Response to Referee #2:

We thank referee #2 for their very helpful comments. Our responses are given below in black with the referee's comments in blue. The new text in the modified manuscript is given in red (italicized).

General comments:

The manuscript "The site-specified primary calibration conditions for the Brewer spectrophotometer" by Xiaoyi Zhao et al., aimed to answer two primary questions related to calibration of the Brewer spectrophotometers. The first question addresses site-specific factors, i.e., why the calibration procedure will not work well at certain locations. The second question, closely-related to the first question, deals with the required conditions to achieve a certain calibration quality.

The answers to these questions are important to the assessment of the measurement quality of the ground-based Brewer (and Dobson) measurement networks, and therefore important to Ozone research. The use of available auxiliary data and the development of a modelling framework using MERRA-2 to answer these questions are done in an innovative and convincing way.

The manuscript is well written and well presented. The methods employed are scientifically robust and they are clearly explained. The figures are clearly illustrated, except for one or two that can be easily improved. This manuscript fits within the scope of AMT. Therefore, I recommend its publication after addressing the comments of Reviewer #1 and some of the minor comments below.

We appreciate the very positive comments from the referee on our work.

Specific comments:

1. The Authors briefly mentioned the ETCs from the ICF (Instrument Constant File). I think it should be clarified in a few words what this is, and how the ETC values in this file are obtained.

We have included a more detailed description as suggested.

> *The documented ETC values from the instrument calibration file (ICF) are shown as green lines with their validation period indicated by vertical black dash lines. Note that these ICF ETCs are the numbers used in each Brewer's ozone data production. Here, the world and regional references instruments' ICF ETCs were acquired via PCM (e.g., Kerr, 1997), while the other instruments' ICF ETCs were acquired via CTM during calibration campaigns (e.g., Redondas et al., 2018a).*

2. Related to comment #1 above, it is my understanding that Brewer spectrophotometers have internal quartz-halogen lamps, which are used for the purpose of monitoring instrument stability and changes in ETC values. It would have been interesting to know the results from these regular lamp tests and how the

lamp test results can be employed to support the calibration method. I would have preferred if there was a discussion about it.

We thank the referee to point out this technical detail, which most non-expert readers would not find. We avoided such details intentionally to make the paper a bit more easily to be absorbed.

The internal halogen lamp test is the SL (standard lamp) test. During such a test, the micrometer position is kept at its operating position for ozone/SO_2 measurements, and the intensities at all six wavelengths are recorded by observing with the lamp's light (not the solar light). The same instrument responses (lamp response F values; F_{lamp}) are calculated with these lamp observations. Such ΔF_{lamp} values (F_{lamp} at calibration subtract F_{lamp} at measurement) are used to adjust the instrument response F values (i.e., $F_{adjusted} = F - \Delta F_{lamp}$). Such "SL correction" is included in both ECCC and AEMET Brewer algorithms (e.g., Fioletov et al., 2005; Savastiouk, 2006; Redondas et al., 2018). This means that the reported TCO data used in this work have already accounted for such SL correction. The instrument responses (F; e.g., see Eqns 3–5 in the revised manuscript) used in the Langly fittings also included the SL correction. Note that, this is equivalent to the expression of "adjusted ETC" for the SL test (Fioletov et al., 2005; Redondas et al., 2018).

In short, the SL test results are already included in this work; SL test can be used to monitor the instrumental stability, which typically has small variations due to short-term instrumental changes (e.g., due to the instrument's ambient temperature changes and other factors, see Fig. R7 as an example of SL correction changes for Brewer #119). However, the SL test results should be interpreted carefully, as the values depend on many factors that do not only come from instrumental issues.

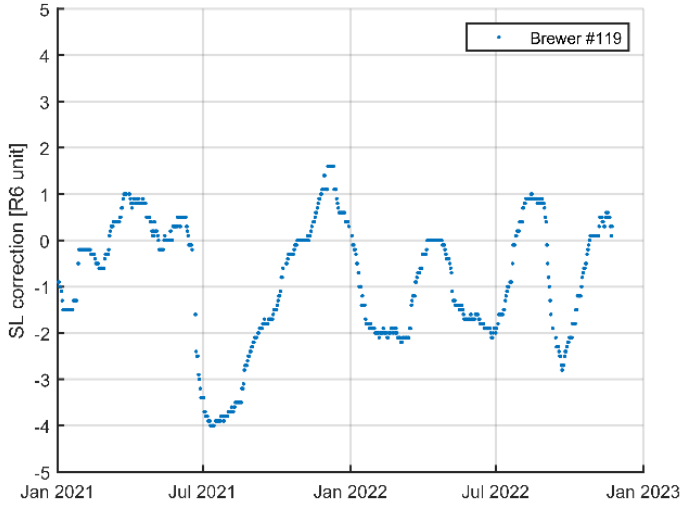

Figure R7. SL correction values from Brewer #119 at MLO from 2021 to 2022.

We included some of this information in the revised manuscript.

> *Here R6 is a measurement-derived double ratio in the actual Brewer processing algorithm, corresponding to the measured slant column ozone (e.g., Savastiouk, 2006; Zhao et al., 2021). Note that in typical conditions (e.g., $\Delta\alpha$ = 0.34 and μ = 2), 5 R6 units is equal to 0.74 DU or about 0.25% for a typical ozone value of 300 DU. Also, in this work, the so-called SL (standard lamp) corrections are implemented in the TCO calculation (Fioletov et al., 2005; Savastiouk, 2006; Redondas et al., 2018). The SL corrections are based on internal lamp tests and they compensate for changes in the instrument characteristics that lead to changes in the instrument's ETC and the Langley fitting results (i.e., the instrument responses (F values) used in Langley fits, have the SL correction included). In general, Fig. 1 shows that the selection of fitting equations will not impact the agreement of the final averaged ETC, as long as an adequate number of individual ETCs are included.*

3. "short-term ozone field": It took me a while to decipher its meaning, perhaps it is a modelling term. I would recommend clarifying it when the term first appears on page 1.

Done. We have modified the sentence in the abstract (on page 1).

> *In practice, these two calibration methods have different physical requirements, e.g., the PCM requires a stable ozone field in the short-term (i.e., half-day), while CTM would benefit from larger changes in slant ozone conditions for the calibration periods.*

Technical Corrections and Suggestions:

As Reviewer #1 already mentioned, "site-specific" is more appropriate than "site-specified", which appears in many places in the manuscript.

Done.

P.5, Line 142: "are can be found" -> are found OR can be found

Done.

> *More details about Mark II and III measurements and other characteristics ~~are~~ can be found in Zhao et al. (2021).*

Fig. 2. Among the nice figures, this is probably the only one I find difficult to decipher. It is an important figure, perhaps this could be improved. The panels are too small. Also, I would suggest to avoid using red and green markers in the same figure.

Done. We have modified the figure as suggested. The figure is laid out such that each colum represents a group of instruments at one site (e.g., the first column is for instruments at Arosa-Davos). If needed, maybe

the text editor could help us to rotate the figure 90 degrees to make it a full-page figure.

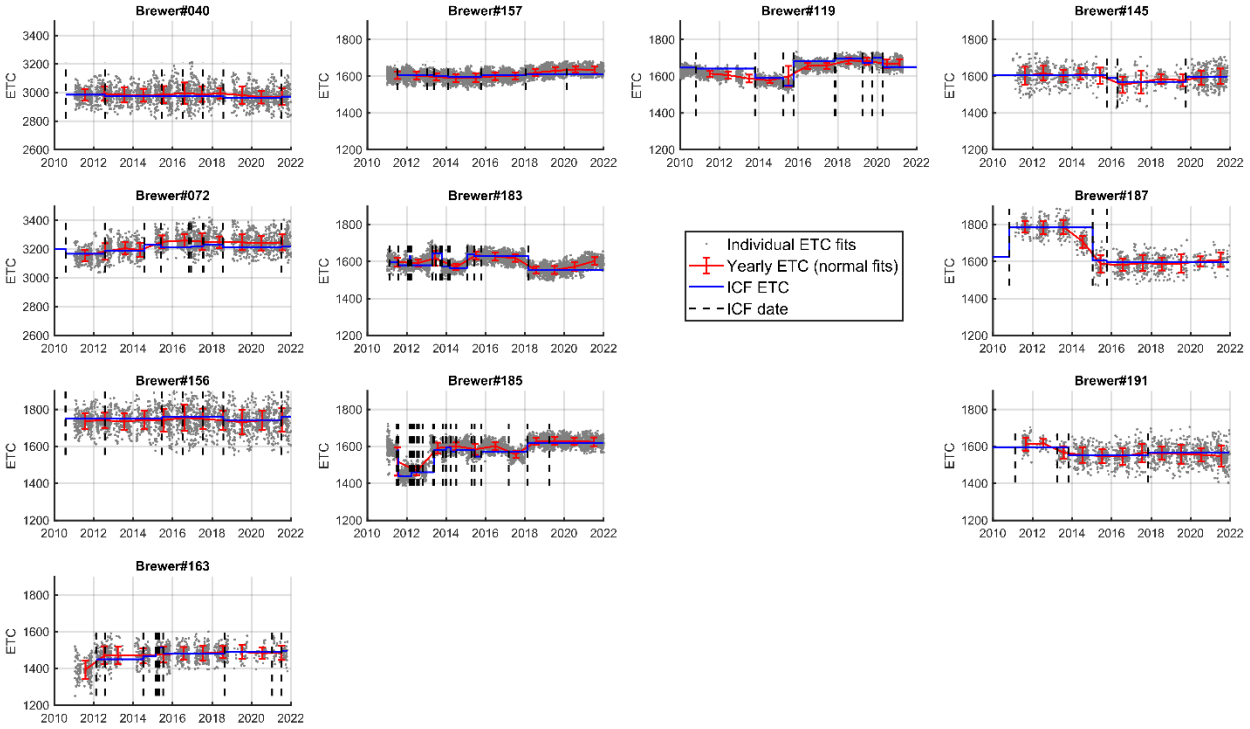

Figure 2. Long-term Langley fits of Brewers from Arosa/Davos (Brewers #040, #072, #156), Davos (Brewer #163), Izaña (#157, #183, #185), MLO (#119), and Toronto (#145, #187, #191). Gray dots are fitted ETCs for individual half-days, red lines are yearly mean values of these individual ETCs (error bars are 1σ values), and blue lines are ETC values obtained from instrument calibration files (ICF; validation periods are indicated as vertical black dash lines).

P. 13, Lines 347-348: This (second to the last) sentence needs to be revised or rephrased.

Done.

> *For low-latitude sites (Izaña and MLO), the individual ETCs (gray dots) have less variability and are more closely distributed around the yearly means (red lines) and ICF ETC (blue lines) values.*

P. 23, Line 599: "… the Brewer spectrophotometer taking the most accurate TCO observations among ground-based instruments …". This claim seems like the Brewer has been validated against a "true" measurement of TCO, and then compared with other ground-based instruments, which are also validated against a "true" TCO. If so, please provide references.

Thanks for pointing this out. We agree with the referee that there is no instrument that has been compared with a "true" TCO. We have modified the sentence.

This is simply due to the Brewer spectrophotometer taking the most reliable TCO observations among ground-based instruments (precision within 1%, corresponding to about 3–4 DU in typical TCO conditions).

References: Multiple papers of some authors are not in chronological order, e.g. papers by Kerr et al.

Done.