Reply to Ref. #1, M. Wenig

First of all we want to thank this reviewer for the positive assessment of our manuscript and the constructive and helpful suggestions.

General comments
The paper describes a very extensive comparison study of several MAX-DOAS HONO measurements from different groups. Using coinciding measurements in Mainz as well as modeled spectra, their comparison can focus on different aspects of the instrument characteristics and retrieval process. The paper addresses relevant scientific questions since the findings can be used for improving existing retrievals and are very valuable for future instrument and algorithm design. It is more an evaluation of existing methods rather than the introduction of a novel concept, but the methods and conclusions are very solid. At some points the paper is a little vague, when only speculations are provided instead of explanations for observed differences. The title is appropriate and the abstract summarizes the presented work well. The paper is a little long, especially in combination with the supplementary material. The paper should be readable without the supplement as it should only provide extra material if the reader wants to go into more details. Overall the manuscript is of high scientific quality and I recommend publication after some minor corrections and clarifications as listed in the following

Author reply:
Many thanks for the positive assessment! We modified the paper based on the comments from you and reviewer 2. In order to make the main text readable without the supplement, we added some important numbers in the main part of the paper (in the parts related to the supplementary figures). We hope the revised manuscript is more smoothly readable.

Specific Comments:

1) L13p2 "uncertainty from different sources" sounds confusing, how about "uncertainty from all the different sources" or "overall uncertainty"?

Author reply: we modified it as "overall systematic uncertainty".

2) L14p2 "the systematic bias of the fitted from the simulated real HONO delta SCDs is" is grammatically challenging as well.

Author reply: the sentence is deleted.

3) L18p2 The sentence "However, systematic uncertainties limit the reliability of the results." is very general. What do you know about systematic errors? If the bias is known, you can correct for it. If not, how reliable are the results?

Author reply: We modified the sentence based on your comment and the comment from the reviewer 1. The new sentence is "In summary for most of the MAX-DOAS instruments for elevation angles below 5°, half daytime measurements (usually in the morning) can be above the detection limit of the HONO delta SCD of 0.2×10^{15} molecules cm^{-2} with an uncertainty of $\sim0.9\times10^{15}$ molecules cm^{-2}."

4) L25ffp5 When you define the different quantities, please provide equations. That would make it easier later on in the paper if you could just refer to the variable. For example what is the difference between delta SCD with sequential FRS and dSCD?

Author reply: We add the equation derivations in the revised manuscript to define the dSCD and delta SCDs clearly in section 3.1.

5) L4ffp6 When you describe Fig. 1 you mention the different fit results for O3, NO2 and BrO but for HONO you just write "The HONO absorption structures are well retrieved using both types of FRS.", but the fit parameters are 5.5 and 6.9 10^15 molec. cm^-2, so quite different.

Author reply: We modified the sentence as "The difference of retrieved HONO dSCDs between the two fits is mainly due to the different HONO absorption in the two FRS. The same reason also leads to the differences of retrieved dSCDs of the other trace gases. The difference is substantially larger for the trace gases with considerable stratospheric contributions, e.g. O_3 and BrO, because the stratospheric light paths around noon for the daily noon FRS are much shorter than those during sunset or sunrise."

6) L31fp6 You write "because of unknown instrumental problems, CMA and Boulder didn't participate in the comparisons of the delta SCDs for a sequential FRS and dSCDs for a daily noon FRS, respectively", you could mention that the other measurements were not affected.

Author reply: we add the clarification of "but other instruments are not affected."

7) L1fp7 In the sentence "all the instruments capture well the diurnal evolution and elevation angle dependence of the HONO delta SCDs." "capture well" sounds very general, you could be more specific by mentioning if they agree within the uncertainties or if there are significant differences.

Author reply: Since a detailed quantive analysis of the inter-comparisons results is given in section 3.3, here we only mention that in general similar HONO results are observed by the different instruments. In the revised paper we modified the sentence as "from all instruments a similar diurnal evolution and elevation angle dependence of the HONO delta SCDs is retrieved." and clarified that "A detailed quantitative analysis of the deviations of the HONO results between the instruments is provided in the statistical analysis in section 3.3."

8) L10ffp8 You use the average of selected instruments as the reference values, but then you compare all instrument to that reference (l13p8,l11p9), that's not very consistent. How about using a median or something similar for reference instead?

Author reply: Because the Boulder and CMA instruments are affected by some unknown instrumental problems, we prefer to only include the more stable instruments. We clarified this point in the revised manuscript as "In addition the selection is also because the Boulder and CMA instruments are affected by some unknown instrumental problems."

9) L2fp10 Again, an "unknown problem" can always be the reason for discrepancies, can you be more spefic?

Author reply: Sorry, we can't give the specific reason. We only can speculate this is due to an instrumental problem, like spectrometer or motor. We specify the problem as "an unknown instrumental problem" in the revised manuscript.

10) L12p10 You mention "Note that there is no random noise added into the synthetic spectra." But in Sec. 4.9 (l4fp17,also table S1) a SNR of 3000 is mentioned that is added to the synthetic spectra, so why not in Sec. 3.4?

Author reply: The synthetic spectra used here are to quantify the systematic difference of DOAS fits implemented by different groupds and to quantify the systematic difference of retrieved results from the truth. Moreover, in section 4.9 the effect of noise is investigated and it was found that the noise added in the synthetic spectra doesn't lead to any considerable systematic deviation. Therefore it is meaningful to investigate systematic effects on the HONO retrieval based on synthetic spectra without noise. We added a clarification in the revised manuscript as "Note that there is no random noise added into the synthetic spectra due to the objective to quantify the systematic difference of the retrieved values from the truth and the negligible effect of noise on the systematic difference concluded in Section 4.9. ".

11) L29p13 You write "The Thalman O4 cross section at 203k orthogonalised to that at 293k is calculated by the orthogonalisation based on Gram-Schmitt's algorithm and is shown in Fig.9." and Fig. 9's caption says "normalized absorption cross sections". How were they normalized or related to that what scalar product was used for the orthogonalization algorithm, e.g. is a polynomial removal part of the scalar product?

Author reply: We modified the description in the revised manuscript as " The Thalman O_4 cross section at 203k is orthogonalised to that at 293k based on Gram-Schmitt's algorithm using a polynomial of two degrees. The orthogonalised O_4 cross section is normalized by an arbitrary factor to be shown in a comparable scale with other cross sections in Fig.9."

12) L29p15 The degree of the polynomial was "arbitrarily selected" as five, why not 3? Fewer parameters to fit usually make the optimization more stable.

Author reply: Because of the weak dependence of HONO results on polynomial degrees in 335-373nm, the parameter can be freely selected between 3 to 5. In addition because in some cases short time variations of the sky conditions might happen in real measurements, we selected a higher polynomial degree, which can better account for such changes. We clarified this point in the revised manuscript: "The effect of the degree of the polynomial on the HONO results in the wavelength range 335-373 nm is small. However because in some cases short time variations of the sky conditions might happen in real measurements, we recommend to select a higher polynomial degree, which can better account for such changes. A fifth degree polynomial is used for HONO retrievals in this study. "

13) L12p16 "In spite of these possible interferences, taking into account typical instrumental problems (like spectrograph straylight), the consideration of an intensity offset correction in the fit is still recommended for the HONO retrieval." I agree, but can you underline this statement with some data or estimates? How do you know which influence is more important?

Author reply: We can not quantify the straylight effect. We clarified this point in the revised manuscript: "The effect of spectrograph straylight can not be quantified here because it needs a sophisticated Lab measurement, which was not operated during the campaign.".

14) L23fp18 What do you mean exactly with "systematic differences [. . .] caused by implementation of DOAS fits in the software packages", different fit functions or numerical implementations of the optimization?

Author reply: The exact reason in the codes of software is unknown here. We can only generally attribute the differences of HONO results to the differences of the codes of DOAS software. We

clarified it in the revised manuscript: “However the exact reason in the codes of software, which cause the difference, is unknown here. We can only generally attribute the differences of HONO results to the differences of the codes of DOAS software.”