Response to Reviewer 2

We thank the reviewer for their valuable comments. We present below the comments in black, our replies in blue, and the changes made to the text of the manuscript in blue italic font. Line numbers correspond to the original manuscript as well as the new revised manuscript.

General: The manuscript investigates the hydrological response of a series of volcanic eruptions during the 2nd century BC, focusing also on the societal impacts of the eruptions in the context on the Egyptian history. In addition to empirical evidence, authors use the output of an ensemble of simulations with a comprehensive Earth System Model, simulating potential trajectories of plausible climatic scenarios in the aftermath of the volcanic eruptions.

The manuscript is very well written, material and methods are comprehensively presented and results are discussed within the context of present literature. Therefore, I think the manuscript should be published with some minor comments addressed below. Most specifically, the comments relate to the modeling and statistical part, including a more nuanced discussion and interpretation of model results in the context of past civilizations. Moreover, I would encourage the authors to reduce the overall length of the manuscript by summarizing dedicated paragraphs or moving parts into the supplementary material whenever possible.

The manuscript is revised as described in detail under the specific replies below. We have reduced the main manuscript length as recommended. Some of the old text has been deleted, while other text has been moved to the supplementary material. We note that the manuscript is still comparatively extensive, but this is in line with other published article in this special issue, which (given their interdisciplinarity) have had to introduce a wider range of methods and contexts than might otherwise be required.

1 Introduction:

l. 114: Linking volcanic eruptions directly to revolts or warfare might be afflicted with high degree of uncertainty: In past societies single upheavals or riots always happened – likewise, a close inspection of ice core records will typically also yield one or two eruptions per decade. Linking both just because of their synchronicity might be co-incidental. The processes of both, the impact of the volcanic eruption on climate and the prerequisites leading to riots or revolts during the period previous to the volcanic eruption can have multiple drivers and causes. Therefore this line of evidence in terms of wiggle matching single historical events with volcanic outbreaks should be handled with care.

We fully agree with the concerns listed by the reviewer. In the manuscript, we have now clarified that our work takes place in the context of the previous results of Ludlow and Manning (2016, 2021) and Manning et al. (2017), which considered in detail the correspondence between explosive eruptions (using the volcanic chronology of Sigl et al. (2015) and independently dated revolts across the Ptolemaic era. This work has identified the correspondence as highly statistically significant, also identifying significant correspondences between dates of phenomena such as sales of hereditary agricultural land, which have been previously hypothesized to occur with greater frequency during times of socioeconomic stress, as might follow years of poor Nile flooding, providing a glimpse into the mechanisms by which the impacts of explosive volcanism on Nile flooding might ultimately provoke revolt.

Scholars have now also recommended that case studies of more specific periods and regions might shed greater insight into the drivers and prerequisites of complex phenomena such as revolts. Our present paper thus intends to provide a foundation for such a case study by offering greater detail about the likely hydroclimatic impacts of a key decade in Ptolemaic Egyptian history, the 160s BCE.

l. 118: “hydroclimatic shocks” should be replaced by “pronounced changes in hydrology”

We have corrected the relevant sentence at line number 118 to state “*pronounced changes in hydrological cycle”*. (Line 152 in Revised Manuscript)

l. 120: The hydrological cycle after very large explosive tropical eruptions is not only driven by the north-south contrast of the monsoon (the African monsoon system is far more complex in this respect), rather than less evaporation caused by lower temperatures according to the Clausius-Clapeyron equation.

We have modified the relevant sentence in line 120, to acknowledge the role of regional factors in addition to the reduced temperature gradient *“….well as reducing the meridional (north-south) temperature contrast that controls the intensity of the African monsoon, alongside other regional factors*”. (Line no. 154-155 in Revised Manuscript)

ll. 134–170: This whole section should be shortened/summarized and focus on the very area of research, as outlined in the section ll. 171–186.

We have moved the a substantial portion of the introductory historical context (lines 134-170) to the supplementary material (now appearing there as Section S1.1 Introduction (Historical Context)). (Reference to supplementary info at line number 175 in the Revised manuscript)

ll. 190–205: This section should also be shortened to the most relevant information introducing the content of the subsections.

We have shortened the section (by deleting some information related to model that is available elsewhere (as per the cited paper) as suggested. (Around the line 200 in Revised Manuscript)

In addition to the points mentioned above there are two additional points that should be mentioned already in the introduction:

1) The importance of natural climatic/hydrological variability in the occurrence of Nile floods and their counterparts. This is important to put the proposed "hydrological shocks" in the aftermath of volcanic eruptions into context of externally undisturbed periods.

This point is addressed in the introduction in lines 90-105. In addition, we have inserted some text and relevant citations to emphasize that indeed the considerable variability of the Nile was well known historically, and that explosive volcanism contributed to "some" of this variability (i.e., it certainly did not drive all the observed variability).

See lines 109 in Revised Manuscript: *"But the Nile summer flood was also famously mercurial, with insufficient flooding often leading to adverse societal impacts (e.g., Bell, 1975; Butzer, 1976, 1984; Said, 1993; Hassan, 1997a, b; Hassan, 2007; McCormick, 2013). Some of this variability was likely driven by explosive volcanism."*

2) A more differentiated introduction of the impact of large explosive tropical volcanic eruptions vs. medium-to-small sized high latitude northern/southern hemisphere eruptions. This relates for instance to the overall amount of cooling, the potential for a dynamical response on natural modes of climate variability (cf. North Atlantic Oscillation/El Nino). Introducing this difference in location and magnitude will also help to better explain the different climatic and hydrological response of the initial tropical eruption E1 and the following high-latitude eruptions E2 – E4 that are presented and discussed further down in the manuscript.

Point 1 and 2 are addressed in the Introduction section where volcanic eruptions and their impacts on the Earth's climate system are summarized (lines 55-75).

In our opening paragraph (lines 52-58 In Revised Manuscript), we thus now state: *"The cooling caused by such events can also reduce net evaporation and hence precipitation over large areas (Lui et al., 2016; Iles et al., 2013), while also potentially leading to a near global-scale dynamical suppression of the northward migration of the inter-tropical convergence zone (ITCZ) during the boreal summer, as the convergence follows the surface area of maximum temperature (Petterson et al., 2000; Chiang and Bitz, 2005; Broccoli et al. 2006; Colose et al. 2016). These changes in precipitation can, moreover, impact river outflow (Oman et al., 2006; Sabzevari et al., 2015; Kostiç et al., 2016)..."*

Additionally, we have added the following in lines 77-84 (Revised Manuscript):

"*Extratropical eruptions generally have a comparatively weaker climate impact than tropical eruptions. This happens* following the background Brewer-Dobson circulation upwelling in the tropics and downwelling at higher latitudes, which *directly affects the stratospheric lifetime of* *volcanic aerosols (Kirtman et al., 2013; Myhre et al., 2013; Schneider et al., 2009). Recent studies though illustrated the potential for extratropical eruptions having disproportionally strong forcing and climate impacts, consistent with past reconstructions (Toohey et al. 2019).*"

2.1 Model Description

l. 226: How is the impact of volcanic eruptions implemented?

We have provided a paragraph on how the eruptions were implemented in the NASA GISS ModelE was originally stated under the section "Experiment Design" (line 272 in original Manuscript), We moved this paragraph from experiment design to discuss how the volcanic eruption has been implemented here (Line 239-252 in Revised Manuscript).

2.2 Experiment Design

l. 233: What is the rationale [for] using the PMIP4 mid-Holocene protocol? Maybe the authors could explore in one or two sentences why especially the vegetation is closer to mid-Holocene conditions rather than the one representing the situation during pre-industrial times.

NASA GISS ModelE's Terrestrial Biosphere Model (TBM) Ent (NASA-GISS Version name) (Kiang, 2012; Kim et al., 2015) is not a full Demographic Global Vegetation Model (DGVM). A key missing functionality is the ability to migrate vegetation, driven by changes in climate. In CMIP5, the lack of a fully dynamic vegetation model GCMs led to a failure to reproduce the mid-Holocene wet Sahara conditions over Africa (Harrison et al. 2013). Our model simulations for the mid-Holocene period using the vegetation distribution based on the PMIP4 protocol produced a more realistic result in terms of a wetter Sahara region. Thus, in the absence of a better approach, we linearly interpolated the mid-Holocene PMIP4 vegetation distribution to the 2.5k period to achieve more accurate background climate conditions.
We have outlined our arguments for the use of the PMIP sensitivity vegetation distribution in our results section while discussing the implications of using these boundary conditions in section 3.1.1, (line 300 in Revised Manuscript and specific discussion is at line 340 onwards).

l. 272: The authors could add some effects on the timing of the eruption, i.e. when the eruption date is set to a summer date, especially for the potential effects on monsoon and the northern hemispheric winter atmospheric circulation. It could also be explicitly stated that it is not possible to decipher the exact timing of the eruption in the annual cycle because of dating uncertainties involved in the ice core reconstructions.

Thank you for pointing this out; we have thus added the line "*Because the exact date of an eruption cannot be directly determined based upon ice-core sulphate deposition data, both because of possible uncertainties in the ice-core chronologies and because of variable time lags between eruptions and the atmospheric circulation of the resulting sulphate and its deposition in the polar ice, we selected a summer eruption date to investigate the impact on northern hemisphere monsoon and wintertime atmospheric circulation.*" (Line 291 in Revised Manuscript)

l. 273: I suggest to move the following section on the implementation of the volcanic forcing at the end of the model description paragraph – also some words on the uncertainties of the sulfate reconstructions based on ice cores would be helpful to indicate that modeling results on the subsequent simulations are dependent on the magnitude of the reconstructed sulfate injected into the stratosphere.

We moved the relevant paragraph to the model description section as suggested, and added following lines at line 295.

"*We also note that the accuracy of our modelling will depend in part upon the accuracy of the ice-core-based volcanic forcing reconstruction being employed. Uncertainties in reconstructed forcing can arise, for example, because of variation in the deposition of sulphate across the polar regions for any given eruption. In this respect, it is important to note that the Sigl et al. (2015) volcanic forcing reconstruction employs several ice-cores from Antarctica and Greenland, but our results can be revisited as reconstructions become more reliable by incorporating larger numbers of ice-cores*" (line 295 in Revised Manuscript)

3. Results

– The header for paragraph 3.1 is missing –

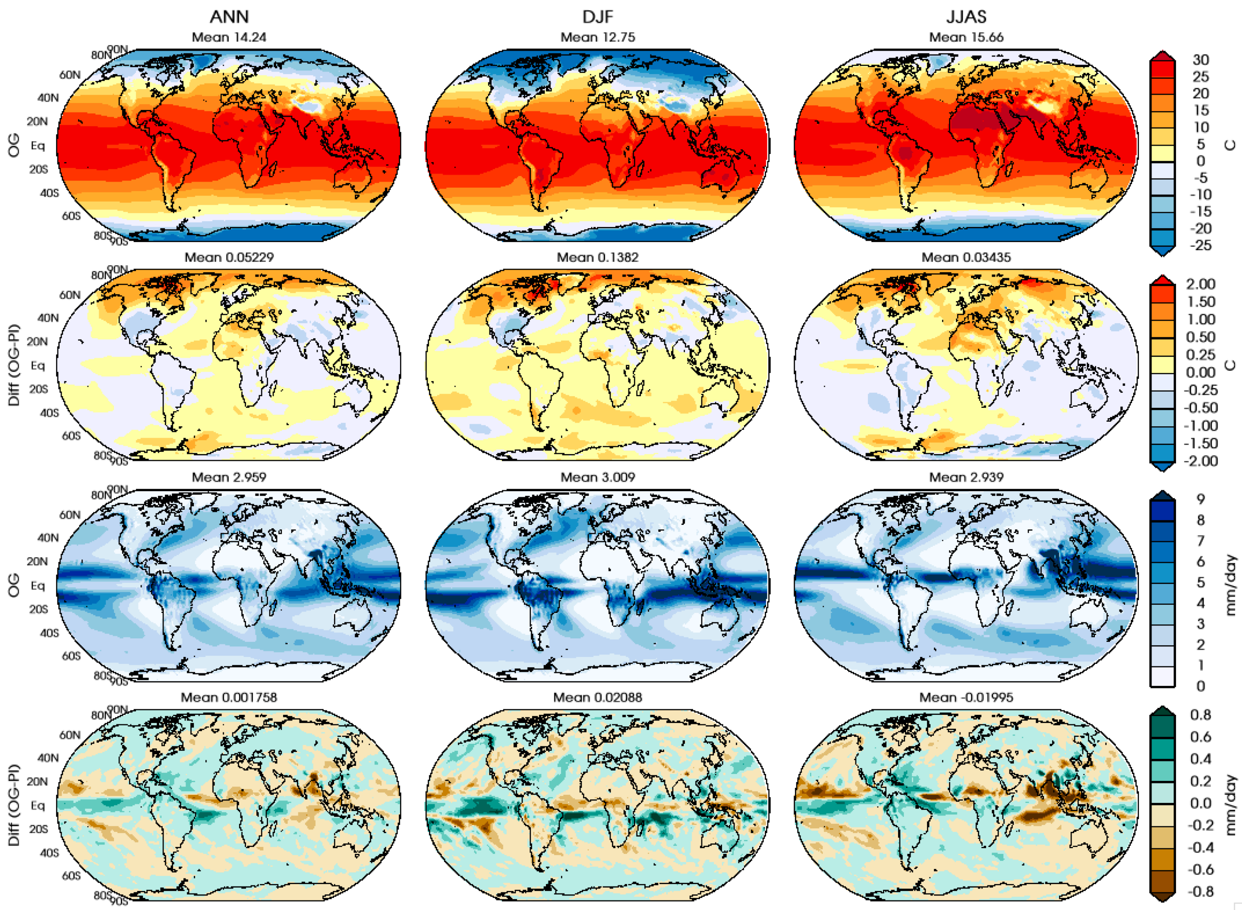
We have introduced the header for 3.1 as "**3.1 2.5ka control runs**".

Changes in orbital forcing – the supposedly most important factor between 2.5 and 6k – were already considerable different at 2.5 k. Therefore I guess that also the classical mid-Holocene pattern is different, even without dynamic vegetation. It would be good to at least indicate those implications when interpreting the 2.5 k pattern in the context of the mid-Holocene 6k climate and vegetation changes.

Another note: Changes due to orbital forcing are mostly effective on a seasonal basis on Holocene timescales, because changes in the inclination of the earth axes do not change the annual amount of radiation received by the sun. An alternative in structuring Fig 1 and Fig 2 is to omit the mean climate states in the upper and middle panel (also for section 3.1.2) and replace them by the patterns for the winter and summer season for the different experiments (together with the annual). This would also show better the impact of the (orbitally induced) background climate conditions between 2.5 k and PI.

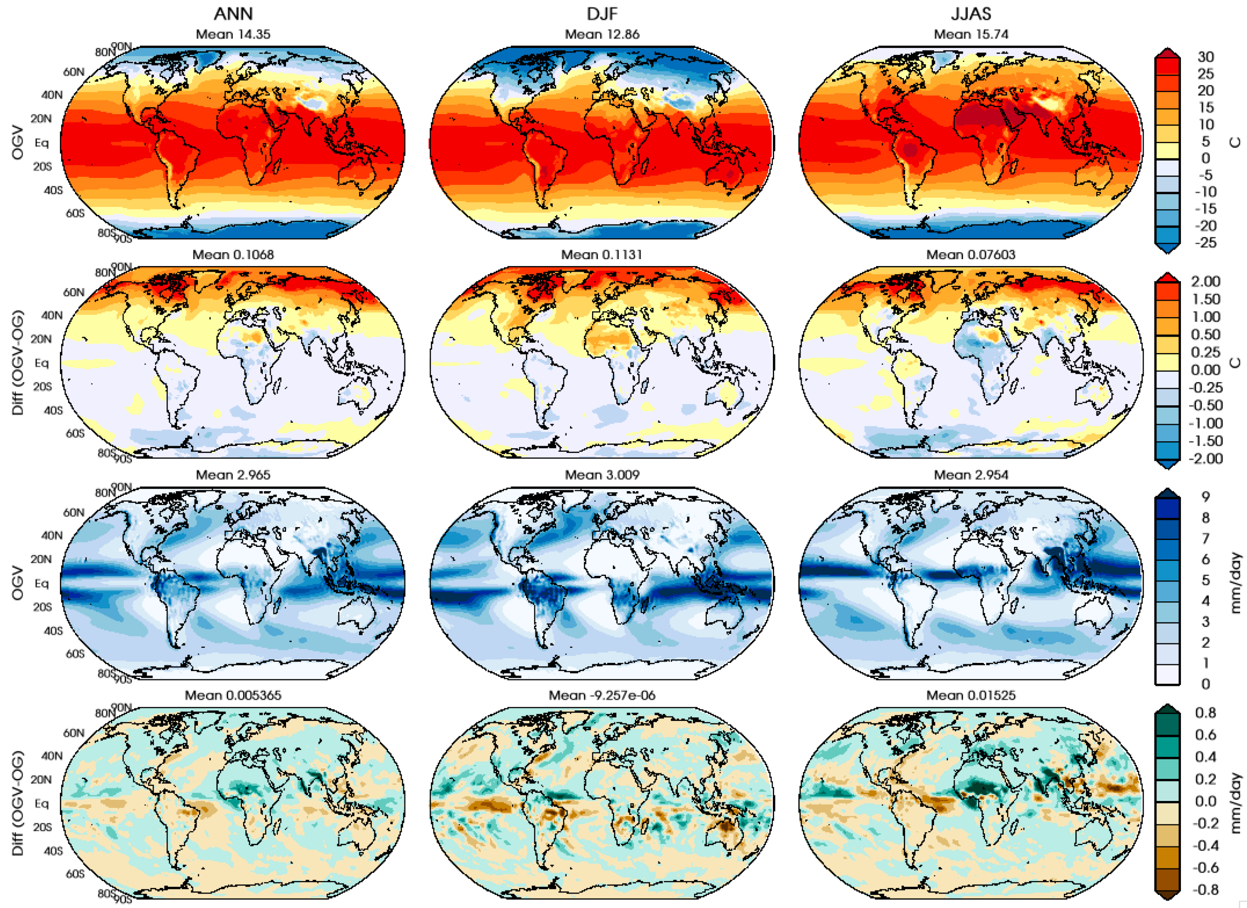
We certainly agree that the impact of orbital forcing for mid-Holocene may be slightly different (cooler for mid-Holocene) than the PI control for both the current (PMIP4-CMIP6) and previous (PMIP3-CMIP5) generation of models (Brierley et al., 2020). We have focused on the North African monsoon season rainfall and the impacts due to the inclusion of PMIP4 vegetations over the region. Since the impact of changes in orbital forcing for a 2.5k period is evident in the surface temperature changes over the northern hemisphere, but rainfall changes over Africa appear more reasonable with vegetation changes only.

We have now modified the plots (Fig 1 & Fig2). We have included the seasonal differences for Annual, DJF, and JJAS seasons along with the mean seasonal climate for the 2.5ka period with GHG and ORB in fig 1 and GHG, ORB, and vegetations in fig 2. We included the mean panels in both plots for the reader to understand the mean climate with the difference with the inclusion of different forcing factors. See the revised Fig 1, below.



"Fig 1. *Seasonal means (Annual, DJF & JJAS) of surface air temperature (top row) for 2.5k period equilibrium run, differences from the preindustrial period (2.5ka-preindustrial) for all three seasons (2nd row from top) and seasonal (Annual, DJF & JJAS) mean precipitation (3rd row from top) and the difference (bottom row) from preindustrial period (2.5ka-preindustrial). The equilibrium run for 2.5k period have the orbital and GHG concentration changes for the 2.5k period (referred as OG), the preindustrial period (as PI), and their difference (OG-PI) as simulated by GISS ModelE2.1*."

For the revised Fig 2, see below:



"Fig 2. Mean surface air temperature for Annual, DJF and JJAS seasons (top row) and seasonal mean precipitation (3rd row from top) for the equilibrium runs with the PMIP4 vegetation for 2.5k period and surface temperature difference (2nd row from top) as well as the seasonal precipitation differences (bottom row) for 2.5k period as simulated by GISS ModelE2.1. We used a short initial notation for forcing to denote the difference (ORB+GHG+VEG = OGV and ORB+GHG= OG)"

3.1.2 2.5Ka ORB+GHG+VEG climate

l. 344: The authors should provide some implications the linear interpolation of vegetation might have on their results (e.g. it is also likely that vegetation changed considerably earlier to preindustrial-like conditions, resulting in a higher albedo due to less forest over the high northern latitudes.)

We have thus summarized the implication of the inclusion of vegetation cover specific to 2.5ka as increased temperature over higher latitudes and northward expansion of the African monsoon during the JJAS season. Although the albedo changes due to introducing vegetation over Africa are not substantial, this enhanced rainfall supports the role of biogeophysical processes in reproducing the wet African conditions over Africa relative to PI period for the mid-Holocene period (Kutzbach et al., 1996; Claussen et al., 2003; Kutzbach and Liu., 1997; Hewitt and Mitchell, 1998). These results also support the importance of having a dynamic vegetation component to represent regional-scale processes and their impact on the climate.

3.3 Volcanic aerosol properties

Concerning the overall length of the manuscript, I suggest to move this section into the supplementary information, as the general content of the manuscript is for an interdisciplinary readership.

We prefer to keep this section in the manuscript because it conveys important information on the model setup, which might get lost in the supplement. This section also complements the description section on how we model sulfate aerosols, aerosol-radiation interaction, and related properties. We also note that the journal does not have strict word limits, and our article is not overly long relative to others in the same special issue. We of course take the reviewer's general point that the manuscript should not be needlessly long.

3.4. Latitudinal temperature response to volcanic aerosol forcing

l. 483: How did the authors estimate their level of significance ? A few words in the supplementary [material] or within the section would be helpful to assess the robustness of the test, using only a limited number of ensemble simulations for the estimation of the level of statistical significance.

Thanks for pointing this out. To highlight this, we have added additional sentences in section 3.4 along with our results:

"The statistical significance level is estimated using the 2-tail student t-test after Deser et al., (2012) and following the assertion that 10 ensembles are sufficient for reasonable estimation of internal variability at a regional scale (Singh and AchutaRao, 2019)." (Line 484 in Revised Manuscript)

3.5 Latitudinal precipitation response to volcanic aerosols

ll. 506: The authors should add one or two sentences on the potential complications [of] investigating the direct output of global and coarsely resolved earth system models. For instance, the simplified parameterizations used for the simulation of precipitation in global models which impact a realistic simulation of tropical convection.

The coarser resolution of Earth system models is a notable cause of uncertainty in the modeling of convective rainfall. However, recent finer resolution models with convective cloud resolving capabilities have shown a significant improvement relative to coarse resolution models. But coarser resolution models can still be successful in simulating large-scale patterns of changes in rainfall. We have thus added these lines:

"*We used a coarser resolution earth system model having a simplified parameterization and was successful in simulating the large-scale patterns of rainfall change*".

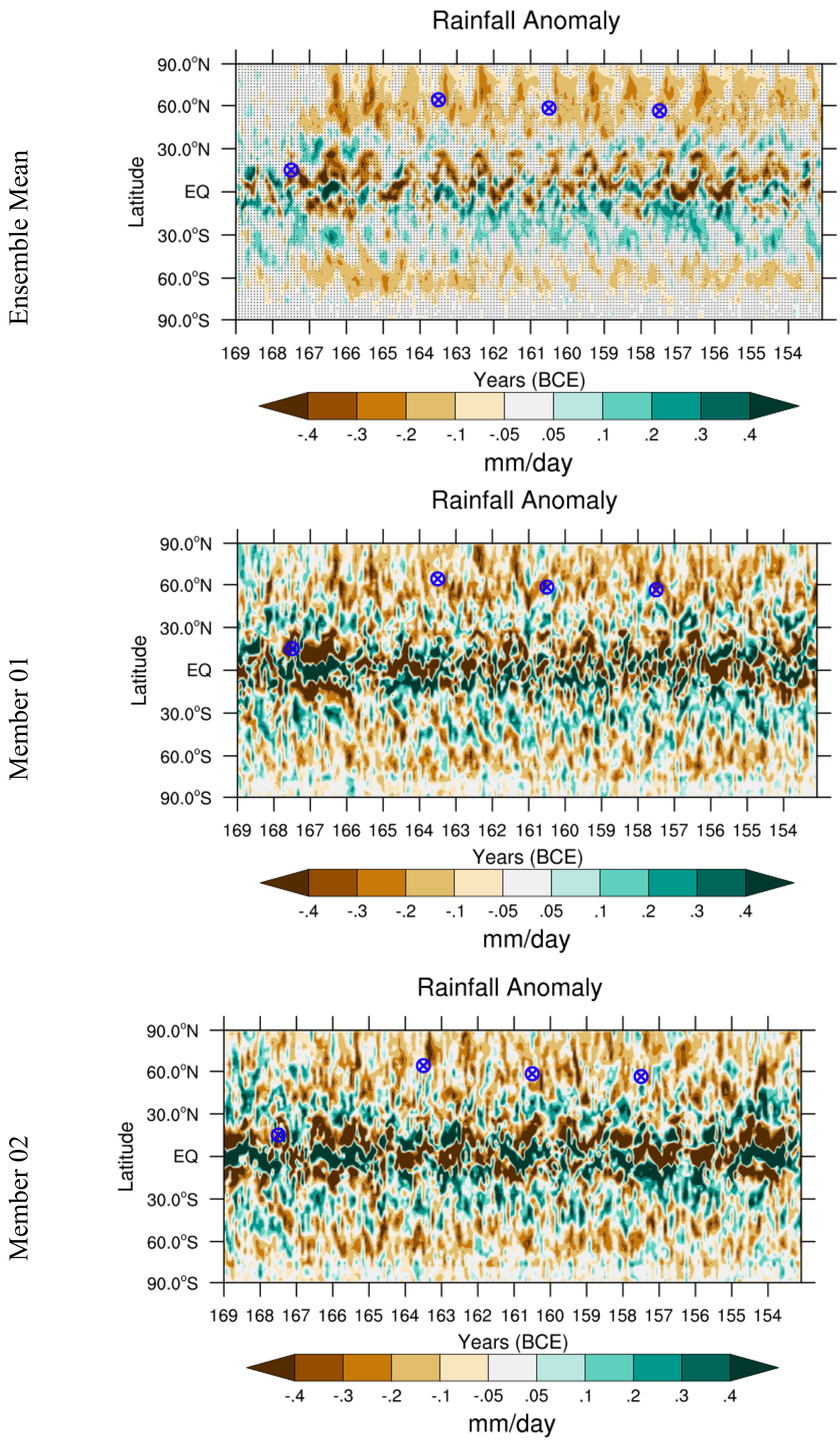
l. 514: This section is one of my critique zones, especially in the context of interpreting climatic trajectories in the context of past societies: The ensemble mean never happens in the real world – if any, a single trajectory compares best to a real world manifestation. Therefore it would also be imperative to show trajectories for single ensembles. This also reflects the bandwidth of potential hydrological changes in the aftermath of volcanic eruptions.

The reviewer's point. The ensemble mean may, for example, be biased by several factors, such as outlier members of the ensemble. One of the reasons for focusing on the ensemble mean is that it is impossible to select the most accurate member as we do not have the historical observations to compare to. However, the spatial representation of ensemble means with accompanying statistical confidence information can be helpful in the interpretation of the robustness of the magnitude and sign of change across ensemble members.

To convey how the results of individual members may differ meaningfully from the ensemble mean with, for example, stronger variability in the signal for a field such as rainfall, we have now plotted two (see below) ensemble members and compared them with the ensemble mean. These two individual members do exhibit some difference in rainfall response in the northern hemisphere tropics after the eruption E1. Still, we feel that in the absence of any indication of which ensemble member is the more accurate, the ensemble mean is the most relevant to focus on in the main text, but we explicitly note that it may be the case that one or more of the ensemble members is more accurate, though we cannot at present tell which. We have added the following text to the manuscript:

"It can be argued that an individual ensemble member can represent the historical period, but it is impossible to select in the absence of observation. Also, the added noise due to natural variability can alter the sign of change at the spatial scale among the individual ensembles. Thus, we selected the ensemble mean with statistical significance to show the response to volcanic eruptions with robustness for the climate variable. "

To additional reflect this uncertainty, we emphasize (see, for example, lines 542-547) cases in which ensemble variability around the ensemble mean is particularly high, thereby alerting the reader to instances in which the ensemble mean may be less parsimonious.


Ensemble Mean
Rainfall Anomaly
Latitude
90.0°N
60.0°N
30.0°N
EQ
30.0°S
60.0°S
90.0°S
169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154
Years (BCE)
-.4 -.3 -.2 -.1 -.05 .05 .1 .2 .3 .4
mm/day
Member 01
Rainfall Anomaly
Latitude
90.0°N
60.0°N
30.0°N
EQ
30.0°S
60.0°S
90.0°S
169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154
Years (BCE)
-.4 -.3 -.2 -.1 -.05 .05 .1 .2 .3 .4
mm/day
Member 02
Rainfall Anomaly
Latitude
90.0°N
60.0°N
30.0°N
EQ
30.0°S
60.0°S
90.0°S
169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154
Years (BCE)
-.4 -.3 -.2 -.1 -.05 .05 .1 .2 .3 .4
mm/day

l. 540: Here again, a more detailed information on the evaluation of the statistical significance would be helpful.

Please refer to our reply to the reviewer's comment on section 3.4, above.

3.6 African monsoon and Nile River response

l. 581: The already mentioned information that a more consistent comparison between model and empirical evidence can only happen at the single simulation level can again [be] picked up here, because in the real world one could not expect a mean response of different simulated trajectories for single events in history.

Please refer to our response to the reviewer's comment on section 3.5, above.

l. 584: Results for the E1 eruption seem convincing and also have a large-scale character that can be attributed to an external event – However, eruptions E2 – E4 show a very inconsistent pattern (even in the ensemble mean). This is also reflected in the statistical test (that presumably uses standard testing techniques that are not taking account the small sample size of n=10 samples). This heterogeneity in the response of the northern hemisphere E2 – E4 eruptions should be more emphasized, also in the subsequent interpretation in the context of their sustained effects on Nile floods.

This is an important observation. We have now inserted two new paragraphs along with the discussion on the possible reasons for heterogeneity in the response over the Nile basin.

(Line 743-785 In the Revised Manuscript)

"*It is evident that the mean surface temperature response in the northern hemisphere is significant at the control period's* $1\sigma_{ctrl}$ *and* $2\sigma_{ctrl}$ *levels. However, while rainfall and river discharge responses are significant at the* $1\sigma_{ctrl}$ *level, they fall within the* $2\sigma_{ctrl}$ *levels, although a few individual members do show significance at* $2\sigma_{ctrl}$ *as well. However, the statistical significance of the rainfall and discharge response may be sensitive to the dearth in the modeling Nile River basin at a relatively coarse resolution of the GISS ModelE, as well as the boundaries chosen to model the Nile basin and its headwaters. In particular, given the complexity of the Nile's hydrology and disparate sources of discharge for the White and Blue Niles. We thus investigated the post-volcanic change in river flow for the southern (White Nile-dominated) and northern (Blue Nile and Atbara river-dominated) parts of the basin by dividing it at 10° N (Fig 13). Annual mean river flow change for the south (blue lines) and north (red lines) of the Nile basin were in broad agreement with a negative flow anomaly after eruption E1. This was most notable in the eruption year and the first year following, with the* 95^{th} *percentile envelopes (dotted lines) deemed significant at the 95% confidence level for both these years (i.e., crossing* *the dashed lines parallel to the x-axis (Fig 13). In contrast, the mean north and south responses disagreed, including in the sign of the observed changes, after the extratropical eruptions (E2, E3 & E4). More specifically, while the mean flow anomalies in the year of E2 were unremarkable and showed little north-south contrast, a more notable divergence was observed in the first year following, with a positive flow anomaly in the south and negative in the north. In the year of E3, flow in the south showed no notable anomaly, while flow in the north was marginally negative. This distinction became more marked in the first year following, mainly due to a larger negative anomaly in the north. In the year of E4, a negative anomaly was again observed in the north, persisting for three post-eruption years, and contrasting with positive or unremarkable anomalies in the south.*

These results are consistent with our earlier-described results (e.g., spatial rainfall variability over the Nile River basin, as per Figs. 10 and 11) and proposed mechanisms, alongside expectations from the literature (e.g., Manning et al., 2017). Thus, tropical eruptions (like E1) may result in a more consistent (negative) north-south flow response due to their more even interhemispheric aerosol burden and associated radiative impact. Extratropical NH eruptions (like E2-E4) that can result in a more asymmetric hemispheric aerosol burden may, by contrast, introduce contrasting flow anomalies by suppressing the northward migration of the ITCZ, negatively impacting flow in the Blue Nile and Atbara rivers by diminishing monsoon rainfall in the Ethiopian highlands, while potentially enhancing flow in the White Nile, fed by rainfall over the equatorial lakes"

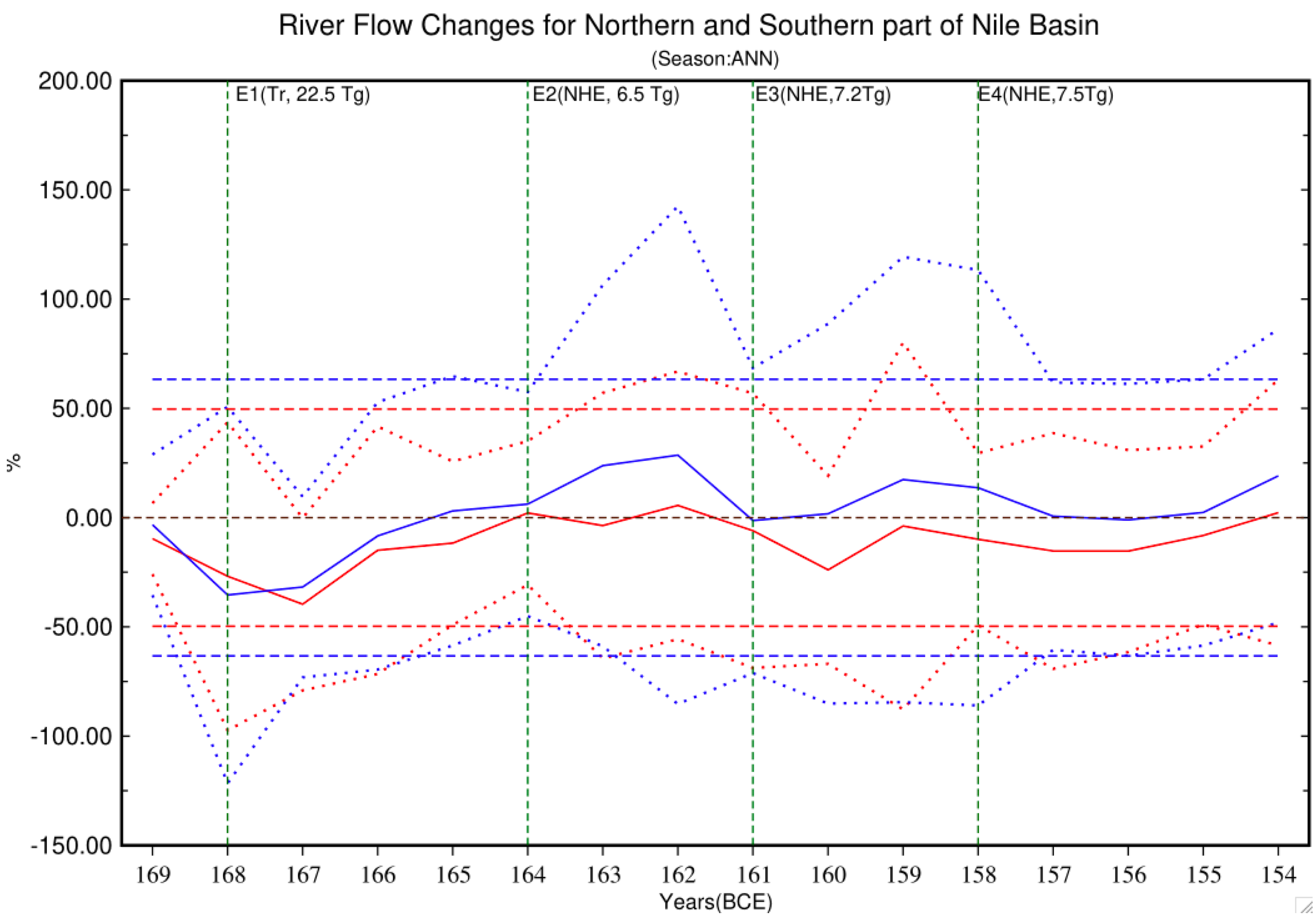


Fig 13. Annual Nile River flow changes averaged over the northern (red) and southern (blue) parts of the basin (divided at 10° N) for the entire simulation period. The solid lines represent the ensemble mean for each part of the basin; the dotted lines are ±1.95σ, where σ is derived from across all the ensembles, and the horizontal dashed lines parallel to x-axis are the ±1.95σ_{ctrl} where σ_{ctrl} is the standard deviation across the 100-year control run. Red and blue lines correspond to the northern and southern parts of the Nile basin, respectively.

Please refers to the explanation for next comment (l. 614 incl. Table 2: Interpreting….)

l. 614 incl. Table 2: Interpreting the Table and the calculation of the according values correctly, the standard deviation is based on the volcanically forced ensemble members and the difference on the mean over all ensemble members minus the climatological mean of the 100 year control?

An alternative is to calculate the annual standard deviation of the 100 year control run and include it as the 1.95*std = 95% confidence interval. This will give an indication how the mean discharge value is outside the natural range. In the present form it gives the bandwidth of the volcanically forced simulations, not taking into account the natural undisturbed variability. The interpretation of the 95% confidence interval based on the control will give an indication how exceptional the respective year after eruption E1 – E4 was in the context of the natural variability.

We thank the reviewer for this suggestion. We have calculated the variability (σ_{ctrl}) for the 100-year control period and compared the statistics in Table 2 against the confidence interval (1.95*σ_{ctrl}) suggested by the Reviewer. The calculated variability (σ_{ctrl}) and 95% confidence interval (1.95*σ_{ctrl}) for annual river flow are 25.20% and 49.155% respectively. It is noticed that annual ensemble mean change is within the 95% confidence interval, but individual ensemble member reaches beyond 95% confidence interval for some years. River flow change over the Nile basin varies up to 3*σ for a few ensembles for some years.

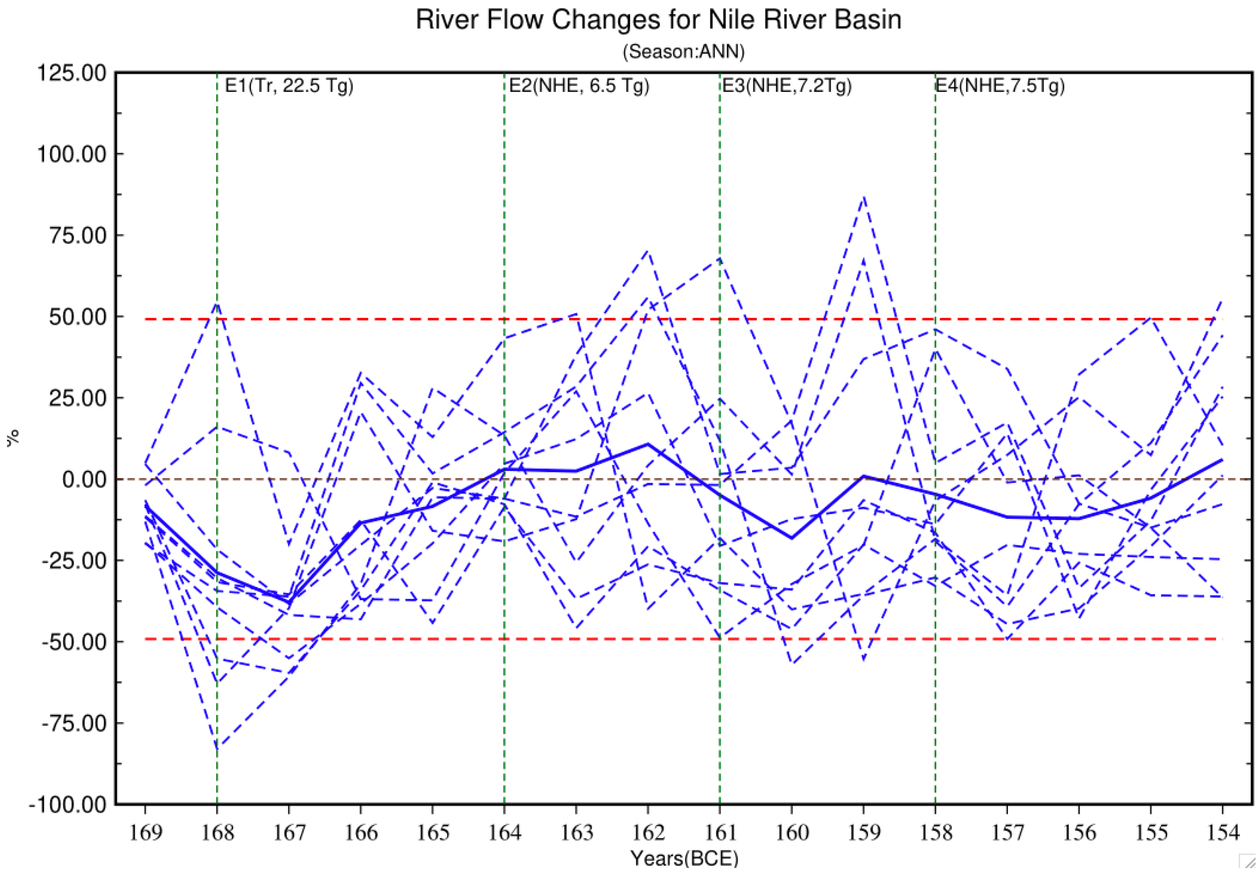


Fig Rev2.1. The solid blue line shows the annual mean change in river discharge over the Nile River basin, and dashed (blue) are the individual ensemble member. The red-colored dashed line parallel to the x-axis represents the 1.95*σ_{ctrl} for the 100-year control run.

Fig 12: This figure contains basically very good information – Similar to Table 2, and to show the exceptional behavior of the different metrics, it would be better to illustrate the 1.95*std of the natural 100 y control run as two lines parallel to the x-axis, together with the individual trajectories of the 10 volcanically forced simulations. When a considerable number of trajectories falls above or below the 95% confidence levels for an individual year, one can speak of a robust response – according to the hypothesis proposed, the discharge trajectories should then fall below the lower boundary for the years after the volcanic eruptions.

In addition, without the green vertical lines it is difficult to decipher the volcanic eruptions based on the evolution of precipitation and discharge, because also other sub-periods with considerable reductions in stream flow appear that are unaffected by volcanic forcing (e.g. year 159). An alternative interpretation that could be hypothesized relates to an increased intra-ensemble variability after volcanic eruptions compared to undisturbed conditions.

We have now inserted the 1σ (Red dashed line) and 2σ lines (Red solid line) on the plot for all three diagnostics. The modified Fig 12 is now shown below.

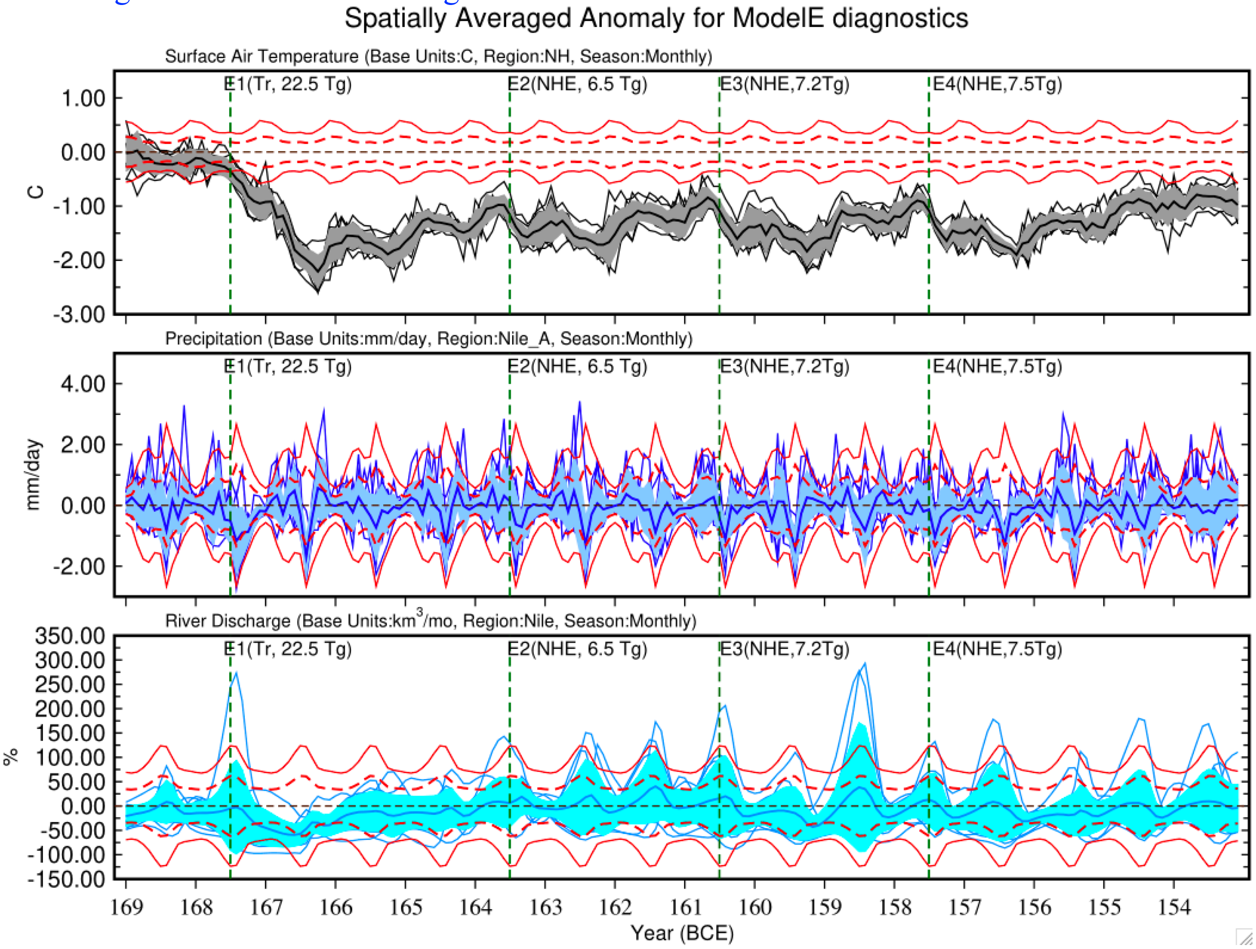


Fig 12: Monthly time series of individual ensemble and mean of surface temperature response (˚C) averaged over northern hemisphere (NH) (top panel), rainfall change (mm/day) over the spatial box over Nile River watershed (Latitude: (5N, 18N), Longitude: (30E, 42E)) (middle panel) and Nile River discharge anomaly (%) at the delta region (grid box centered at 29.0N, 31.25E). The dark solid (Thick) line shows the multi-ensemble mean, individual member (thin line), and the color envelope shows the associated variability (±σ; Standard deviation). The annual cycle of climate variability of control run is shown as $1\sigma_{ctrl}$ (Red dashed line) and $2\sigma_{ctrl}$ lines (Red solid line) along the x-axis for all three variables. The vertical dotted green line shows when each eruption happens

4. Discussion and Conclusions

ll. 712: This paragraph also relates to the interpretation of empirical evidence in the context of earth system and climate model simulation: It is important also taking into account the natural or stochastic nature of historical processes that are not always determined by external environmental forcings. Otherwise a state in the interpretation and explanation of historical events will be reached, where numerous single historical events are only interpreted within the climate-determinism concept, which can be true for severe events, but might be misleading for most medium-to-small size events, especially in the context of volcanic eruptions.

We have now added a more substantial introduction to the Introduction section to address challenges in assessing connections and causality between environmental forcings and historical human/societal events. This also in part addresses similar challenges highlighted by Reviewer 1, and which are addressed more fully in our response to this reviewer.

Additionally, we have clarified that the present paper builds upon the work of Ludlow and Manning (2016, 2021) and Manning et al. (2017), that explicitly address the challenges of testing statistically the association between explosive eruptions and revolts in Ptolemaic Egypt (also noting that it is not assumed that all revolts were "triggered" or otherwise caused by volcanically induced hydroclimatic variability (or indeed hydroclimatic variability of any origin).

The intent of the present paper is, therefore, to provide clarification on the likely hydroclimatic variability experienced in Ptolemaic Egypt during the 160s BCE, a time already well known to historians as one of considerable societal upheaval in Egypt. Our work here should provide a firmer foundation for a planned follow-up study that will more closely examine the available historical evidence for the impacts of these eruptions (now informed by our modelling) and assess their contribution to the turbulent history of the period.

l. 731: For producing a basis for "historical realization" it is again of utmost important to look and investigate the trajectories of individual realizations of ensemble simulations from climate models and not (only) their mean response.

We agree with the reviewer's argument, and would refer back to our earlier responses to this important issue.

l. 791: As the authors state correctly, from a conceptual point of view there is no "best" member, because all members are equal probable under the same set of external forcings implemented. What might be more important is the notion that the combination of external AND internal forcings shape the exact evolution in both, the real and the model world.

We again agree fully with the reviewer here.

Figures and Tables:

In general, Figures and Tables are presented with high quality and an appropriate level of information included. Below just a few minor suggestions how to improve or modify selected items:

Fig 1: As already stated in the main text, Fig 1 and 2 might be combined into one single Figure by representing only the differences for annual, winter and summer (alternative JJAS) mean.

Figures 1 and 2 are now modified to include more seasonal details as suggested.

Fig 5 center panel: The style of the presentation of the single trajectories could be used as template for the precipitation and discharge Figure 12 to show the variability of the different ensemble simulations together with the 95% confidence level of natural variability of the 100 yr control simulation.

Figure 12 has now been modified to include the individual ensemble members along with the $\pm\sigma$ envelope for the volcanically forced ensembles. We have also included the $\pm 1*\sigma_{ctrl}$ and $\pm 2*\sigma_{ctrl}$ annual cycle for the 100-year control run as suggested.

Table TS1: The authors might include also the volcanically forced simulations as an additional column and highlight those simulations that are presented in the manuscript.

We have inserted the column for the volcanically forced ensemble.

Additional Revisions

Apart from the reviewer's comments, we also revised the manuscript sections for some general aspects for better representation of results in the final version of the manuscript and included more relevant references—some of these as mentioned below.

1.) The word "stratovolcanic" has been changed to "Large, strato-volcanic" for more precise reference in abstract line 20 and introduction and 46
2.) in the abstract "NASA GISS ModelE" is changed to "NASA GISS ModelE2.1".
3.) Abstract line 29: "South and East Asian" changed to "South Asian, and East Asian".
4.) Inserted a sentence at line 48 (introduction) "*The sulfate aerosols of the 1991 eruption of ~18 Tg* SO_2 *from Mt. Pinatubo increased the optical depth of the atmosphere from ~0.6 to ~0.75.*"
5.) Inserted reference "Colose et al. 2016" at line 53.

6.) Text inserted at line number 85. "*Effectively, the ITCZ shifts away from the hemisphere with the greatest amount of volcanic aerosol; for tropical eruptions, this movement is typical more southward as well owing to the larger amount of land in the Northern Hemisphere and higher thermal capacity of the oceans*."
7.) Modified the format of conditions stated at line 250-255.

Reference:

Kirtman, B., Power, S. B., Adedoyin, A. J., Boer, G. J., Bojariu, R., Camilloni, I., Doblas-Reyes, F., Fiore, A. M., Kimoto, M., Meehl, G., Prather, M., Sarr, A., Schar, C., Sutton, R., van Oldenborgh, G. J., Vecchi, G., and Wang, H.-J.: Chapter 11 - Near-term climate change: Projections and predictability, edited by: IPCC, Cambridge University Press, Cambridge, 2013.

Ludlow, F. & Manning, J. G. in Climate Change and Ancient Societies in Europe and the Near East: Diversity in Collapse and Resilience (eds Erdkamp, P., Manning, J. G. and Verboven K.) 301-320 (Palgrave Macmillan, 2021).

Ludlow, F. & Manning, J. G. in Revolt and resistance in the Ancient Classical World and the Near East: The crucible of empire (eds Collins, J. J. & Manning, J. G.) 154–171 (Brill, 2016)

Manning, J. G., Ludlow, F., Stine, A. R., Boos, W. R., Sigl, M., and Marlon, J. R.: Volcanic suppression of Nile summer flooding triggers revolt and constrains interstate conflict in ancient Egypt, Nat Commun, 8, 900, https://doi.org/10.1038/s41467-017-00957-y, 2017.

McConnell, J. R., Sigl, M., Plunkett, G., Burke, A., Kim, W. M., Raible, C. C., Wilson, A. I., Manning, J. G., Ludlow, F., Chellman, N. J., Innes, H. M., Yang, Z., Larsen, J. F., Schaefer, J. R., Kipfstuhl, S., Mojtabavi, S., Wilhelms, F., Opel, T., Meyer, H., and Steffensen, J. P.: Extreme climate after massive eruption of Alaska's Okmok volcano in 43 BCE and effects on the late Roman Republic and Ptolemaic Kingdom, Proc Natl Acad Sci USA, 117, 15443–15449, https://doi.org/10.1073/pnas.2002722117, 2020.

Myhre, G., D. Shindell, F.-M. Bréon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang,: Anthropogenic and natural radiative forcing. In Climate Change 2013: The Physical

Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Doschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, Eds. Cambridge University Press, pp. 659-740, doi:10.1017/CBO9781107415324.018, 2013

Schneider, D. P., Ammann, C. M., Otto-Bliesner, B. L., and Kaufman, D. S.: Climate response to large, high-latitude and low-latitude volcanic eruptions in the Community Climate System Model, 114, https://doi.org/10.1029/2008JD011222, 2009.

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., Büntgen, U., Caffee, M., Chellman, N., Dahl-Jensen, D., Fischer, H., Kipfstuhl, S., Kostick, C., Maselli, O. J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D. R., Pilcher, J. R., Salzer, M., Schüpbach, S., Steffensen, J. P., Vinther, B. M., and Woodruff, T. E.: Timing and climate forcing of volcanic eruptions for the past 2,500 years, Nature, 523, 543–549, https://doi.org/10.1038/nature14565, 2015.

Toohey, M., Krüger, K., Schmidt, H., Timmreck, C., Sigl, M., Stoffel, M., and Wilson, R.: Disproportionately strong climate forcing from extratropical explosive volcanic eruptions, 12, 100–107, https://doi.org/10.1038/s41561-018-0286-2, 2019.