving window length determines the number of years included in the regression analysis used to calculate sensitivity. This can also be seen in the methodology of Li et al. (2024), which was followed in this study. In Li et al. (2024), $f = 0.25$ was used to detrend the atmospheric $CO_2$ growth rate and tropical temperature, after which their sensitivity was calculate using a 25-year moving window. Their study also did not specify that the two variables were synchronized, nor did it explicitly require the smoothing scale to match the window length. We therefore do not consider the synchronization of the smoothing scale and the window length to be methodologically essential.

Accordingly, the following sentence was added to the revised manuscript: "The LOWESS smoothing span determined the level of smoothing used to remove the trend from the entire time-series, while the moving window length determines the period used in the subsequent sensitivity regression analysis."

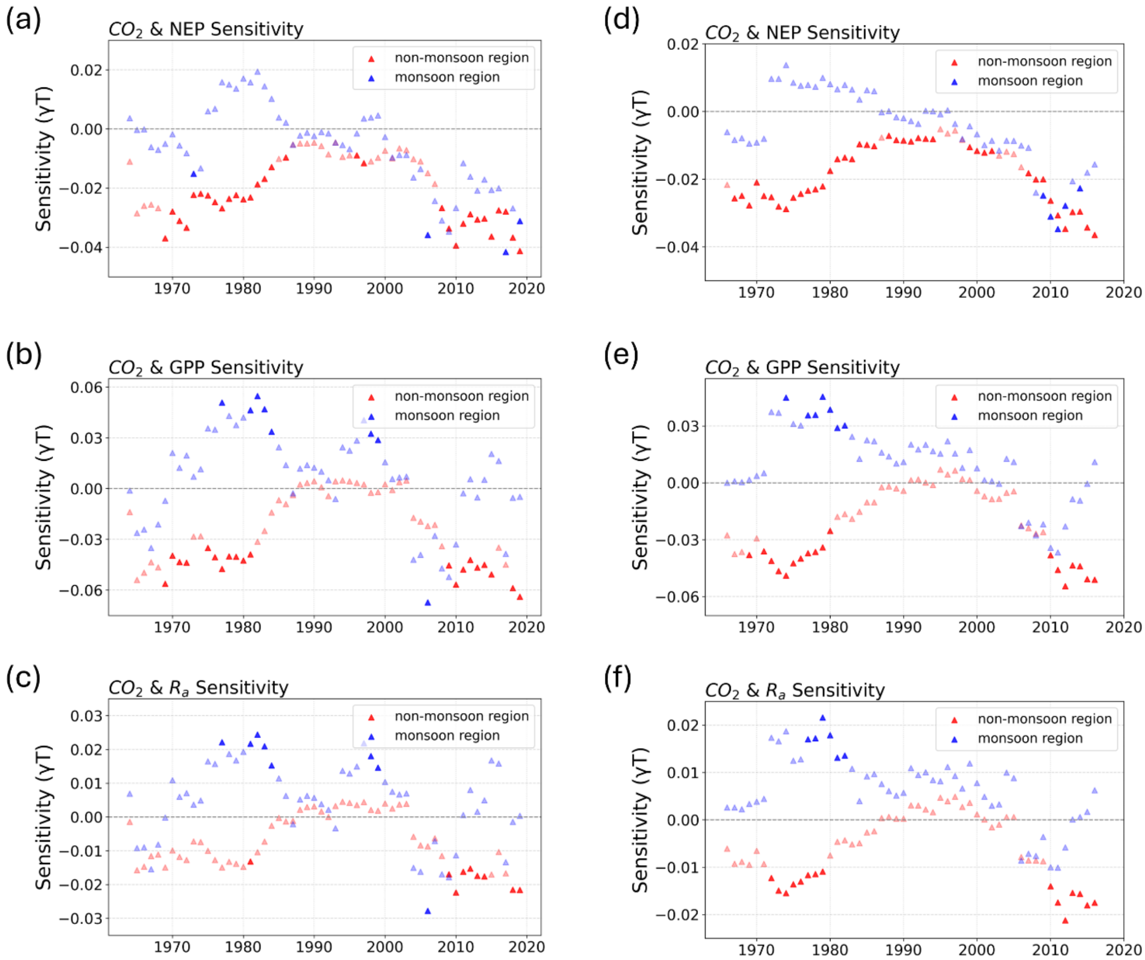


Figure R4. Time series of the sensitivity between atmospheric $CO_2$ growth and (a, d) NEP, (b, e) GPP and (c, f) $R_a$ over the entire period from 1959 to 2023. Panels (a-c) show the results using 10-year moving windows, whereas panels (d-f) show the results using 15-year moving windows. Markers show the ensemble mean of nine TRENDY models for the non-monsoon (red) and monsoon (blue) regions. Dark markers indicate statistically significant regression slope within each moving window, whereas faded markers indicate non-significant slope; $p>0.1$ was considered statistically non-significant.

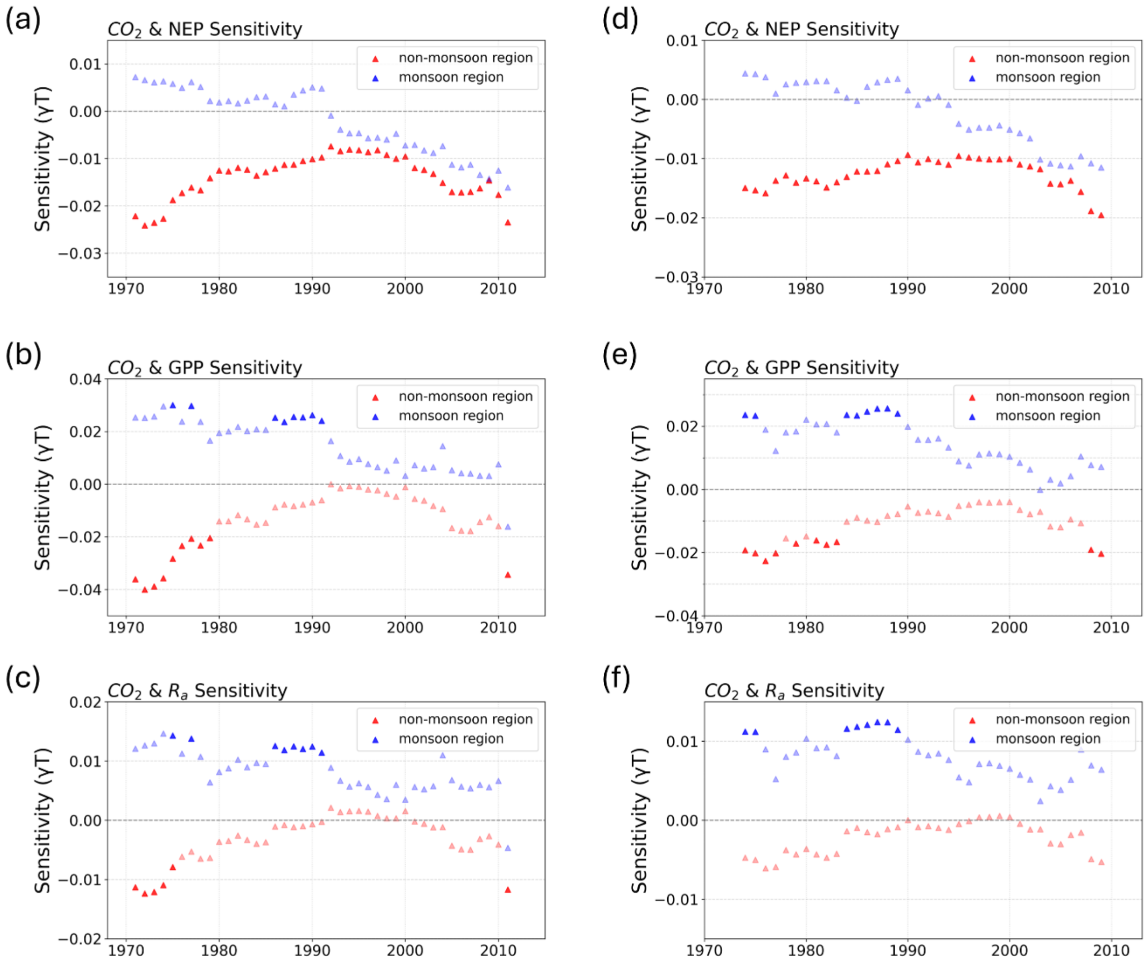


Figure R5. Time series of the sensitivity between atmospheric $CO_2$ growth and (a, d) NEP, (b, e) GPP and (c, f) $R_a$ over the entire period from 1959 to 2023. Panels (a-c) show the results using 25-year moving windows, whereas panels (d-f) show the results using 30-year moving windows. Markers show the ensemble mean of nine TRENDY models for the non-monsoon (red) and monsoon (blue) regions. Dark markers indicate statistically significant regression slope within each moving window, whereas faded markers indicate non-significant slope; $p>0.1$ was considered statistically non-significant.

*3. The authors based their response on monotonic behaviour of the sensitivities. This statement causes disagreement. The nature of the behaviour of the sensitivities individually and their relationships does not allow us to conclude that they are monotonic. In addition to this comment see the next comment below.*

→ Valuable comment! We are sorry for this misunderstanding. The monotonic relationship mentioned in major comment #3 of the previous reply letter refers to statistical characteristics evaluated using Spearman's correlation, and does not refer to the temporal variations in the sensitivity time series. We did not explain that sensitivity changed monotonically over time. Furthermore, the reason we mentioned Spearman's correlation in the previous reply was to provide an additional means of analysis in response to the reviewer's concern regarding the use of linear correlation; this allowed us to confirm the robustness of the Pearson correlation results.

Responding to the reviewer's comments, we revised the following sentence in the revised manuscript: "In addition, Spearman rank correlation analysis was performed to complement the linear correlation analysis of the relationships shown in Figure 4."

*4. In addition to the authors' response, it should be noted that for "ideal" linear relationships f1 = b1*x + a1 and f2 = b2*x + a2, the ratio of their values when x are equal is defined as b3 = b1/b2. Given the stochastic nature of the relationships, it can be argued that the closer the trends f1 and f2 under study are to linear relationships, the closer the values of the statistically determined ratio are to the "ideal" value, and vice versa: a significant difference between these two parameters indicates nonlinearity or non-monotonicity in the trends f1 and f2. Therefore, it is not possible to determine the characteristics of linear correlation between these trends or to use the corresponding tools of linear correlation analysis. Instead, it becomes necessary to investigate the behaviour of the trends in question and to look for suitable tools.*

→ Practical comments! We would like to clarify that *b1/b2* represents the ratio of the slopes of the two linear functions *f1 = b1*x + a1 and f2 = b2*x + a2*. However, this does not generally refer to the ratio of the two functions at the same value of x; the ratio of the values of the two functions is expressed as $\frac{f_1}{f_2} = \frac{b_1x+a_1}{b_2x+a_2}$ . When *f1* is expressed as a function of *f2* for the same value of *x*, the slope of the relationship between the two functions is *b1/b2,* and this can be expressed as follows:

$$f_1 = \frac{b_1}{b_2} f_2 + (a_1 - \frac{b_1 a_2}{b_2})$$

Furthermore, as discussed in major comment #4 of the previous reply letter, while the denominator shares the same variance of ACGR, the numerator contains different covariance; therefore, the relationship between the two cannot be expressed as *b1/b2*, i.e. *dGPP/dSoilMoisture*. Nevertheless, we acknowledge the reviewer's concern regarding the use of linear correlation analysis in cases where the relationship may not be strictly linear. To verify whether our results depend on the assumption of a linear relationship, we additionally performed Spearman's correlation analyses for both the early and late periods, as this method can assess monotonic relationships without requiring a strictly linear relationship (Table R1, Figs. R6&R7). The sign of the correlation was maintained in 11 out of the 12 comparison cases; the only instance where a difference in sign was observed was in a relationship that was close to zero and not statistically significant. However, a few differences were observed. In Figure R6c, d, the positive correlation between PAR and GPP sensitivity observed during the late period was statistically significant with Pearson's correlation, but not with Spearman's correlation. Furthermore, in Figure R7c, d, the positive correlation between PAR and GPP sensitivity was still significant in both periods; however, the temporal intensification observed in the Pearson correlation coefficient was not reflected in the Spearman correlation coefficient. We have therefore reflected these changes in the revised manuscript.

Table R1. Comparison of Pearson and Spearman correlation coefficients for the relationships shown in Figures R6&R7.

| Relationship | Period | Pearson correlation coefficient (significance) | Spearman correlation coefficient (significance) |
|---|---|---|---|
| Figure R6a, b | early period | 0.92 (***) | 0.86 (***) |
| | late period | 0.87 (***) | 0.92 (***) |
| Figure R6c, d | early period | 0.09 | 0.19 |
| | late period | 0.59 (**) | 0.26 |
| Figure R6e, f | early period | 0.89 (***) | 0.82 (***) |
| | late period | 0.85 (***) | 0.89 (***) |
| Figure R7a, b | early period | 0.01 | -0.13 |
| | late period | -0.30 | -0.30 |
| Figure R7c, d | early period | 0.46 (**) | 0.65 (***) |
| | late period | 0.77 (***) | 0.64 (***) |
| Figure R7e, f | early period | 0.13 | 0.20 |
| | late period | -0.50 (**) | -0.72 (***) |

Note: Early and late periods represent 1959-1998 and 1999-2023, respectively. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$; not significant, $p > 0.1$.

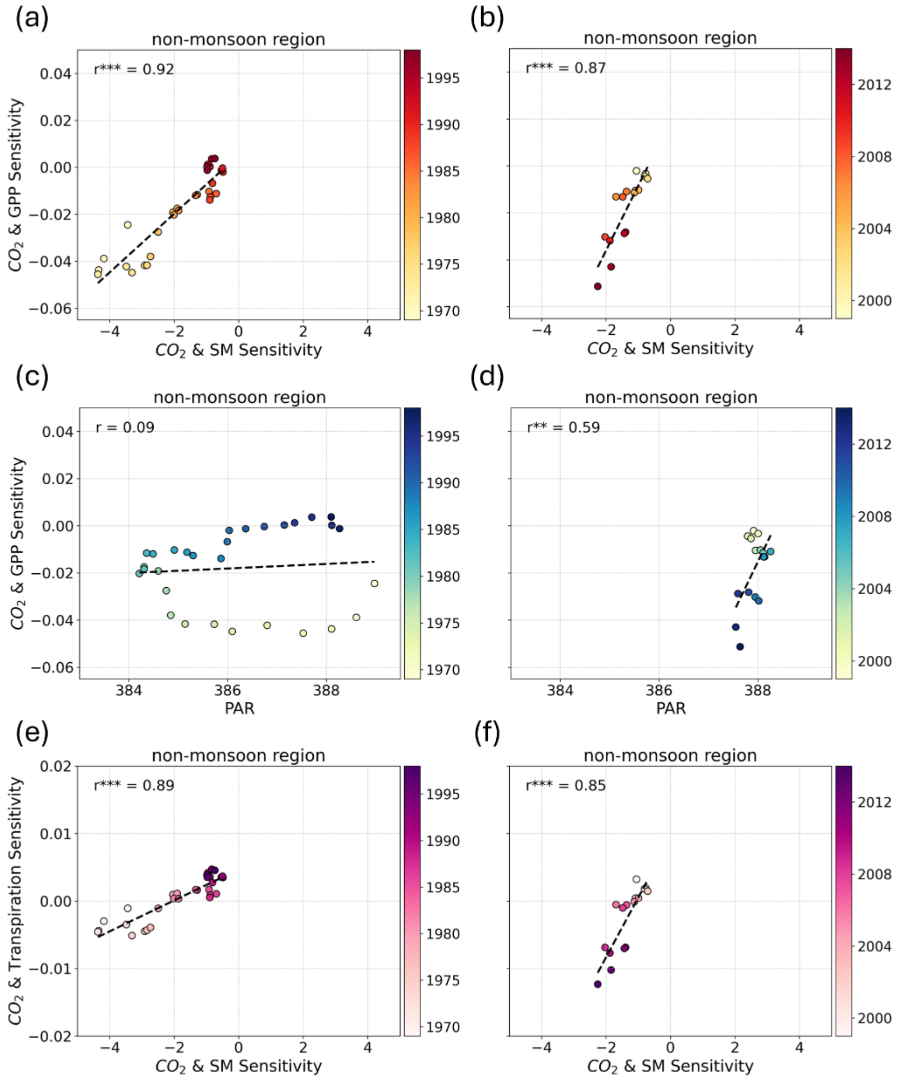


Figure R6. (Figure 5 in the revised manuscript). Correlation between (a, b) $CO_2$&GPP sensitivity (y-axis) and $CO_2$&SM sensitivity (x-axis), (c, d) $CO_2$&GPP sensitivity (y-axis) and PAR (x-axis), and (e, f) $CO_2$&transpiration sensitivity (y-axis) and $CO_2$&SM sensitivity (x-axis) for the non-monsoon region. The left and right panels correspond to moving window years of 1969-1998 and 1999-2014, respectively. Each scatter represents a 20-year moving window, with color changes representing the time period. Asterisks (*) indicate statistical significance levels where applicable: ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$; not significant, $p > 0.1$.

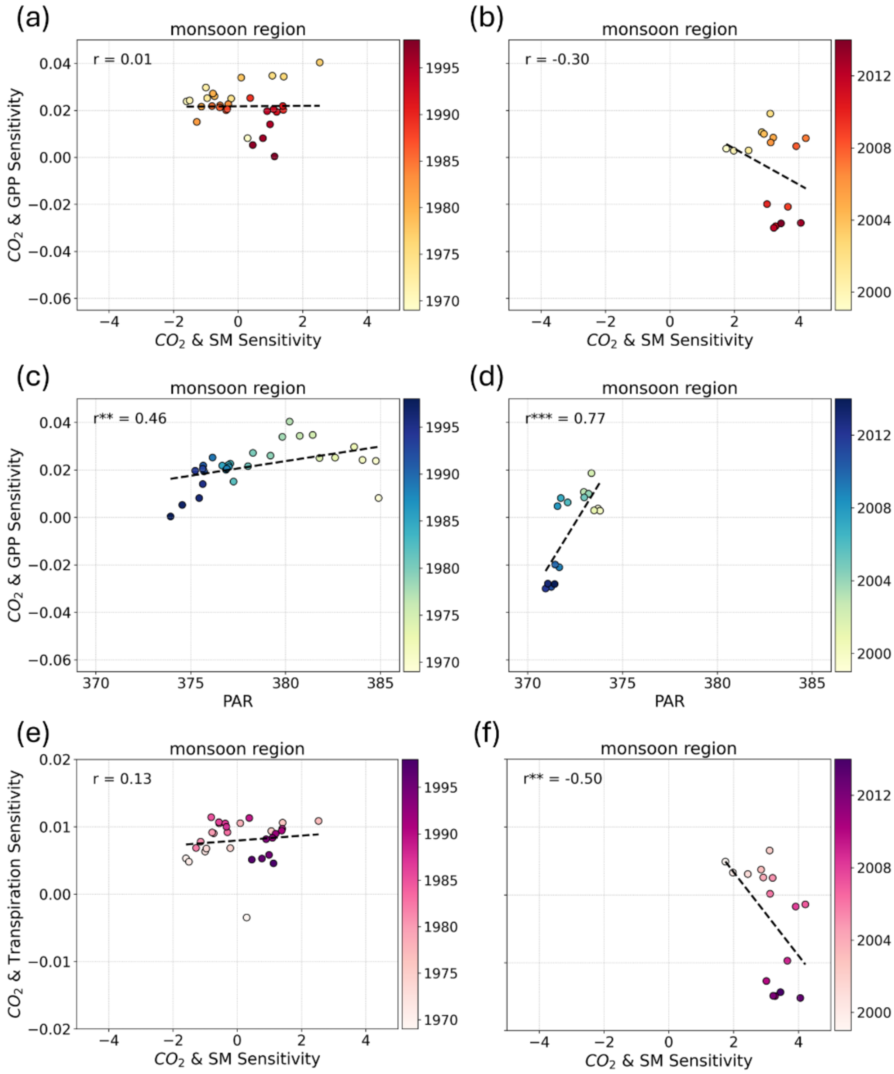


Figure R7 (Figure 6 in the revised manuscript). Correlation between (a, b) $CO_2$&GPP sensitivity (y-axis) and $CO_2$&SM sensitivity (x-axis), (c, d) $CO_2$&GPP sensitivity (y-axis) and PAR (x-axis), and (e, f) $CO_2$&transpiration sensitivity (y-axis) and $CO_2$&SM sensitivity (x-axis) for the monsoon region. The left and right panels correspond to moving window years of 1969-1998 and 1999-2014, respectively. Each scatter represents a 20-year moving window, with color changes representing the time period. Asterisks (*) indicate statistical significance levels where applicable: ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$; not significant, $p > 0.1$.

Responding to the reviewer's comments, we add the following sentences in the revised manuscript:

"In addition, Spearman rank correlation was also conducted for both regions during the early and late periods. Overall, the direction and statistical significance of the principal relationships were generally consistent with the Pearson correlation results. However, in the non-monsoon region, the correlation between PAR and GPP sensitivity during the late period was not statistically significant in the Spearman analysis. In the monsoon region, consistent with

Pearson results, the relationship between GPP sensitivity and PAR maintained a significant positive correlation in both periods; however, the increase in the correlation over time observed in Pearson correlation was not shown in the Spearman analysis. These results show that, while the primary correlations were generally robust across both analysis methods, certain temporal changes differed depending on the method used."

***Minor comments***

*3. Unfortunately, the given information does not reveal the content and significance of the improvements the authors have made to the basic method, so the original comment remains relevant.*

→ Please refer to the above major comment #1 and #2.

*6. One can agree with this statement if GPP sensitivity varies (in either direction) in a small number of regions. Therefore, to support this claim, authors could provide an estimate of the ratio between the number of regions where it remains relatively constant and those where it varies. In addition, it would also be useful and interesting to include additional markings on the map indicating the regions where GPP sensitivity has changed.*

→ Great suggestion! As explained in a previous response, the GPP sensitivity maps for the entire period were not intended to represent the temporal invariance of individual grid cells. Its purpose was rather to provide a representative overview of the overall spatial distribution and to classify the study area into monsoon and non-monsoon regions, following the monsoon boundaries defined in previous studies. Following this regional classification, the present study independently assessed the temporal variability of GPP sensitivity using moving window time-series for each region (Figure 3b in the manuscript).

However, we agree that directly presenting the temporal changes in spatial patterns would clarify this point more effectively. Accordingly, we compared the signs of the entire period's average GPP sensitivity with those of the GPP sensitivity for each of the 46 moving windows across each grid. As a result, 64.73% of the grids showed the same sign as the average for the entire period in more than 70% of the moving windows, while 35.27% of the grids had a consistency rate of less than 70%. Consistent with the regional temporal changes observed in Figure 3b (in the revised manuscript), the sign and spatial distribution of GPP sensitivity were shown to vary over time in some grids. Therefore, we present all GPP sensitivity maps for each of the 46 moving windows to visualize the specific patterns of these spatial variations (see Fig. R8 below). Indeed, GPP sensitivity varied by region over time, and these spatial variations were found to correspond to the distinct temporal variability patterns identified in regional time-series analysis (Figure 3b in the revised manuscript). Therefore, the GPP sensitivity map for the entire period was used as a representative spatial indicator for regional classification, while temporal patterns and regional differences were assessed using time-series.

Accordingly, the following sentence was added to the manuscript: “Although the spatial pattern shown in Figure 2b represents the average GPP sensitivity over the entire period, some grid cells showed temporal variations in sign and spatial distribution (figure not shown). Nevertheless, the overall spatial pattern was used for regional classification. These temporal patterns and regional differences can be identified through time-series analysis.”

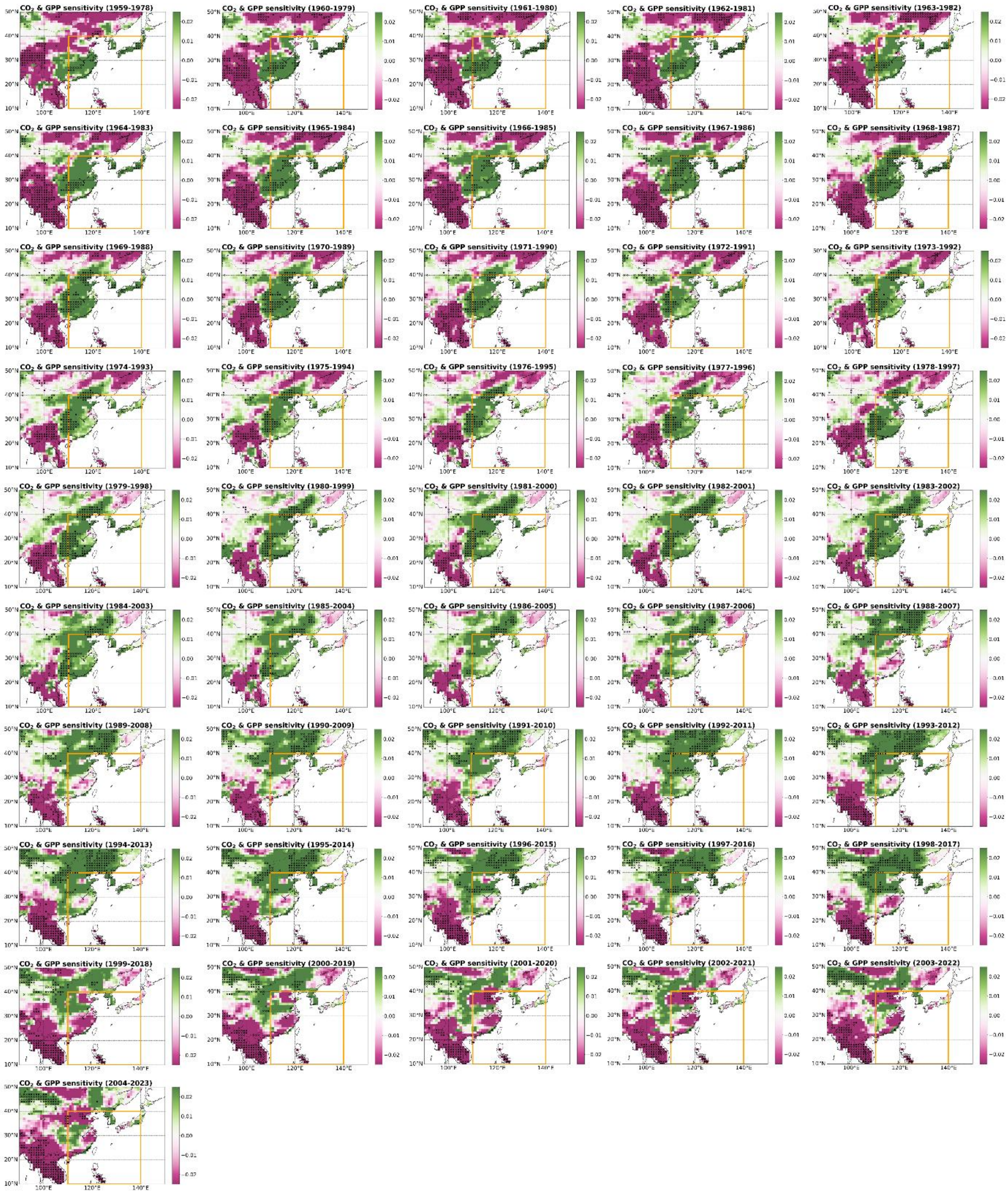

Figure R8. Spatial patterns of the sensitivity to atmospheric $CO_2$ growth and GPP over East Asia, based on the ensemble mean of nine TRENDY models during 1959-2023. Each panel shows GPP sensitivity for each 20-year moving window, with a total of 46 moving windows. An x mark denotes statistically significant relationships at the 90% confidence level.