Review of AMTD paper by R.S. Gao et al.

The authors present a new version of the well-known NOAA-1 ozone instrument described by Proffitt and McLaughlin (1983). This new version NOAA-2 differs from the older one in many features and properties. I am basically happy with the actual manuscript, but have some minor concerns which I would like to see considered before publication in AMT.

Minor concerns

Abstract. I don't understand why you give the specification "polarized" in line 7/8. That the light gets polarized by the beam splitters has no implication, or? I guess, you would like to say that the light path is folded, which is much more important here.

The instrument precision should be given in more detail. Most/many optical instruments are limited by shot noise which is usually given as $\Delta O_3 \sim f^{0.5}$, with f the measurement frequency. This is important here, as most UV photometers have a response time of ~10s. Moreover, the value "O_3 molecules cm^{-3}" is certainly fine, but most readers will first convert this value in ppbv. Thus, please (also) give a precision such as "ppbv at 1 bar and 2 Hz".

Introduction. It is very general and only in the last paragraph you start with UV photometers. Please, add at least a further paragraph listing the many other O_3 instruments for airborne application, inter alia, Kalnais and Avallone, JAOT, 2010, or Hintsa et al., JAOT, 2004.

p.3477 What absorption cells are used (material, dimensions)?

p.3480, l.5 Why turbulence increases noise, due to light scattering? Give an explanation and a reference.

p.3482, l.7 Are the intrinsic precision (5×10^9 O_3 molecules cm^{-3}) and the total precision (11×10^9 O_3 molecules cm^{-3}) completely independent of pressure and the O_3 mixing ratio? In this respect I don't understand the residuals in Fig.8a. For instance, the light blue point at 300 ppbv and 600 hPa. With the above total precision I calculate a precision of 0.7 ppbv. On the plot I see -5 ppbv. The Beer-Lambert equation for inferring the O_3 m.r. is $[O_3] = c(\lambda,T,p,\ldots) * \ln(I_0/I)$ with I_0 and I the intensities without and with O_3. That is, the noise should almost completely be determined by the precision or $\ln(I_0/I)$ and not by the other parameters like the absorption coefficient, absorption length etc. that determine the accuracy.

Give a precision in ppbv (at x Hz) at 1000 hPa and at 100 hPa.

Try to give an explanation for the increasing noise with increasing sample flow.

Write if you mean flow at normal conditions or volume flow.

Section “3.2 accuracy”. You should indicate how your devices compare with the NIST Standard Reference Photometer, see e.g. Viallon et al. Metrologia, 2006. What (T dependent) absorption cross section is used?

Also here, you should cite more references, e.g. Zucco et al., Meas.Sci.Technol., 2003.

p.3484, l.24. What does “negative artefact” mean? The humidity artefact goes in both directions depending if you enter or leave a humid flight section, e.g. a cloud.

p.3485, l.5ff. In my opinion, this test says little. Such a humidity test is usually performed by abruptly switching from dry to humid air or vice versa, comparable with a flight through a cloud. This artefact strongly depends (besides the scrubber) on the cuvette wall material and cleanness, as the wall controls the amount of adsorbed water and thus affects the transmission.

Do you have an explanation for this strange O_3 threshold of 130 ppbv above which no artefact is apparently detectable. At first view this contradicts the common explanation (scrubber, cuvette transmission).

Fig.8 Important is the percental deviation of the two devices. Thus please choose a log-log diagram. What tell the different symbols?

Fig.9 graph a: Except at the beginning and at the end where the pressure is high, in tropospheric air or at low O_3 m.r., respc., the blue line is ever below the red one, at ~74000s by at least 5 ppbv. Why?

The noise of NOAA-2 mostly appears larger than of NOAA-1, especially at ~71500s.

graph b: again, choose a log-log plot. The deviations sometimes appears high, e.g. around 400 ppbv, and basically far too high when considering the precision and accuracy. Is it due to non-perfect synchronisation or a not perfect consideration of the sample air transport time from the atmosphere to the cells? Are these time shifts considered?

I suggest adding a figure c that shows a counting statistics of the deviation between the two instruments (with Gaussian fit and discussion).