Review of the manuscript “A physically-based correction for stray light in Brewer spectrophotometer data analysis” by V. Savastiouk et al.

The paper describes an easy to use new procedure to take into account spectral stray light in the derivation of total column ozone and total sulphur dioxide from measurements of solar irradiance by single Brewer spectrophotometers.

The method uses a single parameter to quantify the impact of stray light on total column ozone (and a second parameter for the SO2 retrieval), which is specific to a Brewer spectrophotometer. These parameters are usually retrieved by comparison to a reference instrument, that is not sensitive to stray light (e.g. a double Brewer for example, or an already corrected single Brewer), even though a method is explained how these parameters can be retrieved without external reference instrument. A detailed uncertainty budget is also provided.

The method is validated by applying it to several single Brewer instruments, that are collocated with a double Brewer used as the reference. Furthermore, a theoretical justification of the stray light correction methodology is provided by modelling the stray-light impact from laboratory based slit function measurements for one Brewer instrument (Brewer #009).

The authors are well-known scientists, and experts of the Brewer spectrophotometers. The paper is very well written and understandable to a broad audience (in my opinion). Prior work is mentioned as needed, and the references are comprehensive and complete. Figures, tables are all of excellent quality.

I commend the authors for this paper that will definitely have an impact on the data quality of the Brewer network.

In my opinion, the paper can be accepted as is, even though I list a number of points below which the authors might want to take into account, or at least answer in their response to this review:

We thank the reviewer for taking the time to revise our manuscript and for their constructive comments.

Our point-to-point reply is given hereafter.

1) The authors claim a “physically-based correction for stray light”. I assume that they base this claim on the fact that they introduce a stray-light correction concept based on slit functions, and validated through a physical model of the Brewer. However the validation is really only qualitative, since the retrieved stray-light factor alpha is off by a factor of 2 when comparing the value obtained from the simulation with the actual factor obtained from the outdoor comparison using real measurements. The authors claim that the discrepancy comes from using outdated slit function measurements of this particular Brewer (#009), and in the meantime this Brewer has changed. The comparison between simulation and measurements would have been more convincing if the authors would have been able to use a Brewer where the slit functions are known and could be trusted. In the current version of the manuscript, the discrepancy leaves the argument hanging. Without a better validation, this method is also only sort of empirical, where the physical justification is well argued, but not substantiated.

The reviewer's point is noted, however we stand by our assertion that PHYCS is physically-based. We call our algorithm physically-based because the mathematics of it is based on the physical principle behind the instrumental internal stray light, namely the additional, "stray", light that contributes to the detected count rate at every measurement. Also, in contrast with other attempts to correct for the stray light in the Brewers, our algorithm only uses measured quantities, i.e. the count rates detected by the counting system. The algorithm does not introduce "external" or calculated values like the air mass factor as many other proposed corrections do. As such, we believe calling our algorithm physically-based is justified.

The absolute values for the correction factors were not the main result of our modelling of the stray light in the Brewers. The most important takeaway from the simulations is the finding that the stray light contribution, in count rate space, is relatively constant across the measured wavelengths, making it possible to apply it in a simple way. This is because the largest contribution to the stray light is coming from longer, less affected by ozone and thus much brighter wavelengths. By applying this correction to experimental data we confirmed that this is working very well.

The laser scan that we used for Brewer #009 shows the "wings" at a level of 10^{-4}. An offset of 10^{-4} was added to the laser scan. With that small addition in absolute sense, the model predicts parameter alpha to be 0.4%, same as our experimental data shows. Please also refer to the response for your point 6), where we describe other possible shortcomings in the laser scanning process.

2) line 304: "orthogonality between alpha and ETC": Without clear definition of this term, it sounds a bit colloquial, and I suggest to replace it by a more precise statement. For example the term "correlated" or "uncorrelated" could be used here.

We agree with the reviewer. Modified: "Since alpha and ETC are uncorrelated, the two unknowns can be determined independently without significant cross-talks"

3) Line 318 : "the results show that Langley plots with measurements affected by stray light make little sense" is a worrying statement, considering that historically, the Brewer network traceability was based on a triad of single Brewers. Could the authors comment on the potential impact to the historical datasets of the Brewer network, and if it would be necessary and possible to apply this stray light correction procedure retrospectively to the whole network?

Thank you for this important comment. The accepted uncertainty of Langley method when applied to Brewers, is +/- 5 units in ETC (roughly +/- 1.5 DU when solar zenith angle is 0 and less when sza>0). We estimate that the error in ETC from the Langley method due to stray light is 10 to 15 units, making this contribution significant enough to be of concern when discussing the quality of calibrations of the primary standards, but still its effect on the ozone data is less than 1% at air mass factors > 1.5, which most of the Brewer data are. We do not see this error warranting recalculation of past absolute calibrations of the primary standards. However, as we mention in the paper, calibration transfers to field instruments were significantly affected by stray light and definitely deserve attention of the Brewer community.

What we wanted to stress by our statement, is that now that a simple algorithm for correcting for stray light exists, there is no justification for not applying it to the data, including those data that are used for the Langley method calibrations. We have replaced the wording "makes little sense" with "lead to errors in the determination of the ETC".

It is indeed not only possible, but also imperative to reprocess past calibration data and then apply our algorithm to past data from all Brewers. While our algorithm is easy to implement, the sheer volume of the Brewer data makes this task both important and time-consuming, requiring additional resources to do it right. We would like to also stress that not all data centres are able to track/identify changes in the submitted data, which will lead to challenges for the data end users when some data have been corrected and some have not.

A comment about the need for calibrations and data reprocessing has been added to the Discussion section.

4) table 3 . it would be helpful to briefly describe the terms of the first column in the table caption. For example I do not understand what DeltaETCO3 stands for.

We appreciate your pointing this out. The following text was added in the table caption: "The rows report, respectively: the serial number of the reference MkIII Brewer, the α and β coefficients used in the correction, the ozone extra-terrestrial coefficient after the application of the stray light correction, the original extra-terrestrial coefficient used in the Brewer, the difference (ΔETC) between the two values above, and the minimum air mass factor available in the analysis."

5) In the uncertainty budget, Table 5, the contribution labelled "Mathematical model", is probably more correctly named "residuals", as the residuals between retrieving the model parameters by fitting the model to the measurements?

"Mathematical model" is the wording used by the Guide to the expression of uncertainty in measurement (GUM; BIPM et al., 2010). The word "residuals" has been added in brackets to avoid misinterpretation.

6) the discussion S1 in the Supplement, and referred to at lines 705-708 on the correctness of using measured slit functions as a basis for a stray-light correction is not very convincing. The authors are correct in stating that the stray light measured through the various slits of the Brewer spectrophotometer can differ between themselves, and that the common orientation of the grating could produce a common feature in the slit functions. However this does not mean that there is some additional mysterious effect that could produce additional stray light that would not be captured by an accurate slit function measurement. To avoid any misunderstanding, I would recommend to rewrite this part of the manuscript (or to skip as it has no impact on the paper itself).

We appreciate the reviewer's position and would like to clarify ours. The issue here is that in order to have "accurate slit function" it needs to be accurately and precisely measured for each of the 6 slits used in the Brewer and it has to be measured at the operating wavelength at each of the slits, which can only be done using a tunable laser as we state in the paper. The scan should preferably be made with a single pass for a slit. However, the dynamic range of the Brewer counting system cannot cover a factor of 10^9 that is needed for an accurate scan of a laser line peak and the wings at the same attenuation level, even if a tunable laser is available.

Our comment about some features that correspond to a particular diffraction grating angle is aimed to stress the sometimes significant differences between instruments, when some Brewers potentially have reflective surfaces that contribute to stray light, but are difficult to characterise in a model. While this effect is not mysterious, and has been detected in several instruments, it poses challenges that are not easily overcome in modelling. PHYCS, when applied to experimental data, works well for Brewers that show this effect, which is what is important in practice.