Response to Reviewers

Manuscript title: "Comparing Multi-Model Ensemble Simulations with Observations and Decadal Projections of Upper Atmospheric Variations Following the Hunga Eruption"

Manuscript authors: Zhihong Zhuo, Xinyue Wang, Yunqian Zhu, Wandi Yu, Ewa M. Bednarz, Eric Fleming, Peter R. Colarco, Shingo Watanabe, David Plummer, Georgiy Stenchikov, William Randel, Adam Bourassa, Valentina Aquila, Takashi Sekiya, Mark R. Schoeberl, Simone Tilmes, Jun Zhang, Paul J. Kushner, Francesco S.R. Pausata

We would like to thank Christopher Smith and another anonymous reviewer for their efforts and insightful comments, which have contributed to improving the clarity and quality of our manuscript. We have carefully revised the manuscript in accordance with the reviewers' feedback. Below, we provide the list of reviewers' questions and comments (in gray italics), along with our responses (in black). Texts in the revised manuscript are highlighted in red.

1. **Response to Reviewer #1**

***Citation**: https://doi.org/10.5194/egusphere-2025-1505-RC1*

This paper introduces a multi-model study (can we call it a MIP?) of the stratospheric projections following the Hunga Tonga-Hunga Ha'apai (HTHH) eruption. Its utility arises from including both the water vapour and the aerosol components of the HTHH injection and using models that resolve ozone loss, therefore allowing (potentially - see comment below on radiative forcing) for the net climate impact of this eruption to be evaluated. It confirms that the water vapour perturbation is longer-lived than the stratospheric aerosol perturbation, potentially leading to a slight surface warming as a consequence (not evaluated). The model results are compared to the MLS data for water vapour and ozone and GloSSAC for stratospheric aerosol, showing good correspondence.

With the end of life of the MLS instrument and uncertainty around a suitable replacement platform, this MIP might be the best source of information available to estimate and project forward the evolution of the HTHH water vapour plume at a very critical time during a gap in the satellite observations. Therefore I fully support this initiative.

Thank you for the positive feedback.

Please if possible could you analyse the effective radiative forcing from these results? It would simply be the global mean TOA net radiation change from 2022 and subsequent years compared to the no-HTHH runs in each model in the fixed SST runs. Splitting out the all-forcing and H2O-only forcing runs where models performed the latter would also be really useful. This help add to the discussion on whether HTHH is a net warming or net cooling climate influence.

Yes, the model simulations used in this study have also been analyzed to assess the effective radiative forcing (ERF) and broader climate response to the Hunga eruption. A coordinated set of papers is currently underway that explores various aspects of the eruption scenario using HTHH-MOC data. These include dedicated studies on the climate impact

(Bednarz et al., 2025, *ACPD*), effective radiative forcing, stratospheric ozone response, and plume transport into the mesosphere (in preparation).

As indicated by its title and scope, the current paper focuses specifically on the upper atmospheric response. Its primary aim is to evaluate model performance against observational constraints and to explore model-based projections of long-term upper atmospheric evolution following the HTHH eruption.

We recognize, however, the broad scientific interest in the radiative forcing associated with this event and would like to address this point in this response letter:

Schoeberl et al. (2024) quantified the radiative forcing contributions from Hunga-induced stratospheric water vapor, aerosols, and ozone changes. They found that most of the net radiative forcing effects had dissipated by the end of 2023, with stratospheric aerosols being the dominant driver of radiative changes during the two years following the eruption. Bednarz et al. (2025, *ACPD*), using CESM2-WACCM6 simulations, identified a small but non-negligible influence on regional surface climate. Simulations using both coupled ocean and fixed-SST configurations suggest differing near-surface temperature responses, though most changes remain within the range of internal variability.

A multi-model analysis of cumulative surface air temperature changes from HTHH simulations supports these findings. Across models, slight warming or cooling responses are observed over 6-month and 1-year timescales, with a consistent net cooling signal across all models over the first two years post-eruption. A dedicated paper focused on effective radiative forcing and surface climate impacts is currently in preparation. These results will also be synthesized in the *Hunga Assessment Report*, scheduled for release at the end of 2025.

Line 42: How does the "anomaly" duration differ from the e-folding duration? The anomaly perhaps being the period of time for which a non-zero increase in stratospheric water vapour is detectable?

Here, the anomaly refers to the period during which stratospheric water vapor concentrations exceed the range of background variability, as indicated by the horizontal gray shading in Fig. 2. The definition of this background range, along with the e-folding time used to characterize the decay of the anomaly, is provided in Section 3.2.1. This section also includes a discussion of the various methods used to evaluate the lifetime of the stratospheric water mass, along with comparisons to previous estimates derived from both observational data and model simulations.

Line 44: "local" cooling: in the sense of ULTS, rather than the surface? Please be explicit. (since this paper will be interesting to readers who are not upper atmosphere focused, like myself).

To avoid misleading, we revised the text to "This prolonged water vapor perturbation leads to significant stratospheric and mesospheric cooling…".

Line 127: there's probably good reasons, but why limit to the 2012-21 climatology from MLS when data goes back to 2004?

We chose to use the 2012–2021 period for the MLS-based climatology in order to align more closely with the time of the HTHH eruption and the period covered by the model simulations, which span 10 years post-eruption. This decision ensures greater consistency when comparing anomalies between observations and models. As noted in Millán et al. (2024), MLS water vapor mass values tend to be lower in earlier years (pre-2012), even though the seasonal cycle remains relatively stable throughout the full MLS record. Including earlier data would shift the climatological baseline and potentially exaggerate the anomaly due to long-term changes in instrument calibration or atmospheric variability. We try to minimize discrepancies in anomaly definitions between observations and models by restricting the climatology to the decade immediately preceding the eruption.

Lines 135-138: I understand from reading the Randel et al (2024) reference why the temperature from MLS was detrended. It kind of makes sense, though not described how it was done, that QBO and ENSO are removed from the record since free-running models will generate their own variability patterns that will unlikely sync with that of the real world. Was the solar forcing (i.e. from CMIP6) an input forcing to the models? I presume not, since this has been taken out of the MLS data. And how did you account for the fact that the solar cycle isn't a monotonic trend?

In both the real world and model simulations, QBO, ENSO, and solar variability occur concurrently and can influence stratospheric temperatures. The detrending of the MLS temperature record was applied specifically to improve comparability between observations and model simulations. Free-running models generate their own internally driven QBO and ENSO variability, which are not phase-locked to the real-world events. Similarly, while solar forcing from the CMIP6 protocol was included in the model simulations, it is present in both the Hunga experiment and its corresponding control runs. Therefore, by calculating the difference between the experiment and control simulations, the modelled Hunga-induced signals are effectively isolated. We did not remove trends, solar or other signals from the various model results, but only calculated differences from the control runs.

In contrast, the observational record lacks a control, necessitating a more explicit detrending approach to isolate the Hunga-related signal. This was done in multiple steps: First, solar cycle effects were removed by regressing MLS temperature data against the observed F10.7 solar radio flux—a widely used proxy for solar activity. Importantly, this approach does not assume a monotonic trend but captures the cyclical nature of the solar signal. Second, a piecewise linear regression was applied to account for longer-term trends in the MLS record. Finally, a multiple linear regression was used to remove the influence of ENSO (using Niño 3.4 index) and the QBO (using 30 hPa and 50 hPa zonal wind indices). The residuals from this process represent isolated anomalies attributed to the Hunga influence.

Section 3.3: could you rename to "Global mean stratospheric air temperature evolution" or similar to reduce ambiguity – just because it's not what I would define as global mean temperature, which commonly relates to the surface or near-surface air temperature. (similar comment to line 44). Similarly I recommend deleting "global mean" in line 362 and in figure 4.

To avoid misleading, we revised the section 3.3 title to "Global-mean air temperature evolution", the text in line 262 to "The upper atmospheric global-mean air temperature anomaly calculated …", and the figure 4 caption to "…global-mean air temperature…".

Line 417: "et la" to "et al"

Revised.

Data availability: please give more information on how to obtain the data from JASMIN: I ask as somebody who is interested in the results.

To access the data from JASMIN, please reach out to the activity leaders (https://csl.noaa.gov/assessments/hthh/contacts.html), i.e. Yunqian Zhu and Graham Mann for detailed information.

2. Response to Reviewer #2

***Citation**: https://doi.org/10.5194/egusphere-2025-1505-RC2*
Zhuo et al. (2025) evaluates multi-model ensemble simulations of the Hunga volcanic eruption's atmospheric impacts by comparing model results with satellite observations from 2022-2024, finding good agreement in stratospheric water vapor, aerosols, temperature, and ozone patterns. In general, I find the paper interesting, however lack of in-depth evaluation expeically on the ozone chemistry. Besides, the discussion on the upward motion of SWV to the mesophere can be improved.

Thank you for the positive and constructive feedback.

As noted in our response to Reviewer 1, the stratospheric ozone response to the Hunga eruption involves complex interactions between chemical and dynamical processes. Given the scope and importance of this topic, a coordinated and dedicated study focusing specifically on the ozone response is currently underway. This will allow for a more comprehensive evaluation than is possible within the constraints of the present paper.

Regarding the upward motion of SWV into the mesosphere, we appreciate the suggestion and have strengthened this section by incorporating additional analysis and discussion. Specific changes and justifications are provided in response to the relevant detailed comments below.

Line 325-224: "Beyond this phase, the upward transport of H2O into the mesosphere above 1 hPa becomes the dominant mechanism for the removal of stratospheric water vapor (SWV)."

Maybe this statement and relevant discussions are correct but could be misleading.

- *This sentence discussed the fate of a tiny amount of HTHH water vapor (<1% of total water). The majority of HTHH water fells down to the troposphere in the mid and high latitudes. Please add one sentence to avoid confusion.*
- *Still, from Figure 3, the water mixing ratio is similar between 1 and 0.1 hPa, which means the mass is one order of magnitude larger at 1 hPa. How we can tell it is the dominant mechanism for the removal of SWV at 1 hPa? A concentration map can be helpful.*

We thank the reviewer for the helpful clarification regarding the interpretation of stratospheric water vapor (SWV) removal mechanisms.

As the reviewer correctly pointed out, the amount of water vapor transported into the mesosphere above 1 hPa accounts for only a small fraction of the total Hunga injected water vapor. Our analysis of the H_2O burden between 1–0.01 hPa confirms that the peak mesospheric burden occurred in mid-2023, with values of approximately 4.6 Tg in the observations, 3.1–3.3 Tg in Earth system model simulations, and up to 4.07 Tg in GSFC2D (Fig. S2a). However, as shown in Figure 2, this constitutes less than 1% of the total stratospheric water vapor burden.

We agree that it is therefore not accurate to describe the upward transport into the mesosphere as the dominant mechanism for SWV removal. To avoid confusion, we have revised the text to:

"Beyond this phase, stratospheric water vapor transport is increasingly dominated by upward flux into the mesosphere above 1 hPa, resulting in a peak mesospheric burden of approximately 3–4 Tg by late 2023 (Fig. S2). However, this mesospheric contribution represents only a small fraction of the total H_2O injected by the eruption (cf. Fig. S2 and Fig. 1). The majority is progressively removed through stratosphere–troposphere exchange, particularly at high latitudes. For instance, in January 2025 (Fig. S3a), a wedge-shaped region just above the tropopause marks a sharp decline in H_2O concentration, indicating a key region where much of the Hunga H_2O is removed from the stratosphere. Above this feature, high-latitude maxima in H_2O in both hemispheres are consistent with enhanced transport driven by the Brewer–Dobson circulation. This behavior is further supported by evidence of pronounced dehydration in the Southern Hemisphere polar stratosphere during winter, as illustrated in July 2025 (Fig. S3b), aligning with Antarctic vortex-induced dehydration mechanisms described in Zhou et al. (2024). These pathways are expected to continue dominating the removal of Hunga-injected H_2O as it is gradually transported downward by the global stratospheric circulation (Fig. 10 in Randel et al., 2024)."

In addition, to better support this discussion, we have included in the Supplementary Materials:

- A time series plot of water vapor mass between 1–0.01 hPa (Fig. S2).
- Latitude-level plots of water vapor concentration in January and July 2025 as examples (Fig. S3).

These additions help clarify the relative importance of vertical and meridional transport processes in the redistribution and removal of the HTHH-injected water vapor.

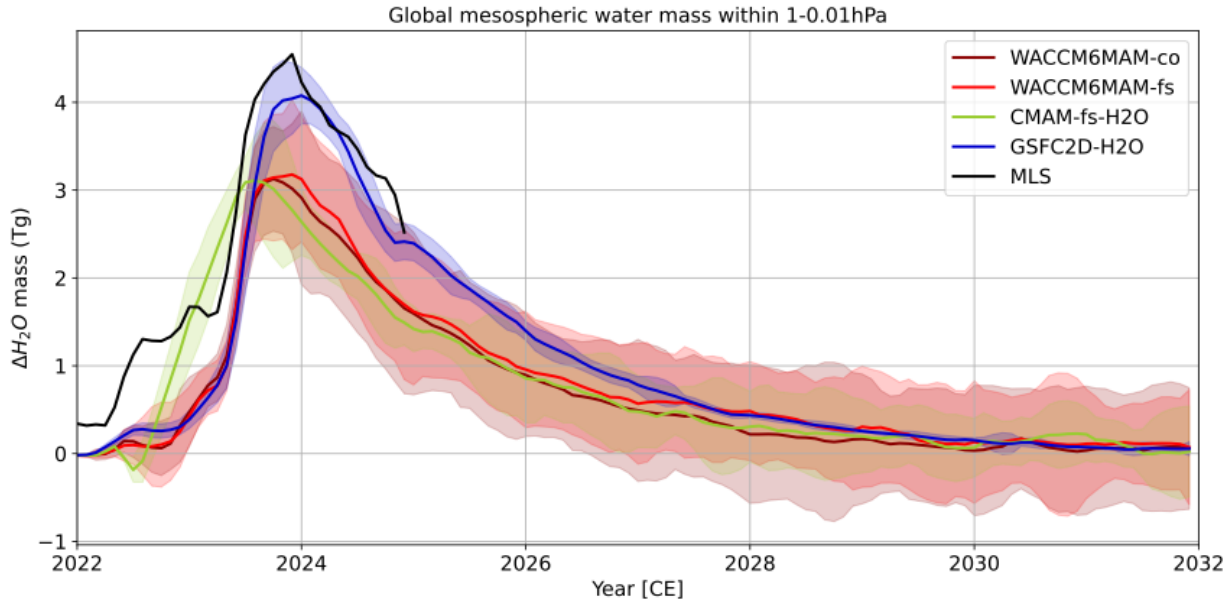
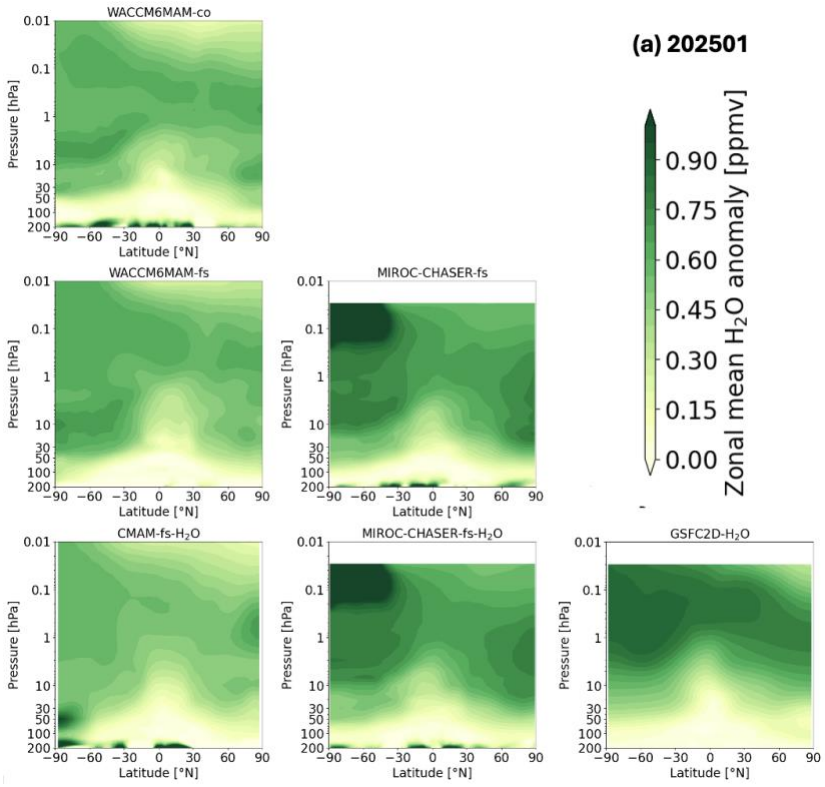


Figure S2. Simulated and observed global stratospheric H_2O mass anomalies within the 1–0.01 hPa pressure range following the Hunga eruption. Colored lines show ensemble-mean anomalies relative to the control simulations for each model, with shading indicating the respective ensemble spreads. The black line represents the observed anomaly derived from Microwave Limb Sounder (MLS) water vapor measurements.



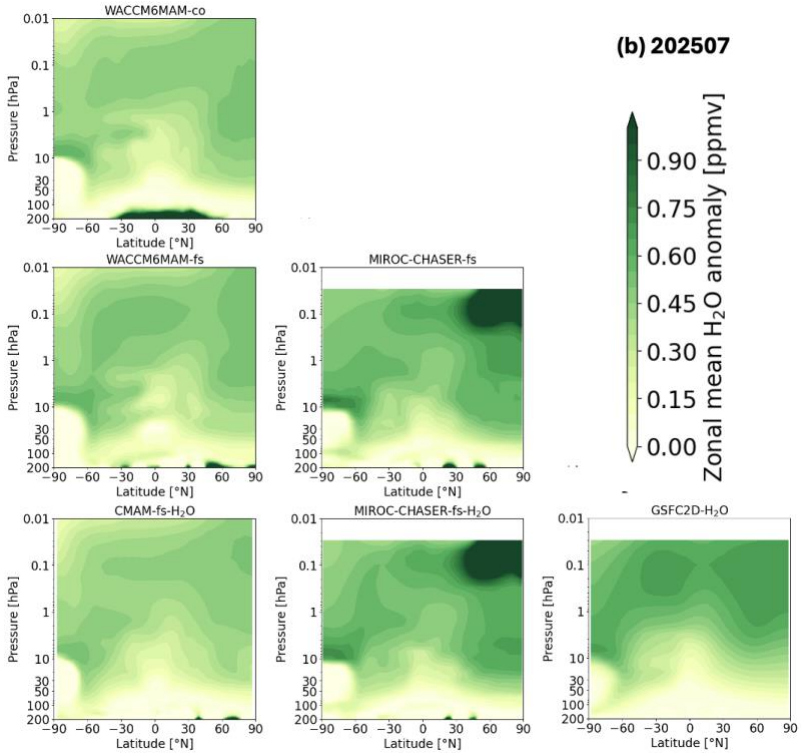


Figure S3: Latitude–pressure distribution of simulated zonal-mean water vapor anomalies in the upper atmosphere following the Hunga eruption, shown for (a) January 2025 and (b) July 2025. Anomalies are computed as differences between the ensemble mean of the experiment and control simulations.

Added Reference:

Randel, W. J., Podglajen, A., & Wu, F.: Stratospheric transit time distributions derived from satellite water vapor measurements. *Journal of Geophysical Research: Atmospheres*, *129*(21), e2024JD041595, 2024.

The interpretation of lower-level warming could be strengthened by overlaying aerosol fields on Figure 4. If the warming is indeed caused by aerosol descent, this should be evident in the aerosol spatial distribution patterns. Furthermore, this hypothesis could be tested by examining SO2-only simulations from 2022, where such warming should be absent due to lack of aerosol formation. Including these additional analyses would help validate the proposed mechanism for the temperature response.

As suggested by the reviewer, we have added aerosol extinction coefficients at 0.3 and 0.6 $10^{-3}km^{-1}$ and water concentration at 1 ppmv contours to Figure 4 to illustrate the spatial distribution of aerosols and to clarify their relationship with the observed temperature changes. The results show that the warming in the lower stratosphere clearly coincides with regions of enhanced aerosol extinction, supporting the interpretation that aerosol descent contributes to the warming.

The H_2O-only simulations, which lack aerosol formation, are included in the bottom panel of Figure 4. Compared to MIROC-CHASER-fs-H_2O, which does not include aerosol injection, the MIROC-CHASER-fs simulation shows stronger water vapor upward transport, weaker cooling in the middle stratosphere and stronger warming in the lower stratosphere. These differences further support the hypothesis that the lower stratospheric warming is primarily driven by the presence of aerosols.

Lines 415-420: The analysis of ozone chemistry (Figure 5) remains largely qualitative. While the authors attribute the simulated ozone changes - depletion in the lower mesosphere and enhancement in the middle stratosphere - to both UV radiation and various chemical pathways, the paper lacks quantitative calculations to support these mechanisms. Including detailed photochemical reaction rates, radiative transfer calculations, or chemical box model simulations would strengthen their arguments about the relative importance of different processes controlling ozone distribution changes.

We agree that a more detailed, quantitative analysis of the ozone response mechanisms would further strengthen the interpretation of the results. In this study, our discussion of the chemical pathways is based on mechanisms established in previous literature, including quantitative analyses performed using the same models employed here (e.g., WACCM6 and GSFC2D). These studies have explored the photochemical processes and radiative impacts of stratospheric water vapor increases in detail.

The primary objective of the present paper is to assess model performance through multi-model comparisons and model–observation evaluation, and to project the long-term evolution and impact of the Hunga eruption in the upper atmosphere. Given the complexity of the ozone response—which involves interactions between dynamics, radiation, and chemistry—we recognize the need for a dedicated, in-depth analysis. A coordinated study specifically focused on the ozone response and underlying mechanisms is currently in preparation as part of the broader HTHH-MOC effort.

We have revised the manuscript to clarify this scope limitation and now include the following statement:

"The ozone response mechanisms discussed here draw on previous single-model studies that conducted detailed photochemical analyses using the same modeling frameworks. While the current study does not include new quantitative calculations of individual reaction rates or radiative effects, a dedicated multi-model analysis of the ozone response and its underlying mechanisms is currently underway."

Line 94, Please more specific, what aspects of short or long-term evolution of HTHH are unclear from previous studies, and why previous studies only provided "limited" findings as you stated in Line 75.

As discussed in lines 77–93, most previous studies focused on short timescales—typically months to two years after the Hunga eruption—and were based on single-model frameworks. An exception is the study by Fleming et al. (2024), which extended to a decadal

timescale but only considered water vapor injection using the GSFC2D model, a two-dimensional model that lacks full representations of aerosol microphysics and chemistry.

To clarify this point and better justify our summary statement, we have relocated and revised the sentence to follow the description of individual studies. The updated text now reads: "These single-model studies, which primarily considered only water vapor injection with limited ensemble sizes or short simulation durations of up to two years (e.g., Zhu et al., 2022; Niemeier et al., 2023; Zhou et al., 2024), provide a limited understanding of the full evolution of the Hunga eruption. Although Fleming et al. (2024) explored decadal-scale impacts, they considered the H_2O injection only and did not include aerosol-chemistry interactions. Therefore, comparisons of multi-model simulations with larger ensemble sizes and longer time horizons are needed to fully understand both the short-term (months to two years) and long-term (multi-year to decadal) evolution of Hunga volcanic emissions and their atmospheric impacts"

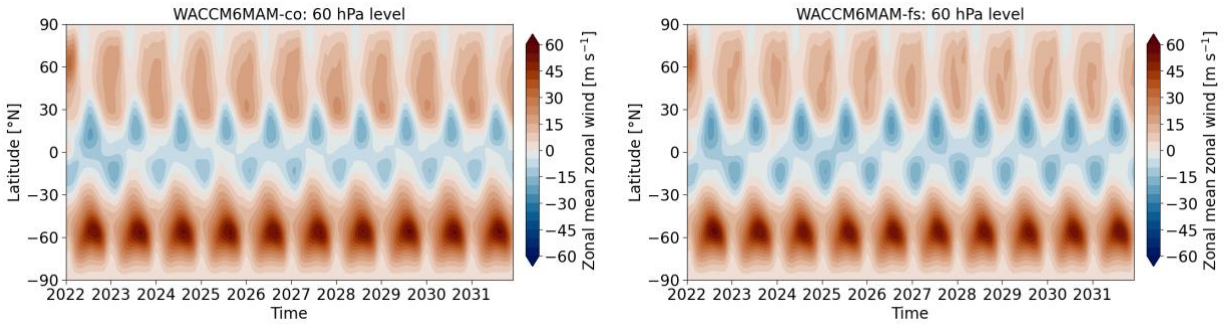
Figure 1 caption says OMPS data is used. Both datasets were used in the GSFC2D model.

We appreciate the reviewer's comment. Both GloSSAC and OMPS aerosol extinction datasets were indeed used as inputs to the GSFC2D model. For Figure 1, the SAOD shown for GloSSAC is taken directly from the observational SAOD data provided in the dataset. In contrast, the OMPS dataset does not include direct SAOD values; therefore, we derived SAOD from the available aerosol extinction data. To clarify this distinction, we have revised the figure caption to include the following sentence:

"The top-right panel shows the observed anomaly from the Global Space-based Stratospheric Aerosol Climatology (GloSSAC), relative to the 2012-2021 climatological period. The aerosol extinction of the GloSSAC data was used in the GSFC2D model. The aerosol extinction of the GloSSAC data was used in the GSFC2D model as their prescribed aerosol field input (Zhu et al., 2025). The bottom-right panel displays the Stratospheric Aerosol Optical Depth (SAOD) calculated from aerosol extinction data obtained from the Ozone Monitoring and Profiler Suite Limb Profiler (OMPS), which was utilized in the GSFC2D model."

Line 229: How does the secondary sAOD peak between 30-60S form?

The secondary SAOD peak between 30°–60°S is primarily influenced by the seasonal evolution of the stratospheric polar vortex. During austral summer (e.g., January 2023), the polar vortex weakens, allowing aerosols to penetrate further poleward from the midlatitudes. In contrast, during austral winter when the vortex is stronger, aerosols are more confined along the vortex edge, circulating within the midlatitudes. This behavior is evident in the zonal wind at 60 hPa—where most of the aerosols are located—shown in the figure from both the coupled ocean and fixed-SST WACCM6MAM simulations.



Line 292: could you please explain why one model injects 750 Tg WV, which is about 5 times of other models?

We appreciate the opportunity to clarify this point. Indeed, the water vapor injection mass of 750 Tg in GEOSCCM is notably higher than in the other models. This discrepancy arises due to differences in model-specific climatological conditions, as well as the treatment of microphysical and chemical processes. As outlined in the experimental protocol (Zhu et al., 2025, *ACP*, accepted), models were not required to use the same injection mass. Instead, they were asked to achieve a common target: approximately 150 Tg of water vapor retained in the stratosphere in Jan 2022, consistent with satellite observations of the H2O cloud following the Hunga eruption (as noted in lines 170–174 of the manuscript).

The 750 Tg water vapor injection in GEOSCCM was empirically determined to achieve the protocol target of ~150 Tg retained in the stratosphere. This relatively large initial mass reflects the model's sensitivity to water vapor loss due to condensation and transport processes, particularly under model-specific temperature conditions. Our tests show that GEOSCCM requires a higher injection mass when water is injected at lower altitudes to heat the environment by latent heat where colder temperatures promote greater condensation losses—an effect also supported by inter-model comparisons.

For example, WACCM6MAM achieved the 150 Tg retention target using the smallest injection mass (150 Tg), distributed over a broader vertical (25–35 km) and horizontal (22°S–6°S, 182.5°E–202.5°E) domain, thereby preventing widespread condensation. In contrast, GEOSCCM and MIROC-CHASER injected over a narrower range (25–30 km, 22°S–14°S, 182°E–186°E), resulting in higher water vapor losses and, thus, a need for greater initial injection. These differences highlight the importance of both injection geometry and background temperature structure. Notably, the results may also indicate a cold bias in GEOSCCM's lower stratosphere compared to other models, which would further increase condensation and reduce stratospheric water retention.

This finding underscores how differences in model configurations and biases can significantly affect the implementation of a common experimental design. It also provides useful guidance for future model development and intercomparison efforts, especially regarding stratospheric temperature biases and water vapor microphysics.

Line 311: Does the model simulate reduced ice cloud formation in MIROC-CHASER-Fs?

In MIROC-CHASER-fs, the ice clouds generated quickly and subsequently sedimented into the troposphere. Although the model did not archive the ice cloud formation rate, we infer a reduction in formation, as radiative heating from sulfuric acid aerosols could slightly warm the region even in the nudged simulation.

Line 319: Could you please elaborate the quantities that determine the lifetime (or e-folding time) of water and key differences between Zhou 2024 and present study?

The lifetime (or e-folding time) of stratospheric water vapor is determined by a combination of sources and sinks, which are governed by both physical and chemical processes. The primary source of stratospheric water vapor is upward transport from the troposphere through the tropical tropopause layer. Once in the stratosphere, water vapor is redistributed via vertical and horizontal transport, largely controlled by the Brewer–Dobson circulation. Unlike in the troposphere, where water vapor is rapidly removed via condensation and precipitation, removal in the stratosphere is much slower, occurring primarily through large-scale transport back to the troposphere, with a small amount loss via Antarctic dehydration and chemical pathways. Key chemical loss mechanisms include photolysis and reactions with excited oxygen atoms ($O(^1D)$), which convert H_2O into OH. These processes can also be modulated by aerosol-related chemistry involving OH and H_2O.

Regarding the differences between Zhou et al. (2024) and this study, several important methodological and conceptual distinctions exist. Zhou et al. (2024) used an offline 3D chemical transport model (CTM) driven by ERA5 meteorology, which does not simulate the internal variability of stratospheric dynamics or microphysical feedback. Their simulation included only water vapor injection, without SO_2 or aerosol-chemistry interactions. As a result, the model's water vapor burden is more stable, as reflected in the smoother time series shown in their Fig. 3. Moreover, their e-folding time was calculated based on the total atmospheric water vapor burden (across all model levels), with tropospheric mixing ratios limited to the constant stratospheric input value. In their framework, water removal begins in July 2023, from which the decay is fit, leading to a projected e-folding time of 48 months.

In contrast, our study employs fully coupled Earth system models with interactive dynamics, microphysics, and chemistry, which represent a more comprehensive set of processes affecting water vapor lifetime. Our focus is specifically on stratospheric water vapor (SWV). Since different models reach their maximum SWV burden at different times post-eruption, we calculate the e-folding time starting from the month of peak SWV for each individual model.

While it is clear that fully coupled models offer a more detailed representation of the system, the methodological differences in calculating e-folding time may not be immediately apparent to readers. To address this, we have clarified the distinction in the revised manuscript with the following sentence:

"Differently, using the offline 3D CTM model, Zhou et al. (2024) projected an overall e-folding decay timescale of 48 months from July 2023. Notably, this timescale reflects the removal of water vapor from the entire atmosphere, rather than from the stratosphere as considered in the present study."