Reply to Reviewer 1

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 comment:
This is an interesting work finding an interpretation and a remedy to one of the serious known limitations of climate models. The manuscript is well written and I have only a limited number of minor comments and questions to be considered by the authors.
Thanks for this positive evaluation of the manuscript!

Specific comments:

Solomon et al. (2010) wrote that variations of lower stratospheric water vapour may account for up to 30% of the greenhouse gases modification of the radiative budget on a decadal scale. They do not write that they account for 30% of the total variation since 1850 as suggested by the sentence on lines 13-14 in the manuscript. Please correct to avoid such confusion.
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%."

It is quite clear that the excessive numerical diffusivity of the CMIP6 models should be related to their spatial resolution and their transport scheme, which display large differences among the ensemble. It is quite frustrating that they are here all put in the same bag without any attempt to draw a distinction. For instance, it would be very interesting to know whether it is the horizontal rather than the vertical resolution that matters as very different choices have been made among the ensemble. In terms of horizontal resolution, the T42 resolution f the EMAC used here put is at the lower end of the CMIP6 ensemble but its vertical resolution puts it at the upper end. This is perhaps an answer to the above question as fig.1 shows it does much worse than the CMIP6 mean although what probably matters is not the total number of levels but the number of those which span the UTLS, a parameter which is badly documented.
We totally agree with the Reviewer that it would be beneficial to know whether it is the vertical or horizontal model resolution that matters more in causing the water vapor bias, and that the resolution in the UTLS matters most. Unfortunately, this question is not easy to answer with the available suite of models and simulations. 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 quiet 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.

My main concern is about the figures and their readability. Black contours and labels on dark blue and red are hardly visible and readable. This is not good on the screen and it is terrible on a printed version. This needs to be improved. There is no reason to use a divergent color map to show water vapour in fig2 (a-c). The two first rows of fig.2 show redondantly PV and U on panels which are overcrowded. Please reorganise these two rows to show only 3 variables in each panel. In figs 1 and 2, some quantities (wind, temperature) would be better displayed as differences between EMAC-ClaMS and EMAC. Adding a grid would overload the figures but ticks can be put on the upper and right sides of the figures to improve readability. In figure 3, I do not think that the temperature gradient contours are very useful, and they have no labels and there is no indication of contour intervals in the caption. I would prefer to have some contours for the quantities displayed in color as it is almost impossible to read the values from the color map (or choose a better indexed color map).
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 more 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 very hard to appreciate from the first two rows of fig. 2 that the isentropic PV gradient is strengthened around the tropopopause and whether vapour contours are following or not the PV structure, although it is quite clear from the third row.
In the revised version, the strengthened PV gradient becomes much clearer from the difference plot in the new

Fig. 2f. The same holds for the relation between water vapor and PV changes around the tropopause. (See also the reply to the comment before regarding the presentation quality of figures).

Figure 3 shows that the zonal wind incease it shifted by 30°E with respect to the water vapour anomaly. So the temperature drop should also be shifted which means that the response to the water vapour is not the simple local process advocated in the manuscript but involves also transport and delay.
Thanks for pointing to that. A short related remark is added at the end of the results section Sect. 3: "Slight displacements between the zonal wind and water vapour changes indicate that likely the response is not entirely local (Fig. 1 and 3)."

It should be noted that the differences between EMAC and EMAC-CLaMs are much smaller than those between ERA5 and EMAC except perhaps for the PV in the lower stratosphere. In this respect it would be useful to see the curves for EMAC and EMAC-ClaMS in fig.4 (c-d) to appreciate the improvement in regard to the current dispersion and bias of the models.
As the model configuration for the EMAC simulations used here is not exactly the same as for the CMIP6 simulations, we refrain from including a too close comparison. Some comparison between the water vapour and circulation biases between these models can be seen in the new Fig. 3 e–f, but due to the above differences in model set-up we don't discuss these more thoroughly.

The manuscript is not totally clear about the effect on the monsoon circulation. It is indicated that the equatorward branch is broadened and strengthened on the eastern side but this is not a mechanism which by itself is able to modify the closed monsoon circulation as dicussed in section 3 since this modification is correlated to its internal PV budget.
To understand the complete non-local circulation response the application of a mechanistic model would be preferred. This is, however, beyond the scope of the present paper but could be a good, important focus of future work. Moreover, the main effect of the increased water vapour in the Pacific UTLS on the monsoon anticyclone is an eastward displacement associated with the enhanced longwave cooling over the Pacific. We are more careful with the related formulations in the revised manuscript (e.g. in the second paragraph of the discussion).

I am unsure the proper way to refer to ERA5 data is a link to Lawrence Livermore National Laboratory.
There was an error in the cited data link for ERA5, which is corrected in the revised version. Thanks for pointing this out!