Reply to Reviewer 2

We thank the Reviewer for the careful reading and evaluation of the manuscript and the good comments which helped a lot to further improve the paper. In the following, we address all comments and questions raised (Reviewer's comments in italics). Text changes in the manuscript are highlighted in color (except minor wording changes).

Besides several specific comments, we see two common main concerns raised by both Reviewer's, regarding (i) the presentation quality and clarity of figures, and (ii) the discussion of specific climate model characteristics. A short overview of the related changes in the revised manuscript is:

(i) To enhance the **presentation quality and clarity** we modified all figures in the revised manuscript. In particular, we chose different color schemes, reduced the number of contours, changed to difference plots for certain quantities and improved the model-correlation visualization (a detailed list of figure changes is in the reply to the specific comment below). We are confident that these changes clearly improve the presentation of our results.

(ii) To relate our results to specific **model characteristics**, we included a new subfigure Fig. 3 (e, f). This figure shows the inter-model correlation for wind velocities averaged over the region of maximum signal, and includes information on horizontal and vertical resolution of the models by highlighting models with relatively high number of horizontal grid points with horizontal lines and models with high number of levels with vertical lines. Based on this simple model classification, no clear relation of resolution differences to the simulated water vapour bias is found. These findings are briefly discussed at the end of the discussion section and described in more detail in the Methods section 5.1.5.

A more detailed reply to all comments and description of the changes in the revised manuscript is given below.

General comments:

Ploeger et al. explore the connection between lower stratospheric water vapour anomalies and dynamical biases previously identified by Charlesworth et al. in CMIP6 models using a combination of CMIP6 model output, observations, and specifically designed simulations using the EMAC model run with different transport schemes. Their finding that Pacific UTLS water vapour anomalies impact regional circulation builds on the findings of Charlesworth et al. by exploring inn more detail local water vapour anomaly-circulation bias connections, particularly with relation to the Asian summer monsoon and transport from Asia into the Pacific. I found the manuscript to be well written and the analysis clear. I feel the paper fits the scope of ACP and explores and important topic. However, I would recommend the authors address the comments below before publication.

Thanks for this overall positive evaluation of the manuscript!

Care should be taken throughout the paper to make a better distinction between water vapour anomalies and biases. To me, the anomaly is the relative abundance of local water vapour with respect to the zonal mean, whereas a bias is specifically a comment about the local abundance of water vapour in the model vs observations. These quantities may be related to an extent, but it is possible to imagine a very dry model that has a large water vapour anomaly over the Pacific having a small bias with respect to observations. I feel throughout the paper these two terms are used interchangeably, and this should be addressed.

The use of the terms "anomaly" and "bias" in the paper should exactly be as the Reviewer here suggests. "Anomaly" is the relative enhancement of water vapour with respect to the zonal mean, while "bias" is the deviation of simulated water vapour from the observations. We added a clarifying sentence in the introduction and looked carefully through the entire manuscript again to ensure that this terminology is used consistently throughout the paper.

In their analysis the authors use different lengths of time, over different years, for the different datasets used in this study. The EMAC model data used for analysis covers 2000-2009, the CMIP6 model data covers 2000-2014, and the ERA5 data covers 2005-2015. Given that the authors speak about the variability of water vapour on decadal timescales in the opening sentences of the introduction, I wonder to what extent the features identified in the paper are dependent on the choice of these relatively short time periods. For the CMIP6 data it is conceivable that averaged over multiple ensemble members 15 years represents something of a robust climatology, but can the same be said of 10 years of ERA5 or EMAC data? If the authors had used a different time period (e.g., 2000-2009), or a longer time period (2000-2020), would the ERA5 data show the same results? Can the authors say anything about this that strengthens the arguments made in the paper?

We checked the robustness of the water vapour and circulation responses in the EMAC model experiment by plotting Figs. 1 and 2 also for the 10 years before (1990-1999), which still guarantees 10 years of model spin up after switching to the Lagrangian transport scheme (see Methods 5.1.2). Regarding ERA5, we checked the robustness of the climatology by switching to the period 1990-1999 and 1979–2020. The figure below in the response letter (Fig. 1) shows that all results are insensitive to the choice of the period (for ERA5 only 1990–1999 is shown, but the figure looks almost the same for 1979–2020). A short statement on this robustness is included at the end of the first paragraph in the Methods section 5.1.2.

Many of the figures are difficult to read and crowded with information. Figures 1 and 2 have a lot going on, and it is difficult to see the black contours and arrows on the very dark red and blue shading. In figure 3 it is very difficult to tell the values form the blue and red shading. I would recommend replotting many of the figures to improve readability, and where possible separating the figures into more panels.

Thanks for this very helpful comment, which was similarly brought up also by the other Reviewer! We agree that the clarity and readability of many figures should indeed be improved. For the revised manuscript we did significant efforts to improve the presentation quality. The changes to the figures are summarized below:

- We reduced the number of contours plotted by removing redundancy and focussing only on the most relevant variables. Therefore, PV contours are not shown in Fig. 1 a–d anymore, but only in Fig. 1 e–f. Here, we now show PV together with the zonal wind distribution at 100 hPa in the two EMAC model simulations to illustrate the circulation differences more clearly. Figure 2 is even more substantially reorganized, by showing the observation and Lagrangian EMAC–CLaMS model water vapor and PV fields together with the model difference between the two EMAC simulations (Fig. 2 c, f). The water vapor plots also show temperature contours (and only these), while the PV plots now show zonal wind contours (model differences in Fig. 2 c, f). With this reorganization of figures still the relevant variables are shown, but with much more clarity.
- The color scheme for anomalies and differences is changed to schemes with somewhat lighter maximum blue and red colors, so that dark contours are still readable. Also contour colors have been changed to have less different colors and still ensure optimal contrast (e.g. only black contours in Fig. 1, only white contours in Fig. 2, tropopause as grey lines in Figs. 2 and 3).
- Circulation differences between the two EMAC model simulations (control vs. modified Lagrangian) are now shown in difference plots in Fig. 2 c, f. These new subfigures show the temperature, zonal wind and PV differences as induced by the stratospheric water vapor differences with much more more clarity compared to the manuscript version before. In particular, the cooling effect as well as the strengthening of PV gradients of the enhanced lowermost stratospheric water vapor can be clearly seen now.
- Regions of relevant temperature gradient changes explaining the wind changes in Fig. 3 a, c are now highlighted with white hatching. Although not including quantitative information on the exact values anymore, the modified figure still shows qualitatively that the relation between vertical wind shear and horizontal temperature gradient changes is consistent with thermal wind balance.

We think that with these changes the presentation of our results has become much clearer.

It is clear from the CMIP6 multi-model mean and figures S3/S4 that the Pacific water vapour anomalies are common across CMIP6 models. However, not all models show large positive anomalies with respect to the zonal mean above the region of the Asian summer monsoon (most notably the CNRM model, some configurations of the CESM model, and the SAM0-UNICON model). Can anything be inferred about links between Pacific water vapour anomalies and regional circulation biases in these models that look quite different to the others? Going further, can anything be said about model structural differences (e.g., resolution, transport scheme) in driving the biases explored in this study, or is it the case that whatever the model resolution, global models are too susceptible to processes like numerical diffusion? I think I'd like to see some discussion exploring model differences in the manuscript.

We like the idea to investigate in more detail whether particular CMIP6 model differences can be related to the development of the moisture bias. However, as explained also in the response to the second comment of the other Reviewer (see also main change (ii), summarized in our general comment above) this question is not easy to answer with the available suite of simulations. As also formulated by the other Reviewer, one interesting question here is whether it is the horizontal or vertical resolution which is most critical for the moisture biases. The water vapor maps for single models in the supplement (Fig. S3) show that indeed the model with highest resolution (MPI-ESM1-2-HR) simulates a well confined monsoon moisture anomaly and only a weak Pacific moisture bias. However, the differences in water vapor bias to other models are not related in a simple manner to differences in vertical or horizontal resolution. For instance, models with either a comparable number of vertical levels (e.g., MRI-ESM2-0) or with a comparable number of latitude/longitude grid points (e.g., BCC-CSM2-MR) show stronger moisture biases than the model with highest resolution (MPI-ESM1-2-HR). Furthermore, as already said by the Reviewer, regarding vertical resolution it is likely the number of levels in the UTLS rather than the total number of levels, and information on that is not well documented. In general, as the transport schemes employed by the different models are quite different, the differences in the simulated water vapor distribution are not simply explaniable by differences in the number of vertical levels or horizontal grid points. To discuss these aspects in the revised manuscript, we included a new subfigure Fig. 3 (e–f) which shows the inter-model correlation between Pacific u- and v-wind with the Pacific water vapor index for the winds averaged over the UTLS region of most significant correlation (see figure caption and Methods). In addition we selected from the CMIP6 models a few models with relatively high horizontal and/or vertical resolution (indicated by the horizontal/vertical lines in the scatter plot). The criteria for selecting these models are given in the Methods section 5.1.4. Clearly, neither differences in horizontal nor in vertical resolution explain the spread of simulated Pacific moisture bias. Hence, a more detailed analysis of model differences, best including sensitivity simulations with the same model and differing horizontal/vertical resolution would be needed to draw robust conclusions on this issue. We discuss these points in the revised manuscript in a short paragraph at the end of the discussion section and in the Methods section 5.1.5.

Specific comments:

L13-14: *"Variations in stratospheric water vapour have been shown to modify past global warming by up to 30% (Solomon et al., 2010)" This sentence is misleading – the Solomon et al. paper says that variations in lower stratospheric water vapour on a decadal timescale can have a significant radiative impact, but this sentence could be read as 30% of past climate change can be attributed to water vapour changes, which is not the case.*
Thanks for pointing this misleading formulation out (as similarly by the other Reviewer)! We changed the sentence to: "Decadal variations in stratospheric water vapour have been shown to modify the radiative budget on a decadal scale by up to 30%."

L19: *My understanding is that the 0.1-0.26 $Wm^{-2}K^{-1}$ range from Banerjee et al. is calculated using CMIP5 models. Is there an estimate from CMIP6 models? I know that a recent paper by Nowack et al. (2023) has tried to constrain this estimate and comes up with a slightly reduced range of 0.086-0.201 $Wm^{-2}K^{-1}$, which may be worth noting here.*
We agree that it would be more appropriate to refer also the most recent feedback estimates. We modified the sentence accordingly and include the Nowack et al. (2023) reference in the revised version.

L36-38: *"Analysis of the distribution of water vapour in the lower stratosphere simulated by a climate model has shown a substantial model bias in the Asian monsoon region" I was not sure here if the authors are talking about analysis of a single climate model, or making a statement that is true for all models.*
This sentence refers to the results of Wang et al. (2018) which are based on a single climate model (WACCM). The sentence is slightly modified to make clear that it refers to a "specific climate model".

L52: *Make it clear the anomalies are with respect to the zonal mean distribution of water vapour.*
We added "as compared to the zonal mean distribution" to make the terminology here more clear. (See also our reply to the first comment regarding the use of a consistent terminology).

L63-64: *Here, when saying the EMAC anomaly is more pronounced than the CMIP6 multi-model mean, I feel it is worth explicitly stating whether the EMAC model is an extreme case, or within the CMIP6 model spread, citing figure S3/S4 as evidence.*
We agree that it is worth noting that the EMAC moist bias lies within the CMIP6 spread, and added this information to the sentence: "..., but overall lies well within the spread of CMIP6 models (Supplement Fig. S3)."

L71-72: *When speaking about temperatures here, it would be great if the authors could show the temperature difference between EMAC and EMAC-CLAMS in a separate figure/panel (especially given general comment above about figure readability).*
We totally agree with the Reviewer that presenting difference plots would enhance clarity. Temperature differences between control and modified–Lagrangian EMAC simulations along with water vapour, PV and zonal wind differences are now shown in the new Fig. 2c. (See also our reply to the above general comment about presentation and figure quality).

L114-116: *"The CMIP6 inter-model correlation between local meridional v–wind velocity in the monsoon anticyclone and a water vapour index measuring the strength of the Pacific moisture anomaly (Methods) shows a significant anticorrelation eastward of 100$\circ$E (Fig. 3d)." I am unsure here what feature I should be looking at. Firstly, I think this should be*

figure 3c, not 3d. If so, is it the statistically significant blue shading between 100-50 hPa at around 150 degrees? Please provide more specific description in the text. Additionally, it is very hard to read from the figure the correlation in this region. Perhaps it could be given in the text?
There was indeed a typo in the referencing and the correct figure-reference here should be Fig. 3c. Thanks for noticing that! As the Reviewer assumes, this sentence refers to the negative correlation around 150 degrees E between about 100-50 hPa. We modified the description in the text to make this more clear, and also think that the improvements made to Fig. 3 (see reply to the above general comment about figure quality), in particular less color contours, help to enhamce readability and clarity. In addition, we included scatter plots for the correlation in the relevant region as the new Fig. 3 e–f, which shows the inter-model correlations in the relevant region even clearer.

L119-120: *"Above the tropopause, this correlation pattern closely resembles the pattern of differences in meridional flow (v–wind) 120 between the control and modified–Lagrangian simulations in the EMAC model experiment (Fig. 3a)" Does it? There is some overlap in the blue shaded region in figure 3a and 3c, but the features themselves have very different shapes, horizontal extents, and extensions into the troposphere.*
We agree that in particular the downward extension of signals into the troposphere shows clear differences between the EMAC model experiment and the CMIP6 ensemble. However, above the tropopause in the lower stratosphere the pattern of negative/positive differences agrees remarkably well with the pattern of negative/positive inter-model correlation, in particular given the fact that also other factors than stratospheric water vapour affect the UTLS circulation in CMIP6 models. We modified the text slightly to make even more clear that we refer here to the signal in the lower stratosphere.

Technical corrections:

L62: *remove 'also'*
Done.

L166: *oxydation should be oxidation*
Done.

Figure caption for S4: *I believe it should say (Fig. S3 continued)*
Done.

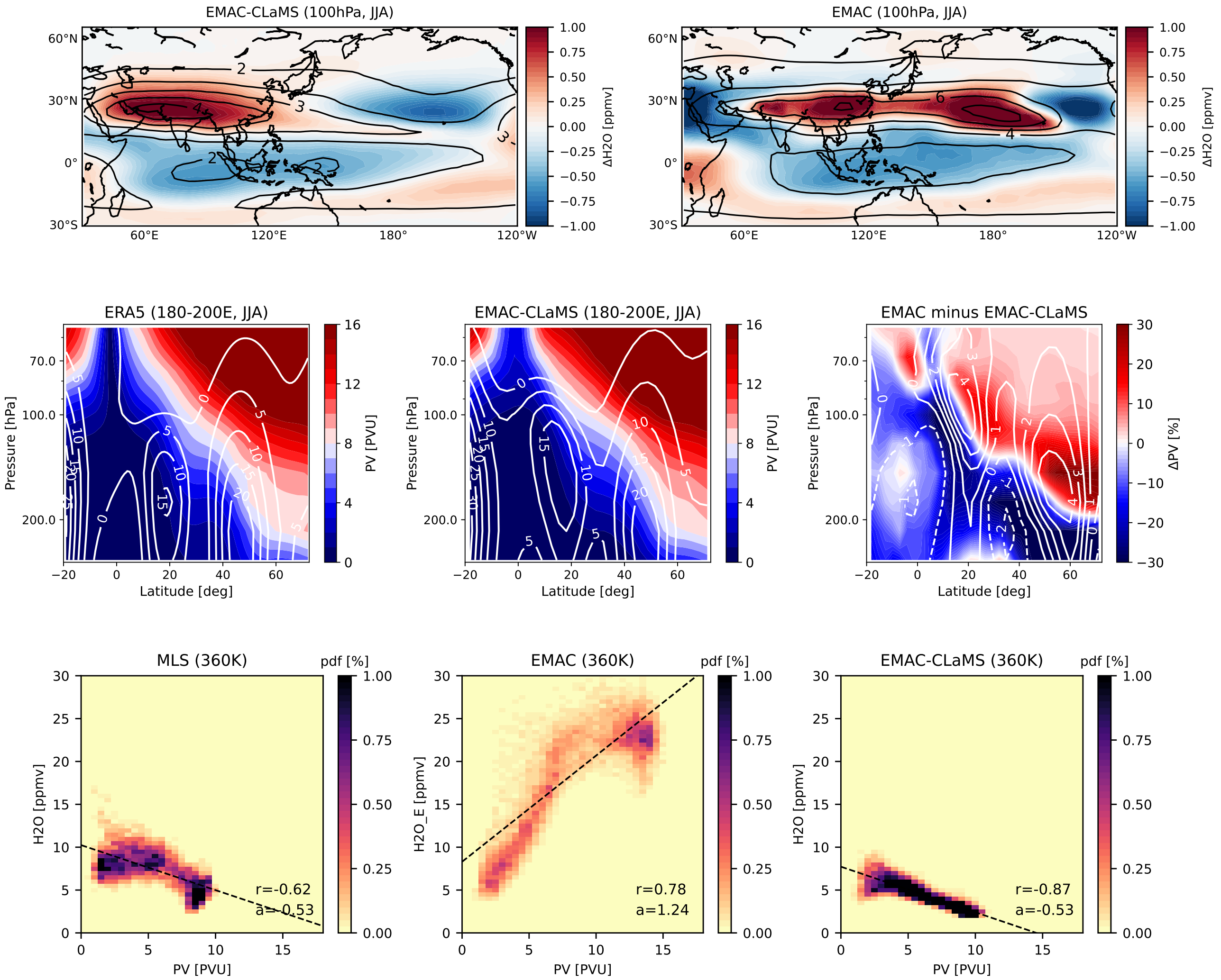

Figure 1: Sensitivity of manuscript figures 1 and 2 to the choice of climatological period, by showing all distributions for the 1990–1999 period (except g). (top) Water vapour zonal anomalies at 100 hPa in boreal summer (June–August) from modified–Lagrangian (left) EMAC–CLaMS and (right) EMAC. (middle) PV cross-sections in the Pacific region (180°-200°E) in boreal summer (June–August), (left) ERA5, (middle) modified–Lagrangian EMAC–CLaMS, and (right) the difference between control EMAC and modified EMAC–CLaMS simulations. (bottom) Correlation between water vapour and potential vorticity at 360 K potential temperature level in the Pacific region (15°-70°N, 140°-240°E) for (g) MLS satellite observations and ERA5 reanalysis PV (here for 2005–2015), (h) control EMAC, and (i) modified–Lagrangian EMAC–CLaMS. The Pearson correlation coefficient r and linear regression slope a are given in each figure and the linear regression fit is illustrated as black dashed line.