Editor:

Dear editor,

You may find below the replies to the three reviewers of the second round of review, together with some replies to your comments. We believe that we have addressed all reviewers' comments, and tried to stress cases where we had already replied to past rounds of reviews sufficiently, which was missed in the second round of reviews. A clear example is reviewer #4 who says that we have not provided point-by-point replies to the first round of reviews, which is obviously not the case.

The fifth reviewer has recommended a rejection, and raised concerns, like the previous three reviewers, about the methodological approach and assumptions behind it.

Thank you for giving us the opportunity to reply. Based on what we see in the manuscript submission page, two reviewers said major corrections (but no rejection), and our replies to them are already online since the last review round. One reviewer said minor corrections and accept, and two reviewers said reject (not three). For the latter three reviews, you can find our detailed replies below. We are also submitting a track-change version of the manuscript, as required by the journal.

All reviews, regardless of their recommendation on acceptance or rejection, were constructive, and helped to greatly improve the manuscript. They asked for explanations of methodology, a restructuring of the results section, and the inclusion of some new results to further support our conclusions. We performed all such requested items.

We would like to highlight a few cases where we respectfully disagreed with the reviewers: reviewer #4 asked for a statistical analysis of air pollution, but we never claimed that we are doing anything related with air pollution in the manuscript; reviewer #5 denied that soil moisture is a proper driver of agricultural productivity, and we made a case on why this is, in our opinion, an incorrect claim; and both reviewers #4 and #5 said that we are not presenting an analysis on model response for GPP/NPP, which in fact is included in the revised manuscript from the first round (the one that both reviewers were reviewing) as a response to the first round of reviews.

We invite you to pay particular attention to the weight of the comments of reviewer #4, which, in our opinion, are not major enough to justify a straight-out rejection. Also worth checking closely is our reply to the last comment of reviewer #5, which summarizes all the changes we have done over the two review rounds.

While we acknowledge that the authors have made an effort to address the comments from the first round of reviews, the revised results are not sufficient to convince independent

referees of the value of the study. There seems to be a fundamental lack of understanding and/or of convincingly illustrating the purpose and methodology of using the drought indices to derive the presented conclusions.
Please see the replies to the first comment of reviewer #5, which outlines why we believe our methodology is, indeed, scientifically sound.

The authors are given the chance to respond to the second round of reviews (#3 through #5), but are reminded that the outcome of a further revision will not necessarily lead to acceptance despite the one positive review, and that as the fifth reviewer also declined to re-review any revised manuscript, any resubmission will likely experience another delay in the manuscript processing due to difficulty in finding more reviewers.

Thank you once more for the opportunity given. We understand that the decision might take longer, given your statement that of the five reviewers one is not willing to re-review the manuscript. We are looking forward to your decision.

Responses to each comment are provided in blue text, with italicized text indicating specific additions to the manuscript and supplementary materials.

Reviewer 3:

Responses to each comment are provided in blue text, with italicized text indicating specific additions to the manuscript and supplementary materials.

Overall this work is very good organized, and the model design are very reasonable. The paragraph figure structure is very reader friendly. So I recommend a technical corrections, for the minor technical issues.

Thank you.

line 164 CH4: use subscription
Done.

line 264-267: this sentence is too long and hard to understand, since PDSI and CMI has been quoted before, you can just use this abbreviation without mentioning the full name or the citation. Even though, this sentence is also too long, please consider to divide it into two.
We modified the sentence:

> “*The limitations of PDSI and CMI in the formulation used for the PET calculation (Thornthwaite, 1948), together with the lack of consideration of land cover types on the water balance, have encouraged the exploration of ETDI for agricultural productivity.*”

line 352-354: is there a missing 'is' before consistent?

Yes, fixed.

Reviewer 4:

Responses to each comment are provided in blue text, with italicized text indicating specific additions to the manuscript and supplementary materials.

This manuscript aims to study the effect of volcanic eruptions on moisture-based drivers of plant productivity. The results indicate a 0.5 °C surface cooling due to the Mt. Pinatubo eruption. However, this is not a major finding, as previous studies have already reported this level of surface cooling following the volcanic activity at Mt. Pinatubo.

Our study does not present the surface temperature response as a novel contribution. Rather, as mentioned as early as in the abstract (quoted below), the primary focus of this work lies in analyzing ecohydrological conditions and their implications. More specifically we focus on drought indices, plant productivity, and their spatiotemporal relevance from seasonal to sub-seasonal (weekly) time scales. Notably, the weekly time-scale analysis provides important insights into agricultural productivity. To the best of our knowledge, this is the first study to offer a comprehensive examination of these critical drivers of plant productivity. The inclusion of surface temperature is intended to support the evaluation of the model’s performance and to provide necessary context for the ecohydrological analysis, and the agreement of our model with past studies demonstrates that the temperature response is well within the bounds of past literature.

“*Here, we will explore the understudied store (soil moisture) and flux (evapotranspiration) of water as the short-term ecohydrological control over plant productivity in response to the 1991 eruption of Mt. Pinatubo.*”

The major flaw of the study is the lack of a direct statistical analysis of the impact of atmospheric air pollution from volcanic activity on climatic and hydroclimatic parameters.

The focus here is on the well-established climatic impacts of stratospheric aerosol injections following explosive volcanic eruptions, which perturb Earth's radiative balance and influence climate and hydroclimatic variables. Air pollution is not a topic we study in this manuscript, so trying to correlate air pollution metrics with climatic or hydroclimatic parameters would be out of scope.

About statistics in general, our analysis examines the climatic response to a volcanically perturbed atmospheric radiative balance, extending beyond surface temperature to include higher order impact metrics such as rainfall, soil moisture, and surface heat fluxes (actual and potential evapotranspiration), which are critical to understanding plant productivity through land–atmosphere interactions. The response is detected using a counterfactual inference approach and evaluated through paired Student's t-tests at a 95% confidence level. Responses not meeting this significance threshold are masked, as already mentioned in several of our figures. We present the full suite of responses—including temperature, rainfall, and higher-order drivers—alongside drought indices (SMDI and ETDI), which reflect soil moisture anomalies and atmospheric moisture demand, respectively, and only where statistical significance is calculated.

Individual anomalies in simulated explanatory variables were inferred as results of volcanic eruptions, but only over 10- 15% of the area. No direct evidence of the decrease or increase in plant productivity is presented.

Direct evidence of plant productivity is inferred through soil moisture–based drivers at both global and regional scales, with the geographic dominance of specific drivers discussed in the conclusion. Additionally, we present changes in gross primary productivity (GPP) as direct evidence in the conclusion (lines 751–763 TC/710-726) and in Supplementary Figure S9. This has also been clarified in response to Reviewer 2.

"In contrast, in the far northern latitudes, water is not the primary driver of plant responses, and productivity is likely to decline. Seasonal-scale changes in gross primary productivity (GPP) confirm the regional trends in plant productivity following the eruption. The simulations show a more pronounced decrease in GPP in the northern high-latitude region and a significant increase in GPP over the European and Mediterranean regions. Additionally, distinct patterns of decrease and increase in GPP are simulated in the tropical northern and southern regions, respectively (Figure S10)."

The estimate of 10–15% of global land area represents a substantial and societally relevant extent.

Terminology in the manuscript is not clearly defined. Section 2.3 on methods is difficult to understand and poorly written.

We modified the text of section 2.3 (quoted below), to clarify the message we were trying to pass.

"

2.3 Methods: *This study investigates the impacts of the Mt. Pinatubo eruption on the major drivers of primary productivity focusing on soil moisture and evapotranspiration related metrics. We followed the counterfactual inference approach to draw causal inference of the Pinatubo eruption. Hereafter, we use 'PCH' (Mt. Pinatubo and Cerro Hudson) to refer to the 'GISS-PIN-SO2' and 'NP' for the 'GISS-NOPIN-SO2' ensembles. We have included the Cerro Hudson eruption in both ensembles since we are focusing on the Mt. Pinatubo-driven climate response.*

2.3.1 Statistical analysis for detecting Mt. Pinatubo-significant regions and anomalies calculations.

We treat the no-Pinatubo ensemble (NP) as a counterfactual climate simulation and utilize it to perform the paired Student's t-test for causal inference. The null hypothesis is that the ensemble means of a quantity of interest (QoI) in a region over a time period are the same between ensembles (i.e. $H_0: \bar{\mu}_{PCH} = \bar{\mu}_{NP}$*). In the subsequent figures in this document, gray regions indicate acceptance of the null hypothesis at the 95% confidence level, while the coloring emphasizes the rejection of the null hypothesis and the significant regions of anomalies relative to 1950-2014 climatology (see Supplementary information section S1.0). However, we also explored the alternate approach of directly comparing the difference between the two ensembles (PCH and NP) for presenting the Pinatubo effect (see Supplement Figure S2). It is concluded that both of the approaches led to the same general conclusions, with only small quantitative differences. Nevertheless, we chose to remain consistent with the baseline requirements for other metrics, and used the historical climatology for the same period 1950-2014 as the baseline for the core of our analysis. Thus, the grey areas indicate no significant differences between the PCH and NP ensembles, while the colored regions represent the statistically significant anomalies, calculated as PCH ensemble mean minus climatology.*

"

The interchangeable use of terminologies adds to the confusion. For example, "counterfactual (sometimes counter-factual) ensemble simulation" and "no-Pinatubo ensemble simulation" refer to the same simulation but are used interchangeably throughout the text.

We apologize for the inconsistent terminology. Since our no-Pinatubo simulation (GISS-NOPIN-SO2) is named "NP" in the manuscript, as mentioned in line 199, we removed the word “counterfactual” from the instances where we were referring to the model simulations and replaced it with “NP”.

Please find below some additional comments:

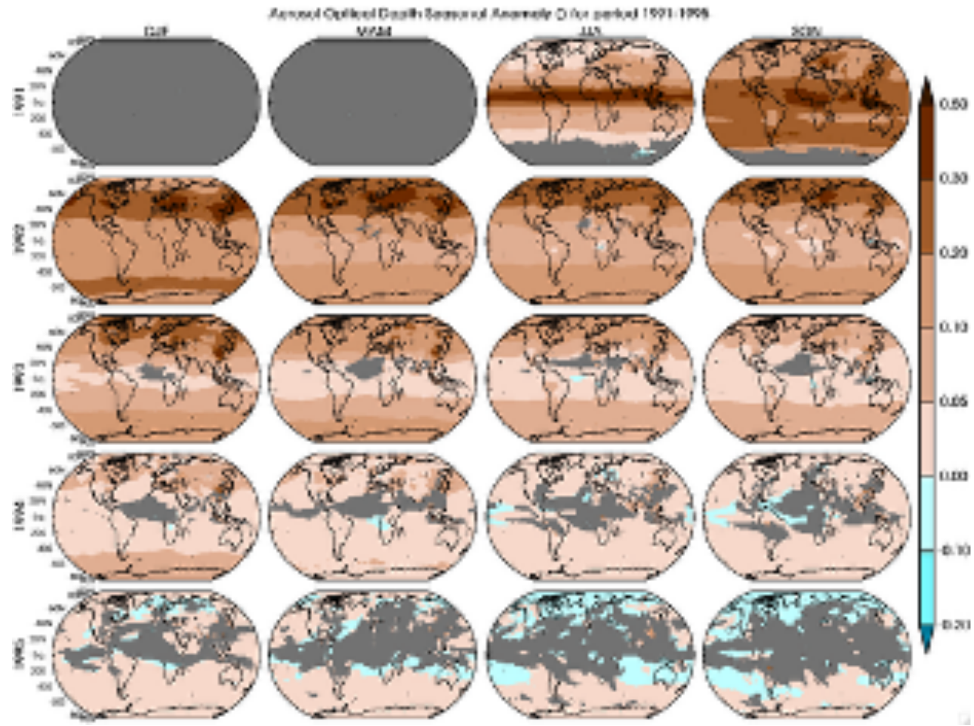
In Figure S3, abbreviations such as LW, WS, and NET are not defined either in the main text or in the figure caption.

We added explanations in captions on what these abbreviations mean.

“*Monthly anomaly of longwave (LW), shortwave (SW), and net (NET) radiative forcing simulated by the GISS ModelE for Mt. Pinatubo (PCH) and without Pinatubo (NP) ensembles*”

Line (337). “The zonal AOD shows the dispersion and transport of aerosol poleward after the eruption.” However, it is not clear from Figure 1 how this conclusion is achieved. It would be better if the grid cell level variation with temporal scale is represented after the eruption event.

We do not exactly understand the reviewer’s comment here. Figure 1 clearly shows an increase in zonally-averaged AOD poleward with time, indicating aerosol transport toward higher latitudes. The initial high AOD blob near the volcano (in the tropics during the second half of 1991) migrates both north and south, and about a half year later significant amounts of it have reached both poles. About 3 years after the eruption the whole atmosphere (in terms of AOD load) has returned to its unperturbed state. All these are clearly visible in the figure and the past literature (e.g., Aquila et al., 2012; Anderson et al., 2015; Singh et al., 2023). Presenting grid-cell level variations over five years (60 monthly-mean differences) is challenging and unlikely to add any meaningful insight, since we are not interested in the longitudinal variability, which is either way very minimal.
A plot similar to what the reviewer has asked for would look like the one below, which shows the seasonal structure of grid point AOD for 5 years.



Line343: QBO is not defined.

Fixed.

"Meanwhile, the phases of QBO (Quasi-Biennial Oscillation) and local heating also play a crucial role in the poleward and vertical dispersion of stratospheric aerosols"

Line 343-355: The spatial connection of AOD's effect on temperature changes cannot be established using Figure 1. Grid-level correlations could provide more insight into their connectivity.

The correlation of temperature response with AOD is not a direct one and not even a perfect one, as seen in Figure 1. Clouds respond to AOD changes as well, even if the aerosol layer is in the stratosphere, and these cloud changes also impact temperature. Other than the complexity of generating the figure requested, which was already addressed in comment about line 337 above, its interpretation would have been a complex task that would have led us outside the scope of this work. The zonal mean plot presented shows that in general higher AOD and lower temperature correlate, but not exactly, and this is a common finding across modeling studies of the same topic. We decided to make no changes.

The seasonal anomalies of temperature and rainfall presented in Figure 2 and Figure 3 are with reference to long-term climate conditions that is from 1950-2014. No rational is provided for this approach either in the result section or in method section. How to account for GHGs radiative effect during this long-term reference period?

The explanation for selecting the reference period 1950–2014 is already provided in lines 207–217 (Quoted below), as well as in Supplementary Section S1 and Figure S1. These illustrate the suitability of this period for capturing the volcanic response with a few tests we performed and explain in the text and supplement. We also addressed this in our responses to both Reviewers 1 and 2.

"*It is concluded that both of the approaches led to the same general conclusions, with only small quantitative differences. Nevertheless, we chose to remain consistent with the baseline requirements for other metrics, and used the historical climatology for the same period 1950-2014 as the baseline for the core of our analysis*"

The anomaly comparison years for temperature are 1991-1995 but for rainfall only two years are presented. No explanation for this inconsistency is available.
The two-year rainfall anomaly comparison was included in order to illustrate the complexity and uncertainty of rainfall response, particularly over land, which motivated us to largely focus on drought metrics instead. The rationale for this is already explained in section 3.3.

Lines 335-335: The results shown in Figure 4 are significant only over forest land (Congo tropical forest and Russian boreal forest), which has a much deeper root zone compared to croplands. In Figure S6, the land mass area has pixels of a similar color to the ocean. Is there any explanation for this?

We assume this is about FIgure 4 and not line 335 which is not relevant.
In Figure 6, the lighter grey shading over land compared to the ocean reflects lower soil moisture levels at a depth of 2–4 feet, due to the presence of non-permeable soil layers. This is explained in text section 3.4.1 (paragraph starting at line 505 and 515 in track changes) and figure captions. We have added that explanation to the Figure 5 and FIgure S6 legend.

"The light grey colored regions represent regions of impermeability."

Since the study is focused on drivers of plant productivity, remove the ocean region from temperature and rainfall figures to have a consistent study area with SMDI and ETDI.

Temperature and rainfall are not solely influenced by land surface processes. Masking ocean regions in surface temperature and precipitation analyses is unconventional and results in the loss of critical information, particularly given the strong land–ocean feedbacks. In contrast, SMDI and ETDI are land-specific indicators that capture the direct influence of land surface conditions. As higher-order impact metrics, they are inherently more representative of land-region responses.

Figure 7 should just be in supplementary materials. However, the rationale for selecting areas in different regions is unclear. It appears that the selection is based on regions with clusters of grid cells showing significant anomalies.

The reviewer is correct that the selection was guided by significant regional anomalies, in order to deeply understand the reason why these regions stand out. The rationale for this analysis is already stated in the abstract and supported by region-specific conclusions (see e.g. Abstract line 29-31, section 3.6 (quoted), paragraph Conclusion section). However, we do agree that this figure can move to the supplement, which is now the new Figure S7 there.

"*Considering the complexity of the representation of spatial features, we selected three distinct regions (shown in Figure S7 and detailed in caption) in the northern hemisphere based on the climate response to Mt. Pinatubo in the seasonal analyses presented in Section 3.0*"

A point-by-point response to the first review was not provided by the author, making it difficult to evaluate whether the major concerns were fully addressed.
The point-by-point responses to both reviewers are publicly available on the journal discussion page (see https://doi.org/10.5194/egusphere-2024-2280-AC1 and https://doi.org/10.5194/egusphere-2024-2280-AC2). The reviewer might have missed the binder icon to the right, which links to our reply.

However, upon reviewing the track-changes-enabled draft, the following anomalies were observed.
Lines (172-177): Why delete all the lines when only change in the original text is the reference to Figure and correction in cited paper?
Apologies for this, we are not sure why the whole sentence appears modified. Indeed we only modified the figure reference and a citation that was incorrect.

It's not clear what revision is made between lines 260-265. It seems like cut and paste.

Apologies for this, we are not sure why the whole sentence appears modified. The only change is the word “to” that became “of”.

Lines between 378-384, All the revision is cut-paste only change is reference to Suppl. Figure.

One more case where for reasons unclear to us the word processor decided that the whole sentence was modified. Apologies.

Revision between Lines 763-775 appears to be selecting alternative word. It does not add value to the text.

This is correct, since one of the comments we received was to improve the text, and this is exactly what we tried to do. As the present reviewer is not recommending any specific changes, we assume that their comment is a statement, rather than a request for further changes.

Reviewer 5:

Responses to each comment are provided in blue text, with italicized text indicating specific additions to the manuscript and supplementary materials.

Currently, this manuscript sits in a weird spot. Methodologically, it didn't to validate the simulated atmospheric response, which is understandable. But the critical problem lies in analysis.
"Moisture-based drivers of plant productivity" is a bad middle ground between hydroglogy and ecophysiology - They could have directly look at GPP/NPP, and examine how temperature/hydrological effects drive those changes; alternatively they could focus on hydrology, which could have broader implications beyond terrestrial ecosystem productivity (e.g. flood risks). Instead, they focus on 2 drought indices (that doesn't offer much extra insight) and made unbacked speculations about how that would impact terrestrial ecosystem/agriculture. From the record of the last round of revision, it doesn't seem that the authors are aware of this major issue.

We respectfully disagree with the reviewer’s comment regarding soil moisture as an unsuitable indicator for linking the hydrological cycle and plant productivity. Soil moisture is a key hydrometeorological variable that directly influences plant health particularly in agricultural systems - by supporting photosynthesis, nutrient transport, and root development (Seneviratne et al., 2010; Dirmeyer et al., 2006; Munoth et al., 2006; Garg et al., 2016). As shown in this study, the precipitation signals across timescales possess a high uncertainty and even short-term soil moisture deficits can significantly reduce productivity. Therefore, soil moisture remains a robust and appropriate driver for assessing agricultural productivity.

Nonetheless, based on the reviewer #1 suggestion, we incorporated this analysis into the revised manuscript. Despite the acknowledged complexities, our conclusions are well supported by the model-simulated GPP response. Our response to reviewer #2 on the topic is repeated here:

"*The aim of this study is to focus on hydroclimate metrics. Transpiration is the most dominant process contributing to AET on land and is strongly correlated with photosynthesis. Thus, an increase in AET serves as a reliable indicator of an increase in GPP. Consequently, we have chosen not to emphasize GPP in this analysis. However, we have revised portions of the manuscript to address this concern. Additionally, we included a plot illustrating the seasonal anomaly of GPP in the supplementary information (Figure S10) and provided a discussion on plant productivity in the conclusions section, along with examples of similar findings from other studies.*" **(Conclusion section)**

Regarding the selected drought indices, these are specifically designed to capture agricultural drought conditions and offer a comprehensive representation of soil moisture storage (SMDI across different soil depths) and key land–atmosphere interactions (ETDI), which reflect the balance between potential evapotranspiration and actual plant transpiration. These indices effectively capture both short- and long-term dry/wet conditions and avoid the known limitations of water balance models used for the PDSI (Narasimhan and Srinivasan, 2005).

As similar issues have be highlighted by earlier referee reports, but still not properly addressed by the authors, I recommend rejection.

The reviewer's comment is a little vague: "...similar issues…not properly addressed…" but without being explicit on which issues and how we failed to address them. This is the summary of the previous review round (major comments only):

- Reviewer #1 single major comment was related with the justification for the reference period. We addressed that comment by demonstrating in the revised manuscript that using other plausible reference periods the answers did not change qualitatively, and our conclusions remained the same regardless of which period was chosen as a reference (Additional figure (S2) is added in supplementary).
- Reviewer #2 had major comments but did not reject the paper: "I therefore recommend publication only after major revisions". Their major comments were: 1) a restructure of the paper to streamline the purpose and message more cleanly, which we had done; 2) discuss more about the model skill and in particular in the light of past ModelE studies, which we have done; 3) a question (not a major comment really) on why we did not use prescribed volcanic forcing and get more ensembles in, which we have answered; and 4) a request to add GPP or NPP anomaly plots, which we have added.

Then from this review round, which reviewer #5 had no opportunity to see before commenting, so only mentioning them here for completion):

- Reviewer #3 was positive with very few minor comments (answers above).
- Reviewer #4 had a number of comments (answers above), none of which were major, in our opinion. The only major point named by the reviewer was the lack of a statistical analysis of air pollution, which is not something we studied here in the first place, so the comment was not relevant.

In summary, we have substantially revised the manuscript to incorporate all suggestions from Reviewers 1 and 2, and we also included an analysis of the model-simulated GPP response, which supports and complements the conclusions drawn from the soil moisture-based drivers.