Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2016-34, 2016

Review of the submitted article:

Uncertainty budgets of major ozone absorption cross-sections used in UV remote sensing applications

Mark Weber, Victor Gorshelev, and Anna Serdyuchenko

General comments:

The submitted article presents a critical review of three published datasets of ozone absorption cross-sections in the Huggins band, including data previously published by the same authors, with a focus on their uncertainty. In that regard it addresses an issue which has often been underestimated. It should bring valuable outputs for the community of scientists involved in ozone monitoring with remote sensing instruments, in providing a sound base to select the most appropriate dataset to finally improve the confidence in measurements. To calculate uncertainties associated with cross-sections at any temperature, the authors have chosen to use Monte-Carlo simulations. This tool is recommended by guidelines on uncertainty calculation in such complex situations, and could be more widely used by the community.
The article is generally well written and organised. The main criticism would be a lack of clarity in the assumptions made by the authors when using datasets published by other teams, and in applying the Monte-Carlo simulations, as detailed in more specific comments below:

Specific comments:

1. The introduction could give more details on the needs. It is explained that remote sensing applications use the reviewed datasets. However this paper is about their uncertainty, and it is not said how this uncertainty is or would be used. It is important because the authors present a good piece of work with the Monte-Carlo simulations, which would not be valuable if users could be satisfied for example with just a conservative relative uncertainty to be applied to all values. This may be straightforward to the authors, but should be developed in the paper.
2. The choice of Monte-Carlo simulation could also be developed already in the introduction. Why this methodology? What are the assumptions? What does it bring? Could it be used for other cases?
3. Section 2 – measurement technique could be better organised. It first introduces the common aspect of all measurements of the cross-section, and then mentions some particularities. This could be introduced, explaining the sources of uncertainties that are expected to be always present due to the Beert-Lamber law, and then additional sources due to experimental choices.
4. Section 3 - review of reported uncertainty is interesting but the reader misses a clear goal. It would be easier to read having in mind the purpose of that review. Is this to provide the inputs for the Monte-Carlo simulation? In that case, why separating type A and type B uncertainties? Monte-Carlo calculation of uncertainties does not make such a distinction. It just needs the PDF associated with all uncertainty sources. One could also think that the users of data sets would need to know which uncertainties are correlated. This would be valuable but this point is never clearly mentioned in that section.
5. Section 3.1 uncertainty budget of Hearn: the goal of that section is unclear, and this subject was already treated in a publication by Viallon *et al.* in 2006 (DOI: 10.1088/0026-1394/43/5/016). It is certainly relevant to note that the BP data set is scaled to the Hearn value, and that BP uncertainties should be at least that of Hearn. However this could be stated shortly within section 3.2. In addition the review of Hearn uncertainty budget contains some inconsistencies and its conclusion is unclear (see line by line comments). It is suggested to either reconsider the analysis or remove this section.
6. Section 4 also needs to clarify the uncertainty treatment associated with the polynomial fit. How are the uncertainties of the polynomial coefficients calculated? Are they outputs of the regressions? If that is the case, what are the uncertainties taken into account in the regression? None, which would justify the move towards Monte-Carlo? But then, what about trying to just use the reported uncertainties associated with each point and perform the regression with those? What would be the difference with Monte-Carlo simulations?

7. Section 5 leaves unanswered questions regarding the Monte-Carlo calculations. Based on the guide cited by the authors (JCGM 101:2008), one would expect some considerations on the measurement equation or process, identifying the input quantities, and explaining the choice of the PDFs associated with each of them. In this paper it is difficult to link uncertainties accounted for in the Monte-Carlo simulation, listed in Table 8, and the input quantities. It is not clear if authors have considered just equation 4 (polynomial) as their model, or a more complex process involving several steps of fitting. In addition, Monte-Carlo simulations to calculate uncertainties implement different version of an algorithm, with different assumptions. Reference to the programme (if external) or the code (if authors did the programme) is missing here. Reference to Wu 1986 found in section 5 seems to be on least-square analysis rather than Monte-Carlo calculations. This is confusing, in particular when the problem seems to be solvable by a least-square code.
8. Validation of the Monte-Carlo method: authors do not mention how they have validated their choice of a simulation package. Have they looked at a few measured cross-sections to compare the output of MCM with the measurement results uncertainty?
9. Section 6 nicely summarises the work done. However conclusions drawn by the authors could go beyond this work. Some comments on the current usage of those datasets, highlighting how uncertainties are commonly neglected, would bring added value to the current study. If the goal is to encourage more careful consideration of the uncertainty associated with those datasets, this should be better emphasized.
10. The authors are kindly asked to review how to express units and quantities. They can refer to the SI brochure for example (mostly available on-line) or the ISO guide 80000-1:2009. Unit names for instance are normally written in roman (upright) type, and they are treated like ordinary nouns. Symbols for quantities are generally single letters set in an italic font.

<u>Line-by-line comments:</u>

Line 28**:** "…the satellite community uses any of the three data sets or other data". Which other data?

Line 32: "Neither of the publications uses the guidelines…". This sounds a bit unfair. Most of the work was published before the JCGM guidelines. The concept of uncertainty was already there, but the vocabulary was not harmonised, and indeed some uncertainty sources were underestimated.

Line 70: the standard deviation of the mean can indeed be used as standard uncertainty in case of white noise, which is often the case in spectroscopic instruments. However what is the implication of this statement? In published papers uncertainty of the mean was often an experimental standard deviation. Do the authors suggest this should be replaced with the standard deviation of the mean?

Line 101-104: what is the rationale behind the choice of a rectangular distribution for the uncertainty published by Hearn? This does not seem consistent with table 1.

Line 104: the term precision brings more confusion than clarity. Is that an uncertainty? What is finally the uncertainty on the Hearn value considered by the authors?

Line 119: "it seems to be an underestimation". Agreed, but what is the decision taken by the authors?

Line 119: it is here stated that the Hearn value is reported with a 1.4% relative standard uncertainty. This was not clear from section 3.1, and is arguable as already noticed. As Hearn provided an uncertainty budget, it seems more reasonable to revise it than using his value with a rectangular distribution of the uncertainty.

Line 153: please explain the "spectral registration". For non-expert, this appears to be a sort of shift applied to the wavelength scale. Is that more complex? How does this process create a statistical uncertainty?

Line 157 "standard deviation (variances)". The variance being the square of the deviation, what is the meaning of the bracket? To clarify that the calculation of uncertainty is performed with the variance, or to state that the standard uncertainty associated with the intensity is the experimental standard deviation? Both? Please clarify.

Line 192 equation (4): the choice of the symbol t for the temperature in °C is rather unusual, as this is normally the time. More important, it may be more robust to always use the temperature in K. it is not clear why this change is needed.

Line 193 “ multivariate linear regression”. Why *linear* when the cross-section dependency versus the temperature is a second order polynomial?

Line 197: use GUM convention in equation (5): $u^2(\sigma_p) = u^2(a_0) + t^2u^2(a_1) + t^4u^2(a_2)$. This equation also assumes that the uncertainty in the temperature is negligible. This should be stated and motivated.

Line 207-212: it seems rather “unfair” for other groups to choose a temperature of 193 K to compare results, as only SG performed actual measurements at that temperature. It may not change the conclusion, but it is suggested to consider another temperature.

Line 213-215: please clarify the purpose of this statement. It is currently difficult to know what to do with it.

Line 222: “..only the uncertainty from the temperature parametrisation using a polynomial”. Does this mean that no experimental uncertainty was included? This should be stated more clearly, and some justification provided for not doing it (in section 4).

Line 253-260: analysis of results: could be put in perspective with uncertainties reported in the literature and choices made for this paper. For instance, the 2% uncertainty on BP certainly reflects the uncertainty from using Hearn value as reference.

Table 1: in Hearn paper, the column “Total SD (RMS)” is simply the combined relative standard uncertainty, calculated from sources listed in other columns. This is consistent on all wavelengths measured by Hearn. This is badly reflected in Table 1.

Table 5: this table tends to associate statistical uncertainty with Type A, and systematic with type B. As pointed earlier in the paper, this is a confusion that should be avoided. Type B uncertainties in particular may not be systematic (or biases), but all uncertainties for which the information was not provided by repeated observations (calibration certificate for instance). It may be the case in the reported experiment that all Type B uncertainties are also systematic. It is then suggested to modify the titles to reflect the type (A or B). However again, one wonders why the distinction is made, when it is no use in Monte-Carlo simulations.

Figures 2 to 4: the information displayed in those figures is too raw, and does not help the reader in analysing the results. It is impossible to compare the 3 plots, difficult to see the level of uncertainties. The choice of those figures should be questioned. Figure 5 only seems to be sufficient in that section.

Figure 4.b) This graph is too busy. What is the goal? If this is to show that a second order polynomial was appropriate, then statistical tools can be used, such as residuals standard deviations. If the authors want to provide the ranges of residuals for each wavelength in the graph, then a table might be more appropriate.

Editorial/technical corrections:

Line 92: $O3$ and $O2$ to be written O_3 and O_2.

Line 136: avoid the symbol @ to mean “at”.

Table 1: the molecular absorption cross-section value misses a multiplication sign and a power notation: 2.4×10^{-19}.

Table 2: the uncertainty in the transmittance determination misses a power notation: “2 in 10^5” and not “2 in 105”.

Table 3: avoid the symbol @ in place of “at”

Figures 2 to 4: axis titles are incomplete: *x* axis should be “λ / nm” with the explanation in the legend that λ is the wavelength.