Response to the Anonymous Referee #1.

We express our sincere gratitude to Anonymous Referee #1 for constructive remarks on the manuscript and the appreciation of our study.

I have three general, rather minor comments that I would like to make. First, even though the focus of this study is the Asian summer monsoon, it is important to remind us that the North American monsoon also plays role in the lower stratospheric water vapor budget. There are some parts of the text that needs to be clarified (see the first specific comment below).

We agree that the North American monsoon may be as important in terms the convective transport into the stratosphere as its Asian counterpart. The respective modifications (following specific remarks) have been implemented into the manuscript.

Second, I would like to see some statements about uncertainties in the trajectory model results. Can we trust 100% of the result? How sensitive is the result to the input variables or dynamical fields?

The following text has been added into Sect. 2.4.:

> **"It should be noted that the trajectory model integrates 1000 backward trajectories per data point along the flight track which are submitted to a random noise equivalent to a diffusion D=0.1 $m^2 s^{-1}$ as in Bucci et al. (2020). As such, the integration is a discretization of the adjoint equation of the advective diffusive equation, which is well posed for backward integration (Legras et al., 2005). Unlike single-trajectory Lagrangian calculations, this method does not generate spurious small-scale features as backward time increases, and can be shown to converge with time for a pure passive scalar. With that, the trajectories from each data point come from several, possibly, many sources and the results presented are a statistics over these 1000 trajectories.**
>
> **Obviously, the results can be affected by biases in the wind field, heating rates and the cloud height product used in this study. The ERA5 is presently considered as the most advanced reanalysis and it was shown to display very consistent transport properties of diabatic versus kinematic trajectories (Legras & Bucci, 2020), which are in excellent agreement with observations (Brunamonti et al., 2018; von Hobe et al., 2021). The main concern in the Asian monsoon region is that ERA5 displays high penetrative convection over the Tibetan Plateau, which might bias the heating rates in the upper TTL over this region (SPARC S-RIP report, 2021). As the trajectories involved in this studied are mostly outside the Plateau, we do not expect any significant impact."**

Third, the figures in the manuscript and the supplement material are outstanding and yet complex. Some figures are probably too complex. I had hard time understanding Figure 1, in particular, as it contains so many variables and colors. It would be helpful to revisit the figures and improve if there is time.

The figures have been modified following the specific remarks as detailed below.

P1, L21 – 'the wettest region' can be replaced by 'one of the wettest regions'

P1, L30 – 'water vapour enhancements' can be replaced by 'water vapour measurements during the campaign'

P2, L40 – 'the primary contributor' could be replace by 'one of the primary contributors'

P3, L82 – 'in the other flights' could be replaced by 'the rest of the flights'

P4, L136-137 – Here, 'distribution' could be replaced by 'distributions' on both sentences.

All suggestions implemented.

P2, L57-60 – It would be helpful to add a sentence summarizing the findings, e.g., is the degree of convective impact different? Or do the source regions differ in different studies?

The following sentence has been added: "In general, there is no consensus regarding the primary convective source regions, nor regarding the net convective effect of deep convection on the CLS water vapour, which points out the complexity of physical processes in the AMA system."

P3, L86-87 –Do Stroh et al. (2021) include the description of the instruments as well?

Stroh et al. (in prep. 2022) include a brief description of the instruments summarized in a table.

P3, L107 – Does 'under UTLS conditions' mean low temperature and low humidity?

Yes, the respective sentence has been modified: "…under UTLS conditions, i.e. low temperature and humidity environment."

P4, L126 – Here, 'They' refers to Singer et al. (2021)?

Yes, the sentence has been modified respectively.

P5, L153 – The meaning of 'evenly distributed' is unclear.

The phrase "evenly distributed" has been removed.

Also, there is a newer version of MLS H2O (v5), which became available more than a year ago. It is also known that MLS H2O (v4.2) has drift issues in the stratosphere.

The satellite data analysis for this study was performed before the MLS v5 retrieval has been issued. We are aware of the drift issues in v4.2 data version however this is of little relevance for this study since we use MLS H2O data mostly to describe the regional and subseasonal variability in the AMA region. A quick comparison of the results presented in Fig. 1 using v4.2 and v5 retrievals did not show any significant differences.

P5, section 2.3 – I think this section provides useful information. It would be helpful to add a sentence explaining the purpose of this section here.

The following sentence was added in the beginning of the section: "In this section, we define the key terms regarding the vertical structure o AMA and physical processes therein."

P5, L178 – It is unclear why the thermal approach to TTL definition is only suitable for this study. Is this related to the fact that water vapor is sensitive to vertical structure of temperature?

The thermal approach to TTL definition is suitable for this study because it relies on the local measurements as those exploited in this study. The text has been reformulated to avoid confusion.

P5, L185-186 – It would be useful to include references for the hydration vs. dehydration processes in the TTL.

References to Jensen et al., 2007 and Schoeberl et al., 2018 have been added here.

P6, L197 – Are there any references for HIMAWARI-8 could be cited here?

The following reference has been added:

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., Kumagai, Y., Miyakawa, T., Murata, H., Ohno, T., et al.: An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites, Journal of the Meteorological Society of Japan. Ser. II, 94, 151–183, https://doi.org/https://doi.org/10.2151/jmsj.2016-009, 2016

P6, L207 – Would 'a specific version of product based on the version 2018.1' be the same as 'version 2018.1'?

Unnecessary or redundant information has been removed from the text

P6, L211 – Is '100 hPa' an arbitrary threshold or based on a statistical analysis?

The 100 hPa threshold was set mostly arbitrarily. The statistical analysis shows a rapid decrease of the number of convective hits above 100 hPa. The lower-pressure threshold does not allow identifying convective sources for the lower-altitude hydration features (e.g. A2), whereas the higher-pressure threshold adds ambiguity to the source identification for higher-altitude features (e.g. B7).

P6, L219 – How is the 2017 Asian monsoon season characterized as a stable anticyclone? Compared to 10-year climatology or compared to previous three years? Some statistics might be helpful here.

The text has been modified: "The 2017 Asian monsoon season was not marked by an anomalous dynamical behavior (Manney et al., 2021), however the campaign occurred during a break - active transition. The strongest convective activity took place above the Southern slopes of Himalayas and the Tibetan plateau during late July and early August as can be inferred from…"

Fig. 1a - It is not easy to tell from Fig. 1a. In Fig. 1a, the water vapor contours can be smoothed (1-2-1 smoothing) or one can use bigger grid boxes to show smooth contours. Also, what do black colors mean in Fig. 1a? I think the wind vectors can be improved here as well (likewise in Fig. 5). For instance, one can put wind vectors every 2.5 degree latitude instead of 5.

Fig. 1b – I think the AMA boundary looks rather too broad here. Adding more contours or choose a smaller threshold of Montgomery stream function might work better

All remarks on Fig. 1 have been implemented

P7, L241 – Is Brunamonti et al. (2018) also based on summer of 2017?

Yes, we have included the respective mention in the text.

P7, L247 – I would recommend using a quantitative adjective than 'striking' here, e.g., 'large' variability.

Replaced by "remarkable".

P7, L261 – Do 'those' refer to the high RHi values?

'Those' refer to the subsaturated cloud occurrences. The text has been rephrased: "Such occurrences are mainly caused by…"

P7, L263 – It would be helpful to add an explanation about 'Lagrangian temperature history' here.

The sentence has been modified: "…depends on the air parcel's (Lagrangian) temperature history".

P7, L274- Is '14%' higher or lower than any statistics or expectations?

The following sentence has been added: "This is consistent with a comprehensive analysis of airborne data from various campaigns by Kraemer et al., 2020, who pointed out a significantly larger amounts of IWC in subsaturated ice crystals above the CPT in AMA compared to that in the surrounding tropical regions, thereby upholding the importance of AMA as the source of LS water."

Comments for Fig. 4 It would be useful to add approximate altitude for Figs. 4a and 4b. Also, in Fig. 4b, high delta D exists as high as 420K potential temperature surface.

Approximate altitude added. The high deltaD values are not totally surprising at this level given the observed hydrated features up to 415 K and their heating rate of more than 1 K/day.

P9, Section 4.2 – I am wondering is there a way to quantify the uncertainty in the derived convective age. Is it sensitive to the type of cloud top data and also meteorological fields?

The following text has been added: "We note that while the 1σ-error of the age estimates is generally less than an hour, the attribution of convective sources largely depends on the cloud top data. In particular, the improved v2018.1 trajectory product coupled with NWF SAF geostationary data analysis provided a qualitatively better correlation between the distribution of convective hits and wet-and-heavy parcels as compared with the product used by Bucci et al. 2020.."

P10, L340 – Does 'mixing ratio enhancement' mean the actual water vapor mixing ratios or only the increased amount of water vapor?

It means water vapour mixing ratio enhancement, as specified now in the respective sentence

P10, L350 – Fig. 6a,b could be replaced by Figs. 6a and 6b

Done.

P10, L350 – Does 'at this level' mean local CPT?

Yes, sentence modified.

P10, L359 – Here, 'such an amount' could be replaced by 'such high amount'. I am also wondering what is the mechanism that enables the convective plume preserve high water vapor for 5 days.

Replacement done. A convective plume that doesn't experience supersaturation can preserve enhanced vapour in the CLS for days and perhaps weeks, provided the weaker mixing at these levels compared to the upper-TTL.

P11, L373 – 'intersecting a large convective system' – It looks like the trajectories lie between the large convective system and the group of small cells to the west.

The locations of convective hits (i.e. where the traced air parcels hit the convective cloud) are indicated in Fig.7a as black-filled red circles. A clarification has been added to the text.

Comments for Fig. 7c I assume temperature means potential temperature in Fig. 7c. Also, it is hard to locate 140 or 160% RHi in this figure as explained in the text.

Figure 7 has been reworked to display the RHi evolution versus the y-axis.

P11, L389 – Does 'across' mean from below to above the CPT?

Yes, the text has been modified accordingly.

P11, L393 – It would be helpful to give the time marks for the presence of subvisible cirrus clouds in Fig. 8a.

There are six occurrences of cirrus clouds during this flight segment. They are tagged and color-coded along the potential temperature time series.

Comments for Fig. 9a Fonts for T1-B7 could be bigger.

Done.

P12, L435 – Instead of 'around the CPT', near or close to might be more accurate.

Changed to "near CPT level"

P13, L448 – Does 'there is typically no more than one case' apply to all the cases referred above?

The sentence has been modified: "…there is typically no more than one case of water vapour enhancement above the tropopause detected during a given field campaign."

P13, L454 – In this sentence, the meanings of 'over these regions' and 'in the summer monsoon anticyclones' are not clear. Are those referring to the North American monsoons?

The sentence has been modified: "..support the role of overshooting convection in maintaining the water vapour maximum in the North American and Asian monsoon anticyclones."

P13, L467 – 'around the tropopause' could be replaced by 'near the tropopause'.

P14, L483 – This sentence could be split into two -> introduction. However,…

All done.

P14, L505 – It would be helpful to add a sentence about what the authors think the future direction or need is in terms of field studies related to StratoClim.

The following paragraph has been added: "Further insights into the AMA gaseous/particular composition and dynamics will be provided by an upcoming airborne campaign within the Asian summer monsoon Chemical and Climate Impact Project (ACCLIP; https://www2.acom.ucar.edu/acclip), which will sample the Western Pacific mode of the monsoon and eastward eddy shedding using NASA WB-57 and NCAR GV aircrafts. The stratospheric impact of overshooting convection in the North American monsoon is a primary target of the Dynamics and Chemistry of the Summer Stratosphere (https://dcotss.org/) project, involving ER-2 high -altitude aircraft.