Answer to Reviewer #1

We would like to start by thanking you for all the time and effort which you spent reviewing our paper. All your comments, suggestions, and questions were taken into account and all the necessary corrections were made in the revised manuscript. Furthermore, we address all your comments and suggestions below, point by point.

General comments:
The paper focuses on the factors affecting the interannual variability of stratospheric water vapor entry in the tropics in observations, CCMI and CMIP6 models. The authors contrast the use of a variety of techniques: multiple linear regression and 3 machine learning methods. Cold point temperatures are the main factor explaining the water vapor variability. They discuss the merits of the different techniques and the relative importance of the QBO and ENSO. They also find non-linear interactions to be important. The comprehensive models, whilst will suffering from a QBO that is not deep enough, have nonetheless improved. The paper is well written and provides an good description of machine learning techniques applied to a geophysical problem. The figures are also mostly clear.

Specific comments:

(1) Make it clear earlier during the introduction that you are looking at interannual variability and not the seasonal cycle.

We clarified this in the introduction.

(2) Some of the CCMI model have multiple ensembles. Do you average over all of them? If so, does this result in less variability and thus make it harder to compare to those runs with only 1 ensemble?

We include each ensemble member separately. We don't average the ensembles together before computing correlations, rather compute for each ensemble member separately. Now clarified.

(3) In the figures, would it be possible to have the models with a nudged QBO labeled in bold text? It would make identifying them easier.

We tried adding this information to the figure legend of figure 1 and 9, but the figures then looked strange. We added it to the caption instead.

(4) On line 4, page 6, you mean ERA5/ERA5.1 I think?

Yes, we corrected this.

(5) On page 6, line 11, "Note that the correlation of the BDC with the QBO is -0.66 (Figure 2), and hence including both in a single regression or ML model can lead to overfitting. " I disagree with this statement. Multicollinearity in your predictors causes a variety of problems but does not specifically cause overfitting. See page 283, Applied linear statistical models 5th edition by Neter et al. (2004). Your validation stage should show if overfitting is an issue.

Indeed, we rectified the sentence to say that multicollinearity can lead to erroneous model interpretation.

(6) Page 10, line 15, the non-linear predictors are interesting but I struggle to relate them to physical processes. Could you give the reader a sense of what $ENSO^2$ might be?

Garfinkel et al 2018 goes into great detail as to why physically La Nina can also lead to a moistening. The short answer is that the region of the cold point moves zonally within the tropics, and even though the lower stratosphere cools, the cold point actually warms.

This has been added to the introduction section where it seems more appropriate than at this point in the text:
"Both La Nina and El Nino can lead to a moistening if the cold point moves zonally within the tropics (to the Central Pacific for El Nino, and to the far West Pacific for La Nina), and even though the lower stratospheric response is opposite for El Nino and La Nina, the cold point warms for both (Garfinkel et al 2018)" We also added a similar sentence to the discussion.

(7) The values in Figure 6 are somewhat hard to read. Could you add a few labeled contour lines please?

Contour lines with labels were added to Fig. 6, Fig. 3 and Fig. 8.

(8) Figure 7 feels unnecessary since the same information can be conveyed with the text.

We removed Fig. 7 from the paper and updated the text to include the SHAP values for each predictor.

(9) In figure 9 (a to c), the text suggests that the solid black lines are observations (and they are not described in the caption) but where are there two parts and at different values? Label the models in 9(a).

We now note the solid black horizontal line is observations, and that entry water is defined separately for CCMI and CMIP (80hPa and 70hPa respectively).

Adding labeling to panel 9a made the figure more visually distracting without any added content, hence we left 9a as is.

Minor comments

Page 1, line 164, Emissions

corrected

Page2, line 5, through the its

corrected

Figure 1. Labels are a bit small and hard to read.

now larger

Figure 4. Are the units of the H20 anomalies correct?

Yes, we clarified it in the caption.

Figure 5 and Figure 9. You use "std" and "std dev". Choose one to be consistent and also explain the abbreviation in the caption.

We have adopted the std.dev abbreviation and updated it in the caption.

Figure 5(a) I am confused about the histogram. Is it normalized? If so, why are the values >1?

The histogram is normalized in a way that the total area of the histogram equals 1. This means that some bars can indeed exceed 1. However, it may be confusing, thus, we replaced the figure with the “probability” normalization where the sum of all the bars equals 1. This changes only the y-axis values and not the shape of the histogram which is more important in the context of this paper.

Answer to Reviewer #2

We would like to start by thanking you for all the time and effort which you spent reviewing our paper. All your comments, suggestions, and questions were taken into account and all the necessary corrections were made in the revised manuscript. Furthermore, we address all your comments and suggestions below, point by point.

General comments:

This paper discusses the importance of the nonlinear interaction between ESNO, QBO, and stratospheric water vapor, based on MLR and advanced machine learning techniques, and analyzes both observational data and chemistry-climate models. The authors conclude that QBO is more important than ENSO^2 than ENSO in predicting entry water vapor. The novel techniques and rigorous analysis of this paper will inspire the whole community, and I recommend this paper be accepted after a few revisions.

1. As the authors mentioned in line 5 and line 13 page 2, ENSO and QBO influences the stratospheric water vapor by influencing the tropical tropopause temperature. Later in Fig. 3, the authors compare the prediction of water vapor from merely tropical tropopause temperature, and from linear/nonlinear combination of ENSO and QBO. Since the ENSO and QBO directly influence tropical tropopause temperature and indirectly influences water vapor, before showing the relationship between 'ENSO, QBO-stratospheric water vapor', additional analysis of how well can linear/nonlinear combination of ENSO and QBO represents the tropical tropopause temperature will make the logic tighter.

Garfinkel et al 2018 and 2021 considered the influence of ENSO on tropical tropopause temperatures in great detail, and we have nothing to add here. We have added more discussion of these papers in the introduction and discussion sections.

The role of the QBO for tropopause temperatures has also been considered extensively in previous work of others, including the papers we cite (Reid and Gage, 1985; Zhou et al., 2001, 2004; Fujiwara et al., 2010; Liang et al., 2011; Kawatani et al., 2014). We don't have much to add here either. The connection is known theoretically (as given by thermal wind balance on an equatorial beta-plane) to be linear.

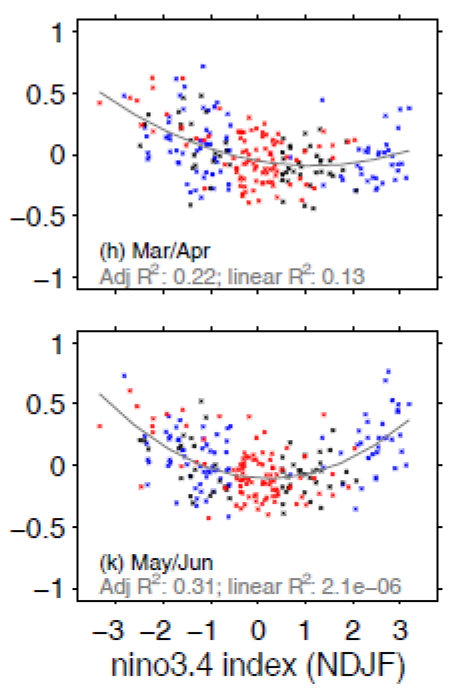
2. It is undoubted that considering the nonlinear process from ENSO and QBO can substantially increase the prediction of stratospheric water vapor, from the statistical analysis of this paper. However, more scientific arguments are needed when showing this result. For example, ENSO^2. The difference between ENSO and ENSO^2 are (1) ENSO^2 always amplifies extreme positive and negative ENSO states; (2) ENSO index has positive and negative values, but ENSO^2 only have magnitude, so extreme EN and LN will have similar ENSO^2 values. The

authors explain (2) in section 3, but lack the necessary analysis of how (1) influences the predictions. Can you add another experiment of, say, absolute(ENSO)? It is possible that the behavior of abs(ENSO) is not as good as ENSO^2, since moderate events are not very important and ENSO^2 emphasizes the importance of extreme events so not necessary to add this experiment into the paper. Then I suggest that can add some more comments on page 13, lines 9-14 on how the two differences between ENSO and ENSO^2 improve the prediction. I also suggest including citations of why choosing ENSO^2 and ENSO*QBO not only in the introductions but also in result sections when discussing the improvement.

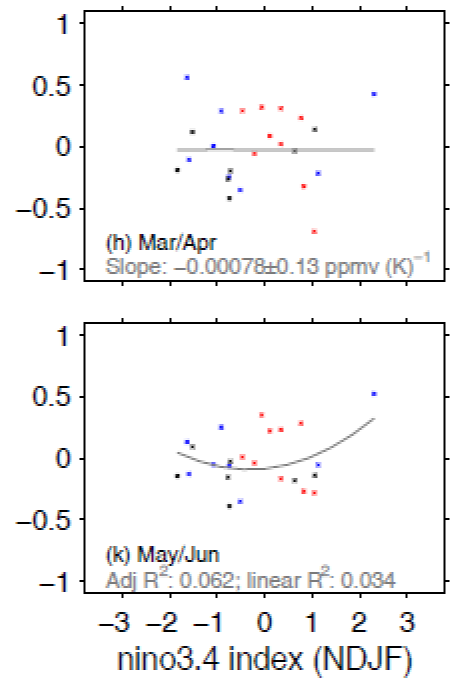
Garfinkel et al 2018 goes into great detail as to why physically La Nina can also lead to a moistening. The short answer is that the region of the cold point moves zonally within the tropics, and even though the lower stratosphere cools, the cold point actually warms. This has been added to the introduction section and discussion section.

Regarding your comment concerning ENSO^2 vs. abs(ENSO), below we copy in relevant figures from Garfinkel et al 2018 showing a scatter plot of the entry water vapor values for different values of the Nino3.4 index.

for the GEOSCCM model:



for the observations:



While one could attempt to discriminate whether abs(ENSO) is "better" than ENSO^2, there simply aren't enough points to make a convincing statistical case either way.

Specific comments:

1. In figures showing the horizontal distributions, i.e., Fig.3, Fig.6, and Fig. 8, since ENSO is one of the most important topics of this paper, I suggest the base map should center at 180° instead of 0°, so the readers can compare the Western and Eastern Pacific more clearly.

We changed the center of these Figs to 180°.

2. 10, please add panel numbers and titles.

We added the panel designations and titles.

3. Page 1, line 15: please include more citations for 'The amount of water vapor that enters the stratosphere is also important for stratospheric chemistry and specifically the severity of ozone depletion, for example, the citations on page 15, line 17.

We added three more.

4. Page 4, line 21: 'In total, more than 2500 year of model output are available' I see no reason to calculate the total years because you are not putting all the model outputs together.

This sentence has been removed.

5. Page 6, line 8: please introduce more about the radiosonde data, for example, is it monthly mean? Is the seasonal cycle included?

The radiosonde data was resampled to monthly means and its seasonal cycle was removed. We clarified this in the text.

6. Page 9, line 22: thanks for sharing, this is helpful to the community!

You're welcome :-)

7. Page 10, line 15: is the 'busts' problem in figure 4 still there in MLR2? 2010, 2015, and 2016 are all ENSO active years or right after so it is interesting to see whether adding ENSO^2 and QBO*ENSO can improve the performance or not.

These busts are present for MLR2 as well, though the error is not any larger than for the ML methods (we added MLR2 to Fig. 4).

8. Page 17, line 15: 'this results' should be 'this result'

corrected