Interactive comments on "A solar signal in lower stratospheric water vapour?" by T. Schieferdecker et al.

My recommendations:

It is appropriate for publication upon major comments below.

General comments:

This paper uses merged time series of stratospheric water vapour from HALOE and MIPAS data to study the influence of solar signal on the temporal variations of stratospheric water vapour. This is an interesting topic regarding a currently debatable issue on long-term water vapour trends in the stratosphere. However, I have several concerns:

One of the concerns is that the analysis is primarily based on a multivariate linear regression of water vapour time series. First, we should be cautious that results from this analysis do not simply imply a causal relation between the solar signal and water vapour variations. Many places in this paper use "control", "influence" and etc., to which should be paid more attention. Second, the water vapour time series used in this study still contains a seasonal cycle, and you should be aware that the large correlation may be dominated by the seasonal cycle, instead of by the real signal. Therefore, I suggest conducting regression analysis on water vapour time series after removing the seasonal cycle. Furthermore, one main conclusion of this paper is that the negative water vapour trends could be obtained when considering the solar cycle impact, which is also based on linear regression including an 11-year solar cycle proxy. It seems that you tried to remove the negative trends intentionally. This may raise a problem that the removal of the water vapor trends may just be a mathematical result rather than a real imprint of solar signal trends. One simple example would be that if you include any other unrelated factors (eg., population) with significant trends instead of the solar cycle, you could also possibly remove the water vapour trends. But this does not imply that the water vapour time series contains signals of these factors. This is an important but confusing result in this paper which you need to clarify.

Specific comments:

- 1. It will be better if you could see the solar signal in the decomposed water vapour time series using the Empirical Orthogonal Function (EOF) analysis.
- 2. **Abstract, page 12354, lines 12-15:** "We conclude from these results that a solar signal generated at the tropical tropopause is imprinted on the stratospheric water vapour abundances and transported to higher altitudes and latitudes via the Brewer–Dobson circulation". Even we could get the former part of the conclusion from this study, the latter part seems to be more likely an assumption.

- 3. **Abstract, page 12354, lines 15-17:** The conclusion that "the tropical tropopause temperature at the final dehydration point of air is also governed to some degree by the solar cycle" is not supported by this study.
- 4. **Page 12360, line 2:** "The third method proved most robust and was finally selected." It needs more explanation why you chose this method. Or you could list literature supporting this method.
- 5. **Page 12361, lines 24-25:** "The residual time series appears to be dominated by a systematic harmonic feature of a period length of about eleven years." How did you find out that it is dominated by an 11-year harmonic feature?
- 6. **Figures 4 and 6:** How statistically significant are the improvements?
- 7. **Page 12363, lines 16-17, and Figure 8:** Larger amplitudes do not imply larger effects. You have largest amplitudes in the tropical tropopause region, which may be due to that the variance of water vapour is largest in this region. I think the amplitudes of the fitted terms that represent the solar cycle in the regression should largely reflect the amplitudes of the water vapour variability. In order to find out the regions where the solar cycle has larger effects, you may need to compare the amplitudes to the water vapour variability.
- 8. **Figure 10:** The figure shows positive values (the phase shift of the solar signal in water vapour the age of stratospheric air) within 50°N/S and between ~15 and 23 km in the tropics, which may be primarily due to the low age of air in these regions. Similarly, the negative values may be due to the high age of air in higher altitudes. You need to elaborate why you want to look at the difference between the phase shift of the solar signal in water vapour and the age of stratospheric air.

Technical comments:

```
Page 12354, line 26: change "Mainly ..." to "It is ... that ...";

Page 12355, line 2: delete "atmospheric";

Page 12360, line 2: change "proved" to "was proved to be";

Page 12360, line 5: change "within" to "for";

Page 12360, line 10: change "was" to "were";
```

Page 12360, lines 21-22: rephrase "they merely help to describe but not to explain the temporal variation."

Page 12360, lines 25: change "not" to "that are not";

```
Page 12361, line 20: change "were" to "was";

Page 12362, line 22: delete "a priori";

Page 12364, line 3: change "which implies" to "implying" or rephrase the whole sentence;

Page 12368, line 3: add a comma after "White et al., (1997)";

Page 12368, line 6: change "which both is" to ". Both of them are";

Page 12368, line 6: change "to" to "with";

Page 12368, line 8: delete "until recently";

Page 12368, line 15: delete the second "a";

Page 12368, line 17: change "attribute" to "attributed";

Page 12368, line 19: change "is" to "was";

Page 12368, line 21: add "anomalies" after "water vapour";

Page 12368, line 23: delete "only";
```

Figure 5: I assume that the green line represents the linear term of the regression, please briefly indicate what the green line represent in the top panel;

Figure 2: change the names of each line to match the text descriptions better, for example,

"HALOE adapted" -> "HALOE avg_ker"; "MIPAS shifted" -> "MIPAS de-biased";

Throughout this paper, there are lots of long sentences using attributive clauses. It will be better to shorten these sentences. In many places, past tense and present tense are not properly used.