Review of "Transport of volcanic aerosol from the Raikoke eruption in 2019 through the Northern Hemisphere" by Zhen Yang et al. (egusphere-2025-4842)

Original referee comments are in blue.

Our responses are in black with regular bold format. Text from the updated manuscript:

Appears in italic and with 0.5 cm indentation and with the modified parts in red.

General comments

Yang et al. investigate the transport pathways of the volcanic plume emitted by the Raikoke eruption in 2019, with focus on the Asian summer monsoon anticyclone (ASMA) region. The analysis is based on a combination of aerosol backscatter measurements by COBALD sondes flown in Lhasa (China) and model simulations using the CLaMS model. Two types of simulations are presented: backward trajectories, aimed to trace back the observed aerosol plumes to the Raikoke eruption, and global tracer simulations, investigating transport in the model and its sensitivity to different parameters. Aerosol number density profiles from POPS sondes flown in Boulder (USA) are also used to evaluate the results of the tracer simulations.

The paper is well written and fits well the scope of ACP. The observational data are of high quality and nicely combined with the model simulations, including the use of TROPOMI SO_2 satellite retrievals to define the initial state of the eruption. The results are presented clearly and concisely, although some of the assumptions made would require in my opinion further elaboration. The figures are excellent.

I support the publication of the paper. However, I have some comments that I think should be addressed to improve and/or clarify a few aspects of the study, mainly related to the backward trajectory analysis and the use of COBALD data to define the volcanic plume, as well as the overall impact of the paper.

We sincerely thank the reviewers for their encouraging comments and comprehensive revision suggestions, which have significantly enhanced the quality of the manuscript.

1. The "empirical" definition of the regions of enhanced BSR_{455} attributed to the Raikoke eruption (orange shadings in Fig. 2), which are used to initialize the backward trajectories, requires some clarification. I generally agree with the identified regions, but I think the criteria used for their classification should be discussed in more detail and could be at least partly quantified.

The main issue is the separation of the volcanic signal from aerosol features related to ATAL. There are enhanced BSR_{455} features in the profiles shown in Fig. 2 that are not classified as volcanic plume, likely because they are related to ATAL, but no explanation is given on why these regions are not considered. For example, on 10-08-2019, an enhanced BSR455 layer starts at ~140 hPa, well below the marked onset of the volcanic plume (84 hPa). Other small features can be seen on 01-08-2019 (~80-130 hPa), 06-08-2019 (~120-140 hPa), and 12-08-2019 (~120-140 hPa). If my interpretation is correct and these features are related to ATAL, I would suggest to discuss this in the text and to highlight them accordingly in Fig. 2 (e.g., by a different color shading). If not, why are these regions excluded?

Secondly, the visual identification of the plume boundaries becomes difficult when the BSR_{455} gradient is smooth, as in the last three flights (Sept-Nov). Here, I have the impression that the lower boundary of the plume does not match the onset of the BSR_{455} enhancement (roughly at 100 hPa), but rather some given threshold in BSR_{455}. Is there a reason to define the lower boundary this way?

One possibility to quantify these criteria could be the use of the COBALD color index (CI), defined as the ratio of the 940-to-455 nm aerosol BSR (i.e., BSR – 1). Using information from both wavelengths, the CI is a proxy of particle size that was used in several studies to separate clouds from aerosols (e.g., Vernier et al., 2015; Brunamonti et al., 2018; Hanumanthu et al., 2020). Here, the CI might help to distinguish the volcanic plume from ATAL, assuming they have different size distributions, and to define its boundaries more accurately. Have you considered looking into this?

We thank the referee for this helpful suggestion. Distinguishing Raikoke aerosol from ATAL using in situ backscatter profiles alone is challenging, because humidity alone cannot distinguish volcanic aerosol from ATAL, and depolarization data are not available for these COBALD measurements. Our revised approach therefore combines BSR, RH_{ice} and CI, and we note that the CI based separation should be regarded as supportive rather than definitive. In the revision, we clarify and partly quantify how the Raikoke related layers (orange shading in Fig. 1 of this reply) were selected for initializing the backward trajectories, and how we separate them from ATAL related enhancements.

We now apply a consistent set of thresholds for the classification. First, we exclude cirrus using $BSR_{455} > 1.2$, $RH_{ice} > 70\%$, and $CI > 7$. In the remaining cloud free parts, aerosol layers are identified using a uniform threshold of $BSR_{455} > 1.1$. When multiple BSR_{455} enhanced layers occur in the same profile, we use the COBALD color index as an additional indicator, applying $CI > 6$, to support the separation between the upper Raikoke layer and lower layers that are more consistent with ATAL.

We also clarify that CI is not used as a stand alone boundary definition. In very clean air where BSR values approach 1, CI can become unstable because it is computed from (BSR−1) at two wavelengths, and it can also be sensitive to small baseline or calibration offsets. We therefore evaluate CI only within BSR_{455} enhanced, cloud free layers and interpret it as a supporting indicator.

These updates are reflected in the Fig. 1 of this reply, where ATAL like enhancements are now highlighted separately from the Raikoke plume. In addition, we provide a supplementary plot (Fig. 2 of this reply) showing box plots of BSR_{455} and CI for the two categories, which supports the separation.

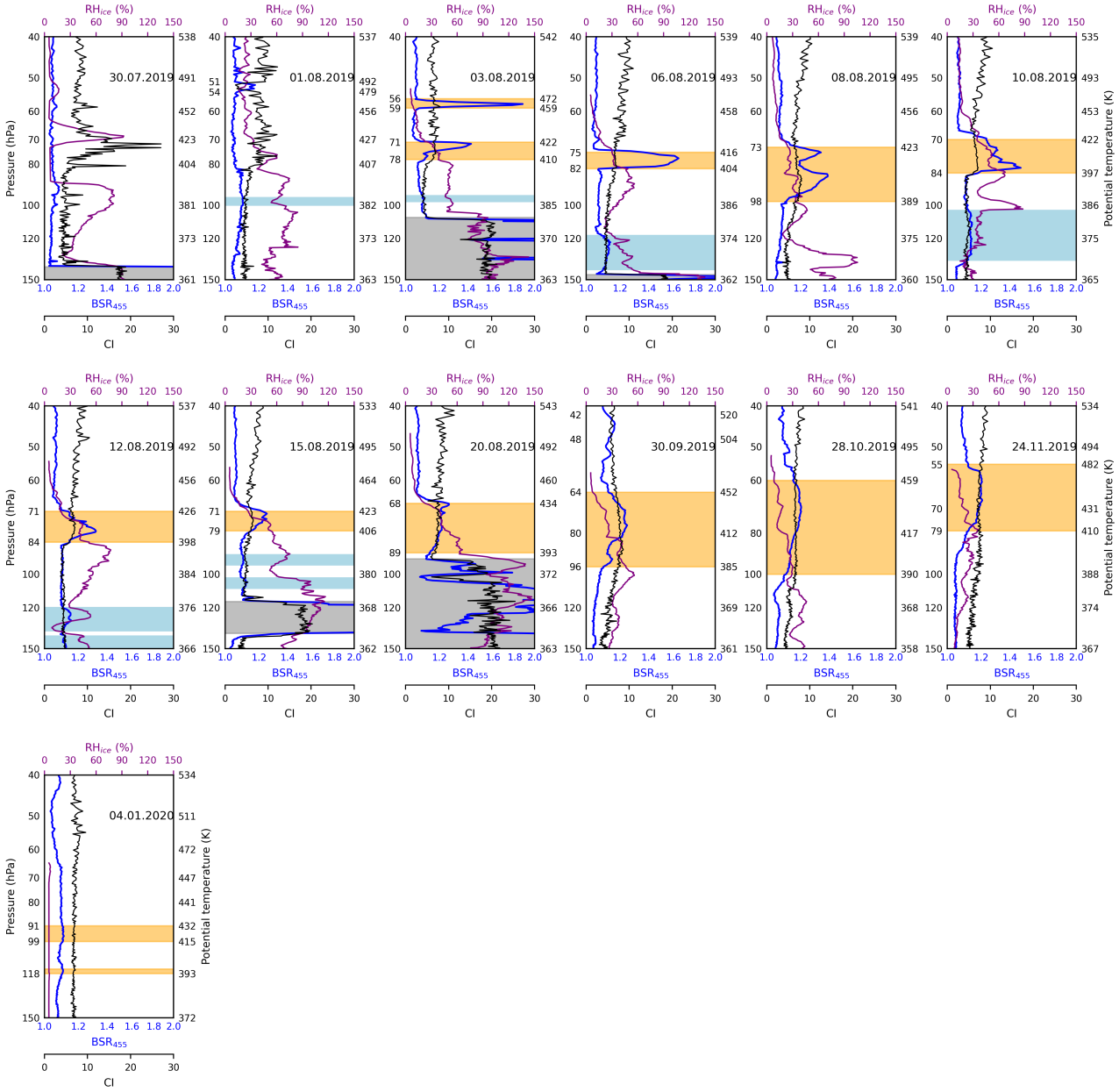

Figure 1 (Figure 2 of the revised manuscript). COBALD and CFH profiles above Lhasa from 30 July 2019 to 4 January 2020. COBALD backscatter ratio at 455 nm (BSR_{455}, blue) and COBALD color index (CI, black) are shown together with CFH relative humidity over ice (RH_{ice}, magenta). Pressure is shown on the left axis and the corresponding potential temperature on the right axis. Orange shading marks the Raikoke-influenced aerosol layers used to initialize the backward trajectories. Light blue shading indicates lower-altitude aerosol enhancements attributed to the ATAL that are not used for plume initialization. Gray shading indicates cirrus conditions.

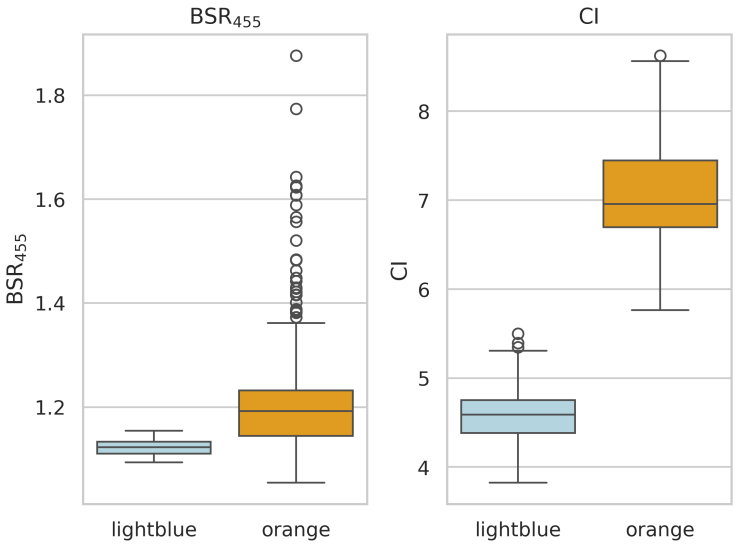

Figure 2. Box plot summary of COBALD BSR_{455} and color index (CI) values sampled within cloud free aerosol layers (BSR_{455} > 1.1) classified as ATAL type (light blue shading in Fig. 1) and Raikoke plume (orange shading in Fig. 1). Boxes show the median and the middle 50% of the data. The whiskers show the typical range, and the circles show unusual values.

2. Since the Boulder profiles are not included in the backward trajectory analysis, the paper focuses almost entirely on transport of the volcanic plume in the ASMA region, rather than the whole Northern Hemisphere (as hinted by the title). The Boulder data are only used to evaluate the results of the tracer simulations, and although the agreement with the model is remarkable (Fig. 8), the discussion of these measurements remains very limited (a few lines on pages 7 and 14). At the same time, the tracer distribution map in Fig. 6 shows that the filament advected over Lhasa represents only a small fraction of the volcanic plume, while most of it stays outside of the ASMA. Therefore, I find that the general discussion is not properly balanced in this respect. I would suggest to try better integrating the Boulder data into the "storyline" of the paper, for example, by adding backward trajectories initialized from these flights into Fig. 5 (or an additional figure). This would allow to frame the entire transport pathway analysis in a more general context, and would be in my opinion a great addition to the paper. Otherwise, it should be at least pointed out more clearly that the ASMA pathway only accounts for a minor fraction of the entire volcanic plume emitted by the Raikoke eruption.

We thank the referee for this insightful comment. We have revised the manuscript according to the reviewer's advice.

To maintain consistent identification of volcanic plumes throughout this paper, we initialize trajectories at the potential temperature interval where the in-situ particle number density exceeds 150 cm^{-3}. This criterion captures the main volcanic plume signal.

Specifically, we now include backward trajectories for the plume layers observed on 7 August 2019 (373–442 K), 27 August 2019 (385–429 K and 480–505 K) in Figure 3 of this reply, and we add the corresponding global three-dimensional tracer simulations for these days on the American continent in Figure 4 of this reply. These additions provide a broader hemispheric context for the transport pathways and clarify how the ASMA-related transport compares with pathways outside the ASMA.

In addition, we revised the discussion to clarify the evolving role of the ASMA in the three-dimensional tracer distributions: during August, during the peak ASMA season, only a minor fraction of the Raikoke tracer is located inside the ASMA, while the majority remains outside; by September, the tracer has become substantially mixed and the difference between the ASMA and non-ASMA regions in the regional distribution is no longer clearly evident.

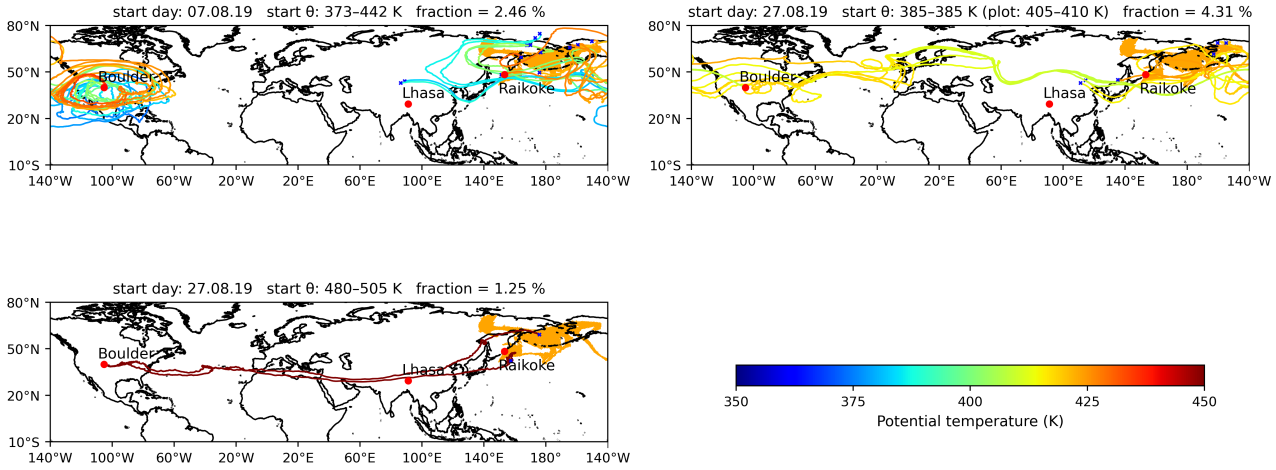

Figure 3 (Figure 6 of the revised manuscript). Backward trajectories for the Boulder flights (7 and 27 August 2019). Only trajectories that intersect the TROPOMI SO_2 mask during 24 June 2019, 22:46 to 25 June 2019, 03:50 UTC are shown. Trajectories are coloured by potential temperature, and the SO_2 mask is shaded in orange. Panel titles give the start day, start θ range, and the fraction of trajectories that satisfy the SO_2 mask criterion. For the 7 August 2019 back trajectories, the fraction is computed for 385 to 429 K, while the plotted trajectories are restricted to 405 to 410 K to better illustrate the transport pathway. Locations of Boulder, Lhasa, and Raikoke are marked.

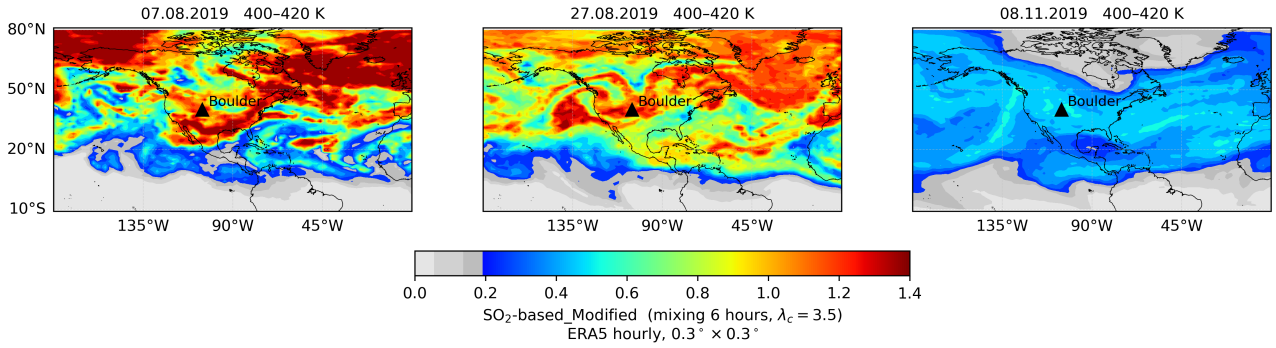

Figure 4 (Figure 8 of the revised manuscript). SO_2-based tracer fractions within the 400–420 K layer over the American sector. The black triangle indicates the location of Boulder.

Specific comments

Page 1, line 15: this sentence requires more context (what does "additional" uplift refer to?).

We removed the statement from the abstract and addressed it consistently regarding comments on pages 16, lines 304–309 (see response below).

Page 2, line 32: consider omitting the definition of VEI (not relevant to this study).

We removed the definition and discussion of the Volcanic Explosivity Index (VEI). We retain the "eruption magnitude" factor but describe it more directly in terms of SO_2 and ash release, and add a brief note on submarine eruptions (e.g. Hunga Tonga 2022).

Several key factors govern the climate impact of volcanic aerosols from an eruption: (1) eruption magnitude; (2) injection height and self-lofting; (3) eruption latitude; and (4) dynamical evolution. ~~*(1) The Volcanic Explosivity Index (VEI) serves as a proxy for eruption intensity (Newhall and Self, 1982), and events with VEI ≥ 4 can inject vast quantities of SO_2, water vapor, and ash, causing marked climate perturbations.*~~ *Eruption magnitude can be characterized by the amount of SO_2 and ash released, which largely controls aerosol formation potential and radiative forcing. Notably, substantial perturbations of stratospheric water vapour and aerosol microphysics have also been reported for the 2022 Hunga Tonga eruption despite modest SO2 injection (Carn et al., 2022; Zhu et al., 2022). (2) Eruption products injected directly into the stratosphere, as well as volcanic plumes in the upper troposphere that self-loft into the lower stratosphere through radiative heating, can persist much longer than plumes confined to the troposphere (Toohey et al., 2025). (3) Aerosols from tropical eruptions are transported most efficiently via the Brewer–Dobson circulation (Brewer, 1949; Dobson, 1956; Butchart, 2014), whereas mid-latitude eruption aerosols can still reach the tropics through Rossby-wave breaking or transport by the Asian Summer Monsoon Anticyclone (ASMA) (Konopka et al., 2009; Kloss et al., 2021; Wu et al., 2023). (4) UTLS jet streams, cyclones, anticyclones, and the stratospheric circulation govern dispersion patterns and dilution rates.*

Page 2, line 44: add some references for the chemical trapping in the ASMA, e.g. Park et al. (2007), Randel et al., (2010) (see reference list below).

Yes, the relevant references have been added.

Deep convection injects pollutants into the UTLS, where the ASMA's strong anticyclonic circulation acts as a dynamical transport barrier, confining these air masses over Asia during ascent into the stratosphere (e.g., Park et al., 2007; Randel et al., 2010; Fadnavis et al., 2014; Santee et al., 2017; Vogel et al., 2019).

Park, M., Randel, W. J., Gettleman, A., Massie, S. T., and Jiang, J. H.: Transport above the Asian summer monsoon anticyclone inferred from Aura Microwave Limb Sounder tracers, J. Geophys. Res., 112, D16309, https://doi.org/10.1029/2006JD008294, 2007.

Randel, W. J., Park, M., Emmons, L., Kinnison, D., Bernath, P., Walker, K. A., Boone, C., and Pumphrey, H.: Asian Monsoon Transport of Pollution to the Stratosphere, Science, 328, 611–613, https://doi.org/10.1126/science.1182274, 2010.

Fadnavis, S., Schultz, M. G., Semeniuk, K., Mahajan, A. S., Pozzoli, L., Sonbawne, S., Ghude, S. D., Kiefer, M., and Eckert, E.: Trends in peroxyacetyl nitrate (PAN) in the upper troposphere and lower stratosphere over southern Asia during the summer monsoon season: regional impacts, Atmos. Chem. Phys., 14, 12 725–12 743, https://doi.org/10.5194/acp-14-12725-2014, 2014.

Santee, M. L., Manney, G. L., Livesey, N. J., Schwartz, M. J., Neu, J. L., and Read, W. G.: A comprehensive overview of the climatological composition of the Asian summer monsoon

anticyclone based on 10 years of Aura Microwave Limb Sounder measurements, J. Geophys. Res., 122, 5491–5514, https://doi.org/10.1002/2016JD026408, 2017.

Vogel, B., Müller, R., Günther, G., Spang, R., Hanumanthu, S., Li, D., Riese, M., and Stiller, G. P.: Lagrangian simulations of the transport of young air masses to the top of the Asian monsoon anticyclone and into the tropical pipe, Atmos. Chem. Phys., 19, 6007–6034, https://doi.org/10.5194/acp-19-6007-2019, 2019.

Page 3, line 79: which model of iMet radiosonde was used?

We used an International Met Systems (InterMet) iMet-1-RSB radiosonde. We have added the specific radiosonde model in the instrument description.

Balloons were equipped with an electrochemical concentration cell (ECC) ozonesonde, a cryogenic frostpoint hygrometer (CFH), a compact optical backscatter aerosol detector (COBALD), and an International Met Systems (InterMet) iMet-1-RSB radiosonde (GRUAN Lead Centre, 2025). ECC measurements are not analyzed in this study.

GRUAN Lead Centre: InterMet iMet-1 (iMet-1-RSB) — GRUAN instrument information (radiosonde models), https://www.gruan.org/instruments/radiosondes/sonde-models/intermet-imet-1, accessed 13 Jan 2026, 2025.

Page 3, line 80: I suggest adding a table to summarize date, time, location and payload of all analyzed balloon flights, and possibly to introduce a sequential numbering of the flights (e.g. F1, F2, ...).

Site	Flight ID	Date	Mid-ascent time (UTC)	Key instruments	Observed plume layer: θ (K)	Observed plume layer: p (hPa)
Lhasa	F01	30.07.19	17:58	COBALD CFH	–	–
	F02	01.08.19	15:45	COBALD CFH	–	–
	F03	03.08.19	15:23	COBALD CFH	410–422; 459–472	78–71; 59–56
	F04	06.08.19	15:26	COBALD CFH	404–416	82–75
	F05	08.08.19	15:10	COBALD CFH	389–423	98–73
	F06	10.08.19	16:05	COBALD CFH	397–422	84–70
	F07	12.08.19	15:49	COBALD CFH	398–426	84–71
	F08	15.08.19	17:24	COBALD CFH	406–423	79–71
	F09	20.08.19	17:14	COBALD CFH POPS	393–434	89–68
	F10	30.09.19	15:34	COBALD CFH	385–452	96–64
	F11	28.10.19	15:24	COBALD CFH	390–459	100–60
	F12	24.11.19	15:09	COBALD CFH	410–482	79–55
	F13	04.01.20	15:29	COBALD CFH	393–399; 415–432	118–115; 99–91
Boulder	F14	28.06.19	17:05	POPS	–	–
	F15	07.08.19	17:08	POPS	373–442	122–73
	F16	27.08.19	16:38	POPS	385–429; 480–505	117–80; 59–52
	F17	08.11.19	17:58	POPS	389–435	107–76
	F18	03.12.19	18:05	POPS	–	–

Table 1 (Table 1 of the revised manuscript). Summary of analyzed balloon flights at Lhasa and Boulder, including sequential flight ID, date, mid-ascent time (UTC), key instruments used in this study, and the potential temperature (θ) and pressure (p) ranges of the observed volcanic plume layer. The plume layer was identified from COBALD measurements where BSR_{455} > 1.1 and CI > 6, and from POPS measurements where the particle number concentration exceeded 150 cm^{-3}. Mid-ascent time denotes the midpoint of the ascent period from launch to balloon burst. The "Key instruments" column lists only the instruments analyzed in this work; additional payload components were flown but are not listed here. A dash (-) denotes that no volcanic plume layer was identified for that flight.

Page 5, lines 98-100: I cannot find the source of the COBALD uncertainties given here in Vernier et al. (2015), neither the "maximum BSR uncertainties" of 1.3 % at 940 nm and 0.2 % at 455 nm at ground level, nor the 5 % at 940 nm and 1 % at 455 nm at 10 km altitude. How are these numbers obtained? Vernier et al. (2015) only estimate a 5 % uncertainty for the entire profile, due to physical constraints, and 1 % precision in the UTLS, without differentiating between the two channels. The same uncertainty of 5 % is also reported by other studies using COBALD data in the lower troposphere (Brunamonti et al., 2021) and UTLS (Reinares Martínez et al., 2021). Considering that the retrieval algorithm of COBALD BSR involves empirically-determined instrumental parameters as well as the measured temperature and pressure to calculate the molecular extinction profile, I doubt that such high accuracies can be achieved.

We agree that our previously stated "maximum BSR uncertainties" specified separately by wavelength channel and altitude are not supported by Vernier et al. (2015). We have therefore removed these values and revised the text. In the revised manuscript, we adopt the commonly used COBALD uncertainty characterization: an absolute error interval of ~5% for the BSR profile and a precision better than ~1% in the UTLS (Vernier et al., 2015), consistent with other COBALD-based studies (e.g., Brunamonti et al., 2021; Reinares Martínez et al., 2021).

~~*The maximum BSR uncertainty is 1.3 % at 940 nm and 0.2 % at 455 nm at ground level. At 10 km altitude, uncertainties increase to 5 % and 1 % at 940 nm and 455 nm, respectively (Vernier et al., 2015).*~~ *COBALD BSR uncertainty is typically characterized by an absolute error interval of about ~5% for the profile, while the precision is better than ~1% under UTLS conditions (Vernier et al., 2015; Brunamonti et al., 2021; Reinares Martínez et al., 2021).*

Page 5, lines 106-107: I would be more conservative with the CFH uncertainty. Vömel et al. (2016) state that the uncertainty "may be" as low as 2 % in the lower troposphere and 5 % at the tropical tropopause, under good operating conditions of the mirror temperature controller. The mirror temperature controller is the largest source of uncertainty in CFH measurements and oscillations around the real frostpoint can be up to ±0.5 K, corresponding to ±10 % in H2O mixing ratio at UTLS conditions (e.g., see Poltera et al., 2025). Based on Fahey et al. (2014), I think a more realistic estimate of the CFH uncertainty in the stratosphere is ±10 %, unless the performance of the mirror temperature controller is evaluated specifically for each flight.

We revised the manuscript to reflect that the 2 % (lower troposphere) and 5 % (tropical tropopause) values in Vömel et al. (2016) apply only under good operating conditions. Given that the mirror temperature controller can dominate the uncertainty in the UTLS, we now adopt a conservative ±10 % uncertainty for CFH water vapor measurements in the UTLS and

lower stratosphere, consistent with Fahey et al. (2014) and the discussion in Poltera et al. (2025).

> ~~*The measurement uncertainty is approximately 2 % in the lower troposphere and increases to 5 % near the tropical tropopause (Vömel et al., 2016).*~~ *CFH uncertainties may be as low as ~2 % in the lower troposphere and ~5 % near the tropical tropopause under good operating conditions (Vömel et al., 2016). For UTLS conditions we use a conservative uncertainty of ±10 % (Fahey et al., 2014; Poltera et al., 2025).*

Page 6, line 138: is temperature really needed to define the starting positions?

Temperature is not an independent variable to define the starting positions. Since our trajectory analysis is performed in isentropic coordinates, we use potential temperature as the vertical coordinate, which is derived from the measured pressure and temperature. We clarified this in the revised manuscript.

> *Diabatic backward trajectories are initialized every second along the balloon's vertical ascent profile, using the in-situ measurements of temperature (for deriving potential temperature), pressure, time, longitude, and latitude to define the start positions.*

Page 6, line 152: add a short explanation of the physical meaning of the critical Lyapunov exponent. Does a higher λc correspond to more or less mixing?

We added a short explanation of the physical meaning of the critical Lyapunov exponent. In the CLaMS mixing scheme, parameterized mixing is triggered when the integral deformation between neighboring air parcels exceeds the critical value $\gamma_c=\lambda_c \Delta t$, where Δt is the mixing interval. Thus, for a fixed Δt, a larger λ_c (larger γ_c) requires stronger deformation to trigger mixing and corresponds to less frequent parameterized mixing (and vice versa).

> *To assess how different mixing intensities influence the reconstruction of volcanic plume transport processes, two simulations were conducted:*
>
> *(i) a control simulation with mixing every 24 hours and $\lambda_c = 1.5$;*
>
> *(ii) a modified simulation with mixing every 6 hours and $\lambda_c = 3.5$.*
>
> *In the CLaMS mixing scheme, parameterized mixing is triggered when the integral deformation between neighboring air parcels exceeds an empirical critical deformation $\gamma_c = \lambda_c \Delta t$, where λ_c is the critical Lyapunov exponent and Δt is the advective time step (mixing interval). For a given Δt, a larger Δt (thus a larger γ_c) requires stronger deformation to trigger mixing and therefore corresponds to less frequent parameterized mixing (and vice versa) (Konopka et al., 2004, 2007). In our setup, the modified simulation corresponds to enhanced parameterized mixing compared to the control simulation. Throughout most of the paper we show results from the modified simulation, as these agree better with the observations. Sensitivity to parameterized mixing intensity and comparisons with the control simulation are discussed in Sect. 5.1. The mixing configurations and the additional sensitivity runs (rectangular mask and coarser ERA5 input) are summarized in Table 2.*
>
> *Konopka, P., et al., 2004: Mixing and ozone loss in the 1999–2000 Arctic vortex: Simulations with the three-dimensional Chemical Lagrangian Model of the Stratosphere (CLaMS), J. Geophys. Res., 109, D02315, doi:10.1029/2003JD003792.*

Konopka, P., et al., 2007: Contribution of mixing to upward transport across the tropical tropopause layer (TTL), Atmos. Chem. Phys., 7, 3285–3308, doi:10.5194/acp-7-3285-2007.

Page 7, lines 159-160: the cloud-filtering criteria used here (BSR_{455} > 1.2, RH_{ice} > 70 %) are those derived by Yang et al. (2023), which are, to my understanding, a modified version of the criteria used in previous studies (Vernier et al., 2015; Brunamonti et al., 2018; Hanumanthu et al., 2020), without taking the color index (CI) into account. As I already mentioned, the spectral information contained in the CI is crucial to make a physically-based (rather than empirical) discrimination, since it allows to distinguish size effects (change in BSR and CI) from number density effects (change in BSR but no change in CI). Therefore, I think it would be very interesting to investigate the CI here, as this may provide a quantitative basis for a more accurate definition of the volcanic plume and its boundaries.

We agree that CI provides important spectral (particle-size) information. We revised the manuscript and now identify cirrus using a combined criterion (BSR_{455}> 1.2, RH_{ice}> 70%, and CI > 7). Different studies use different threshold combinations and we require all three conditions to be met to classify a data point as cirrus.

Cirrus clouds are identified using a combined criterion of BSR_{455}> 1.2, RH_{ice}> 70% and CI > 7 (Vernier et al., 2015; Brunamonti et al., 2018; Hanumanthu et al., 2020; Yang et al., 2023).

Page 7, line 163: if "coexist" means that an aerosol layer and a cirrus cloud overlap in altitude, then the two signals cannot be distinguished (rather than "it becomes difficult"). If they coexist in the same profile but on different altitude levels, then the visual identification may become more difficult, but the signals can still be isolated quantitatively (e.g., using the CI). Please clarify.
We clarified the meaning of "coexist". If cirrus and aerosols are vertically separated within a profile, the aerosol signal can still be isolated quantitatively (e.g., using BSR and CI); however, if they overlap at the same altitude, the aerosol contribution cannot be reliably separated because the cirrus signal dominates.

~~*When cirrus clouds and aerosols coexist, it becomes difficult to isolate the aerosol signal because cirrus BSR455 values are significantly higher than those of aerosols.*~~ *When cirrus and aerosols occur within the same profile, aerosol and cloud signals can generally be separated if they are vertically distinct (e.g., using BSR and CI), whereas if a cirrus cloud overlaps the aerosol layer at the same altitude, the aerosol contribution cannot be reliably isolated because cirrus BSR_{455} typically dominates. Thus, the aerosol cannot be reliably detected or quantified under such conditions.*

Page 7, line 165: I suggest "determined by visual inspection" instead of "empirically highlighted".

In the revised manuscript, the highlighted regions are identified using threshold criteria. Specifically, we first exclude cirrus using BSR_{455} > 1.2, RH_{ice} > 70%, and CI > 7, then define aerosol layers by BSR_{455} > 1.1, and finally classify them using CI = 6 (Raikoke: CI > 6; ATAL/background: CI < 6). We have revised the text accordingly.

~~*Regions showing enhanced BSR_{455} — most likely due to Raikoke aerosols — are empirically highlighted in orange in Fig. 2.*~~ *Using these criteria, we exclude cirrus-contaminated layers. Remaining enhancements with BSR_{455} > 1.1 are treated as aerosol layers and then classified using a CI threshold of 6: layers attributed to the Raikoke plume (CI > 6) are highlighted in*

orange in Fig. 2, while layers more consistent with ATAL aerosol (CI < 6) are highlighted in light blue. We note that CI is used here as an additional indicator within aerosol layers and should be interpreted with caution, and that the ATAL identification in our dataset is limited. In very clean air where BSR values at both wavelengths approach 1, CI can become unstable; we therefore apply CI only within aerosol layers with BSR_{455} > 1.1 after excluding cirrus.

Page 7, lines 171-172: I presume the ATAL profile from 2013 shown in Fig. 3a is the COBALD profile from Lhasa by Vernier et al. (2015). Is this correct? Please add a citation.
Yes, this is correct. The 2013 ATAL profile shown in Fig. 3a corresponds to the COBALD measurements over Lhasa presented by Vernier et al. (2015). We have added the citation in the manuscript.

The ATAL profile from 2013 shown in Fig. 3a is taken from the COBALD measurements over Lhasa reported by Vernier et al. (2015).

Page 7, line 174: how much does 33 K potential temperature correspond in altitude (roughly)?
Using the θ–altitude relationship from the 30 July 2019 background sounding, a separation of 33 K around the tropopause corresponds to roughly ~1.7 km. We have added this approximate altitude difference to the manuscript.

Page 10, line 199: quantify "extreme" BSR_{455} values.
We quantified "extreme" by reporting the peak and 95th-percentile BSR_{455} within the identified plume layer (orange shading in Fig. 2). Peak BSR_{455} decreases from up to 1.88 in early August (p95 up to 1.77) to ≤ 1.25 during 30 September–24 November (p95 ≤ 1.24).

Over the following three months, as the ASMA weakened seasonally, air from lower potential-temperature levels increasingly influenced the Lhasa profiles. During this period, peak BSR_{455} within the identified plume layer (orange shading in Fig. 2) decreased from values up to 1.88 in early August 2019(95th percentile up to 1.77) to ≤ 1.25 from 30 September to 24 November 2019 (95th percentile ≤ 1.24), indicating progressive dilution of the volcanic aerosol layer by relatively aerosol-poor air from the lower troposphere.

Page 10, lines 204-206: why should the filtering criterion be considered "highly selective"? Is this related to the spatial/temporal extent of the mask, or its "patchiness"? Would it help to use a more compact domain (e.g., the rectangular mask used in Section 4.3), or to extend the considered time window? And what are the source regions of the > 90 % trajectories that are not shown? This is a key point of the paper, so I think some more elaboration is required.

We use the full TROPOMI-observed SO_2 plume footprint during the satellite overpass as the filtering criterion because it represents the actual observed situation. In this sense, Fig. 5 of the revised manuscript provides a conservative estimate: trajectories are selected only if they intersect the observed plume footprint within the overpass time window. This makes the criterion "highly selective" mainly because both the spatial footprint and, in particular, the temporal window are narrow, so only a small fraction of trajectories satisfy the constraint.

To assess the sensitivity to the spatial and temporal definition of the filter, we performed two additional tests. First, we approximate the TROPOMI footprint by a broader rectangular domain (137–215°E, 42–73°N; Fig. 5 of this reply) while keeping the original overpass time window (24 June 2019, 22:46 UTC to 25 June 2019, 03:50 UTC) in Fig. 6 of this reply.

Second, we apply the same rectangular domain but extend the time window to 7 days (21 June 2019, 18:00 UTC to 28 June 2019, 18:00 UTC) in Fig. 7 of this reply. With the original TROPOMI footprint and overpass window, the fractions of trajectories reaching the eruption region are ~3–10% across the analyzed layers/dates, whereas the relaxed rectangular domain and 7 days window criterion increases these fractions to ~12–28%. We interpret the latter range as an upper bound under a looser, yet still reasonable, filter. The remaining trajectories largely represent background transport pathways during the specified time window.

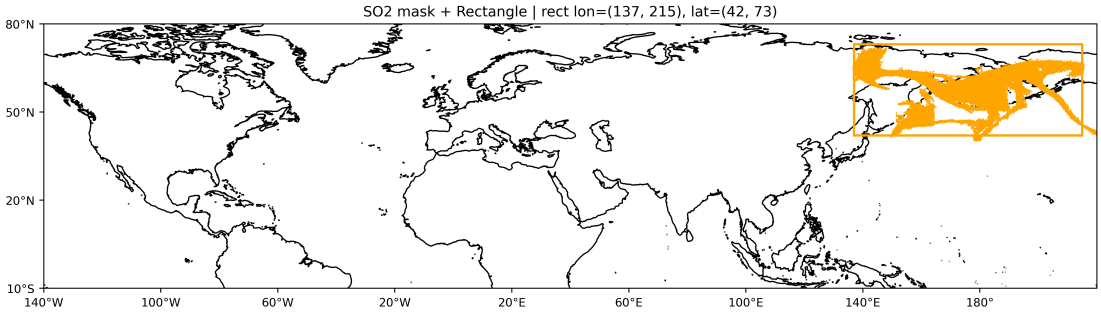

Figure 5. TROPOMI SO_2 plume footprint used for trajectory filtering (orange shading) and its rectangular approximation (137–215°E, 42–73°N; orange box).

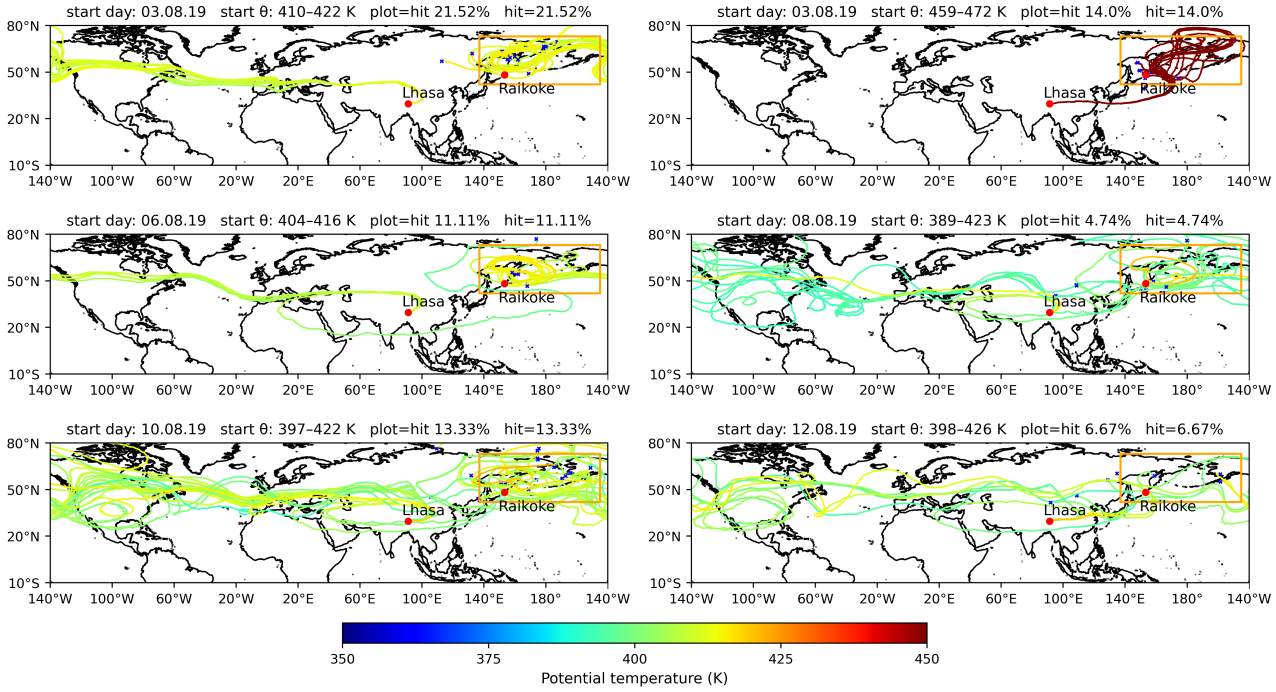

Figure 6. Backward trajectories filtered by intersection with the rectangular domain (137–215°E, 42–73°N) within the original TROPOMI overpass time window (24 Jun 2019, 22:46 UTC–25 Jun 2019, 03:50 UTC).

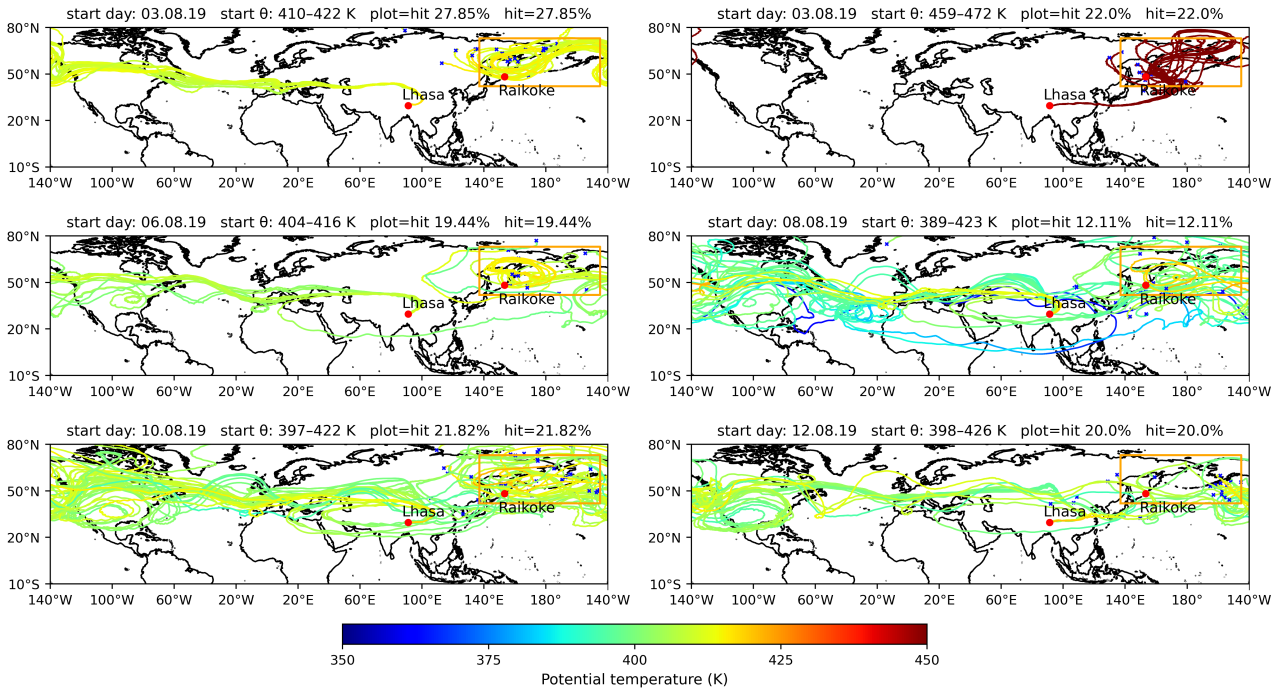

Figure 7. Same as Fig. 6, but using an extended 7-day time window (21 Jun 2019, 18:00 UTC–28 Jun 2019, 18:00 UTC) to test sensitivity to the temporal definition of the filter.

A sensitivity test shows that the fraction of backward trajectories reaching the eruption region is ∼3–10% when using the original TROPOMI SO_2 footprint and satellite-pass time window, but increases to ∼12–28% when applying a broader rectangular domain (137–215° E, 42–73° N) and extending the time window to 7 days. The retained trajectories still follow the same main transport pathways shown here, providing confidence in the identified transport patterns.

Page 10, line 209: which satellite? Is this statement referring to Khaykin et al. (2022)?

We clarified the satellite instrument by TROPOMI on Sentinel-5P and reworded the sentence to specify that the consistency with Fig. 5 refers to the backward trajectories initialized in the 3 August plume layer (459–472 K).

~~*Satellite tracking shows that the primary VVP was entrained in summertime easterlies around 20–25 July, circled the globe three times, and passed south of the Tibetan Plateau on 31 July. This pathway aligns with the backward trajectories in Fig. 5. Furthermore, the potential temperature measured by the satellite during its ASMA transit also closely matched the altitudes of enhanced BSR455 (Gorkavyi et al., 2021; Khaykin et al., 2022).*~~ *Satellite observations from TROPOMI on Sentinel-5P indicate that the VVP core was entrained into the summertime easterlies around 20–25 July (Gorkavyi et al., 2021; Khaykin et al., 2022). The backward trajectories initialized in the 3 August 2019 plume layer (459–472 K) in Fig. 5 are consistent with the late-July phase of this satellite-tracked pathway. The potential temperature of the VVP during its transit through the ASMA inferred from satellite detections also closely matches the altitudes of enhanced BSR_{455} (Gorkavyi et al., 2021; Khaykin et al., 2022).*

Page 11, lines 218-221: why are the backward trajectories of the September-November flights not included?

We did not include backward trajectories for the September–November flights (also the backward trajectories for Boulder on 8 November 2019) because they require tracing the air

parcels much further back in time. Although we have adopted the high-resolution ERA5 data as the input data, the reliability of trajectory reconstruction inevitably decreases over time. In addition, by September–November the Raikoke plume is substantially diluted and more thoroughly mixed with background air, so coherent pathway information becomes less distinct. We therefore focus the trajectory analysis on the July–August period, when plume structures are still relatively coherent and the trajectories provide clearer transport-pathway information.

Page 12, line 248: I suggest adding a table to summarize the different model runs and their main characteristics (mixing parameters, injection region/height, etc.). I would also recommend introducing a more compact notation for these simulations (e.g., "SO2-based_control" → CTRLSO2), and perhaps a more meaningful name for the "modified" scenario (e.g. "mixing-enhanced", "MIX-ENH" or similar).

Thank you for the suggestion. We added a new table (Table 2 in the manuscript) summarizing the model runs and their key settings (injection region/layers, mixing interval, and the critical Lyapunov exponent λc). We still retained the original descriptive simulation names (SO_2-based_Control, SO_2-based_Modified, and rectangle-based_Modified), because they are more informative and allow readers to infer the main differences directly from the naming.

Simulation	Injection region	Injection layer	Mixing interval	λ_c	Notes
All simulations listed below are driven by ERA5 at $0.3^\circ \times 0.3^\circ$ *spatial and 1 hour temporal resolution. Figure A2 shows the* SO_2*-based_Control simulation repeated with coarser ERA5 input (*$1^\circ \times 1^\circ$ *spatial, 6 hours temporal resolution).*					
SO_2-based_Control	satellite SO_2 plume mask	380–400 K; 400–420 K; 420–440 K	24 hours	1.5	reference run
SO_2-based_Modified	satellite SO_2 plume mask	380–400 K; 400–420 K; 420–440 K	6 hours	3.5	main run; best agreement with observations
rectangle-based_Modified	(163°E–170°W, 49°N–62°N)	380–400 K; 400–420 K; 420–440 K	6 hours	3.5	sensitivity to injection-region definition

Table 2 (Table 2 of the revised manuscript). Summary of tracer simulations using different injection-region definitions (satellite SO_2 plume mask and a rectangular domain), including injection layer, mixing interval, and critical Lyapunov exponent (λc).

Page 12, lines 250-252: any idea why is the tracer distribution in the control run more fragmented?

The more fragmented tracer distribution in the control run mainly reflects the different mixing settings: in CLaMS, parameterized mixing is triggered when the integrated deformation exceeds γc=λc Δt, and γc is larger in the control run than in the modified run, so mixing is activated less often and in larger, discrete steps, yielding a patchier and more fragmented tracer field.

Page 13, line 259: the quantity "BSR – 1" is usually termed "aerosol BSR" (ABSR: Cirisan et al., 2014) or particle BSR (PBSR: Reinares Martínez et al., 2021), as it represents the ratio of aerosol-

to-molecular backscatter coefficient (since BSR = (βaer + βmol) / βmol). The term "enhancement" instead typically refers to elevated values over a given background. Please revise this definition.

We agree. We replaced "BSR enhancement" with the standard terminology and now refer to $BSR_{455}-1$ as the aerosol backscatter ratio ($ABSR_{455}$).

> *To quantify model–measurement agreement for each profile, we compute (i) the Pearson correlation coefficient r between the area-normalized tracer-fraction profile and the COBALD aerosol backscatter ratio at 455 nm ($ABSR_{455}$), defined as BSR_{455}-1 (e.g., Cirisan et al., 2014; Reinares Martínez et al., 2021), over the main plume layer (375–450 K; 375–475 K on 30 September, 28 October, and 24 November 2019).*

Page 13, line 268: I cannot see "extreme" values in Fig. 7. Please quantify.

We replaced it with a quantitative statement and now report the peak-altitude mismatch in the control simulation.

> *Although the control simulation (orange lines) occasionally aligns with observed peak altitudes, its overall agreement is weaker and ~~it even generates unexplained extreme values on 12 August and 24 November.~~ it shows large peak-altitude biases on 12 August and 24 November 2019 ($|\Delta\theta\_peak| \approx 25$ K and ≈ 36 K, respectively; Fig. 7).*

Page 13, figure 6: it would be nice to show the tracer distribution over the Boulder site, since these simulations are also compared with the POPS data. Consider expanding the X-axis of Fig. 6 to include the entire Northern Hemisphere, or adding an extra figure focused on the American continent.

We thank the referee for this suggestion. As shown in our response to the major comments, we now include the tracer distribution over the Boulder site in the revised manuscript (Fig. 8).

Page 16, lines 304-309: I struggle to follow this paragraph and the argument of the "4-5 km additional lofting". If I understand correctly, this refers to the difference between the peak injection height of ~11 km derived from Cai et al. (2022), and the injection level inferred from the tracer simulation that best matches the observed profiles (400-420 K, i.e. 15-16.5 km). But how to be sure that this is due to radiative heating and not some other artifact/discrepancy between the different techniques? And what would be the time frame and the associated heating rates of the lofting? Since the plume has a large vertical extent (5-15 km), I would speculate that heating rates have a complex altitude dependency, so just comparing the peak height is not necessarily a good assumption. Unless this argument can be supported by a more detailed analysis, I think it should be presented as a speculation rather than a finding of the paper.

We agree and revised the paragraph accordingly. We now present radiative self-heating as a possible explanation for the vertical offset between the literature injection peak (~11 km) and the tracer level matching the observations (~15–16.5 km), but we clarify that injection-height estimates differ across techniques and are uncertain, so this comparison only provides a speculative interpretation rather than evidence for radiative lofting. Accordingly, we also removed the corresponding wording from the abstract and the conclusion, where it had been phrased as a finding.

> ~~*Based on the Lagrangian reconstruction by Cai et al. (2022), approximately 1.5 Tg of SO2 was initially injected between about 5 and 15 km, with a peak at around 11 km. If we use this central*~~

~~*injection altitude to estimate plume self-heating effects, the cloud may then be lofted by an additional ~ 4–5 km. Therefore, the height at which our SO2-based tracers best match observations will exceed the true injection center altitude. However, different observational techniques often report varying estimates of the injection center altitude (e.g., Kloss et al., 2021; Cai et al., 2022; Vernier et al., 2024). This also leads to uncertainty in our assessment of radiative lofting heights in this study.*~~ *Cai et al. (2022) estimate a peak injection altitude near 11 km, whereas the tracer level that best matches the observed profiles corresponds to ~15–16.5 km (400–420 K). We hypothesize that this ~4–5 km offset could be consistent with radiative self-heating lofting of the SO_2-rich cloud, but it may also reflect uncertainties in injection-height estimates across techniques and model-related discrepancies (e.g., Kloss et al., 2021; Cai et al., 2022; Vernier et al., 2024).*

Page 18, lines 333-334: what are exactly the "small-scale" and "often subgrid-scale" mixing processes?

We have removed this sentence as follow.

Our simulations show that ~~*small-scale atmospheric mixing processes are critical for dispersing the volcanic plume along its pathway through the stratosphere, and that representing these (often subgrid-scale) processes in models is essential for reliable simulations of volcanic plume transport.*~~ *atmospheric mixing processes, as parameterized in CLaMS in relation to flow deformation, are critical for dispersing the volcanic plume during stratospheric transport.*

Page 18, lines 336-337: I would appreciate a few statements on the overall relevance and fate of the volcanic aerosols entrained in the ASMA vortex, beyond the fact that this "may play an important role" in dispersing the aerosols. Is there a special relevance of this filament compared to the rest of the volcanic plume, e.g., in terms of more efficient aerosol transport to the stratosphere via the ASMA dynamics (see Vogel et al., 2019), higher cold-point tropopause, etc.? If so, this would be interesting to discuss.

We added a brief discussion clarifying that the plume fraction entrained into the ASMA is small. Based on established ASMA dynamics (e.g., Vogel et al., 2019), confinement and summertime diabatic uplift within the ASMA can potentially transport this fraction to higher altitudes in the UTLS, followed by redistribution and dilution. We revised the conclusions accordingly.

~~*In particular, our findings show that the ASMA may play an important role in dispersing aerosols from mid-latitude volcanic injections throughout the global stratosphere.*~~ *Our results suggest that a small fraction of the Raikoke plume becomes entrained into the ASMA. Within the ASMA, confinement and summertime diabatic uplift can potentially transport these plume fractions to higher altitudes (e.g., Vogel et al., 2019).*

Technical comments

Page 2, line 40: the full name of the volcano is "Hunga Tonga" or "Hunga Tonga-Hunga Ha'apai".

We revised the text to use the full volcano name ("Hunga Tonga") at first mention.

Page 3, line 72: replace "Section 7" with "Appendix A".

We have corrected “Section 7” to “Appendix A” .

Page 3, line 75: delete coordinates of Lhasa (already given).

Done

Page 5, line 105: the quantity e_{ice} (saturation vapor pressure over ice) is not defined.

where e_{ice} (T) is the saturation vapor pressure over ice at temperature T (in Kelvin), T_{mirror} is the measured frost-point temperature, and $T_{environment}$ is the ambient air temperature.

Page 5, line 114: delete "13 October" (not relevant).

We have removed the specific launch date (“13 October”)

TROPOMI, the satellite instrument aboard ESA’s sun-synchronous Sentinel-5P platform launched ~~on 13 October 2017~~ in 2017, is a hyperspectral imaging spectrometer …

Page 6, line 128: "an isentropic coordinate aligning layers" check grammar.

In the vertical direction it employs an isentropic coordinate ~~aligning layers with constant potential temperature θ,~~ that aligns the model layers with surfaces of constant potential temperature (θ) making the model well suited for stratospheric processes.

Page 11, line 211: add "of the plume" after "potential temperature".

~~Furthermore, the potential temperature measured by the satellite during its ASMA transit also closely matched the altitudes of enhanced BSR455 (Gorkavyi et al., 2021; Khaykin et al., 2022).~~ The potential temperature of the VVP during its transit through the ASMA inferred from satellite detections also closely matches the altitudes of enhanced BSR_{455} (Gorkavyi et al., 2021; Khaykin et al., 2022).

Page 12, lines 248-229: use the abbreviation of the Lyapunov exponent (λc) defined in Section 2.3.

Done

References (note: only papers not cited in the original manuscript are listed)

Brunamonti, S., Martucci, G., Romanens, G., Poltera, Y., Wienhold, F. G., Hervo, M., Haefele, A., and Navas-Guzmán, F.: Validation of aerosol backscatter profiles from Raman lidar and ceilometer using balloon-borne measurements, Atmos. Chem. Phys., 21, 2267–2285, https://doi.org/10.5194/acp-21-2267-2021, 2021.

Fahey, D. W., Gao, R.-S., Möhler, O., Saathoff, H., Schiller, C., Ebert, V., Krämer, M., Peter, T., Amarouche, N., Avallone, L. M., Bauer, R., Bozóki, Z., Christensen, L. E., Davis, S. M., Durry, G., Dyroff, C., Herman, R. L.,

Hunsmann, S., Khaykin, S. M., Mackrodt, P., Meyer, J., Smith, J. B., Spelten, N., Troy, R. F., Vömel, H., Wagner, S., and Wienhold, F. G.: The AquaVIT-1 intercomparison of atmospheric water vapor measurement techniques, Atmos. Meas. Tech., 7, 3177–3213, https://doi.org/10.5194/amt-7-3177-2014, 2014.

Park, M., Randel, W. J., Gettelman, A., Massie, S. T., and Jiang, J. H.: Transport above the Asian summer monsoon anticyclone inferred from Aura Microwave Limb Sounder tracers, J. Geophys. Res., 112, D16309, https://doi.org/10.1029/2006JD008294, 2007.

Poltera, Y., Luo, B., Wienhold, F. G., and Peter, T.: The “Golden Points” and nonequilibrium correction of high-accuracy frost point hygrometers, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2025-2003, 2025.

Randel, W. J., Park, M., Emmons, L., Kinnison, D., Bernath, P., Walker, K. A., Boone, C., and Pumphrey, H.: Asian Monsoon Transport of Pollution to the Stratosphere, Science, 328, 611– 613, https://doi.org/10.1126/science.1182274, 2010.

Vogel, B., Müller, R., Günther, G., Spang, R., Hanumanthu, S., Li, D., Riese, M., and Stiller, G. P.: Lagrangian simulations of the transport of young air masses to the top of the Asian monsoon anticyclone and into the tropical pipe, Atmos. Chem. Phys., 19, 6007–6034, https://doi.org/10.5194/acp-19-6007-2019, 2019.