Referee comments on:
"Measurement of low-ppm mixing ratios of water vapor in the upper troposphere and lower stratosphere using chemical ionization mass spectrometry" by T. D. Thornberry et. al.

We thank the reviewer for his/her thoughtful comments and have addressed the points raised below and, as indicated, in the text.

Specific comments:
P. 385 L. 15, *Please add a missing reference (Neuman et. al. 2000)! You should also be consistent using brackets, e.g. (Neuman et. al., 2000) and sometimes [xxx et.al., 2000].*

Missing Neuman et al. (2000) reference added. [] have been corrected to ().

P.385 L.23*, you only use one detector channel, could the other one be used for some other tracer gas simultaneously with water vapour measurement?*

This would be possible with some reengineering. The current H_2O inlet/ion source design uses most of the available space in the pylon.

P. 385 -> *Inlet system: All the details are described in great detail but it lacks the information of the laboratory setup if it is identical. The authors start describing the inlet system and go into detailed flight mode description and then come back to the beginning of the instrument development such as choosing tubing materials and testing of the laboratory setup. I would like to see a clear segregation between the lab test and the flight mode in 2. Instrumental Description.*

The only portion of the text not related to the flight instrument inlet is section 2.2.2, which specifically states the tests were conducted with a prototype instrument. Text has been added to section 2.2.2 to clarify this.

P. 387 L. 25. *"The near-ambient pressure region of the inlet from the sampling point to the flow control valve (BV1) had a volume of 4.7cm3 and presented the dominant source of sample residence time." But you add calibration gas 15 cm downstream of the inlet tip this means you don't include this region where the sample spends most of its time in the calibration at all? The calibration methods are described in so many sections (P. 387, P.392, P.397) it's hard to follow what's happening.*

The calibration flow needed to be added to the inlet sufficiently downstream of the zero air addition that none of the calibration flow would escape the inlet. Text was added to section 2.2.2 to specify the short time constants observed for the material tests (< 5 s), and given that water is not permanently lost to inlet surfaces (as other species, e.g. O3, could be), the 12 cm of smooth tubing between the two does not pose a risk of artifact, and only a small additional contribution to the inlet time constant. This latter is supported by the similarity in observed time constants

between cal off and zero air on transitions. Text regarding this has been added to section 2.2 and section 3.2.3 discussing this issue.

P. 390, *Ionization Method: You could add a schematic picture to clarify the reactions. This would be easier for the reader to follow with the text.*

A schematic representation of the primary ionization reaction channels has been added as Fig. 4.

P. 397, Uncertainty: *The paragraph describes the sources of uncertainty but not any numbers are given. It would be nice to in this paragraph what the overall uncertainty is for this method.*

This section was intended as a general introduction to the sources of potential error/uncertainty in the measurement—the quantitation of the error contributions under flight conditions was addressed in the MACPEX performance section.

Table 1 and text P.389 L. 7: *The limit of detection (LOD) is usually determined from the standard deviation of the blank (zero) sample and it is not the lowest calibration point so don't call these values LOD. The lowest calibration point might actually be closer to your LOQ (limit of quantification). Also you give the limit a range of 0.5 – 0.8 ppm but where does it come from (also P. 396 L. 11) is it difference in the calibration gases or something else?*

Your point is well taken. "Detection limit" has been changed to "Quantification limit" as that does set the lowest value we could report (no extrapolation). A note to this effect has been added to the table.

Table 2: *I would like to see the summary of these errors here, so put the total uncertainty also here.*

Overall measurement uncertainty has been added to Table 2.

Figure 8: *In the calibration region there is three spikes standing out, where do they come from? You could also point out if the measured points are averaged over 1 s?*

Text added to the caption for Fig. 8 (now Fig. 9): "The spikes that appear in three of the calibration steps in region IV are caused by a burst in the H_2 flow when the 3-way solenoid valves switch the calibration flows from the higher pressure in the dump line to the catalyst. The signal from the bursts decays rapidly ($\tau < 2$ s) and does not affect the average value calculated during the last 10s of the calibration step."