

Response to the review by RC1:

First and foremost we would thank both reviewers for the constructive criticism. We have carefully reviewed the comments and would include the following changes to the original manuscript:

Regarding the major comments (1 - 7):

1. Reviewer: *Some important details of the datasets in this study are left out. For example, what is the horizontal and vertical resolution of the GLORIA retrievals? There should also be at least a short section to describe CLaMS model output and ECMWF ERA5 data. In Figure 3, do you initialize your trajectories from the same vertical level (e.g., flight level, one or multiple levels of the GLORIA retrievals) inside the first hexagon?*

Authors: We have added estimates for the horizontal and vertical resolutions of GLORIA. We say estimates, since the retrieval grid is irregular and thus no unified resolutions can be given. It should agree with the spirit of the comment. We have specified the resolution of the ECMWF data and added a section explaining the CLaMS model in more detail. We have clarified that the trajectories in Fig. 3 are initialized at the positions of retrieval grid, and thus at different levels.

Relevant text: 134-138

(+)The resolution of these 3-D tomographic retrievals generally depends on the retrieval process itself. The retrieval grid used in this study is irregular and chosen such that as many tangent points of the observations as feasible are contained. The vertical resolution of either retrieval is 250\,m, the horizontal resolution lies between 28.6\,km zonally, 22.24\,km meridionally or 38.28\,km for the nearest diagonal point.

Relevant text: 140-158

(+)2.2 Lagrangian simulations and surface-origin tracer
CLaMS (e.g. McKenna et al., 2002; McKenna et al., 2002) is a chemistry transport model that includes irreversible mixing and can resolve fine-scale tracer structures and gradients, particularly at the tropopause or in the vicinity of the ASMA (e.g. Ploeger et al., 2017b; Vogel et al., 2025). While originally designed to simulate the stratosphere, CLaMS was later extended to the troposphere (Konopka et al. (2010a); Konopka and Pan (2012); Pommrich et al. (2014a)and references therein). Results of a CLaMS simulation driven by ECMWF forecasts are used as a priori values for the GLORIA retrieval, whereas surface-origin tracers from a global CLaMS simulation driven by ECMWF ERA5 reanalysis (Hersbach et al., 2020) are used to diagnose the origin of air masses measured by the GLORIA instrument. Here, ERA5 is used in a downscaled version with a horizontal resolution of $1^\circ \times 1^\circ$ and a 6-hourly temporal resolution (similar to Ploeger et al. (2021); Vogel et al. (2024)). This simulation was started on 1 May 2023 and run over the course of the 2023 ASM season (for more details, see Vogel et al. (2025)). Surface-origin tracers are initialized every day in the model boundary layer ($\approx 2\text{--}3$ km above Earth's surface, considering orography) and are subsequently transported into the free atmosphere over the course of the simulation. The percentage of a surface-origin tracer in an arbitrary air parcel indicates the extent to which the considered air parcel originates from the respective region.

The following surface-origin tracers in the ASM region (Fig. 1) are considered in our study: “Northern Indian Subcontinent” (NIN), “Indian Subcontinent” (IND), “Tibetan Plateau” (TIB), “Eastern China” (ECH), “South East Asia” (SEA), “North-West Pacific” (NWP), and “Tropical Western Pacific” (TWP). Furthermore, the sum of the following surface-origin tracers, referred to as the “South Asia” tracer, is used as a proxy for air contributing to the ASMA: “ECH” + “TIB” + “NIN” + “IND” + “BoB” + “NIO” (for more details, see Vogel et al. (2025)).

(-) Lines 151-160

Relevant text: Figure 3 (Caption)

Figure 3. Self matching experiment during flights PH13 and PH14. The first hexagon is marked in red, the second hexagon is marked in green. Flight paths are shown in black lines, flight orientation is indicated by the arrows. CLaMS forward trajectories are shown as gray lines. They begin inside the first hexagon at 01:35 UTC and end at 01:50 UTC on the following day. The trajectories were started at every point on the retrieval grid and were driven by ERA5 (1°x1°) reanalysis. Potential temperature of end points are shown in color-code. Trajectories were calculated for every point of the retrieval grid

2. Reviewer: *I cannot fully evaluate the result because some key details of the datasets are missing. Even if the result is robust at all, I would like the authors to reconsider to bring your findings to a bigger picture. The authors analyzed a case study only. And certainly the GLORIA is not widely used yet. So how could your findings benefit to those who study Asian monsoon?*

Authors: Regarding the robustness of the result: We have included a section in the appendix investigating the robustness of the classification by evaluating the variability of parameters over multiple initializations. This should outline the robustness of the method. Regarding a "bigger picture": The double-flight experiment was conducted to demonstrate the capabilities of the (at the time) proposed Earth Explorer 11 candidate CAIRT, which at the time of the submission was not selected for the mission. However, the similar instrument STRIVE was approved and will be available in a few years. The methodology presented here therefore outlines how similar measurements will be possible in the near future with observations by STRIVE. Any other measurement of multiple trace gas species would be equally suited for the classification approach, but it is mostly limited to measurements since simulated data likely would not have to rely on it. These statements have been included in the manuscript.

Relevant text: 621-632

(+) B5 Robustness of BGMM parameters

An instance of a trained (converged) BGMM generally converges towards the same state. Small variability may be introduced by different initializations or ambiguity of the clusters contained with the data. To illustrate the robustness of the BGMM shown here an ensemble

of 100 instances was trained and the variability of the BGMM parameters such as means (cluster centers) and covariances (cluster shapes) were determined. Fig. B7 summarizes the results. The means of each cluster are indeed rather robust, with the largest variability found for H₂O and the continental tropospheric class (blue) in general. Since water vapor has the largest intrinsic variability this result is partially expected. The mixing weights (panel b) vary little, indicating indeed well defined clusters with sufficient local point density and global variability. The covariances also are subject to very little variability, with the most uncertainty found for the correlation between H₂O and HNO₃ in the blended class and the variance in H₂O in the maritime tropospheric class such that the cluster shape along these axes (and subsequently along the contributions of those axes to the principal components) may vary. The covariances shown here were calculated from the untransformed clusters (in 5-D) based on clusters found in the reduced 3-D space.

Relevant text: 547-552

(+) This self-matching experiment was conducted to demonstrate the measurement capabilities of the Earth Explorer 11 candidate CAIRT ESA (2023, 2025), which at the time of writing of this study was not selected for the mission. The wide range of trace gas species observable by CAIRT and its capability to perform 3-D resolved measurements would have been ideally suited to study mixing processes. A conceptionally similar instrument STRIVE is in development and will be deployed no earlier than 2030 NASA (2026). It will be capable of performing similar experiments and will be able to provide additional insights into the mixing processes in the UTLS region and the ASM.

3. Reviewer: *I am not familiar with the Gaussian Mixture Model that you propose in this study. You spent several pages to explain your whole data pipeline. But if you want to convince readers to use your method, you have to clearly show the limitation of the (simpler) traditional method first. For example, first plot concentration of some species and compute correlation using the traditional method. Then state its disadvantages. Overall, for your section 2.4, it will be much easier to provide some numbers/examples for readers to understand your method. For example, in your PCA subsection, does M represent the three datasets and does N represent the total number of observation made in each dataset?*

Authors: We have added a short classification using the traditional approach using water vapor and ozone to the appendix. This should illustrate the shortcomings of the conventional approach. We had already included the specific numbers for our use case in brackets. We have revised all symbols used in the sections in question to make sure they are consistent. In the (old) notation M denotes the number of observations (here: 4117 for hexagon 1) and N the number of features (here: 5 trace gas species).

Relevant text: 602-620

(+) B4 Comparison to conventional classification

The comparison between the BGMM classification and conventional methods is illustrated in Fig. B6. The tracers H₂O and O₃ have been chosen since most conventional methods use them. Any other combination of tracers will yield a more chaotic correlation. The BGMM classification is shown in Fig. B6a in color-code. The first conventional method used the tropopause similarly to Ungermann et al. (2016) to determine the tropospheric and stratospheric regimes. For either regime a linear regression is performed to determine the air types. Air parcels are included into either tropospheric/stratospheric regime if the regression suggests at least 66% prediction confidence. The classification fails to provide complete classification for either tropospheric/stratospheric air, indicated by the green regimes above the stratospheric band and to the sides of the tropospheric band. Mixed air (in the form of mixing lines between the tropospheric/stratospheric band) is underrepresented compared to the other classifications. This classification would allow a total of four regimes to be defined (in either quadrant of the bands), but the section corresponding to the ASM air in Fig. B6a is generally not considered in this approach. The second conventional approach uses percentile thresholds similarly to Cohen et al. (2023) to identify tropospheric/stratospheric air. Linear regressions through either regime are performed and the air parcels classified accordingly. This classification is most similar to the BGMM classification, but only yields three significant regions. The third conventional approach uses commonly used fixed thresholds similarly to Ma et al. (2022). We use a threshold of 20 ppmV H₂O for the tropospheric regime and a threshold of 0.2 ppmV O₃ and perform linear regressions. This classification yields a stratosphere most similar to Fig. B6a, and a reasonable definition of the troposphere. However, it is only capable of identifying three distinct regimes. Neither conventional method can reasonably perform a classification into more than three classes. Even between different conventional approaches significant differences persist.

4. Reviewer: *In Figure 5, what's your value for the dynamical tropopause? If it is not the usual 2 pvu that I see in the literature, please state your value and probably explain it a little bit, instead of referring to Kunz et al. (2015). You refer to it as tropopause folding so I am curious what I can see using normal lapse rate or cold point tropopause.*

Authors: As stated in the paper the dynamical tropopause is calculated based on the combination of potential temperature and PV based on the climatologies by Kunz et al. 2015. We have added such an explanation for a better understanding. Regarding normal lapse rate/cold point tropopause: as shown in Fig. B2 the lapse rate tropopause lies mostly outside of the retrieval volume and does not reflect the tracer fields well at all.

Relevant text: Figure 5, caption:

(+). The dynamical tropopause adapted from Kunz et al. (2015), which determines the PV-value depending on the potential temperature based on climatological data, is indicated as black line in all panels.

5. Reviewer: *Generally, all 3D plots (e.g., Figures 7, 10, and 12) do not offer much more information than multiple 2D plots. Consider to re-plot them.*

Authors: We have evaluated alternative illustrations and find that none are more beneficial than the currently used ones. We see this as a technical limitation of visualizing 3-D structures in 2-D. We have therefore created interactive 3-D plots in html-format which we provide as supplements for better visualization. We have attached these 3-D plots as a supplement and will provide a suitable host for them for the final publication.

Relevant text: None

6. Reviewer: *Could you provide any uncertainty statistics for the results from the Gaussian Mixture Model?*

Authors: As per request we have performed a small robustness analysis for every parameter of the GMM, which we included in the appendix (see comment above).

Relevant text: 621-632

7. Reviewer: *There are some lightning events shown from GOES-18. But are they enough to quantitatively explain the HNO₃ concentration difference?*

Authors: The amount by which lightning events can increase HNO₃ mixing ratios varies in the literature, but we find a value of around 300 pptv in Tie et. al. 2001, which lies in the correct order of magnitude. We included this statement in the manuscript.

Relevant text: 492-495

(+) . Tie et al. (2001) state that lightning events can lead to increases of up to 300 pptV in HNO₃ compared to the background, which lies in the order of magnitude of observed increase.

8. The comment states that there are multiple minor mistakes with the script, of which many were also noted by RC2. We have addressed these points. Typos have been corrected, including hyphens. Incorrect citation commands (resulting in missing brackets) have been corrected.

Regarding some more specific (minor) comments:

Reviewer: *I don't understand this sentence. The ASM contributes to the ATAL?*

Authors: We have reviewed the sentence and added sources indicating the significance of the ASM for the ATAL.

Relevant text: 51-53

Reviewer: *L122: For the Curtis-Godson approximation, you should also cite the Godson 1953 QJRMS paper.*

Authors: We have included the reference at the relevant place.

Relevant text: 121

Reviewer: *Figure 2: You should state the time before UTC.*

Authors: The measurement flights take between 8 and 10 hours, the hexagons themselves take up around 2 hours. We would abstain from including an hour, as we fear it would lead to confusion, but recognize the critique. We did include the start and end times of the outflow trajectories, which subsequently indicates the retrieval time for either hexagon.

Relevant text: Figure 3 (Caption)

First and foremost we would thank both reviewers for the constructive criticism. We have carefully reviewed the comments and would include the following changes to the original manuscript:

RC2 does not differentiate between minor and major comments, but implied that their comments are of minor nature. We will address them here as such:

1. Reviewer: *Abstract line 3 - "... classification methods based on a larger number of tracers ..."*
Larger than what?
Authors: Changed to „many tracers“ as clearer language
Relevant text: 3
2. Reviewer: *Lines 85, 86 - Brackets missing to enclose the references*
Authors: Included the missing brackets.
Relevant text: 85-86
3. Reviewer: *Section 2.2 -- It would be good to describe Clams model in more detail. Further, the Clams model is focused on stratospheric transport modelling - How good is the model fidelity while considering tropospheric air parcels and transport below the tropopause where many processes can play different roles?*
Authors: We have included a dedicated section for the CLaMS simulations. We have included sources which investigate the fidelity of CLaMS as well.
Relevant text: 140-157

(+) 2.2 Lagrangian simulations and surface-origin tracer

CLaMS (e.g. McKenna et al., 2002; McKenna et al., 2002) is a chemistry transport model that includes irreversible mixing and can resolve fine-scale tracer structures and gradients, particularly at the tropopause or in the vicinity of the ASMA (e.g. Ploeger et al., 2017b; Vogel et al., 2025). While originally designed to simulate the stratosphere, CLaMS was later extended to the troposphere (Konopka et al. (2010a); Konopka and Pan (2012); Pommrich et al. (2014a) and references therein). Results of a CLaMS simulation driven by ECMWF forecasts are used as a priori values for the GLORIA retrieval, whereas surface-origin tracers from a global CLaMS simulation driven by ECMWF ERA5 reanalysis (Hersbach et al., 2020) are used to diagnose the origin of air masses measured by the GLORIA instrument. Here, ERA5 is used in a downscaled version with a horizontal resolution of $1^\circ \times 1^\circ$ and a 6-hourly temporal resolution (similar to Ploeger et al. (2021); Vogel et al. (2024)). This simulation was started on 1 May 2023 and run over the course of the 2023 ASM season (for more details, see Vogel et al. (2025)). Surface-origin tracers are initialized every day in the model boundary layer ($\approx 2\text{--}3$ km above Earth's surface, considering orography) and are subsequently transported into the free atmosphere over the course of the simulation. The percentage of a surface-origin tracer in an arbitrary air parcel indicates the extent to which the considered air parcel originates from the respective region.

The following surface-origin tracers in the ASM region (Fig. 1) are considered in our study: “Northern Indian Subcontinent” (NIN), “Indian Subcontinent” (IND), “Tibetan Plateau” (TIB), “Eastern China” (ECH), “South East Asia” (SEA), “North-West Pacific” (NWP), and “Tropical Western Pacific” (TWP). Furthermore, the sum of the following surface-origin

tracers, referred to as the “South Asia” tracer, is used as a proxy for air contributing to the ASMA: “ECH” + “TIB” + “NIN” + “IND” + “BoB” + “NIO” (for more details, see Vogel et al. (2025))

4. Reviewer: *What is a "filament of air" exactly? Could you be more precise when you first introduce this? The discussion jumps directly into "filaments of ASM air" without explaining them*

Authors: We have included a brief introduction to the wording. It should now be much clearer.

Relevant text: 66-67

(+) ... the edge of a coherent, elongated structure of air with enhanced content originating in the ASM region relative to its surroundings – in the following referred to as a filament of ASM air – associated...

5. Reviewer: *Line 154 - The first sentence should appear later. It is a reference to Fig 3 sandwiched between two detailed discussions on Fig 2.*

Authors: We have switched the sentence around to improve the structure.

Relevant text: (-) 168-169, (+) 177

6.

7. Reviewer: *7. Fig 3 caption says "trajectories were calculated using ERA5". Which Lagrangian particle trajectory was used? Please clarify (probably Clams)*

Authors: We have clarified that the trajectory is indeed by ClaMS.

Relevant text: Figure 3 (Caption)

8. Reviewer: *8. Line 170 "discussed in the following. " Following what? Last word of the sentence missing*

Authors: We have added the missing word.

Relevant text: 186

9. Reviewer: *9. Line 270 - 250 km vertical resolution. That's quite a poor resolution :)*

Authors: We have provided the correct unit to the resolution :)

Relevant text: 286

10. Reviewer: *10. Figure 6 caption should mention that the values are averaged over longitude within the hexagon (I am guessing that is the case)*

Authors: We have clarified that these values are not averaged over longitude, but rather are interpolated along a constant longitude of 160°W. They correspond to the green cross-

section in Fig.10

Relevant text: Figure 6 (caption)

11.Reviewer: 11. *Line 348 onwards - This is a very important paragraph. I think it should be emphasised more and the arguments should be strengthened why this particular class is clearly associated with ASMO air. I was not entirely convinced.*

Authors: We have rewritten the paragraph to better underline that this air type corresponds to the ASM with the available arguments. We hope this version is more convincing. We also included a source of a different paper investigating the different origins.

Relevant text: 363-372

(+/-) Asian Summer monsoon outflow class: the air associated with the orange label shows moderately low values in both tropospheric and stratospheric tracers. Its composition is clear distinct from either tropospheric or stratospheric air and therefore cannot be either. To preempt the succeeding analysis this class is localized precisely in those section of the hexagon where the filament was predicted during flight planning. The regions most aligned with this air type are the East China (ECH) region and the Tropical West Pacific (TWP) region, further suggesting a link to the ASM outflow. Additionally, comparing the surface-origin tracers to the localization of this class (see. Fig. B4) it shows that this class contains the highest mixing ratios of relevant surface-origin tracers. The strong signatures of both ECH and TWP indicate a mixing of these air masses even before the formation of the filament. The CLaMS simulations do indeed show this mixing of TWP air with the outer regions of the ASMA ((see Vogel et al., 2025)). The chemical composition is consistent with such an interpretation. We refer to this type of air as ASMO air (A).

12.Reviewer: 13. *Line 419 - "Elements with very few elements" – typo*

Authors: We have fixed the typo.

Relevant text: 419

Other changes include additional spelling corrections and minor changes in phrasing. Since Reviewer 1 requested better visualization of the 3-D plots we will provide interactive 3-D plots for better visualization as a supplement in the final version. We have attached these in this reply if you wish to inspect them.