Authors' Response to Referee Comments #2

We would like to thank the Referee for the comments and suggestions. Below, we give detailed responses (in blue) where appropriate.

General comments: It should be clearly stated that the development and testing of this sensor is far from complete if the authors claim measurement capability in the lower stratosphere. The sensor must undergo much more detailed testing under LS conditions. The large humidity bias shown at below 10..15 ppm is inadequate for measurements in the lower stratosphere. The claim in the title that this sensor is for UTLS measurements is misleading. Many other research groups have spent incredible amounts of effort into sensor development, incl. multi-year inter comparison campaigns at climate chambers (cited in this manuscript).

We are sorry that the Reviewer was misled by the title/abstract of our manuscript and we regret not meeting his/her expectations. However, the title highlights the key aspects of our manuscript: It presents a novel instrument that has been developed specifically for the application in the UTLS. In addition, it was deployed aboard a meteorological balloon to reach the lower stratosphere (28 km altitude). The instrument has been found fully operational during the entire flight (in both ascent and descent) and, despite being just a prototype instrument, it completed two consecutive flights within two days.

We fully acknowledge the significant efforts of other research groups and we are well aware of the highly challenging measurement situation, but exactly this supports the fact that our unique mid-IR instrument and its performance demonstrated under flight conditions represents a substantial scientific progress and thus fully justifies the title of our manuscript. We do not claim a final and completely characterized instrument, but rather communicate our development, findings, and observations, while clearly pointing out the current limitations. Again, we are grateful for the reviewer's concern and have rephrased the abstract to avoid any misleading of the reader: "*At higher altitude, the quality of the spectral data remained unchanged, but outgassed water vapor within the instrument enclosure was hindering an accurate measurement of the atmospheric water vapor. Despite this limitation, these test flights demonstrated the successful operation of a laser spectrometer in the UTLS aboard a low-volume meteorological balloon, with the perspective of future highly resolved, accurate and cost-efficient soundings.*"

We fully agree that this is not the end of the journey, but more effort has to be invested to validate the laser spectrometer e.g. in concerted intercomparison campaigns. These activities will hopefully lead to much better agreements with the reference method and minimize the observed biases. We are currently working on technically sound solutions, which will be implemented in the near future.

To avoid potential misunderstanding, we also removed the term "conclusive" in L237 P11, changing the sentence to: "*As a first assessment of the novel QCLAS instrument under realistic conditions, ...*"

In addition, we add a separate subsection called "*Outlook*" at the end of the manuscript, containing the recommended further developments.

Specific comments: Title: I believe that the title (and paper) should not state that the presented instrument is for measurements in the UTLS, but rather for tropospheric soundings. The results presented do not warrant lower-stratospheric measurements.

The motivation for the instrumental development with its specific design was driven by the need, complexity, and importance of UTLS water vapor measurements, -- a need which is clearly shared by the reviewer. Our focus is on the description of the development, characterization and very first field-testing of the instrument. This justifies the title, especially since we transparently discuss the current limitations, their implication for future research, and potential solutions to the issues we faced.

The data presented in the paper were repeatedly recorded under UTLS conditions. Although we have identified a bias at high altitude, the system remained fully operational in the UTLS, and the quality of the retrieved absorption spectra remains comparable throughout the whole flight. This alone represents a formidable achievement that is worth to be communicated. Undoubtedly, further work is required to improve, characterize and fully establish the credibility of our instrument in the scientific community.

P1, L9 The sentence "An open path design reduces the risk of contamination, allows fast response..." is contradictory of the results presented. Measurements presented here are indeed "contaminated".

In our statement, we meant the intuitive fact that any enhancement of the volume-to-surface ratio in the detection volume, as well as any enhancement of the sample flow leads to a reduction of ad/desorption-caused biases of the measurement. In this context, an open-path arrangement is the best possible solution.

Indeed, our results do show a wet bias in the stratosphere. However, on P13, L261, the sources of this bias is discussed in detail, delivering profound arguments for being (1) balloon-caused and (2) related to the internal volume of the instrument (gap between laser/detector and cell), which is unrelated to the open-path construction of the sampling setup.

In the revised version, we try to emphasize the difference between these two aspects. They are again mentioned in the Subsection "*Outlook*" that is added to the revised manuscript.

P3, L67. the number of molecules within the light path can be..

These sentences are now rephrased for better understanding*: "In this application, we exploit the fast tunability of a quantum cascade laser to record transmission spectra at kHz repetition rate that cover several wavenumbers. From these spectra, the number density of target molecules is deduced using the Beer-Lambert-law, which describes the attenuation of radiation in absorbing media.* "

P3, L73 which values from a spectroscopic database?

For clarity, we rewrite the sentence at P3, L73: "*The wavenumber dependence of the absorption coefficient σ(ν) can be approximated by a Voigt profile V (ν;α(T),γ(p,T)). The broadening parameters α and γ are calculated as a function of the gas pressure p and temperature T using the coefficients from a spectroscopic database.* "

The next sentence is moved towards the end of the paragraph for sake of structure.

P4, L90 The integrated line strength?

No change. According to the HITRAN naming convention, S_{ij}, i.e., the spectral integral of the absorption coefficient $\int \sigma \, d\nu = S_{ij}$ is called "spectral line intensity".

Fig. 2 indicate parameters such as pressure, concentration, altitude

Done.

Fig. 4 Explain in the main text what the problem of the fitting procedure was.

We added to the text at the end of Section 2.3.3: "*The laser temperature variation is derived by its frequency shift, using the absorption line position as reference (red trace). At the tropopause, where the line contrast is lowest, i.e., the absorption line is still broad but has a rather small amplitude; the determination of the central frequency is most difficult and thus exhibits the highest noise level.*"

P10, L222 Fig. 6(a) should be Fig. 6(c)

Corrected.

P10, L231 the agreement of the slope is fine, but the intercept is really important, too. Especially when the goal of this instrument is to accurately measure extremely low H2O. This is probably difficult to transfer from the weak absorption line probed for this test to the strong one due to different baseline shape, which might highlight a weak point of this experiment.

We agree that the intercept is important to qualify the accuracy. However, in our spectroscopic approach, assuming a zero intercept is well motivated. This is because the absence of target molecules can be determined with an uncertainty that is given by the Allan deviation. In our case, the uncertainty amounts to 0.01 % (for 100 s averaging), which can safely be considered as zero in this context.

Using a deliberately chosen small absorption line to determine the performance at relatively high concentrations may be surprising. However, it avoids the extreme challenges of creating very low water vapor concentrations and determining these with a reference instrument at exactly the same place and time. Given the limitations of this more established approach, using well-defined (higher) concentrations and a totally representative spectroscopic setup in combination with a very weak absorption line is arguably a highly valuable and reliable concept.

We are, however, well aware that this is a preliminary study we have performed using the tools available and it does not replace a more rigorous comparison in a more representative setting. This is again mentioned in the added section "Outlook".

Fig. 6 "A precision of 0.11 % at 1Hz is.." 0.11 % of what? Also Rˆ2 = 1.018 should either state the slope or the correlation coefficient I assume.

We have adapted the manuscript and now explicitly name these values "*relative Allan deviation*", i.e., relative to the measured mixing ratio; and "*coefficient of determination* R^2".

The calculation of the coefficient of determination is revised; the fact that the offset parameter is forced to remain zero is now taken into account in the calculation.

P13, L278 Figure 8(a) should be Figure 8

Done.

P16, L351 No, you were not successful to measure LS water vapor.

We have identified and explained a bias at high altitude, the system remained operational in the UTLS and there is no indication of inferior spectroscopic performance than in the troposphere, which we consider an outstanding achievement. However, we agree, that the term "successful" may be interpreted differently at this point. Therefore, we removed this terminology from the Conclusions section.