

Response letter to Referee #2 on the manuscript "Monte Carlo method for determining uncertainty of total ozone derived from direct solar irradiance spectra: Application to Izaña results"

Authors: Anna Vaskuri, Petri Kärhä, Luca Egli, Julian Gröbner, and Erkki Ikonen

Article reference: amt-2017-403

Authors: We thank Anonymous Referee #2 for carefully reading our AMT Discussion paper and giving constructive criticism. We will later answer point by point all details raised, but at this stage we clarify the scientific concerns raised.

Comments on Introduction

"First, the paper was probably written very quickly and important information is missing, which makes the understanding quite difficult for the inexperienced reader:

- a complete introduction about the importance of ozone measurements and the assessment of their uncertainty should be included (it is only the first line, so far);
- it is not explained why spectral data should contain correlations (physical basis);"

Authors: We revised Introduction and included new paragraphs as the referee suggested. After these amendments, the introduction is:

"Atmospheric ozone has been defined as an essential climate variable in the global climate observing system (GCOS-200 (2016)) of the World Meteorological Organization (WMO). Its long-term monitoring is necessary to document the expected recovery of the ozone layer due to the implementation of the Montreal protocol (UNTC (1987)) and its amendments on the protection of the ozone layer. Atmospheric ozone, first discovered by Fabry and Buisson (1913), protects the humans, the biosphere, and infrastructures from adverse effects of ultraviolet (UV) radiation by shielding the Earth surface from excessive radiation levels (McElroy and Fogal (2008)). Since the 1970's, it is known that human-produced chlorofluorocarbons (CFCs) destruct atmospheric ozone (Molina and Rowland (1974)) and have led to recurring massive losses of total ozone in the Antarctic in the form of the ozone hole (Farman et al. (1985); Solomon et al. (1986)). An unprecedented ozone depletion has also been recently observed in the Arctic (Manney et al. (2011)), while in the middle-latitudes, moderate ozone depletion has been observed (Solomon (1999)). The Montreal protocol and its amendments have been successful in reducing the emission of ozone-depleting substances (Velders et al. (2007)). Nevertheless, recent studies give conflicting results with respect to the observation of a recovery of the ozone layer, and model projections have shown the recovery to occur not before the middle of the 21st century (Ball et al. (2018); Weber et al. (2018)). Therefore, careful monitoring of the thickness of the ozone layer with uncertainties of 1% or less is crucial in verifying the successful implementation of the Montreal Protocol and the eventual recovery of the ozone layer to pre-1970's levels.

Response letter to Referee #2

Article reference: amt-2017-403

"Traceability for atmospheric total column ozone (ATMOZ)" was a three-year project funded jointly by the European Metrology Research Programme (EMRP) and the European Union (ATMOZ project (2014– 2017)). The goal of this project was to produce traceable measurements of total ozone column (TOC) with uncertainties down to 1%, by a systematic investigation of the radiometric and spectroscopic aspects of the methodologies used in retrieval. Another objective of the project was to provide a comprehensive treatment of uncertainties of all parameters affecting the TOC retrievals using spectrophotometers. This paper presents outcome of the work on studying the uncertainty of TOC obtained from spectral direct solar irradiance measurements, taking correlations among wavelengths explicitly into account.

TOC can be determined from spectral measurements of direct solar UV irradiance (Huber et al. (1995)). We have developed a Monte Carlo (MC) based model to estimate the uncertainties of the derived TOC values. One frequently overlooked problem with uncertainty evaluation is that the spectral data may hide systematic wavelength dependent errors due to unknown correlations (Kärhä et al. (2017b, 2018); Gardiner et al. (1993)). Omitting possible correlations may lead into underestimated uncertainties for derived quantities, since spectrally varying systematic errors typically produce larger deviations than uncorrelated noise-like variations that traditional uncertainty estimations predict. Complete uncertainty budgets for quantities measured are necessary to understand long term environmental trends, such as changes in the stratospheric ozone concentration (e.g. Molina and Rowland (1974)) and solar UV radiation (e.g. Kerr and McElroy (1993); McKenzie et al. (2007)).

Physically, the correlations may originate, e.g., from lamps or other light sources used in calibrations. If their temperatures change e.g. due to ageing or current setting, a spectral change in the form of Planck's radiation law is introduced. Non-linearity in the responsivity of detector causes systematic differences between high and low measured values. The effect of the introduced spectrally systematic, but unknown changes on the TOC values derived from the affected spectra may deviate significantly from the uncertainty calculated assuming that the error behaves like noise. The presence of correlations in measurements can be seen in many ways. For example, problems have occurred when new ozone absorption cross-sections have been taken into use (Redondas et al. (2014); Fragkos et al. (2015)). Derived ozone values may change significantly because different systematic errors are included in the different cross-sections. Also, TOC estimated from a measured spectrum often depends on the wavelength region chosen, although the measurement region should not affect the result much.

In this paper, we introduce a new method for dealing with possible correlations in spectral irradiance data and analyse uncertainties in ozone retrievals for three different spectroradiometers used in a recent intercomparison campaign at Izaña, Tenerife, to demonstrate how it can be used in practice. One of the instruments is QASUME (Gröbner et al. (2005)) that is the World reference UV spectroradiometer at the World Radiation Center (PMOD/WRC). The second one is an array-based high-quality spectroradiometer BTS2048-30-UV-S-WP (BTS) from Gigahertz-Optik (Zuber et al. (2017a, b)), operated by PTB. The third one is an array-based spectroradiometer AvaSpec-ULS2048LTEC (AVODOR) from Avantes, operated by PMOD/WRC. The field of view of the spectroradiometers has been limited so that they measure direct spectral irradiance of the Sun, excluding most of the indirect radiation from the remainder of the sky. "

New references were included:

Ball W. T., Alsing J., Mortlock D. J., Staehelin J., Haigh J. 5 D., Peter T., Tummon F., Stübi R., Stenke A., Anderson J., Bourassa A., Davis S. M., Degenstein D., Frith S., Froidevaux L., Roth C., Sofieva V., Wang R., Wild J., Yu P., Ziemke J. R., and Rozanov E. V.: Evidence for a continuous decline in lower stratospheric ozone offsetting ozone layer recovery, *Atmos. Chem. Phys.*, 18, 1379–1394, 2018.

Fabry C. and Buisson H.: L'absorption de l'ultra-violet par l'ozone et la limite du spectre solaire, *J. de Phys.*, 3, 196–206 (1913).

Farman J. C., Gardiner B. G., and Shanklin J.D.: Large losses of total ozone in Antarctica reveal seasonal ClOx/NOx interaction, *Nature* 315, 207–210, 1985.

GCOS-200: The Global Observing System for Climate: Implementation Needs (World Meteorological Organization, 2016).

Manney G. L., Santee M. L., Rex M., Livesey N. J., Pitts M. C., Veefkind P., Nash E. R., Wohltmann I., Lehmann R., Froidevaux L., Poole L. R., Schoeberl M. R., Haffner D. P., Davies J., Dorokhov V., Gernandt H., Johnson B., Kivi R., Kyrö E., Larsen N., Levelt P. F., Makshtas A., McElroy C. T., Nakajima H., Concepción Parrondo M., Tarasick D. W., von der Gathen P., Walker K. A., and Zinoviev N. S.: Unprecedented Arctic ozone loss in 2011, *Nature*, 478, 469–475, 2011.

McElroy C.T. and Fogal P.F., *Ozone: From Discovery to Protection*, *Atmosphere-Ocean*, 46, 1–13, 2008.

Redondas A., Evans R., Stuebi R., Köhler U., and Weber M.: Evaluation of the use of five laboratory-determined ozone absorption cross sections in Brewer and Dobson retrieval algorithms, *Atmos. Chem. Phys.*, 14, 1635–1648, 2014.

Solomon S., Garcia R. R., Rowland F. S., and Wuebbles D. J.: On the depletion of Antarctic ozone, *Nature*, 321, 755–758, 1986.

Solomon S.: Stratospheric ozone depletion: A review of concepts and history, *Reviews of Geophysics*, 37, 275–316, 1999.

UNTC, United Nations Treaty Collection, Chapter XXVII, Environment: 2.a Montreal Protocol on Substances that Deplete the Ozone Layer, Montreal, 16 September 1987.

Velders G. J. M., Andersen S. O., Daniel J. S., Fahey D. W., and McFarland M.: The importance of the Montreal Protocol in protecting climate, *Proc. Natl. Acad. Sci. U.S.A.* 104, 4814–4819, 2007.

Weber M., Coldewey-Egbers M., Fioletov V. E., Frith S. M., Wild J. D., Burrows J. P., Long C. S., and Loyola D.: Total ozone trends from 1979 to 2016 derived from five merged observational datasets – the emergence into ozone recovery, *Atmos. Chem. Phys.*, 18, 2097–2117, 2018.

SPECIFIC COMMENT 1 - Retrieval model

"It should be explained why the retrieval method (Eq. 8) was chosen. An obvious drawback is that this method is not invariant for spectrally-constant factors ("full correlation" or systematic errors in the absolute calibration, as hypothesised for AVODOR). It should be stated what networks / instruments use this model. For example, it is false that "This approach is consistent with the ozone measurements with Brewer" (p. 7, line 17): the Brewer algorithm is invariant for "full spectral correlation" (therefore, the "full correlation" term would not make sense with a different retrieval method, e.g. for a method where the offset is also included in a DOAS-like fit)."

Authors: Omitting an offset factor in our model was a mistake from our side, as we did not take into account how easily full correlations appear in solar UV irradiance measurements. As the results of AVODOR show, it is clearly needed. We thus modified the atmospheric model as

$$E_s(\lambda) = c \cdot E_{\text{ext}}(\lambda) \cdot \exp[-\alpha_{\text{O}_3}(\lambda, T_{\text{eff}}) \cdot \text{TOC} \cdot m_{\text{TOC}} - \tau_{\text{R}}(\lambda, P_0, z_0, \phi) \cdot m_{\text{R}}, \quad (1\text{R}) \\ -\tau_{\text{AOD}}(\lambda) \cdot m_{\text{AOD}}]$$

$$\tau_{\text{AOD}}(\lambda) = \beta \cdot \lambda^{-\alpha}, \quad (2\text{R})$$

where the parameters are as in Eqs. (3)–(7) in the AMT Discussion manuscript and new multiplier c is the scaling factor set as an additional free parameter. Thus, there are three free parameters; TOC, $\beta \geq 0$, and c to be fitted. We set $\beta \geq 0$ since to our understanding aerosols attenuate the direct solar UV spectrum. The offset factor will correct for any wavelength independent deviation in the measurements. The results of the paper change due to this modification. The large offset of AVODOR will diminish. The other devices will also have small changes in their results. We give below discussion of the changed results.

We now call $E_s(\lambda)$ as the modelled direct spectral irradiance at the Earth surface, and β as Ångström turbidity coefficient, although Ångström (1964) himself called β as extinction coefficient.

We also studied the effect of using a linear approximation of aerosol optical depth in TOC values estimated as it was used e.g. by Huber et al. (1995). In this test, Eq. (2R) was replaced with

$$\tau_{\text{AOD}}(\lambda) = a \cdot (\lambda - 340 \text{ nm}) + b, \quad (3\text{R})$$

where a and b are free parameters without bound constraints. When linear AOD model in Eq. (3R) is used with (1R), the multiplier needs to be removed by setting it to $c = 1$, because the free parameter b produces almost similar multiplier ($e^{-b \cdot m_{\text{AOD}}}$). Based on our analysis, the linear AOD model gave almost similar results as Eq. (2R), so we prefer the Ångström AOD model and an offset factor c , as we believe this approach is more physical.

"Also, since this method gives more weight to lower irradiances, the authors should carefully explain how the "wavelength region where the signal is above the noise floor" was determined (5E-3 and 1E-5 noise floors);"

Authors: It is not obvious which least squares minimization method works best, as the method of obtaining TOC from a measured spectrum is very critical to the way that the least squares fitting is performed. We have now studied different options how to perform the weighting. To facilitate discussion, we rewrite Eq. (8) in our AMT Discussion manuscript as

$$S = \sum_{i=1}^n w(\lambda_i) [E_s(\lambda_i) - E(\lambda_i)]^2, \quad (4\text{R})$$

where $w(\lambda_i)$ is the weight.

In our AMT Discussion manuscript, we used relative errors in calculation, i.e., $w(\lambda_i) = E(\lambda_i)^{-2}$. This method was also used in the article by Huber et al. (1995). This method works very well with high-accuracy double-monochromator instruments giving several orders of magnitude of useful data down to $10^{-6} \text{ Wm}^{-2}\text{nm}^{-1}$ at UV region, such as the one used in the original study by Huber et al. (1995), or the QASUME used in our experiment. However, with large zenith angles in the morning and in the evening, the spectra have to be cut at some reasonable threshold level even with monochromator-based instruments when the relative least squares minimization is used.

With array spectroradiometers that suffer from stray light and high baseline noise as presented in Fig. 1R, the dynamic range can easily be less than two orders of magnitude. Figure 1R shows the fitted spectra for all three devices, and the residuals of the fits. The residuals show clearly how the stray light is present in both BTS and AVODOR. The threshold level where the data is cut affects the TOC results at large zenith angles significantly, when the relative least squares minimization is used as demonstrated in Figure 2R.

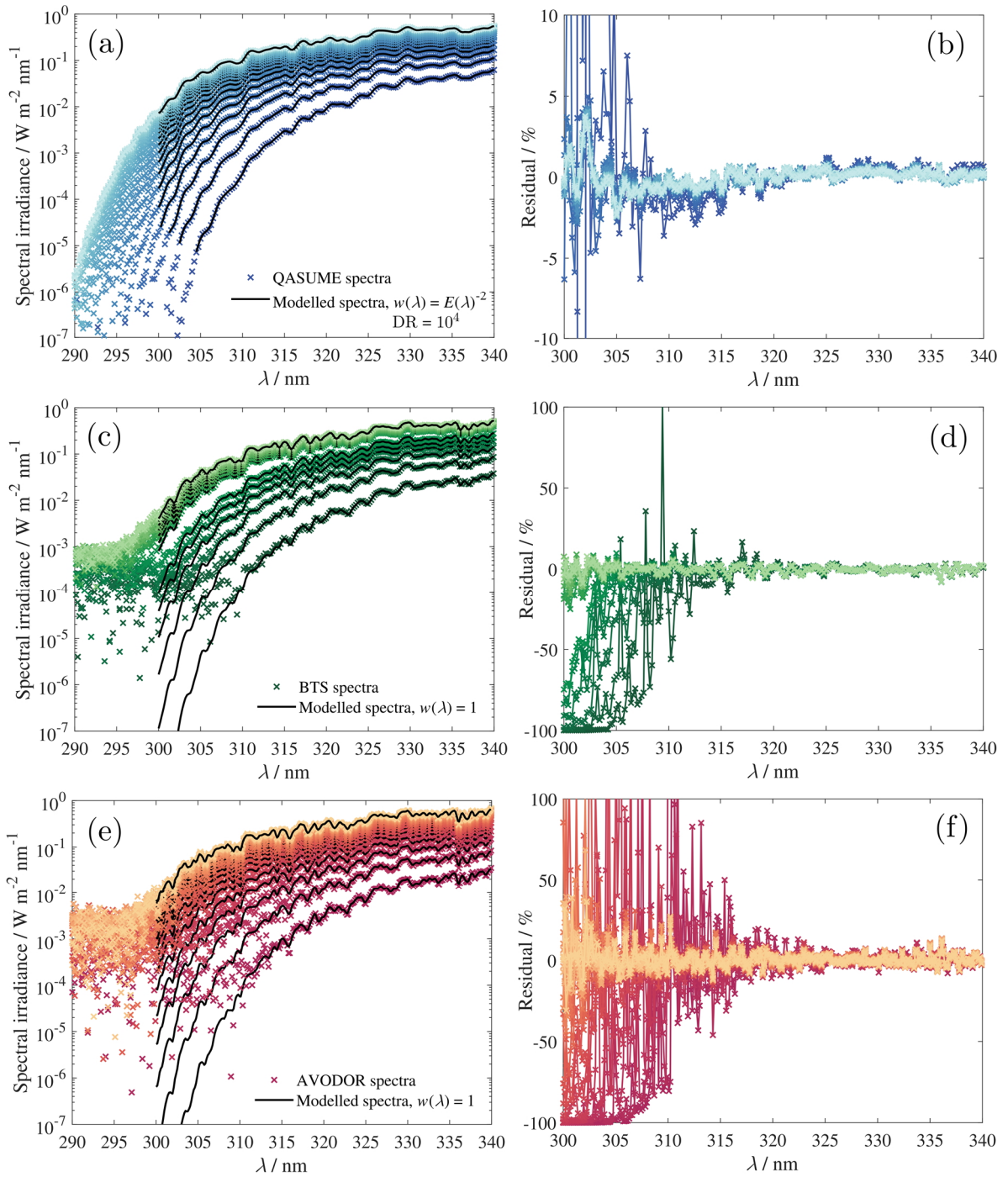


Figure 1R. Examples of fitting the atmospheric model to the direct ground-based solar UV spectra between 300 nm and 340 nm for QASUME (a–b), BTS (c–d) and AVODOR (e–f). In figures on the left hand side, the coloured symbols indicate measured spectra, and the black solid curves indicate modelled spectra. Figures on the right hand side show the relative spectral residuals of the fits. In (a), the abbreviation DR refers to the dynamic range of QASUME data used in the least squares fitting.

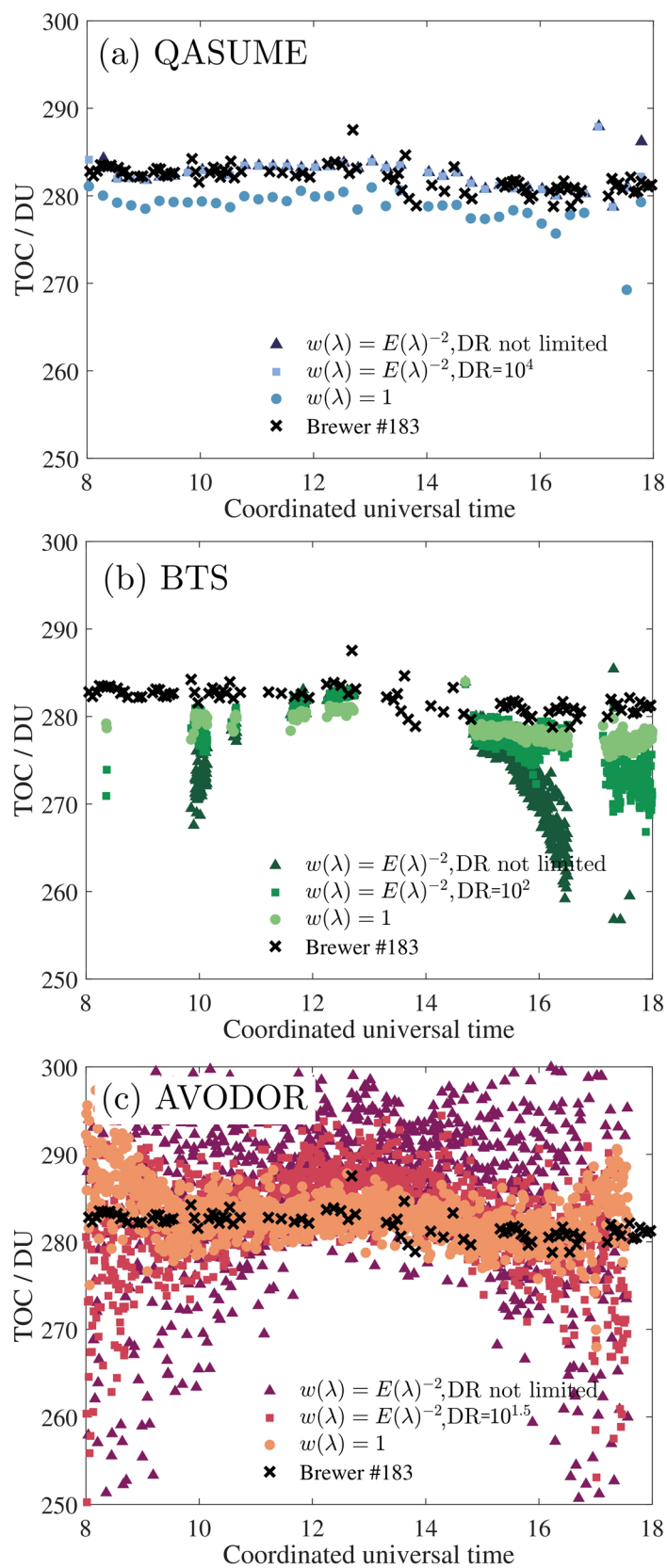


Figure 2R. TOC values estimated using different weighting in least squares minimisation and using Ångström AOD model for QASUME (a), BTS (b) and AVODOR (c). TOC values for Brewer #183 are plotted as black crosses for comparison. The colour codes and the associated figure legends denote the weighting used (Eq. (4R)) and the dynamic range DR used.

Figure 2R shows TOC values calculated from spectra throughout the day using various minimization methods. As can be seen, using relative least squares fitting causes significant inverse U-shape to the results obtained with BTS and AVODOR, which is not present in the QASUME data (triangle and square symbols). In the case of BTS and AVODOR, the TOC results change significantly by changing the threshold level. This can be seen in the curves, where the dynamic range of the data was 10^2 for BTS and $10^{1.5}$ for AVODOR. However, the stray light still causes U-shape to the curves. In our opinion, the most objective method for analysing the results of the array based spectroradiometers is to use absolute least squares minimisation by setting the weight to $w(\lambda) = 1$ (light green and orange circles in Fig. 2R). The absolute least squares minimization does not give much weight to the baseline noise tail. As can be seen in Fig. 2R, with the absolute least squares fitting there is very little variation throughout the day and we do not have to subjectively limit the dynamic range the spectra.

As we know from the results of Brewer and QASUME e.g. in Fig. 2R, the TOC was rather constant during the measurement day. Thus, the inverse U-shape notable for most curves is a property of the analysis method originating from the stray light and the baseline noise that our atmospheric model cannot account for. Herman et al. (2015) have also reported the inverse U-shape in TOC due to stray light of array spectroradiometers.

Our conclusion is that the relative least squares minimization (Eq. (4R) with $w(\lambda) = E(\lambda)^{-2}$) is only applicable to spectroradiometers free from stray light. For our paper, we choose to use this approach to QASUME with the dynamic range of four orders of magnitude. With high-accuracy instruments this is justified, because such weighting uses all data available. For both array spectroradiometers, BTS and AVODOR, we use absolute least squares minimization (Eq. (4R) with $w(\lambda) = 1$). This approach is justified, as it gives less weight to the low irradiance values distorted by stray light.

According to Fig. 2R the array spectroradiometers operate well at noon, even if analysed using relative least squares fitting. Comparing results at noon shows that using absolute least squares fitting underestimates the TOC values by 2 DU as compared to relative least squares fitting. The same offset, independent of zenith angle can be seen with QASUME.

Figure 3R shows the final comparison results. As can be seen, the inverse U-shape has diminished in both BTS and AVODOR data as compared to Fig. 5 of our AMT Discussion paper, and the agreement is rather good for all instruments.

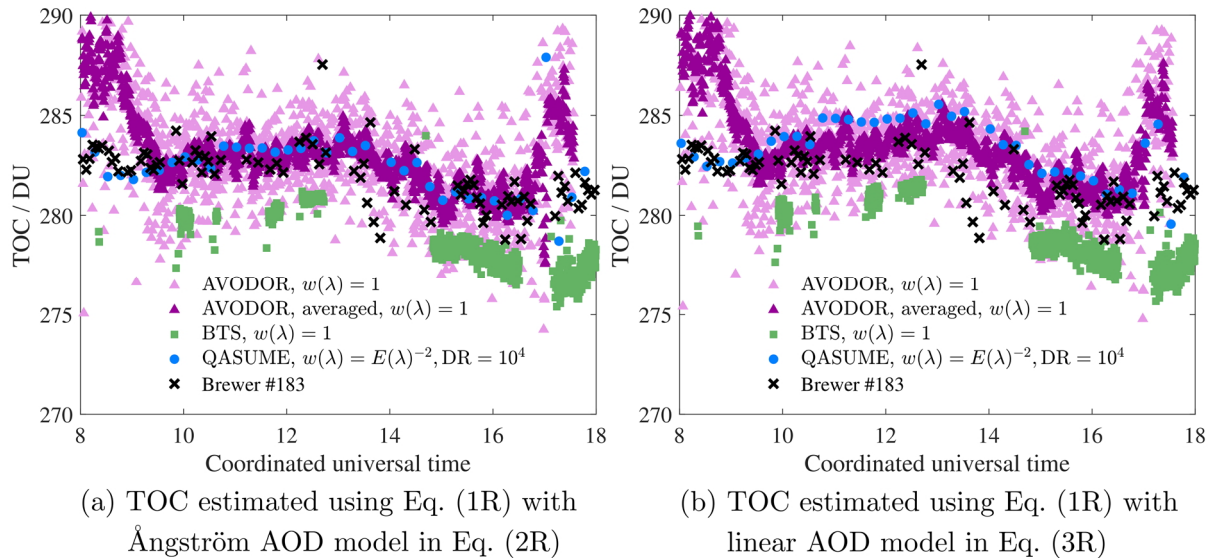


Figure 3R. Absolute TOC values estimated for the spectroradiometers studied. Average of 10 neighbouring values has been included for AVODOR to show the spectral shape behind the noise.

In our revised manuscript, we will include the presented Figures 1R – 3R, and their descriptions. Relative least squares fitting is used with QASUME using the dynamic threshold of four orders of magnitude which can be justified by the extremely low noise level as seen in Figure 1R. For the array spectrometers, we use the absolute least squares fitting. In addition, to avoid confusion we removed the sentence stating that our approach is consistent with Brewer algorithm.

A new reference was included in the manuscript:

Herman J., Evans R., Cede A., Abuhassan N., Petropavlovskikh I., and McConville G.: Comparison of ozone retrievals from the Pandora spectrometer system and Dobson spectrophotometer in Boulder, Colorado, *Atmos. Meas. Tech.*, 8, 3407–3418, 2015.

“For example, it is false that “This approach is consistent with the ozone measurements with Brewer” (p. 7, line 17): the Brewer algorithm is invariant for “full spectral correlation” (therefore, the “full correlation” term would not make sense with a different retrieval method, e.g. for a method where the offset is also included in a DOAS-like fit).”

It is true that with our revised MC uncertainty method, “full correlation” does not contribute, to the uncertainty of TOC. Full correlation in the ozone absorption cross section is an exception to this, since variable c cannot completely compensate for errors in the exponent. However, it is still meaningful to divide the components to full, unfavourable and random correlations. The part that is fully correlated does not contribute to uncertainty, thus the resulting uncertainty will be smaller correspondingly. For example, in Table 4 of our AMT Discussion manuscript, the spectral uncertainty arising from full spectral correlation of the extra-terrestrial spectrum does not contribute to the uncertainty in TOC, but it contributes to the total uncertainty of the extra-terrestrial spectrum.

SPECIFIC COMMENT 2 - Uncertainty model

"The uncertainty model (Eq. 9-10) is very complex. However, for the most part of the paper, only three terms of the Fourier series are used. Thus, I wonder why such a complex initial framework must be described."

Authors: The Fourier series approach is intended as a generic tool that can be applied to any quantity derived mathematically from a spectrum. This is explained on lines 19–20 on page 10 of the AMT Discussion manuscript. In the publication by Kärhä et al. (2018), it is demonstrated that many national metrology institutes have systematic wavelength dependent deviations in their spectral irradiance scales. The Fourier series can reproduce such spectral shapes. The main benefit of using the Fourier series is that the variance of the systematic spectral deviation can easily be controlled, as introduced in Section 4.1 of our AMT Discussion manuscript.

An updated reference:

Kärhä P., Vaskuri A., Pulli T., and Ikonen E.: Key comparison CCPR-K1.a as an interlaboratory comparison of correlated color temperature, J. Phys.: Conf. Ser. 972, 012012, 2018. doi:10.1088/1742-6596/972/1/012012

"Also, it should be explained why deviations in measurements should follow this model. Coming to the "unfavourable correlation", "The obtained TOC value is affected most by spectral distortion that mimics the spectral shape of the ozone absorption... The first combination of constant offset and one sinusoidal function ... is closest to this extreme"."

Authors: As it is explained on lines 19–20 on page 10 of the AMT Discussion manuscript, the model does not assume any particular spectral shape of the deviation. Fourier series has a property that in its full form it can mimic any shape of deviation. This takes place at the Nyquist criterion, where N is equal to half the number of wavelengths available. Also with smaller values of N , the MC parameters can account for unknown spectral shapes of lower complexity.

We sum the base functions starting with the simplest forms, $N = 1, 2 \dots$. The logic here is that when one studies which kind of correlations there are present in typical measurements, correlations such as these appear commonly. This can be seen in (Kärhä et al. (2018)) and (Woolliams et al. (2006)) which examine deviations in the spectral irradiance measurements of national standards laboratories.

A reference:

Woolliams E. R., Fox N. P., Cox M. G., Harris P. M. and Harrison N. J.: Final report on CCPR-K1-a: Spectral irradiance from 250 nm to 2500 nm, Metrologia, 43, Technical Supplement 02003, 2006.

"However, from Eq. 10, this term varies with λ_1 and λ_2 , so the width of the sinusoid is different for the three instruments and comparison of the results is difficult."

Authors: This is true. We now changed our analysis so that all instruments use the same wavelength region 300 nm – 340 nm and remodelled the results.

"Also, why should this term be a full sinusoid cycle? Why should the period of the oscillation be the same for all uncertainty components (calibration, measurement, cross sections, etc.)?"

Response letter to Referee #2
Article reference: amt-2017-403

Authors: As it is explained on lines 22–25 on page 18 of the AMT Discussion manuscript, simpler forms of deviations could be used after the general solution of the problem is known.

In our method, adding half a sinusoid or a slope as a base function in the general solution would lead into more complicated calculations when determining the rest of the base functions, as all the base functions need to be orthogonal with respect to each other. It is possible to use other orthogonal sets of functions instead of sinusoidal functions, but that would involve more complicated mathematics.

“Finally, it is obvious that ϕ has an enormous role for the "unfavourable correlation" term: depending on ϕ 's value (i.e. similarity with ozone spectral cross section), the induced error could be huge or negligible. The role of ϕ should be explained better. Physically, what does ϕ represent? It is expected to randomly change in one instrument or not?”

Authors: Parameter ϕ_i is used as an equally distributed MC variable between 0 and 2π (see page 9 lines 29–30 of the AMT Discussion manuscript). It is needed in the Fourier series so that it can reproduce all possible spectral shapes.