Anonymous Referee #1

The author would like to thank the reviewer for the constructive comments and suggestions that were taken into account. Please find our response to your comments (in bold). The responses are listed below each question. We have made changes to the original manuscript and the changes were written in red italics. We have also included a version of the paper with all changes highlighted.

General comments: This paper provides comparison of three data retrieval software available to Brewer spectrophotometer users. As the data retrieved with different software might sometimes be compared to each other or used side by side without knowing the original software used, the information about any possible biases or false trends or other discrepancies any algorithm might produce is important. This is especially true for Brewer instruments that are, together with Dobson spectrometers, the most reliable source of total ozone column data. The way each software tracks changes and drifts in the instrument are considered in the paper. Mean biases in comparison to other software are defined. However there is no discussion if using specific software could affect the trends in any way.

The analysis of the trends was included. The additional sections on that issue were added in Method (2.6) and Results (3.4):

2.6 *Trend analysis*

To assess whether a specific software could affect the trend, we estimated the trend from the annual mean anomalies. We applied the same methodology proposed by Fountoulakis et al., (2016). Climatological ozone values for each day were calculated over the period under study. The daily anomaly with respect to the daily climatological value was calculated. Afterward the monthly anomalies were determined by averaging the daily anomalies for each month provided that at least 15 days of data were available. Finally the monthly anomalies were averaged to determine the annual mean anomalies. The trend among the three codes was expressed as a percentage variation per decade and used in their comparison. The statistical significance of the trends is derived from the Mann–Kendall test with statistical significance set at p≤ 5%.

3.4 Comparison among the trends estimated by the three processing software ozone retrievals

The detected trends in ozone series calculated by using the three processing software are reported in Table 6. The trends were quantified over the period 2005-2015 for Rome to be consistent with the EUBREWNET ozone data coverage, and 2007 -2015 for Aosta. Ozone data which showed large differences among the codes were not included in the trend analysis.

The QBO and solar cycle effects were not filtered in the ozone series. The former was found small at the mid-latitude station (Fountoulakiset al., 2016), whereas the latter was not taken into account due the length of the analysed ozone series (< 11 years). All trends were found to be not statistically significant (p-value is 0.05).

It is clear from Table 6 that there are not significant differences in the trends among the three codes, when data affected by rapidly changes in R6 or the spectral response of the instrument shows a persistent drift, were removed.

***Table 6**. The total ozone linear trends derived by the processed ozone values using three different processing codes*

	period	*BPS (% per decade)*	*O3Brewer (% per decade)*	*EUBREWNET (% per decade)*
Rome	*2005-2015*	*-0.23 ± 0.18*	*-0.32 ± 0.20*	*-0.34 ± 0.21*
Aosta	*2007-2015*	*0.07 ± 0.35*	*0.04 ± 0.34*	*0.00 ± 0.38*

While the differences between the software are quite nicely quantified, the analysis does not go very deep into thinking what could cause the differences. Only source of difference considered in more detail was the way to apply standard lamp test information. The standard lamp test being the way to follow the changes in the instruments spectral sensitivity that affects the ozone retrieval significantly. However even for this variable there was no quantification of its effect; if it explains all the difference or not. Also using only the daily averages produced by the software, the information is lost if the differences are due to different way of selecting "good" measurements or because of something that happens when processing a single measurement. There should be more detailed analysis if the standard lamp correction makes all the difference or if there is more reasons and what the reasons could be. Also comparison of data rejection rules is required when comparing daily average values.

We decided to rework the paper performing additional analysis on both daily means and individual calculated ozone values to better investigate the differences found among the three processing software. The last version of BPS and O3Brewer was applied taking into account the same rejection criteria on ozone values used by EUBREWNET, i.e. maximum standard deviation of 2.5 DU and maximum ozone air mass of 3.5. TOC.

This issue was specified as follows in Section 2.3 Measuring instruments and sites:

In this study we analysed individual DS values and daily averages of Rome and Aosta stations, generated by BPS version 2.1.1 updated to 2017/02/14 (Fioletov and Ogyu, 2007), by O3Brewer software packages version 6.0 updated to 2018/03/14, and EUBREWNET level 1.5 ozone products. Level 1.5 individual TOC values are discarded when the standard deviation is above 2.5 DU and the maximum ozone air mass is above 3.5. In addition ozone values less than 100 DU and greater than 500 DU are also rejected. The stray light correction was not applied because it requires a calibration against a double monochromator Brewer and an instrumental characterization (Karppinen et al., 2015, Redondas et al., 2016) which was not available. Level 1.5 TOC values were downloaded from EUBREWNET platform over the period 2005-2015 at Rome and 2007-2015 at Aosta.

We set in the configuration file of BPS and O3Brewer software, where it is suitable, the same rejection criteria used in EUBREWNET, i.e. maximum standard deviation of 2.5 DU and maximum ozone air mass of 3.5. TOC.

The rejection criteria on ozone values are hardcoded and consist on three sequential checks: 1) if raw counts are less than 2500, the value is rejected; 2) if calculated ozone for DS/ZS is less than 50 DU, the value is rejected 3) if observation is in the DS mode and the calculated ozone is between 50 and 100 DU, the value is rejected (Ogyu, personal communication 2018).

The maximum calculated ozone is indeed configurable in the BPS setup and was set to 500 DU.

The limits on the calculated ozone are not configurable in the O3Brewer setup. In the latest version used in this study, the standard lamp maximum for applying of ETC correction from SL test results is now configurable. Here we used the default limit of 500 units for the difference between R6 and the reference $R6_{ref}$.

I found the paper quite well structured in general but there were some irregularities that are highlighted in specific and technical comments. The language was heavy to read at times, when too much information was being compressed into a single sentence. This was highlighted by extensive use of parentheses.

We eliminated the use of parentheses as much as possible.

Specific comments:

Abstract line 32: if the difference between software is in order of the instrument uncertainty is it a good result? I would expect different software that calculate the same thing to be well within the uncertainty of the measurement itself.

The paragraph in the abstract was rephrased:

The overall agreement of the BPS and O3Brewer TOC data with EUBREWNET data is very good and within the estimated total uncertainty in the retrieval of total ozone from Brewer spectrophotometer (1%). However differences can be found depending on the software in use. Such differences become larger when the instrumental sensitivity exhibits a long-term drift which can affect the ozone retrievals significantly. Besides that reason, if daily mean values are directly generated by the software, differences can be observed due to the configuration set by the user to process single ozone measurement and the rejection conditions applied to data to calculate the daily value.

page 3 line 51 inaccurate phrasing, maybe "...to measure ground level spectral intensities of solar ultraviolet radiation attenuated by ozone absroption. Form these spectra it is possible...

The phrase was modified as:

The most common ground-based instruments to measure TOC are spectrophotometers which are designed to measure ground level spectral intensities of solar ultraviolet radiation attenuated by ozone absorption. From these spectra, it is possible to retrieve the TOCs.

line 108 "by measuring irradiances of the direct sunlight,..." there are also measurement mode for focused sun (Josefsson, W. A. P. (1992), Focused sun observations using a Brewer ozone spectrophotometer, J. Geophys. Res.,C297(D14), 15813–15817, doi:10.1029/92JD01030.) Perhaps the other modes are entitled to some reference if the global irradiance one is?

The reference Josefsson, W. A. P. (1992), was acknowledged as well as those concerning the other modes (zenith sky light and the moon light) and the references were also acknowledged:

The Brewer instrument is a spectrophotometer designed to retrieve the total ozone column by measuring irradiances of both direct sunlight (Kerr et al., 1981) and polarized radiation scattered from the zenith sky (Brewer and Kerr, 1973, Muthama et al., 1995). Total ozone can be also derived from focused sun measurements, commonly employed at high latitudes (Josefsson, 1992). It is also possible to measure total ozone by using the moon as a light source (Kerr et al., 1990), or the global irradiance method (Kerr and Davis, 2007) in the UV region.

Brewer, A.W. and Kerr, J. B.: Total ozone measurements in cloudy weather, Pure appl. Geophys. 106-108, 928-937, 1973.

Josefsson, W. A. P.: Focused sun observations using a Brewer ozone spectrophotometer, J. Geophys. Res.,97(D14), 15813–15817, doi:10.1029/92JD01030,1992.

Kerr, J.B., McElroy C.T., Wardle D.I. and Dorokhov V.: Measurements of arctic total ozone during the polar winter, Atmosphere-Ocean, 28:4, 383-392, 1990.

Kerr, J. B., McElroy, C. T., and Olafson, R. A.: Measurements of total ozone with the Brewer spectrophotometer, in Procs. of the Quadrennial Ozone Symposium, edited by J. London, 74–79, Natl. Cent. for Atmos. Res., Boulder, Colo.11., 1981.

Kerr, J.B., McElroy C.T., Wardle D.I. and Dorokhov V.: Measurements of arctic total ozone during the polar winter, Atmosphere-Ocean, 28:4, 383-392, 1990.

Muthama N.J., Scimia U., Siani A.M., Palmieri S.: Toward optimizing Brewer zenith sky total ozone measurements at the Italian stations of Rome and Ispra, J. Geophys. Res.,100, 3017-3022, 1995.

line 113 Don't both these papers conclude the ds accuracy of 1%?

We agree and it was specified in section 2.1 Theory of direct sun measurements with Brewer spectrophotometer that:

It was shown (Vanicek, 2006) that the accuracy of measurements taken with a well-maintained Brewer spectrophotometer is 1% in the DS mode and 3-4% in the ZS mode. The random errors of individual measurements were found to be within ± 1% for all measurements (Fioletov et al., 2005).

line 118 the slit information is very specific so you need to introduce the operation of the slitmask before. Maybe leave out the specific slit number and just say wavelengths are selected by rapidly rotaing slitmask and photon counts are registered by a photomultiplier.

We added that: *The wavelengths are selected by a rapidly rotating slit-mask and raw photon counts for each slit-mask wavelength position are registered by a photomultiplier. During each measurement run cycle the slit-mask is rotated 20 times. The operational wavelengths correspond to slits 2 to 5 in the rotating slit-mask.*

line 120 Maybe highlight that dark count and dead time are characteristics of the photomultiplier to help people not familiar with Brewers to have a clue what these are.

We agree with the suggestion and we included that: *The raw photon counts are then converted into count rates and are corrected for the characteristics of the photomultiplier (dark count and dead time) and for the internal Brewer temperature (Kerr, 2010).*

line 124 you have only introduced four wavelengths so no need to say "longer" here line 143 suggested change "weighted ozone absorption coefficient" to" differential ozone absorption coefficient"

“weighted ozone absorption coefficient” was replaced by “*differential ozone absorption coefficient*"

line 162 change to "conditions with small..."

done

line 184 mentioning the slit mask here also makes it even more important to introduce its meaning earlier in the text

We agree and the meaning of the slit mask was introduced.

line 207 Highlight that the reference value is determined at every calibration.

The above suggestion was inserted

line 211 Hard to interpret but i think i finally understood it. Suggest to give out the normal case first (abs(r6ref - r6) <= 250 units).

We rephrased many parts of the section (2.1) describing the SL correction applied by the processing software, as follows:

Depending on the processing software used by the station operator, ΔSL is computed in different ways, not always clearly explained by the software documentation:

•*In the BPS, the reference value* $R6_{ref}$ *is determined with a triangular smoothing filter of SL-test values over the 15 day period immediately following the calibration date. There should be at least one good SL-test value per day. If the corresponding B-files are not available, the program is not able to establish the reference SL level and the ETC will be not adjusted. Notice that for other processing software* $R6_{ref}$ *is based on the SL-test values during the calibration campaign. If the abs(*$R6_{ref}$ *-R6)* ≤ 250 *units, then the median of daily averages from all R6 data before 15 days and after 15 days for a particular day is used for the correction. The median is used because it is less influenced by single invalid R6s. If the abs(*$R6_{ref}$ *-R6) is above 250 units then ETC is adjusted taking into account the difference between the* $R6_{ref}$ *and the present daily mean values of R6.* That correction is reported in the file named "o3data" produced by the BPS. The threshold and the time window are however not adjustable by the users (Fioletov personal communication, 2018).

•O3Brewer adjusts the ETC using a Gaussian smoothing filter on R6 values (Stanek M., 2016). *There should be SL measurements 10 days before and 10 days after the selected date period*-The software creates the smoothed R6 time series (hereafter named $R6_{smooth}$) which is used for ETC adjustment. *It means that there should be at least one SL test per day. There was a limit between R6 and the reference* $R6_{ref}$ *for applying of ETC correction from SL test results which is configurable in the latest version* (Stanek personal communication, 2018). The time window is however not adjustable by the users. If this difference exceeds the threshold, *then the software can remember the last day with good SL test and will apply that correction (*Stanek personal communication, 2018). This option can be turned off and then the daily mean values for SL are used for the correction of the ETC.

•Level 1.5 total ozone column data from EUBREWNET are recalculated with the ΔSL correction determined applying a triangular moving average over the daily median values of R6 within a seven days window (default time window). The correction is applied if the difference between $R6_{ref}$ and the calculated value exceeds 5 units. Level 2.0 are 1.5 observations validated with a posterior calibration. If the reference constants of a posteriori calibration do not differ significantly from the values in use then level 1.5 products are not reprocessed and represent the most reliable products (http://rbcce.aemet.es/dokuwiki/doku.php).

At the present time, tools for Level 2.0 are developed but not yet implemented. A complete description of the processing can be found on the EUBREWNET website (2017).

The way BPS determines the r6 reference value may already introduce offset as for others the r6 is given by hand after the calibration based on the sl test values during calibration campaign. Offset propably very small though but should be looked into.

Thanks to the reviewer for this further important remark which was analysed. Section 3.1 was completely re-written and also attached at the end of this document. Concerning the BPS offset we included the following paragraph:

The discrepancy between the two codes could have been caused by the offset introduced by the way BPS determines the R6 reference value as for the other code the $R6_{ref}$ *is obtained during the calibration campaign and set manually in the configuration. The BPS* $R6_{ref}$ *is computed with a triangular smoothing filter of SL-test over the 15 day period after the calibration and it is calculated "on fly" from daily mean SL values and it is not stored (Fioletov, personal communication 2018). To look into the possible effect of the BPS offset we estimated* $R6_{ref\_BPS}$*, for each day over the 15 days after the calibration by subtracting the correction (reported in the file o3data.txt) to the corresponding R6 value. Then the average over the 15* $R6_{ref\_BPS}$ *values was compared with* $R6_{ref}$ *(given by hand after the calibration). The estimated offset introduced by BPS with respect to* $R6_{ref}$ *is very small, ranging between -19 to 6 units at Rome and between -10 to 2 units at Aosta. Consequently the BPS offset appears not to be responsible for the ozone difference that can be attributed to the calculation method of the standard lamp correction.*

line 228 There is a lot of information here not relevant of how the sl test is introduced in EUBREWNET algorithm.

Only relevant information about Sl test was left, see above our previous answer to this issue.

These differences of processing software specific rejection rules should be stated especially where they differ but this is not the right position for them as this paragraph was supposed to be about sl-test. Could you add data rejection criteria to a more suitable place in text?

As suggested the rejection rules were moved and included in Section 2.3

Level 1.5 individual TOC values are discarded when the standard deviation is above 2.5 DU and the maximum ozone air mass is above 3.5. In addition ozone values less than 100 DU and greater than 500 DU are also rejected. The stray light correction was not applied because it requires a calibration against a double monochromator Brewer and an instrumental characterization (Karppinen et al., 2015, Redondas et al., 2016) which was not available. Level 1.5 TOC values were downloaded from EUBREWNET platform over the period 2005-2015 at Rome and 2007-2015 at Aosta.

We set in the configuration file of BPS and O3Brewer software, where it is suitable, the same rejection criteria used in EUBREWNET, i.e. maximum standard deviation of 2.5 DU and maximum ozone air mass of 3.5. TOC.

The rejection criteria on ozone values are hardcoded and consist on three sequential checks: 1) if raw counts are less than 2500, the value is rejected; 2) if calculated ozone for DS/ZS is less than 50 DU, the value is rejected 3) if observation is in the DS mode and the calculated ozone is between 50 and 100 DU, the value is rejected (Ogyu, personal communication 2018).

The maximum calculated ozone is indeed configurable in the BPS setup and was set to 500 DU.

The limits on the calculated ozone are not configurable in the O3Brewer setup. In the latest version used in this study, the standard lamp maximum for applying of ETC correction from SL test results is now configurable. Here we used the default limit of 500 units for the difference between R6 and the reference $R6_{ref}$.

line 335 By using daily mean values you include the effect of different rejection criteria also. Would be good to see if there is more or less perfect agreement when comparing simultaneous

measurements or if there are differences even then. Maybe there should have been a separate comparison of individual measurements and the resulting daily values? Comparing the individual measurements might have given more clue of the origin of the differences.
Usually daily values (or even more sparse time grid) are used for time series analysis, so it is important to see if any software introduces nonexistent drifts or biases to the data. Still, when comparing methods together it would be good to make more detailed analysis of where the differences come from.

We analysed the time series of TOC daily means and individual ozone values. The whole section 3 was completely re-written, additional figures on individual ozone values were inserted.

Figure 1 Why are there no points in 2008 summer in EUBREWNET data?

It was specified in the caption of Figure 1 that*: At Aosta the EUBREWNET L1.5 ozone values were not generated between May 24 and September 8, 2008, because the standard lamp got burned out since May 2008 and was replaced in September 2008.*

Figure 2 upper panel I dont understand. Here the cut-off for R6smooth is for sure lower than 500 units which was stated to be the threshold earlier.

In the new section 3.1 the cut off was better explained:

Looking at R6 behaviour (Fig. 6 upper panel), it can be noticed that the sensitivity of the instrument at Rome has changed mainly in two periods (between 1994 and 1995, and between 2006 and 2007). $R6_{smooth}$ becomes a constant offset when the sensitivity of the instrument starts to change. The cut off is not exactly equal to the threshold set in the configuration (in this case 500 units), but lower, because the filter looks 10 days before and 10 days after the date when SL R6 is calculated. If the cut off remains constant, it means that the last calculated correction which passes through rejection criteria, is taken into account, the same situation is experienced when there is no valid SL test (Stanek personal communication, 2018). Consequently, the temporal behaviour of $R6_{smooth}$ during these time intervals appears as a plateau. In this case SL correction is not applied since it is too high. Once a new calibration is performed (i.e. new references of R6 and the ETC are defined) R6 and $R6_{smooth}$ show a similar behaviour again.

Figure 2 lower panel I assume the R6 presented in the figure 2 are daily averages. I am not sure though. I am just wondering how many sl-tests there were when such spikes appear.

It was specified that: *It is worth noticing that the number of standard lamp test per day is on average from 4 to 6 at Rome, and from 2 to 4 in winter and from 8 to 10 in summer at Aosta and that only the daily means of BPS correction and $R6_{smooth}$ are stored. The latter is calculated if at least one standard lamp test is performed.*

I think it is a bit weird that the algorithm (BPS) has been made to pick up spikes so easily and use that to mistake them as valid r6 measurements. However I am also surprised that the results may be better than with the other software during those spikes.

That's true $R6_{BPS}$ follows the behaviour of R6 even during the spikes.

There should be information of standard lamp changes also. Or maybe they were changed only at calibration. Usually drifts like in Rome 2006 are caused by lamp being at the end of its lifetime but when looking at the corrected data it is apparent that the spectral response of the instrument really changed that dramatically and thus R6smooth can not follow the changes. Changes this big are rare and probably it should be considered alarming sign if R6 changes more than the threshold of O3Brewer?

We described in Section 3.1 the new analysis conducted when differences between BPS and O3Brewer ozone data exceed more than the DS accuracy: 1. $R6_{BPS}$ lower than $R6_{smooth}$, 2. $R6_{BPS}$ higher than $R6_{smooth}$, 3. R6BPS similar to $R6_{smooth}$

Figure 3 Could you address the amount of ozone difference because of difference in SL R6? Maybe not change these figures but in addition to this information. Just take the standard lamp part of equation (5).

In the new section 3.1 the following information was added:

Slight ozone difference took place when R6BPS was lower than R6smooth (at least 100 units), then the difference in ozone daily means was between -3% and 21% and in case of individual values from -3% up to 27 %, at Rome. At Aosta there was only one episode (2011/6/18) in which the O3Brewer daily mean differed about 30% from BPS determined.

……

Large negative ozone differences occur when $R6_{BPS}$ is higher than $R6_{smooth}$ (at least >100 units). This causes a variation between the daily means generated by the codes from -5% till -50% at Rome and from -51% till -91% at Aosta: Considering individual values a mean percentage difference between -3.1% and -57% was found at Rome, and of the same magnitude as that of daily means at Aosta.

line 399 I think the sl-corredction should be used especially on those days because the etc has dramatically changed. Now the O3Brewer with its cut-off does not follow the changes and the result of this is seen in figure 1 where O3Brewer data is very different than other around 2006-2007. Now if this change in r6 would have been because of a rapidly changing lamp irradiance then these values of r6 should not be used.

In the revised analysis we did not use data belonging to periods in which R6 produced drift or spikes, in the comparison with EUBREWNET, OMI and in the trend analysis.

line 400-402 Many of these other reasons can be checked from the raw files. I think anomalous R6 values should not be used in processing the data. Smoothing filter somewhat helps avoid these spikes. I think O3Brewer might do well if there was no cut off at 500 units (or whatever the cut off is).

In the last version of O3Brewer used for this revised analysis, the standard lamp maximum for applying of ETC correction from SL test results is configurable. We used first the default limit of 500 units for the difference between R6 and the reference R6ref. instrumental Then, we processed Rome ozone data using O3Brewer by setting the SL maximal limit to higher value to assess whether the smooth correction can properly process ozone data when large changes occurred in the response. The SL maximal correction limit was to 3000 units keeping identical conditions for the air mass and the standard deviation of the previous processing. This was still explained in the new section 3.1

line 406 Why? No other sources of disparity between the sofware are really addressed than R6. Does it explain all the differences? It is stated that there are discrepancies in "good data" also but no explanation or theory or a guess what would be the reason. There should be a case study of good measurements that differ greatly to address other sources of differences.

Another source of discrepancy was addressed which occurred when R6BPS similar to R6smooth. This case was analysed in Section 3.1 :

A different number of observations can be taken into account in the determination of the daily means by the two codes generating differences that can be significant in some cases. Such difference can be due to the fact that there are no filter conditions on the minimum and the maximum ozone values calculated by O3Brewer. Consequently, the daily means generated by this software are determined including anomalous values.

Two examples were provided in the revised manuscript:

We showed individual ozone values generated by both codes on 23/06/2001 at Rome with a daily average of 335 DU for BPS and 375.4 DU for O3Brewer (Fig.11, upper panel). It is clearly visible that the high individual ozone value generated by O3Brewer (618.7 DU) affecting the daily average provided by this code. Another example is provided for Aosta (Fig. 11, lower panel). On 5/1/2010 the daily average is 323.5 DU for BPS whereas it is 208.4 DU for O3Brewer. It is found that very low ozone values generated by O3Brewer, not discarded in the determination of the daily means, affect the quality of its value.

line 486 and table 5 Does it makes sense to think about the change in RMSE in case of O3Brewer as it has been shown in figure 1 that in special cases it does not follow the changes

in spectral sensitivity of the instrument correctly. Hopefully no one uses these software so loosely that they don't check their data in case of large drifts.

The new Table 5 reports the summary of the statistics of the comparison between OMI and ground-based data taking into account only periods not belonging to the three cases analysed in Section 3.1

line 491 It was stated earlier (page 12 line 298-) that the use of daily value is fine because ozone is so stable but here it is noted that it might have an effect.
The phrase was cancelled

line 501 The drift still needs to be quite fast and dramatic to exceed the O3Brewer threshold between two calibrations (1-2 years).
The above issue was specified in the conclusions .

line 505 Which one is most "correct". Does the BPS not follow the outliers a bit too closely? Usually the spikes are false R6 and should not be followed. The spectral sensitivity of the instrument is not expected to chance rapidly back and forth. For sure in the case of drifts it is not a good option to do the O3Brewer way and cut off but on other cases I would not want to follow every bump and spike in the R6 data.

The conclusion was modified as follows:

When R6 exceed the default value of the cut off (550 units) set in the configuration of the O3Brewer software during an occasional spike, the correction is not applied, whereas the BPS correction does. This could generate false high/low ozone values. In latest version of O3Brewer it is possible to set the cut off to higher value that is useful when there a large R6 drift is experienced. However anomalous ozone values can be still observed, since in O3Brewer there are no filter conditions on the minimum and the maximum ozone values. Similarly, the current Level 1.5 in the EUBREWNET can produce erroneous ozone recalculations when anomalous R6 values are experienced. The issue is expected to be solved in Level 2.0 products, when they will be released. The BPS ozone recalculations seem to be less affected in the case of R6 drift. However when serious changes in the spectral sensitivity of instrument is experienced, a solution consists in dividing the periods of R6 drifts into shorter time intervals and for that period a new set of constants ($R6_{ref}$ and ETC) could be established by the user as the averages of R6 ratios in that time interval. This process ("synthetic calibration") allows the user to introduce standard lamp corrections larger than the software hardcoded thresholds. In any case the synthetic constants in use must be confirmed at the next calibration with the reference instrument.
Here we decided to discard the periods with drifts or occasional abrupt changes in R6, and a good overall agreement is found between BPS, O3Brewer and EUBREWNET (MPE about <1%).
……
As a final remark, it is important to underline that for sake of consistency and comparability between the results from different stations which send ozone products to international data

centres such as WOUDC or others, it is important to know the processing software used to generate individual ozone values, the time behaviour of the instrumental stability, the method applied for the standard lamp correction as well as the adopted rejection criteria to determine the daily means.

line 515 I agree on the responsibility of the instrument operators. I also agree that there could be ways to work round some problems regarding to software behaviour. But also I think if there are behaviour in the instruments that the software dont handle well, the software should be changed accordingly if possible. I wonder if there was a way to get rid of the cut off in O3Brewer for the revised version.

The following phrase was included in the conclusions

When R6 exceed the default value of the cut off (550 units) set in the configuration of the O3Brewer software, the correction is not applied. In latest version of O3Brewer it is possible to set the cut off to higher value that is useful when there a large R6 drift is experienced. However anomalous ozone values can be still observed, since in O3Brewer there are no filter conditions on the minimum and the maximum ozone values.

Technical/typographical:

All incorrect words and typos were changed and formatted in italics

page 1 line 24 loose the parentheses, maybe "Italian stations Rome and Aosta"
Done

page 1 line 26 can you loose parentheses for example EUBREWNET level 1.5 product
Done

page 1 line 31 remove clearly and (as expected)
Done

page 3 line 60 This sentence should be rephrased. This sentence should be rephrased. "Satellite... made by using the solar UV light backscatterd..."
The statement was modified as**:** *Satellite-based ozone measurements are made by using the solar UV light backscattered from the Earth's atmosphere.*

page 4 line 73 could this be rephrased to not use brackets
The brackets were removed : *Even though all available processing software packages use the same TOC retrieval algorithm, which is based on the Bouguer-Lambert-Beer law, slightly different implementation can trigger some differences in the processed TOC data.*

page 4 line 74 implementation
Done

page 4 line 84 Brewers
Done

line 87 suggestion to lose the parenthesis and write ... packages: the Brewer Processing Software, hereafter called BPS,
Done

line 89 confirm title for Mr Stanek
the title was modified in *Ing*

line 90 inconsistent way of using parentheses inside a single sentence, could it "be EUBREWNET level 1.5 ozone product."
Done

line 94 to what extent
Done

line 95 change to "no other collocated TOC measurements were available" ? Somehow this sentence needs to be simplified
The sentence was modified as: *The OMI data were used since no other collocated TOC measurements were available.*

line 97 Paragraph starting here could be rewritten so that there are full stops between the sentences. The information is there but somehow the structure makes it hard to read.

We re-wrote the paragraph:
This paper is structured as follows: Section 2.1 briefly describes the theory on the ozone estimates from Brewer direct sun (DS) measurements. In Section 2.3, the procedure used by three software packages to process ozone data is presented. Section 2.4 describes the Brewer stations under study. Section 3 is dedicated to the comparison among the three TOC data retrievals and to understand the causes responsible for the differences among processed ozone values. Additional comparison between ground-based data and OMI products is also carried out. Ozone trends are estimated to investigate if using specific software could affect the results. Finally, conclusions are drawn in the last section.

line 107 suggest to leave out the (DS) from the header and introduce it in the text and "... spectrophotometer"
Done

line 110 is this a paragraph change or not?
It was changed

line 169 suggest change to "time series of the internal standard lamp tests, described in the following section."
Done

line 180 "to verify that"
Done

line 181 "and to follow the changes" (probably you can not really control them too much)
we replaced with *to track*

line 185 rephrase so there is no need for brackets, and "using an internal 20 W quartz-halogen lamp as the light source"
done

line 187-188 Rephrase so there is no need for parentheses.
here the rephrased statement *The DT test measures the dead-time of the photomultiplier and the photon-counting circuitry and the result of the test value should be within 5 ns with respect to the instrument constant. Also during the DT test, the halogen lamp is turned on.*

line 189 rephrase so there is no need for parentheses. Maybe give hg test its own paragraph?
here the rephrased statement*: For the Hg test a mercury lamp is used. This test ensures the correct wavelength alignment of the Brewer due to the internal temperature changes.*

line 195 no need to repeat the lamp power in my opinion, suggest to change "is performed using the internal quartz-halogen lamp as the light source"
done

line 198 rephrase so you can lose the parentheses
here the rephrased statement*:In this way changes with respect to $R6_{ref}$ are constantly tracked.*

line 201 rephrase to lose the parentheses
here the rephrased statement*: If a change in R6 is experienced, this results in a corresponding change in the ETC assuming that the relative lamp intensities at the four wavelengths do not change. Consequently, a correction in the reference ETC should be applied to determine the ozone values in between each calibration, ...*

line 227 suggest to start sentence with "Level 1.5 total ozone column data were..."
Done

line 278 "stray light"
Done

line 280 "is not available" to "was not available"
Done

line 307 "The Aura satellite describes a..." to "The Aura satellite travels in a..."
Done

line 308 suggest to start a new paragraph from "Two algorithms..."
Done

line 313 Could this be simplified to something like: "Here we used OMI-TOMS because it has been shown to have a better agreement with the ground based Brewer and Dobson instruments. (Balis et al., 2007)" ?
Done

line 323 Mean Bias says bias already so "(or bias)" is not needed
Done

line 439 missing a full stop.
Done

line 462 This should be stated in the caption of the picture!
New figures were included

figure 5 More detailed caption needed! What are the panels? It was actually in the text but the caption needs to be more detailed.
In the new Figure 5 (now Fig.16) more information was added.

line 472 rephrase to loose the parentheses
We calculated the scaled correlation coefficient as suggested by referee 2 so the following statement was included.

In general, the scaled correlation is, for both sites, on average RHOs= 0.8 which represents how the series are well connected in the short term.

line 476 A bit confusing way to put a sentence together. Also, can it be "about less than 1%"? It is either less than 1% or about 1%.

OMI products show a systematic underestimation with respect to ground-based data. At Rome satellite data are less than 1 % for both O3Brewer and EUBREWNET whereas at Aosta about 2.5%; 1.2% (Rome) and 2.5% (Aosta) in the case of BPS data.

Anonymous Referee #2

The author would like to thank the reviewer for constructive comments and suggestions that were taken into account. Please find our response to your comments (in bold). The responses are listed below each question. We have made changes to the original manuscript and the changes were written in red italics in the revised manuscript. We include also a version of the paper with all changes highlighted.

General comments

The article provides a comparative study of the main public software packages for Brewer data processing. The paper is well structured, but the language is probably too much technical for readers outside the Brewer spectrophotometer users community. A very nice set of Brewer data is used, with an impressive calibration history. The methodology is well explained, but the results needs a better analysis in order to explain the main differences found in each software. The results are very useful for the evaluation of ozone trends, once most part of the Brewer data available is processed by one of these software packages, allowing significant differences for the same measurements.

We decided to rework the paper performing additional analysis taking into account both daily means and individual calculated ozone values to better investigate the differences found among the three processing software.

Specific comments

Line 117: It should be mentioned that each single count rate is set after a number of scan cycles (nominaly 20) for slits 1 to 6.

In the revised manuscript (section 2.1) the following statement was included: *"The wavelengths are selected by a rapidly rotating slit-mask and raw photon counts for each slit-mask wavelength position are registered by a photomultiplier. During each measurement run cycle the slit-mask is rotated 20 times. The operational wavelengths correspond to slits 2 to 5 in the rotating slit-mask.*

Line 123: "Fi" must be defined as the instrumental count rate (counts per second) measured during the direct sun spectral irradiance observation for the slit number "i". The meaning of "i" is the slit number corresponding to each one of the 4 wavelengths referred on lines 117 and 118.

In the revised manuscript (section 2.1) the following statement was included: "*A linear combination (F) of the count rates (Fi) measured during the direct sun spectral irradiance observations for the i-th slit is computed …"*

Line 125: The weighting coefficients wi were chosen in order to minimize not only the effect of the aerosol scattering but also its absorption. So the best sentence should be "in order to minimize the effect of aerosol attenuation" or "in order to minimize the effect of the aerosol optical depth".

In the revised manuscript it was reported that: "*The weighting coefficients are chosen in order to minimize the effect of the aerosol extinction, to eliminate the effect of the sulphur dioxide absorption (Kerr et al., 1981; Kerr, 2010) and all factors independent of the wavelength (flat factors):*

Line 190: The HG test "ensures the correct wavelength alignment of the Brewer", could be completed with ", due to the internal temperature changes"

In the revised manuscript it was specified that: "*This test ensures the correct wavelength alignment of the Brewer due to the internal temperature changes*.

Line 471: The "excellent" agreement with OMI is mainly due to the seasonality of TOC. A more interesting analysis could be if seasonality and trend were removed from the series.

In the revised manuscript the agreement between OMI and Brewer data was assessed by calculating the scaled correlation (RHO) which excludes the possibility that the source of the correlation is a common cycle (e.g. the annual cycle). That calculation is performed splitting the series of the ozone daily values in short intervals (here K=30 days) and for each interval RHO coefficient is determined. In this way the high frequency component (<30 days) common to Brewer and OMI series are revealed.

An additional paragraph was included in the section 2.5 (statistical metrics) of the revised manuscript and the results were included in Table 4. We found that "*In general, the scaled correlation is, for both sites, on average RHOs= 0.8 which represents how the series are well connected in the short term."*

Examination on total ozone column retrievals by Brewer spectrophotometry using different processing software.

Anna Maria Siani[1*], Francesca Frasca[2], Francesco Scarlatti[3], Arianna Religi[4], Henri Diémoz[5], Giuseppe.R. Casale[1], Massimiliano Pedone[6], Volodya Savastiouk[7]

[1]Department of Physics, Sapienza Università di Roma, Rome, Italy
[2] Department of Earth Sciences, Sapienza Università di Roma, Rome, Italy
[3]Independent scientist
[4]Institute of Services Science, University of Geneva, Geneva, Switzerland
[5]Aosta Valley Regional Environmental Protection Agency (ARPA), Saint-Christophe, Italy
[6]Infosapienza Settore per i Sistemi Centrali e per l'Office Automation, Sapienza Università di Roma, Rome, Italy
[7]International Ozone Services Inc., Toronto, Ontario, Canada

*Correspondence to: Anna Maria Siani (annamaria.siani@uniroma1.it)

Abstract. The availability of long-term records of the total ozone content (TOC) represents a valuable source of information in studies on the assessment of short and long-term atmospheric changes and their impact on the terrestrial ecosystem. In ~~addition~~ *particular*, ground-based observations represent a valuable tool to validate satellite-derived products. To our knowledge, details about ~~processing~~ software packages to process Brewer spectrophotometer measurements and to retrieve the TOC are seldom specified in studies using such datasets. ~~although some discrepancies can arise from the use of different algorithms and implementations~~. ~~The deviations among retrieved TOCs from the Brewer instruments~~ *The sources of the differences among retrieved TOCs from the Brewer instruments located at the Italian stations Rome and Aosta, using* three freely available codes (Brewer Processing Software, O3Brewer software and EUBREWNET *Level 1.5 products*) are here investigated. Ground-based TOCs are also compared with the Ozone Monitoring Instrument (OMI) TOC retrievals used as an independent dataset since no other instruments near the Brewer sites, are available.

~~Although the overall agreement of the BPS and O3Brewer TOC data with EUBREWNET data is clearly very good (as expected) and in most cases within the Brewer declared uncertainty less than 2%, it is worth noticing that slight differences have been seen depending on the software in use.~~

~~Such differences become larger when the instrumental sensitivity exhibits a long-term drift and even in short-term episodes due to the different algorithm for the standard lamp correction.~~

The overall agreement of the BPS and O3Brewer TOC data with EUBREWNET data is within the estimated total uncertainty in the retrieval of total ozone from Brewer spectrophotometer (1%). However, differences can be found depending on the software in use. Such differences become larger when the instrumental sensitivity exhibits a fast and dramatic drift which can affect the ozone retrievals significantly. Moreover, if daily mean values are directly generated by the software, differences can be observed due to the configuration set by the users to process single ozone measurement and the rejection conditions applied to data to calculate the daily value.

This work aims to provide useful information both for scientists engaged in ozone measurements with Brewer spectrophotometry and for stakeholders of the Brewer data products available at web-based platforms.

Key words: ozone, Brewer spectrophotometry, standard lamp correction, processing software, calibration

1.INTRODUCTION

Although ozone (O_3) is present in small amounts in the terrestrial atmosphere, it plays a crucial role in the attenuation of solar ultraviolet (UV) radiation (200 - 400 nm) reaching the surface and in radiative processes controlling the energy balance on the Earth (Ramanathan and Dickinson, 1979; Dessler, 2000; Bordi et al., 2012; WMO, 2015).

The cumulative amount of stratospheric and tropospheric ozone represents the total ozone column (TOC). The most common ground-based instruments to measure TOC are spectrophotometers which are designed to measure ~~ground level intensities of attenuated incident solar ultraviolet radiation in the ozone absorption spectra, from which it~~ *ground level spectral intensities of solar ultraviolet radiation attenuated by ozone absorption. From these spectra, it is possible to* retrieve the TOCs. The first TOC observations were recorded using the Dobson spectrophotometer (Dobson and Harrison, 1926) in the late 1920s~~ but only in a few places~~. Since then, a growing number of sites were equipped with the Dobson spectrophotometer and later in the 1980s with the automated Brewer spectrophotometer (Brewer, 1973). Nowadays, both the Dobson and the Brewer spectrophotometers are used all over the world and ~~if properly maintained and calibrated they provide TOC data within 1-2% accuracy (Fioletov et al., 2005, Vanicek, 2006).~~ *the accuracy of measurements taken with a well-maintained Brewer spectrophotometer is 1% in the direct sun (DS) mode (Vanicek, 2006).*

~~Satellite-based ozone measurements are made by use of the sun UV light backscattered from the Earth's atmosphere. These measurements have the advantage of quasi-global coverage by one and the same instrument. On the other hand, ground-based instruments regularly undergo calibrations with an absolute reference instrument and have longer lifetimes.~~

It should be pointed out that high-quality TOC retrievals from ground-based stations are necessary not only in support of the validation of satellite-derived products (Tzortziou et al., 2012) but also for the assessment of the long-term ozone trend and the verification of the effectiveness of the Montreal Protocol on substances that deplete the ozone layer. Moreover, ground–based TOC data are also necessary to calibrate the parameters in the global climate models used to predict the expected behaviour of the ozone layer in the future (Stübi et al., 2017).

The above issues show the importance to measure the ozone amount from ground-based stations with a very good performance.

~~Even though the same TOC retrieval algorithm, based on the same and acknowledged physical principle (i.e. Bouguer-Lambert-Beer law), is adopted by all available processing software~~

~~packages, slightly different implementations can trigger some differences in the processed TOC data.~~

Even though all available processing software packages use the same TOC retrieval algorithm, which is based on the Bouguer-Lambert-Beer law, slightly different implementations potentially trigger some differences in the processed TOC data.

The largest part of the total ozone column data analysed in the current/available scientific literature is extracted from the WOUDC data archive (World Ozone and Ultraviolet Radiation Data Centre) i~~n which detailed information on the used processing software is not always available. T~~o our knowledge, the processing software of Brewer TOC data varies from site to site, the processing algorithm and the data rejection rules are seldom specified. *WOUDC ozone files (2017) do not include information on the software used to process ozone data, the version of such software as well as the adopted data rejection rules. The same information is usually not reported in the studies related to ozone monitoring, trend detection and satellite validation.* This can be due to the fact that a standard processing software of Brewer raw data has currently not been adopted. *For this reason,* the COST Action ES1207 "A European Brewer Network" (EUBREWNET) was established aiming at defining, among the others, a standard procedure for processing the raw Brewer data, thus ensuring the quality of the data and harmonizing the products from the European *Brewers* (EUBREWNET, 2017).

~~The purpose of the present study is to: 1) investigate the differences among the TOCs retrieved by three different processing software packages (the Brewer Processing Software, hereafter called BPS) developed by Dr Fioletov V. and Ogyu A. (Environment Canada), O3Brewer software developed by Dr Stanek M. (Solar and Ozone Observatory of CHMI/International Ozone Service) and the EUBREWNET products (ozone Level 1.5). To the purpose of the intercomparison, we tested the mentioned software on the datasets collected by the Brewer instruments located at Rome and Aosta, Italy; 2) compare Brewer ozone recalculations with the Ozone Monitoring Instrument (OMI) TOC retrievals to investigate at which extent the ground-based and satellite-based retrievals are similar.~~

The purpose of the present study is to investigate the differences among the TOCs retrieved by three different processing software packages: the Brewer Processing Software, hereafter called

BPS, developed by Dr Fioletov V. and Ogyu A. (Environment Canada), O3Brewer software developed by Ing Stanek M. (Solar and Ozone Observatory of CHMI/International Ozone Service) and the EUBREWNET level 1.5 ozone product. To the purpose of an intercomparison exercise, we tested the mentioned software on the datasets collected by the Brewer instruments installed at Rome and Aosta, Italy. *Then,* Brewer ozone recalculations were also compared with the Ozone Monitoring Instrument (OMI) TOC retrievals. The OMI data were used *since no other independent collocated TOC measurements were available.*

~~This paper is structured as follows: the theory on the ozone estimates from Brewer direct sun (DS) measurements is first briefly described (Section 2.1); furthermore, the methods to correct the ozone data using the three different ground-based processing software packages are presented in Section 2.2 and the measuring instruments sites in Section 2.3 and 2.4; then, TOC retrievals by the processing software are compared with the purpose to understand the reasons of the differences in ozone retrievals; finally a comparison between ground-based data and OMI products is carried out to investigate at which extent the ground-based and satellite-based retrievals are similar (Section 3); the last section summarizes the main conclusions.~~

This paper is structured as follows: Section 2.1 briefly describes the theory on the ozone estimates from Brewer direct sun (DS) measurements. In Section 2.3, the procedure used by three software packages to process ozone data is presented. Section 2.4 describes the Brewer stations under study. Section 3 is dedicated to the comparison among the three TOC data retrievals and to understand the causes responsible for the differences among processed ozone values. Additional comparison between ground-based data and OMI products is also carried out. Moreover ozone trends are estimated to investigate if using specific software could affect the results. Finally, conclusions are drawn in the last section.

2. DATA AND METHOD

2.1 *Theory of direct sun measurements with Brewer spectrophotometer*

The Brewer spectrophotometer is an instrument designed to retrieve the total ozone column ~~by means of measurements of direct sunlight, zenith sky light, focused moonlight or using~~

~~the global irradiance method (Kerr and Davis, 2007) in the UV region.~~ *by measuring irradiances of both direct sunlight (Kerr et al., 1981) and polarized radiation scattered from the zenith sky (Brewer and Kerr, 1973, Muthama et al., 1995). Total ozone can be also derived from focused sun measurements, commonly employed at high latitudes (Josefsson, 1992)*. It is also possible to measure total ozone by using the moon (Kerr et al., 1990), or the global irradiance method in the UV region (Kerr and Davis, 2007), as a light source.

The most accurate method to determine the total column amount of an atmospheric gas is based on the direct sun (DS) measurements. ~~It was shown that the accuracy of TOC with DS measurements taken with a well-maintained Brewer spectrophotometer is better than 2% (Fioletov et al., 2005, Vanicek, 2006).~~ *It was shown (Vanicek, 2006) that the accuracy of measurements taken with a well-maintained Brewer spectrophotometer is 1% in the DS mode and 3-4% in the ZS mode. The random errors of individual measurements were found to be within ± 1% for all measurements (Fioletov et al., 2005).*

The algorithm to retrieve the total ozone column from the Brewer in DS mode is based on a differential measurement method involving 4 selected wavelengths in the ozone absorption spectra, nominally 310.1, 313.5, 316.8 and 320.1 nm. ~~A photomultiplier registers photon counts of radiation that pass through the exit slits from 3 to 6 corresponding to the operational wavelengths.~~

The wavelengths are selected by a rapidly rotating slit-mask and raw photon counts for each slit-mask wavelength position (from 3 to 6) are registered by a photomultiplier. During each measurement run cycle the slit-mask oscillates 20 times. The raw photon counts are then converted into count rates and *are corrected for the characteristics of the photomultiplier (dark count and dead time) and for the* internal Brewer temperature (Kerr, 2010). In addition, a correction for the spectral transmittance of the attenuation filters can be added depending on the filter used, if the respective characterisation is available.

A linear combination (F) of the base-ten logarithms ~~the measured spectral direct irradiances at the four longer wavelengths~~ *of the count rates (F_i) measured during the direct sun spectral irradiance observations for the i-th slit* is computed by weighting the F_i with coefficients (w_i=-1, +0.5, +2.2, -1.7). The weighting coefficients are chosen in order to minimize the effect of *the aerosol extinction*, *to* eliminate the effect of the sulphur dioxide absorption (Kerr et al., 1981; Kerr, 2010) *and all factors independent of the wavelength (flat factors)*:

$$F = \sum_{i=1}^{4} w_i \log F_i \qquad (1)$$

F_i is also compensated for the effect of the Rayleigh scattering by subtracting:

$$\frac{p}{p_o} \mu_R \sum_{i=1}^{4} w_i \beta_i \qquad (2)$$

where p is the climatological pressure at the measurement site and p_o is the pressure at the sea level; μ_R is the Rayleigh air mass factor (i.e. the slant path of direct radiation through air), calculated for a thin layer at 5 km altitude, β_i is the Rayleigh scattering coefficient at the wavelength, λ_i.

According to the Bouguer-Lambert-Beer law, it is possible to retrieve the total ozone column (TOC) as:

$$TOC = \frac{Fo - F}{\Delta\alpha\mu} \qquad (3)$$

where $\Delta\alpha$ is the *differential ozone absorption coefficient*, i.e. the linear combination of the ozone cross sections using the same weighting coefficients employed for F. $\Delta\alpha$ is calculated after performing a specific test using spectral lamps providing the precise operational wavelengths and applying the convolution with the slit function characterised for each individual spectrophotometer. Then $\Delta\alpha$ is obtained for these wavelengths using Bass-Paur ozone absorption spectrum (Bass and Paur, 1985) at the fixed temperature of -45°C (Kerr, 2010).

The standard Brewer algorithm assumes that the ozone is concentrated in a thin layer at the altitude of 22 km, thus the air mass factor (μ) is expressed by:

$$\mu = \sec\left[\arcsin\left(\frac{R_E}{R_E + 22}\sin Z\right)\right] \quad (4)$$

where R_E is the Earth's radius and Z is the solar zenith angle. F_o is also expressed as the linear combination of the extraterrestrial irradiance at the operational Brewer wavelengths with the same weighting coefficients used for F. F_o corresponds to F at the top of the atmosphere and it is usually named "extraterrestrial constant" (ETC), *a specific factor different for each Brewer, and determined through a calibration procedure.*

There are two methods to determine the ETC. The first is based on the use of the Langley plot technique i.e. plotting F versus μ, and then the ETC value is extrapolated at zero air mass. This method is used for the calibration of primary standards and requires to be carried out under stable atmospheric conditions and low pollution concentrations. The second method is based on transferring the calibration from a reference Brewer instrument with a known ETC to a candidate instrument during field campaigns. This latter technique is the most common way for regularly calibrating the instruments which belong to the Brewer network. In between the calibration audits with a travelling standard, the TOC data are processed adjusting the ETC according to the changes of the radiometric sensitivity of the instrument, if needed. The correction uses *time series of the internal standard lamp tests, described in the Section 2.2.*

Direct-sun measurements are carried out at specific solar zenith angles throughout the day depending on the user schedule (a sequence of commands written by the operator), allowing the Brewer to make observations continuously and automatically. During a DS measurement, *a group of* five consecutive sub-measurements are taken in less than five minutes. Then the mean and the standard deviation of the five ozone values are computed and associated to that DS measurement.

~~An individual TOC value is considered acceptable if the standard deviation of the five measurements is lower than 2.5 DU. In this case, the value is included in the number of accepted DS measurements to provide the daily TOC mean.~~ *The standard deviation is used to determine the acceptability of each TOC measurement. An individual TOC value is normally considered acceptable if the standard deviation of the five measurements is lower than 2.5 DU or 3 DU.*

2.2 Standard lamp correction

Several tests are performed on a daily and weekly basis *to verify that* the Brewer operates correctly and ~~to take under control~~ *to track* the changes in instrumental properties. The main standard tests included in the diurnal operational schedule are: shutter motor run/stop (RS), photomultiplier dead time (DT), mercury lamp (Hg) and standard lamp (SL).

The RS test verifies that the slit-mask motor is operating properly. It calculates the ratio of irradiances at the operational wavelength using ~~(using as the light source a quartz-halogen lamp of 20 W)~~ an internal 20 W quartz-halogen *lamp as the light source* in a dynamic mode and in a static mode. This ratio should be as close as possible to unity.

The DT test measures the dead-time of the photomultiplier and the photon-counting circuitry and the result of the test value should be within 5 ns with respect to the instrument constant. Also during the DT test, the halogen lamp is turned on.

~~The Hg test (in which a mercury lamp is used)~~ *For the Hg test a mercury lamp is used. This test ensures the correct wavelength alignment of the Brewer due to the internal temperature changes.* This test is usually carried out several times every day.

The ~~standard lamp~~ SL test is used to monitor the stability of the instrument response after the calibration with the reference spectrophotometer. The test is performed ~~by the use of a quartz-halogen internal lamp (20 W)~~ *using the internal quartz-halogen lamp* as the light source. The photon counts are recorded at the same operational wavelengths employed in the DS measurement and the result of the SL test, the so-called R6 ratio which corresponds to a fictitious value of ozone column density, is determined using Eq.(1). In this way changes *with respect to the reference R6 value ($R6_{ref}$), determined during the calibration with the reference instrument, are constantly tracked* ~~in the instrument response are constantly tracked (i.e. changes with respect to $R6_{ref}$ and hence to the corresponding ETC, both established during each calibration campaign)~~.

If a change in R6 is experienced, this results in a corresponding change in the ETC ~~(~~assuming that the relative lamp intensities at the four wavelengths do not change~~)~~. *Consequently, a correction in the reference ETC should be applied to determine the ozone values in between each calibration, as follows:*

$$TOC = \frac{ETC - F + \Delta SL}{\Delta \alpha \mu} \quad (5)$$

where ΔSL is the correction factor measuring the difference between $R6_{ref}$ ~~(from the last intercomparison)~~ *which is determined at every calibration* and R6 for a specific day.

Depending on the processing software used by the station operator, ΔSL is computed in different ways, *not always clearly explained by the software documentation:*

- ~~the BPS adjusts the ETC taking into account the difference between the $R6_{ref}$ (calculated with a triangular smoothing filter of SL-test values from 15 consecutive days since that calibration) and the present daily mean values of R6, if the difference between the $R6_{ref}$ and the current value is ≥ 250; if the difference is ≤ 250 units then a median of R6 data before 15 days and after 15 days is used for the correction. That correction is reported in the file named “o3data” produced by the BPS. The threshold and the time window are however not adjustable by the users.~~
- *In the BPS, the reference value $R6_{ref}$ is determined with a triangular smoothing filter of SL-test values over the 15- days period immediately following the calibration date. There should be at least one good SL-test value per day. If the corresponding B-files are not available, the program is not able to establish the reference SL level and the ETC will be not adjusted. Notice that for other processing software $R6_{ref}$ is based on the SL-test values during the calibration campaign. If the abs($R6_{ref}$ -R6) ≤ 250 units, then the median of daily averages from all R6 data before 15 days and after 15 days for a particular day is used for the correction. The median is used because it is less influenced by single invalid R6s. If the abs($R6_{ref}$ -R6) is above 250 units then ETC is adjusted taking into account the difference between the $R6_{ref}$ and the present daily mean values of R6.* That correction is reported in the file named “o3data” produced by the BPS. The threshold and the time window are however not adjustable by the users (Fioletov personal communication, 2018).

-

- O3Brewer adjusts the ETC using a Gaussian smoothing filter on R6 values (Stanek M., 2016). ~~The program reads the R6 daily means of the SL test 10 days before and 10 days after the selected date period~~, ~~O3Brewer applies the Gaussian low-pass filter when the difference between R6 and the reference $R6_{ref}$ does not exceed a certain threshold (500 units, Stanek personal communication, 2016)~~. *There should be SL measurements 10 days before and 10 days after the selected date period.* The software creates the smoothed R6 time series (hereafter named $R6_{smooth}$) which is used for ETC adjustment. *It means that there should be at least one SL test per day. The application of the ETC correction is done when the difference between the reference $R6_{ref}$ and R6 from SL test results, does not exceed is a certain value (the default value is 500 units) . This threshold is now configurable in the latest version 6.0* (Stanek personal communication, 2018). The time window is however not adjustable by the users. If this difference exceeds the threshold, *then the software can remember the last day with good SL test and will apply that correction* (Stanek personal communication, 2018). This option can be turned off and then the daily mean values for SL are used for the correction of the ETC.

- ~~The EUBREWNET architecture is based on three different data quality/processing levels of TOC estimates from DS measures. Level 0: the TOC is taken directly from the Brewer files (named Bfiles) as calculated by the standard algorithm (Eq. (3)); Level 1: the TOC is recalculated with the standard algorithm applying the set of constants verified by the operator and the spectral attenuation of each filter is added in Eq. (5); Level 1.5: the TOC is filtered for the standard deviation of five consecutive observations (default value is 2.5 DU) and the maximum ozone air mass (the default maximum value is 3.5). Additionally, the wavelength alignment of the spectrometer must be within ±2 microsteps (valid Hg tests) before and after the ozone measurement to ensure the quality of TOC measurements. In addition, TOC values less than 100 DU and greater than 500 DU are discarded. The TOC is calculated taking into account Eq. (5) and adding the spectral attenuation of the filters and, if available, the stray-light correction is applied (Karppinen et al., 2015; Redondas et al., 2016). The ΔSL correction is determined applying a triangular moving average over the daily median values of R6 in a window of seven days (default time window). The correction is applied if the difference between $R6_{ref}$ and the calculated value exceeds 5 units. Level 2.0: ozone products are consistent with Level 1.5 products validated with a posterior calibration. If the reference constants of a posteriori calibration do not differ significantly from the values in use then level 1.5 product is not reprocessed and it represents the most reliable product.~~

•Level 1.5 total ozone column data from EUBREWNET are recalculated with the ΔSL correction determined applying a triangular moving average over the daily median values of R6 within a seven days window (default time window). The correction is applied if the difference between $R6_{ref}$ *and the calculated value exceeds 5 units. Level 2.0 are 1.5 observations validated with a posterior calibration. If the reference constants of a posteriori calibration do not differ significantly from the values in use then level 1.5 products are not reprocessed and represent the most reliable products (http://rbcce.aemet.es/dokuwiki/doku.php).*

At the present time, tools for Level 2.0 are developed but not yet implemented. A complete description of the processing can be found on the EUBREWNET website (2017).

2.3 Measuring instruments and sites

Brewers MKIV serial numbers 067 and 066 have been operating at the Solar Radiometry Observatory of Sapienza University of Rome (hereafter Rome) and at the headquarter of Aosta Valley Regional Environmental Protection Agency (ARPA) at Aosta-Saint Christophe (hereafter Aosta), respectively. The former has been recording TOCs since 1992 (Siani et al., 2002) whereas the latter since 2007 (Siani et al., 2013).

In this study the above-mentioned sites were selected because both Brewers belong to Sapienza University of Rome and have been calibrated with the same reference spectrophotometer since their installation, submitting regularly data to the WOUDC and taking part to the COST Action ES1207 EUBREWNET. The station characteristics are reported in Table 1.

Table 1. Characteristics of the two Italian Brewer sites

Station name (GAW ID)	**Brewer Serial number**	**Coordinates** (latitude, longitude, elevation (m above sea level)	**Observation period**	**Environmental context**
Aosta (AST)	066	45.7°N, 7.4°E, 569 m a.s.l.	29/01/2007 - 31/12/2015	semi-rural
Rome University (ROM)	067	41.9°N, 12.5°E, 75 m a.s.l.	01/01/1992 - 31/12/2015	urban

Since their installation, both Italian Brewers have been calibrated every one or two years by intercomparison with the traveling reference Brewer 017 from International Ozone Services Inc. (IOS), (2017). This Brewer is in turn calibrated against the World Brewer Reference Triad in Toronto (Fioletov et al., 2005). In this way the ozone calibration of Italian spectrophotometers is also traceable to the Brewer Reference Triad.

The calibration history of the two instruments used in this study is reported in Table 2. Although zenith sky and global irradiance measurements were available, only DS measurements were selected in this study because they have a lower uncertainty compared to the other types of measurements (Fioletov et al., 2005).

~~Individual DS observations for each Brewer were recalculated with BPS (Fioletov and Ogyu , 2007), O3Brewer software packages (Stanek, 2016), satisfying the standard deviation criteria ≤ 2.5 DU and air mass factor ≤ 4. TOC real time values Level 1.5 were also downloaded from EUBREWNET platform over the period 2005-2015 at Rome and 2007-2015 at Aosta. The stray light correction was not applied because it requires the calibration against a double monochromator Brewer and the instrumental characterization (Redondas et al., 2016) which is not available.~~

~~Daily means were then calculated from all available data sets (hereafter named TOC BPS, TOC O3Brewer and TOC EUBREWNET). We used daily TOC averages because the applied ETC correction is the same for all individual measurements within the same day.~~

In this study we analysed individual DS values and daily averages of Rome and Aosta stations, generated by BPS version 2.1.1 updated to 2017/02/14 (Fioletov and Ogyu, 2007)*, by O3Brewer software packages version 6.0 updated to 2018/03/14, and EUBREWNET level 1.5 ozone products. Level 1.5 individual TOC values are discarded when the standard deviation is above 2.5 DU and the maximum ozone air mass is above 3.5. In addition ozone values less than 100 DU and greater than 500 DU are also rejected. The stray light* correction was not applied because it requires a calibration against a double monochromator Brewer and an instrumental characterization (Karppinen et al., 2015, Redondas et al., 2016) which *was* not available. Level 1.5 TOC values were downloaded from EUBREWNET platform over the period 2005-2015 at Rome and 2007-2015 at Aosta.

Table 2. Calibration history of Brewer 066 and 067. In brackets it is reported the month of the calibration for Brewer 067 (*The recalculation of the constants was performed by IOS after the calibration on July 2009). In one case the calibration of Italian Brewers was performed in Arosa (Switzerland) at the Lichtklimatisches Observatorium during the Seventh Intercomparison campaign of the Regional Brewer Calibration Center Europe (WMO-GAW, 2015). In 2013 the calibration of both Brewers was carried out at Aosta.

Year	Period	Location (Brewer 066)	Location (Brewer 067)
1992	January		Rome
1993	September		Rome
1995	May		Rome
1996	April		Rome
1997	May		Rome
1998	July		Rome
1999	September		Rome
2000	September		Rome
2002	March		Rome
2003	September		Rome
2006	September		Rome
2007	April	Aosta	Rome
2009	July	Aosta	Rome
2010*	January	Aosta	Rome
2011	August (July)	Aosta	Rome
2012	August (July)	Arosa	Arosa
2013	May (June)	Aosta	Aosta
2014	July		Rome
2015	July	Aosta	Rome

We set in the configuration file of BPS and O3Brewer software, where it is suitable, the same rejection criteria used in EUBREWNET, i.e. maximum standard deviation of 2.5 DU and maximum ozone air mass of 3.5. TOC.

The rejection criteria of ozone values are hardcoded in the BPS software and consist on three sequential checks: 1) if raw counts are less than 2500, the value is rejected; 2) if calculated ozone for DS/ZS is less than 50 DU, the value is rejected 3) if observation is in the DS mode and the calculated ozone is between 50 and 100 DU, the value is rejected (Ogyu, personal communication 2018).

The maximum calculated ozone is indeed configurable in the BPS setup and was set to 500 DU.

The limits on the calculated ozone are not configurable in the O3Brewer setup. In the latest version used in this study, the standard lamp maximum for applying of ETC correction from

SL test results is now configurable. Here we used the default limit of 500 units for the difference between R6 and the reference $R6_{ref}$.

2.4 Satellite TOC data

The Ozone Monitoring Instrument (OMI) products were used as an ancillary dataset with the purpose of helping understand the difference among the investigated Brewer retrievals and the comparison should not be regarded as exhaustive validation exercises of satellite total ozone data. Daily averages of the Brewer TOC were compared with satellite ozone values obtained during the overpass. The use of daily means instead of Brewer TOC observations taken close to the OMI overpass is reasonable since it allows to compare a large number of pair measurements (Antón et al., 2009; Vaz Peres et al., 2017) because there are only one or two daily satellite values.

Satellite overpass data at Rome and Aosta were derived from OMI, on board NASA EOS-Aura spacecraft launched in July 2004. The OMI instrument is a nadir-viewing spectrometer measuring solar reflected and backscattered light from the Earth atmosphere and surface in the wavelength range from 270 nm to 500 nm, providing global daily coverage with a spatial resolution of 13×24 km^2 in nadir. The Aura satellite ~~describes~~ *travels in* a sun-synchronous polar orbit, crossing the equator at 13:45 local time. *Two algorithms, OMI-TOMS (Total Ozone Mapping Spectrometer) and OMI-DOAS (Differential Optical Absorption Spectroscopy), are used to produce OMI daily total ozone datasets.*

In our study OMI-TOMS ozone overpasses based on TOMS V8.5 algorithm (Bhartia and Wellemeyer, 2002) at the stations under study over the period 01/10/2004-31/12/2015 were downloaded from the NASA –AURA validation data center platform. *Here we used OMI-TOMS since it has a better agreement with the ground-based Brewer and Dobson instruments (Balis et al., 2007).*

~~Here we used OMI-TOMS for the reason that the comparison between ground-based Brewer and Dobson data and OMI satellite ozone data showed an agreement of better than 1% for OMI-TOMS and better than 2% for OMI-DOAS data (Balis et al., 2007).~~

2.5 Statistical parameters

The following statistical parameters are used with the aim to quantify the differences among the TOC series: nonparametric Spearman coefficient (RHO), Mean Bias (MB), Mean Percentage Error (MPE), Root Mean Square Error (RMSE). RHO was used to measure the correlation between two variables without making any assumption about their distribution. MB represents the systematic differences between two selected datasets; MPE provides the average of percentage errors with respect to TOC values taken as the reference. RMSE is an estimate of the standard deviation of the difference (residuals) between two dataset.

$$MB = \frac{1}{N}\sum_{1}^{N}(y_i - y'_i) \qquad (6)$$

$$MPE = 100 * \frac{1}{N}\sum_{1}^{N}\frac{(y_i - y'_i)}{y'_i} \qquad (7)$$

$$RMSE = \sqrt{\sum_{1}^{N}\frac{(y_i - y'_i)^2}{N}} \qquad (8)$$

The previous equations show the formulas of the mentioned statistical parameters, where y_i is the i-th TOC value (O3Brewer, or OMI) value, $y'_i$is the i-th TOC value of the BPS (or EUBREWNET) series, N the number of all the possible data pairs analysed. *The uncertainty of MB and MPE is characterized by the standard deviation.*

In the comparison between Brewer and OMI data the scaled correlation (RHO) was calculated (Diémoz et al.,2016) to exclude the possibility that the source of the correlation is a common cycle (e.g. the annual cycle). That calculation is performed by splitting the series of the ozone daily values in short intervals (here K=30 days) and for each interval RHO coefficient is determined. Then RHOs are given by:

$$RHOs = \frac{1}{K}\sum_{i=1}^{K} RHO_i \qquad (9)$$

In this way the high frequency component (<30 days) common to Brewer and OMI series are revealed.

2.6 *Trend analysis*

To assess whether a specific software could affect the trend, we estimated the trend from the annual mean anomalies. We applied the same methodology proposed by Fountoulakis et al., (2016). Climatological ozone values for each day were calculated over the period under study.

The daily anomaly with respect to the daily climatological value was calculated. Afterward the monthly anomalies were determined by averaging the daily anomalies for each month provided that at least 15 days of data were available. Finally the monthly anomalies were averaged to determine the annual mean anomalies. The trend among the three codes was expressed as a percentage variation per decade and used in their comparison. The statistical significance of the trends is derived from the Mann–Kendall test with statistical significance set at p≤ 5%.

3. RESULTS AND DISCUSSION

The time series of TOC daily means *generated by* BPS, O3Brewer and calculated from EUBREWNET individual ozone values, are presented in Figs. 1 (upper panel Rome, lower panel Aosta). *Individual measurements are distinctly plotted for each site in Fig.2 and Fig.3.*

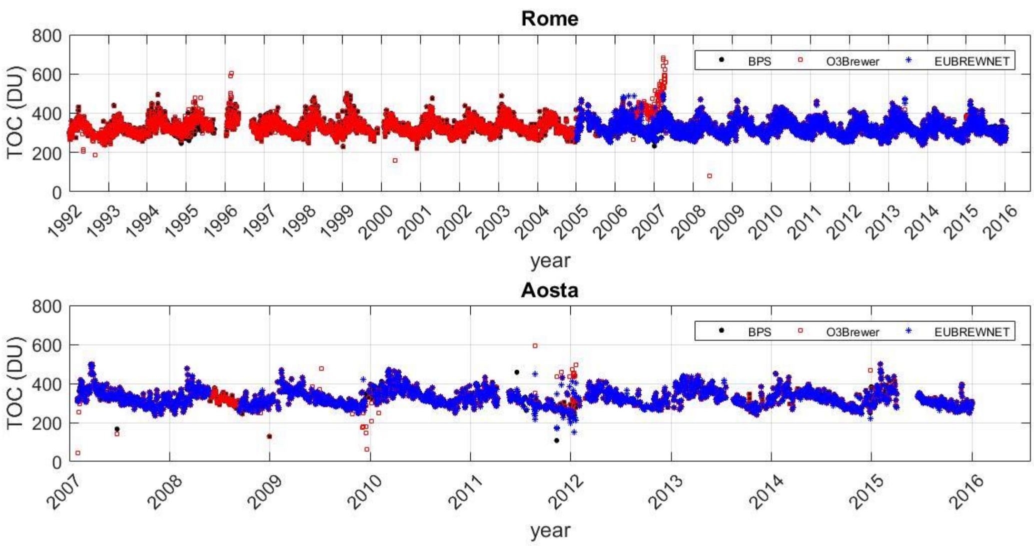

Figure 1. Time series of TOC daily means from BPS (black), O3Brewer (red) and EUBREWNET (blue) at Rome (upper panel) and at Aosta (lower panel). *At Aosta the EUBREWNET L1.5 ozone values were not generated between May 24 and September 8, 2008, because the standard lamp got burned out since May 2008 and was replaced in September 2008.*

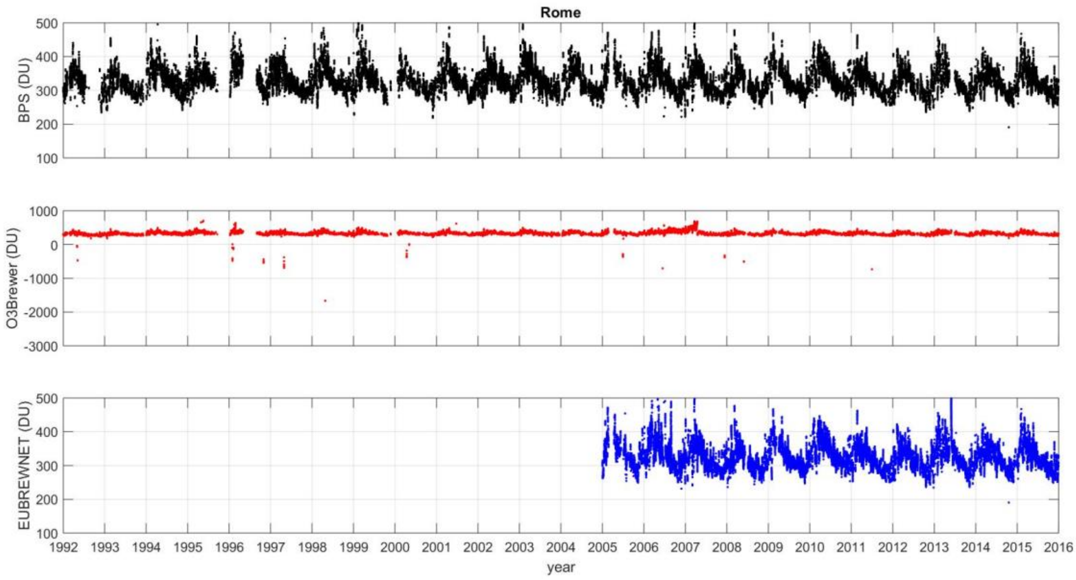

***Figure 2**. Individual TOC values generated by BPS (black), O3Brewer (red) and EUBREWNET (blue) at Rome.*

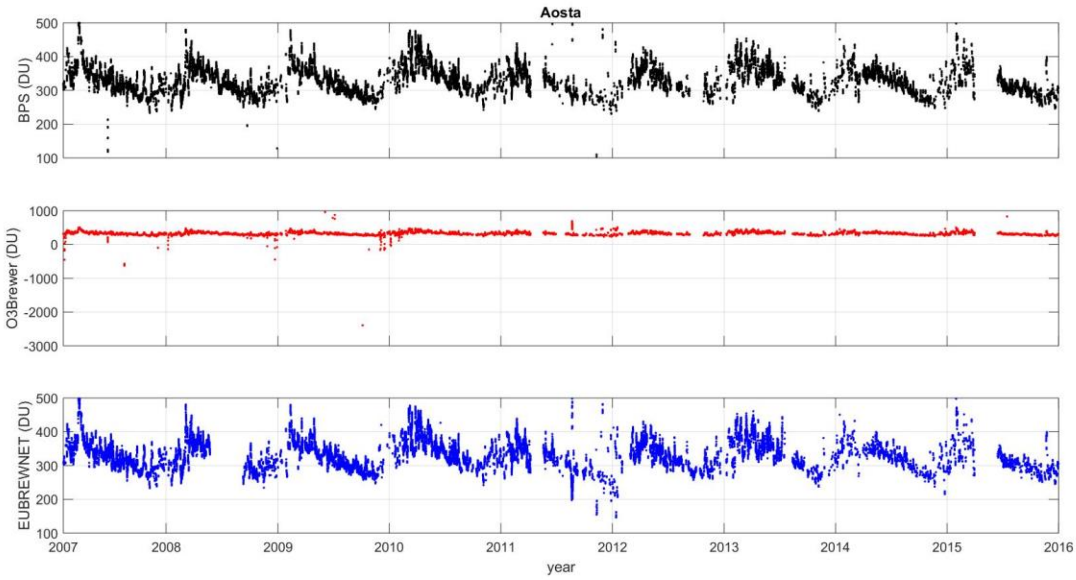

***Figure 3**. Individual TOC values generated by BPS (black), O3Brewer (red) and EUBREWNET (blue) at Aosta.*

It is worth noticing that ozone seasonal cycles show an overall similarity between the two sites with maximum value in late spring and minimum in late autumn, both on daily means and on individual ozone series. *The seasonal behaviour of O3Brewer is not easily distinguishable since the y-axis range has flatted it due to negative recalculated ozone values. However it is clearly visible that there are some periods in which TOC daily means as well as individual measurements* *obtained by the three processing software, are different* (mainly between 2006 and 2007 at Rome and at the end of 2011 at Aosta).

~~It is worth noticing that ozone seasonal cycles show an overall similarity between the two sites with maximum value in late Spring and minimum in late Autumn. However it is clearly visible that there are some periods in which TOC daily means, obtained by the three processing software are different (e.g. between 1994 and 1995, and between 2006 and 2007 at Rome).~~

~~With the aim at controlling the stability of the Brewer instruments, the R6 ratios are plotted in Fig. 2. In the same figure $R6_{BPS}$ (obtained as the sum of BPS correction and $R6_{ref}$), $R6_{smooth}$ series and the $R6_{ref}$ established during the calibration campaigns, are also shown.~~

In order to understand *where the differences come from, we analysed both individual TOC observations and the resulting daily values processed by BPS and O3Brewer*. Afterwards we compared both TOC retrievals with EUBREWNET data. Finally, the processed Brewer data were compared with OMI products.

3.1 Comparison between BPS and O3Brewer TOC retrievals

~~Looking at the standard lamp test results (Fig. 2), it can be noticed that the sensitivity of the instrument at Rome has changed mainly in two periods (between 1994 and 1995, and between 2006 and 2007). The problem turned out to be the deterioration of the filter (NiSO4/UG11) which was replaced during the calibration visits both in 1995 and 2007. Brewer 066 (Aosta) exhibited a better stability except in some occasional cases, where unusual R6 ratios were experienced. $R6_{BPS}$ shows a very similar behaviour to R6 at both stations due to the calculation method of the standard lamp correction by the BPS, whereas $R6_{smooth}$ time series displays a different trend with respect to R6. In particular, at Rome (Fig. 2, upper panel) $R6_{smooth}$ becomes a constant offset when the sensitivity of the instrument starts to change. This is due to the fact that the Gaussian low-pass filter in O3Brewer software is not applied when the difference between the reference $R6_{ref}$ and R6 exceeds a certain threshold (500 units, Stanek personal communication, 2016). In this case the correction is equal to the $R6_{ref}$ plus 500. Consequently, the temporal behaviour of $R6_{smooth}$ during these time intervals appears as a plateau. Once a new calibration is performed (i.e. new references of R6 and the ETC are defined) R6 and $R6_{smooth}$ show a similar behaviour again. At Aosta the $R6_{smooth}$ temporal evolution (Fig. 2, bottom panel) shows a stable behaviour.~~

~~A better visualization of the effect of the correction factor on TOCs is provided plotting the difference between the TOC retrievals (TOC BPS – TOC O3Brewer) as a function of the difference between $R6_{BPS}$ and $R6_{smooth}$ (Fig. 3). Large deviations between the two reprocessed TOC daily means appear when there is a large difference between $R6_{BPS}$ and $R6_{smooth}$, as expected.~~ ~~A better visualization of the effect of the correction factor on TOCs is provided plotting the difference between the TOC retrievals (TOC BPS – TOC O3Brewer) as a function of the difference between $R6_{BPS}$ and $R6_{smooth}$ (Fig. 3). Large deviations between the two reprocessed TOC daily means appear when there is a large difference between $R6_{BPS}$ and $R6_{smooth}$, as expected.~~

Fig. 4 shows the temporal behaviour of the ozone differences between BPS and O3Brewer taking into account both daily means whereas Fig. 5 shows individual values. It can be noticed that in several cases large differences can be attributed to wrong negative ozone recalculations by O3Brewer as also shown in Fig. 2 and 3. The minimum and maximum differences in the daily means are -278.1 DU and 567.9 DU at Rome, -332.3 DU and 532.0 DU at Aosta, respectively. The differences between BPS and O3Brewer individual ozone values range from a minimum of -304.4 DU to a maximum of 90.6 DU at Rome, from -435.6 DU to -157.7 DU at Aosta.

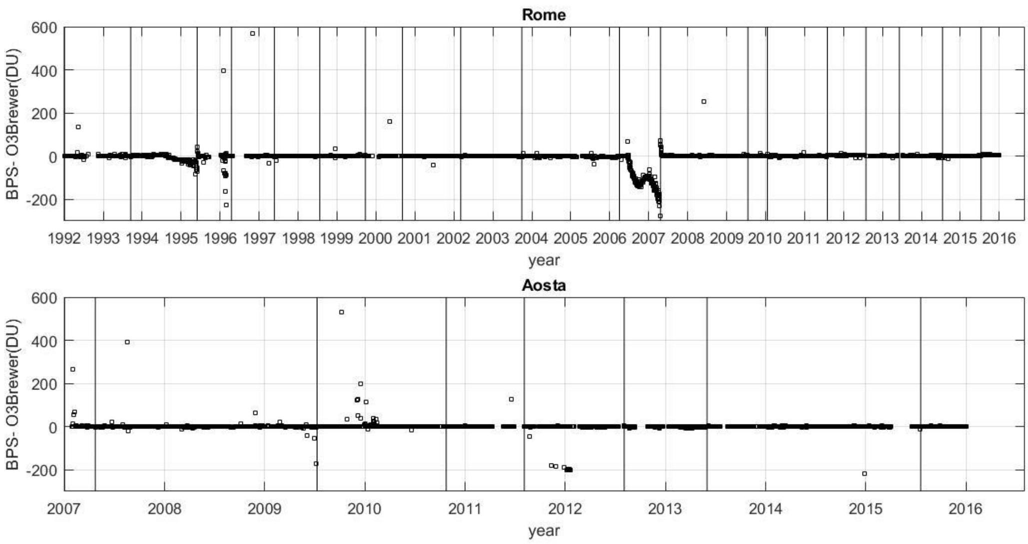

Figure 4. *Time plot of the differences between BPS and O3Brewer daily means at Rome (upper panel) and at Aosta (bottom panel). Vertical lines represent the date of the calibration campaigns.*

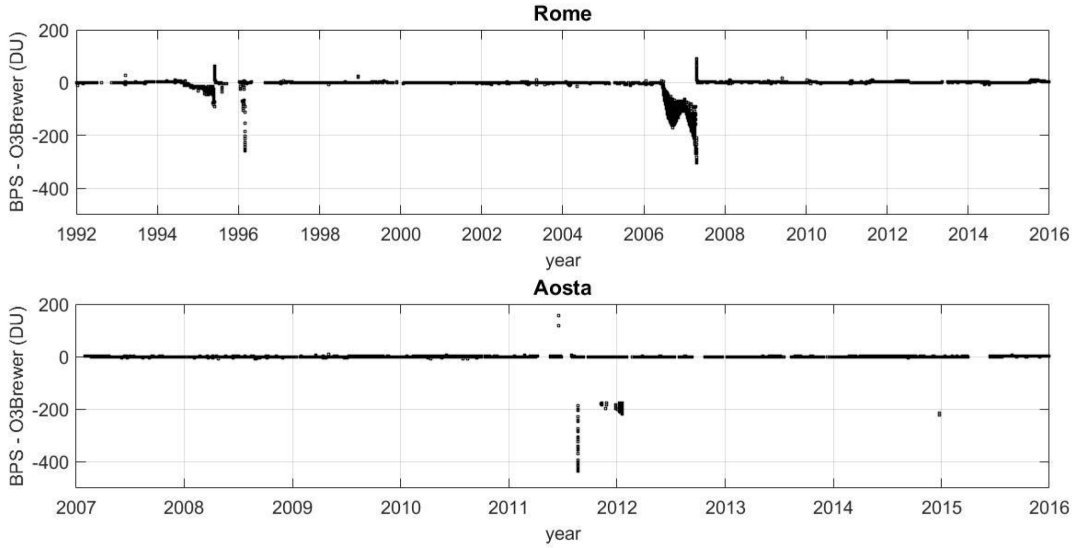

Figure 5. *Time plot of the differences between BPS and O3Brewer individual ozone data at Rome (upper panel) and at Aosta (bottom panel).*

We took into consideration the spectral sensitivity of both Brewer instruments through the R6 ratio time behaviour (Fig. 6). In the same figure it is also plotted how each software ($R6_{BPS}$ and $R6_{smooth}$) tracks changes in the spectral sensitivity of the instrument. $R6_{BPS}$ is obtained as the sum of BPS correction and $R6_{ref}$. $R6_{ref}$ values established during the calibration campaigns are also plotted. *It is worth noticing that the number of standard lamp test per day is on average from 4 to 6 at Rome, and from 2 to 4 in winter and from 8 to 10 in summer at Aosta and that only the daily means of BPS correction and $R6_{smooth}$ are stored. The latter is calculated if at least one standard lamp test is performed.*

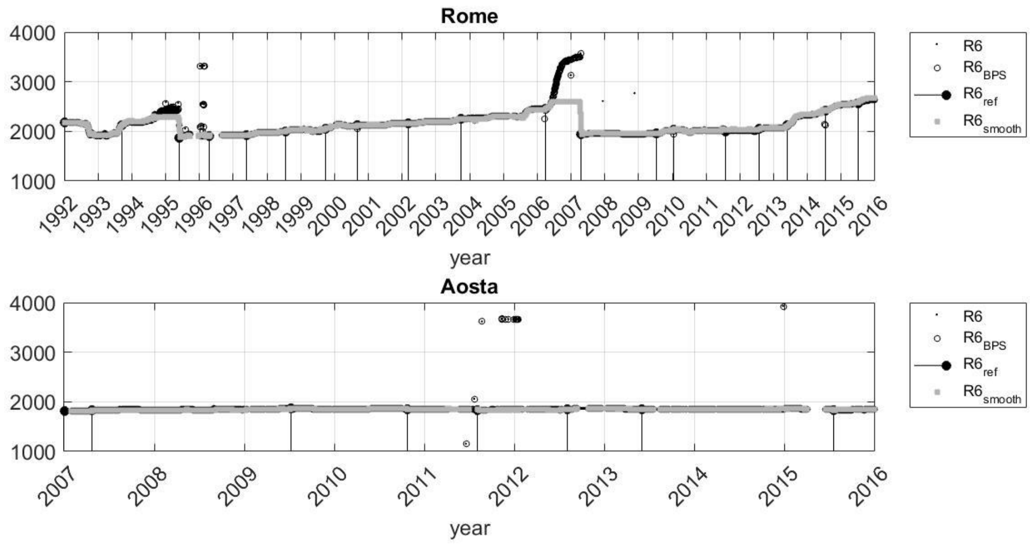

Figure 6. Daily series of the ratios R6, $R6_{BPS}$ and $R6_{smooth}$ at Rome (upper panel) and at Aosta (bottom panel). Vertical lines represent $R6_{ref}$ established during each calibration campaign.

Looking at R6 behaviour (Fig. 6 upper panel), it can be noticed that the sensitivity of the instrument at Rome has changed mainly in two periods (between 1994 and 1995, and between 2006 and 2007). $R6_{smooth}$ *becomes a constant offset when the sensitivity of the instrument starts to change. The cut off is not exactly equal to the threshold set in the configuration (in this case 500 units), but lower, because the filter looks 10 days before and 10 days after the date when SL R6 is calculated. If the cut off remains constant, it means that the last calculated correction which passes through rejection criteria, is taken into account, the same situation is experienced when there is no valid SL test (Stanek personal communication, 2018). Consequently, the temporal behaviour of* $R6_{smooth}$ *during these time intervals appears as a plateau. In this case SL correction is not applied since it is too high. Once a new calibration is performed (i.e. new references of R6 and the ETC are defined) R6 and* $R6_{smooth}$ *show a similar behaviour again.*

Brewer 066 (Aosta) exhibited a better stability except for some R6 spikes (Fig. 6, bottom panel) whereas $R6_{smooth}$ time series shows a stable behaviour with respect to R6. $R6_{BPS}$ shows a similar behaviour to R6 at both stations due to the calculation method of the standard lamp correction by the BPS.

~~Two distinct periods were found at Rome belonging to the first condition (3rd October 1994 - 10th June 1995; 27th June 2006 - 24th July 2007), due to the deterioration of photomultiplier filter which was replaced during the calibration visit both in 1995 and in 2007. In those cases the standard lamp correction should not be applied. Some days that belong to anomalous cases were found at Aosta. Occasional anomalous R6 ratios can occur for several reasons, such as wrong wavelength selection by the micrometer, communication problems or incorrect zenith drive position in relation to the lamp.~~

A better visualization of the effect of the correction factor on TOCs is provided by plotting the difference between the TOC daily means (BPS – O3Brewer) as a function of the difference between $R6_{BPS}$ and $R6_{smooth}$ (Fig. 7). Large deviations between the two reprocessed TOC daily means appear when there is a large difference between $R6_{BPS}$ and $R6_{smooth}$. However large differences occur even if $R6_{BPS}$ does not differ too much from $R6_{smooth}$.

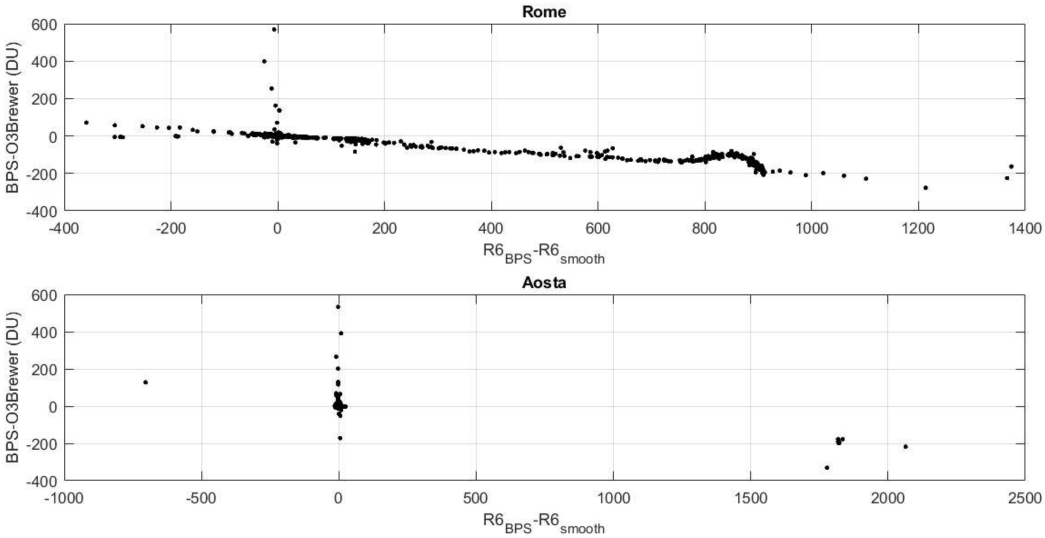

Figure 7. Differences between BPS and O3Brewer TOC daily means vs $R6_{BPS}$-$R6_{smooth}$ at Rome (upper panel) and at Aosta (bottom panel).

Three circumstances are here analysed when differences between BPS and O3Brewer ozone data exceed the value of the declared DS accuracy: 1. $R6_{BPS}$ lower than $R6_{smooth}$, 2. $R6_{BPS}$ higher than $R6_{smooth}$, 3. $R6_{BPS}$ similar to R6smooth.

1. $R6_{BPS}$ lower than $R6_{smooth}$.

Slight ozone difference took place when $R6_{BPS}$ was lower than $R6_{smooth}$ (at least 100 units), then the difference in ozone daily means was between -3% and 21% and in case of individual values from -3% up to 27 %, at Rome. At Aosta there was only one episode (2011/6/18) in which the O3Brewer daily mean differed about 30% from BPS. In that case, O3Brewer average was derived by three individual ozone values that showed the same difference with respect to the BPS ones. In this case, a large negative correction was applied to ozone values, thus generating a falsely high ozone case. The spike in the R6 value was originated by the two wrong SL test carried in that day caused perhaps by the micrometer in a wrong position, noisy communication, incorrect zenith drive position, or lamp aging. Consequently, the negative BPS correction generated high ozone values with a large standard deviation, whereas $R6_{smooth}$ was not applied to individual TOC data that resulted consistent with ozone values before and after that date.

At Rome the conditions in which $R6_{BPS}$ was lower than $R6_{smooth}$ occurred during the calibrations in 1995, 2006, 2007 and 2014. The discrepancy between the two codes could have been caused by the offset introduced by the way BPS determines the R6 reference value as for the other code the $R6_{ref}$ is obtained during the calibration campaign and set manually in the *configuration. The BPS $R6_{ref}$ is computed with a triangular smoothing filter of SL-test over the 15 day period after the calibration and it is calculated "on the fly" from daily mean SL values and it is not stored (Fioletov, personal communication 2018).*

To look into the possible effect of the BPS offset we estimated $R6_{ref\_BPS}$, for each day over the 15 days after the calibration by subtracting the correction (reported in the file o3data.txt) to the corresponding R6 value. Then the average over the 15 $R6_{ref\_BPS}$ values was compared with $R6_{ref}$ (given by hand after the calibration). The estimated offset introduced by BPS with respect to $R6_{ref}$ is very small, ranging between -19 to 6 units at Rome and between -10 to 2 units at Aosta. Consequently, the BPS offset appears not to be responsible for the ozone difference that can be attributed to the calculation method of the standard lamp correction.

2. $R6_{BPS}$ higher than $R6_{smooth}$

Large negative ozone differences occur when $R6_{BPS}$ is higher than $R6_{smooth}$ (at least >100 units). This causes a variation between the daily means generated by the codes from -5% till -50% at Rome and from -51% till -91% at Aosta: Considering the individual values a mean percentage difference between -3.1% and -57% was found at Rome, and of the same magnitude as that of daily means at Aosta.

Two long periods were found at Rome belonging to this condition (29st October 1994 - 5th May 1995; 26th June 2006 - 16th April 2007). *The large drift in R6 turned out to be the deterioration of the filter (NiSO4/UG11) which was replaced during the calibration visits both in 1995 and 2007.* In both cases it can be observed the cut off in $R6_{smooth}$ and hence the O3Brewer recalculation provides uncommon TOC values. *Then, we processed Rome ozone data using O3Brewer by setting the SL maximal limit to higher value to assess whether the smooth correction can properly process ozone data when large changes occurred in the instrumental response. The SL maximal correction limit was to 3000 units keeping identical conditions for the air mass and the standard deviation of the previous processing. In addition, a further ozone processing was carried out by switching off the smoothing filter. Fig. 8 shows the time series of the ratios R6, $R6_{BPS}$ and $R6_{smooth}$ (setting the limit to 3000 units) at Rome. It can be noticed that the $R6_{smooth}$ has now similar behaviour as $R6_{BPS}$, nevertheless in some circumstances its behaviour is noisier than both $R6_{smooth}$ (setting the limit to 500 units) and $R6_{BPS}$.*

Fig.9 shows individual TOC data processed by O3Brewer 1) without applying the smoothing filter, 2) with the SL limit correction set to 500 and 3) with the SL limit correction set to 3000 units over the period of the R6 drift (2006-2007) at Rome. In the same figure, individual

BPS recalculations without modifying the set up are also plotted. A better agreement with BPS ozone data is visible when ozone data were processed without the smoothing filter and with higher cut off in R6, however there are still anomalous ozone values due SL correction, whereas ozone values calculated without the correction seem not be not affected.

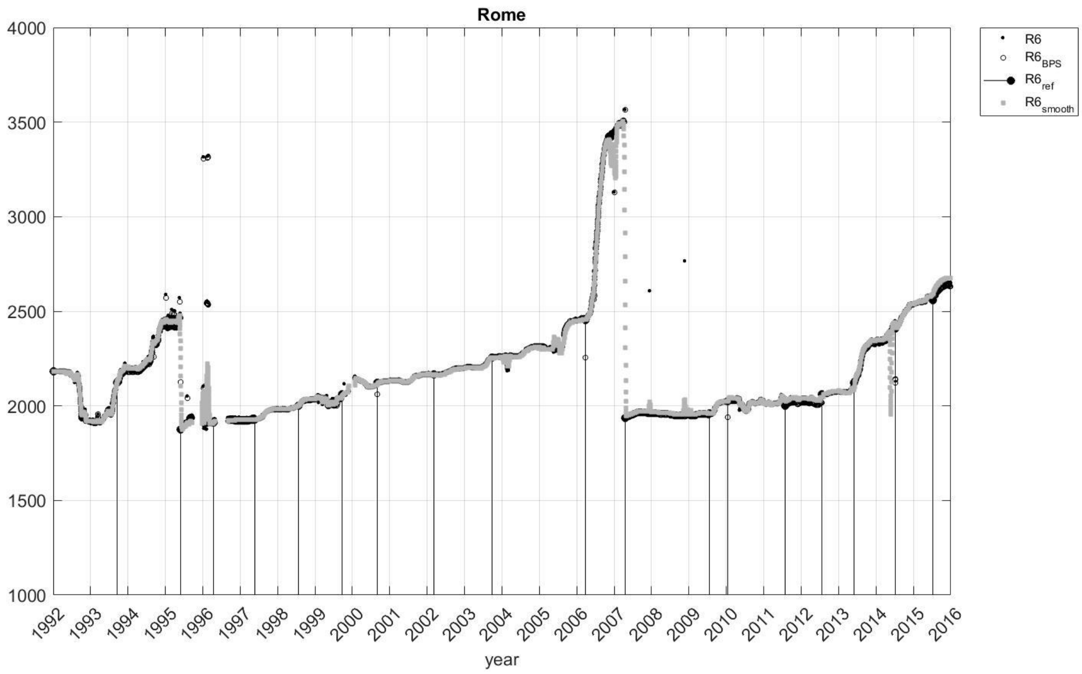

Figure 8. *Daily series of the ratios R6, $R6_{BPS}$ and $R6_{smooth}$ at Rome. Vertical lines represent $R6_{ref}$ established during each calibration campaign.*

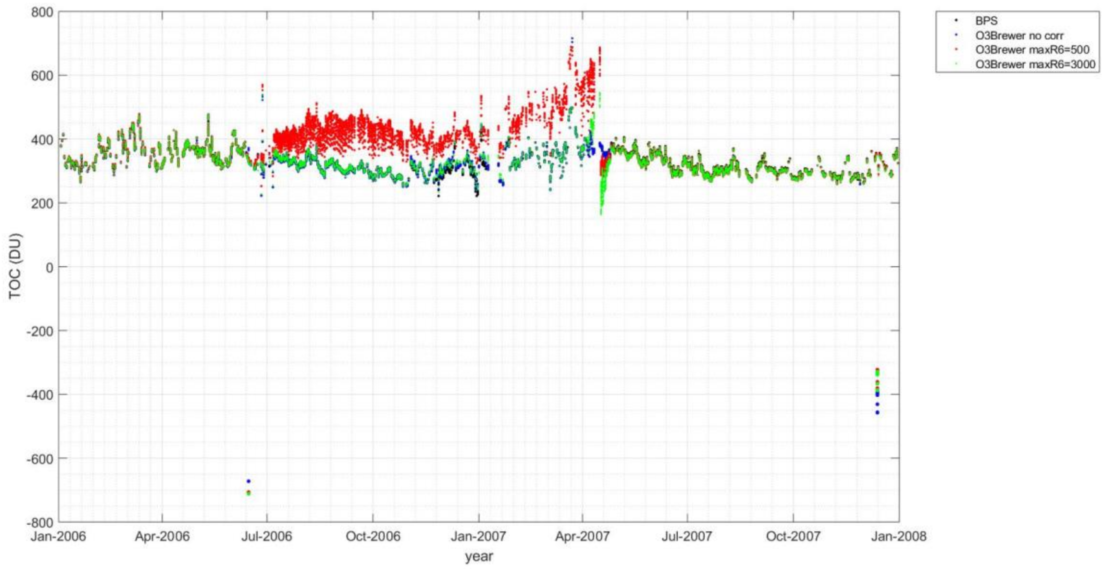

Figure 9. *Individual ozone values calculated by the BPS (black), O3Brewer without the filter smoothing correction (blue), O3Brewer with the cut off at 500 units (red), O3Brewer with the cut off at 3000 units (green) over the period of the drift in 2006 -2007 at Rome.*

The occasional anomalous R6 ratios occurred at Aosta, most of them in 2011 and at the beginning of 2012. Wrong wavelength selection by the micrometer, communication problems or incorrect zenith drive position in relation to the lamp could have caused the R6 spikes. In this case the algorithm of O3Brewer (with the cut off at 500 units) did not follow the abrupt change. The correction was not applied resulting in large over- or under-estimation of TOC or with uncertain data quality.

3. $R6_{BPS}$ similar to $R6_{smooth}$

A different number of observations taken into account in the determination of the daily means by the two codes can generate significant differences in some cases. The total number of individual calculated total ozone values by O3Brewer is 104666 at Rome and 50088 at Aosta, the number of those calculated by BPS is 100352 at Rome and 46617 at Aosta. Fig. 10 shows the number of individual ozone values calculated by O3Brewer which is, in some days s, higher than that of BPS.

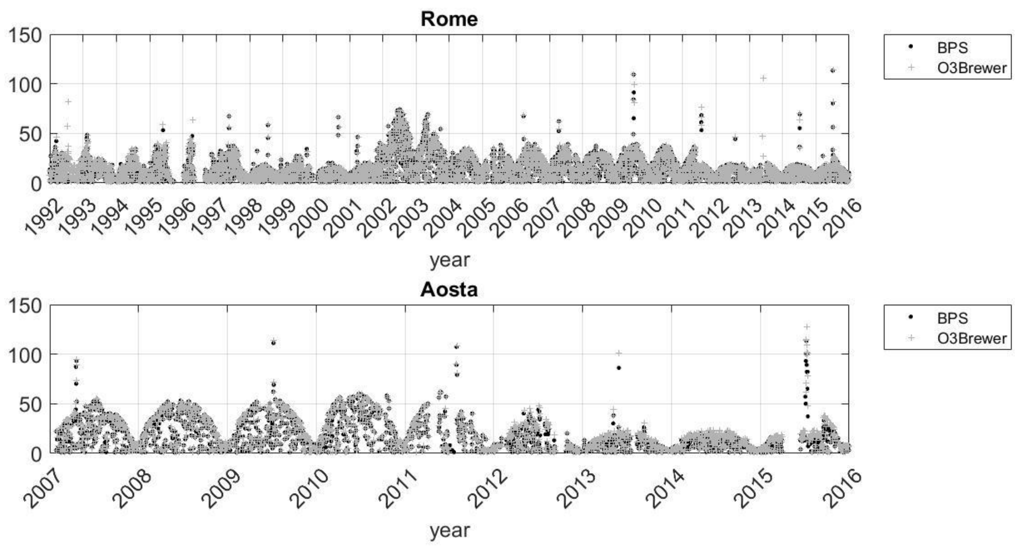

Figure 10. *Time plot of the number of individual ozone values per day.*

Such difference can be due to the fact that there are no filter conditions on the minimum and the maximum ozone values calculated by O3Brewer. Consequently, the daily means generated by this software are determined including anomalous values. The case of $R6_{BPS}$ similar to $R6_{smooth}$ responsible for significant ozone differences in the daily means (>5%) falls in this conditions.

As a specific example of the above case, we showed individual ozone values generated by both codes on 23/06/2001 at Rome with a daily average of 335 DU for BPS and 375.4 DU for O3Brewer (Fig.11, upper panel). The high individual ozone value generated by O3Brewer (618.7 DU) affecting the daily average is clearly visible. Another example is provided for Aosta (Fig. 11, lower panel). On 5/1/2010 the daily average is 323.5 DU for BPS whereas it is 208.4 DU for O3Brewer. It is found that very low ozone values generated by O3Brewer, not discarded in the determination of the daily means, affect the quality of its value.

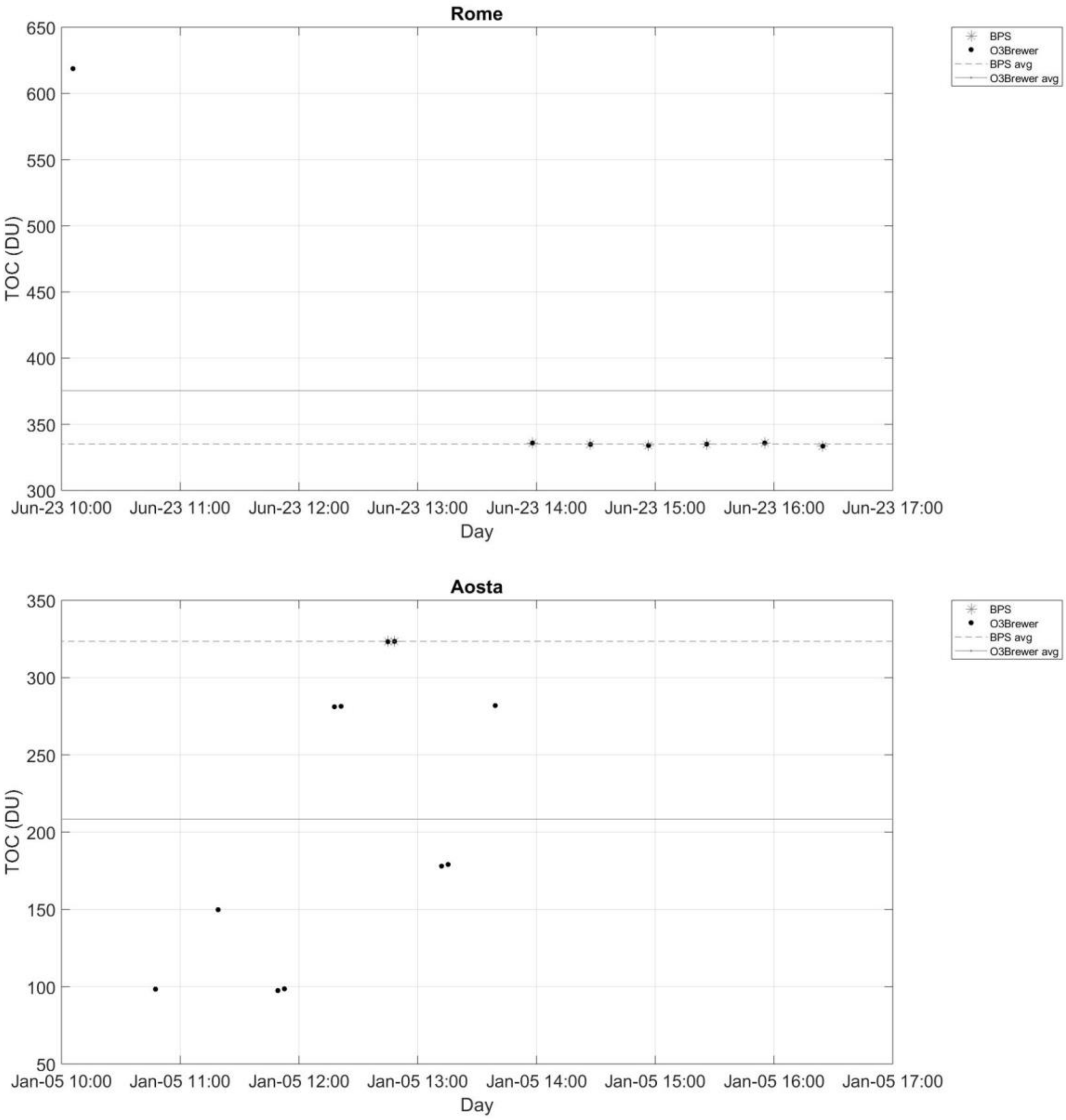

Figure 11. *Individual TOC values generated by BPS and O3Brewer on 23/06/2001 at Rome (upper panel) and on 5/1/2010 at Aosta (bottom panel) taken as examples where large difference between occurred although the SL correction is similar. Horizontal lines (dashed for BPS; solid for O3Brewer) represent the daily average (avg).*

~~Then, TOC daily means with daily std ≥50 DU were also discarded since large daily variability often occurs in case of ozone spikes in that particular day. This high threshold was chosen because the diurnal TOC variability can reach 40-50 DU at mid- to high-latitude locations during late spring and summer (Siani et al., 2002; Tzortziou et al., 2012). TOC daily means without R6 values (no SL test was performed in that day) were also discarded.~~

Table 3 shows the statistical comparison between O3Brewer (with cut off at 500 units) and BPS individual data and daily means, after data belonging to three groups described in the previous section, have been discarded. TOC data without R6 values (no SL test was performed in that day) were also discarded. The temporal behaviour of the differences between O3Brewer and BPS individual calculated ozone values, are plotted in Figure 12 showing a variability in general within ±25 DU at Rome and ±10 DU at Aosta.

A good overall agreement is found both on individual values and daily means when data belonging to the above conditions were removed, the correlation is close the unity at both stations; MPE is not significant on both individual values and daily means at Rome as well as at Aosta.

Table 3. *Summary of the statistics O3Brewer vs BPS at both sites (N= number of data; RHO= Spearman correlation; MB =Mean Bias, MPE=Mean Percentage Error, RMSE =Root Mean Square Error, the uncertainty of MB and MPE is characterized by the standard deviation).*

O3Brewer_vs_BPS	*N*	*RHO*	*MB (DU)*	*MPE (%)*	*RMSE (DU)*
Rome					
Individual values	*89273*	*0.997*	*-0.6±2.1*	*-0.2±0.7*	*2.18*
Daily averages	*6304*	*0.997*	*-0.8±2.4*	*-0.2±0.7*	*2.47*
Aosta					
Individual values	*44117*	*0.999*	*0.1±0.8*	*0.03±0.30*	*0.83*
Daily averages	*2381*	*0.999*	*0.004±1.700*	*0.001±0.600*	*1.70*

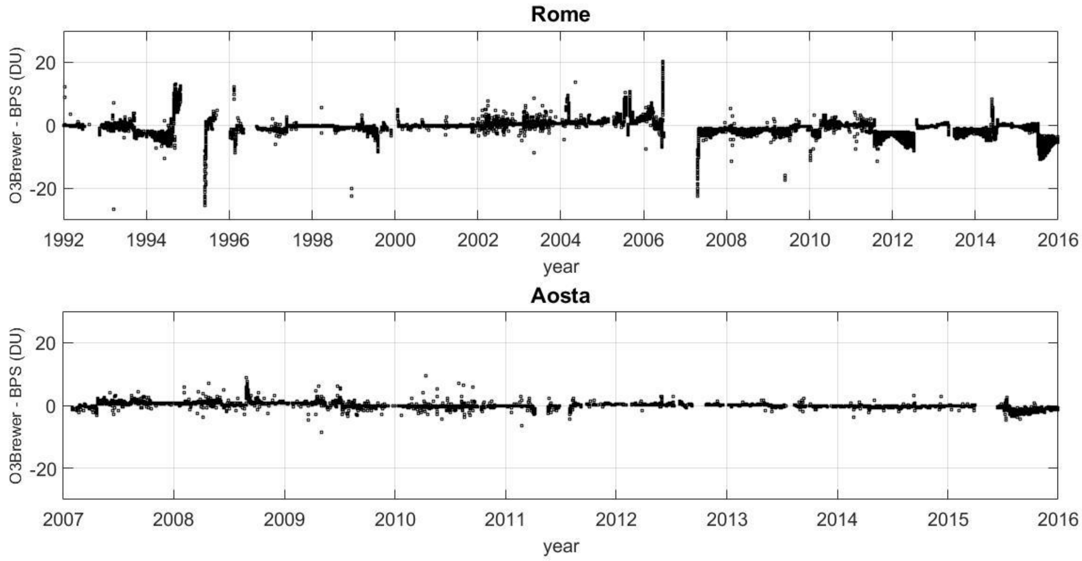

Figure 12. *Difference between individual TOC values generated by BPS and O3Brewer at Rome (upper panel) and at Aosta (bottom panel) when anomalous values were discarded.*

3.2 Comparison of BPS and O3Brewer TOC retrievals with EUBREWNET data

The TOC individual values and daily means retrieved by O3Brewer and BPS data were compared with those derived from EUBREWNET retrievals ~~(including also the questionable data).~~ *The analysis was performed by removing data belonging to the three periods mentioned in Section 3.1 from all series.*

~~There is no significant difference among the TOC retrievals (less than 1%) except in the comparison O3Brewer *vs* EUBREWNET in the case of "all" at Rome in which MPE is 2.5%. This is mainly due to the TOCs over the period 27th June 2006 – 24th July 2007, when the O3Brewer processing software applied a constant correction value. Although the overall agreement of the BPS and O3Brewer TOC data with EUBREWNET data is clearly very high (as expected), it is worth noticing from RMSE results that slight differences are still experienced depending on the software in use and, specifically, on the standard lamp correction algorithm.~~

Table 4 shows the statistical results of the two processed TOC data sets against the EUBREWNET data set. It was found that the difference among the TOC retrievals is less than 1%.

Table 4. Summary of the statistics O3Brewer vs BPS at both sites (N= number of data; RHO= Spearman correlation; MB =Mean Bias, MPE=Mean Percentage Error, RMSE =Root Mean Square Error, the uncertainty of MB and MPE is characterized by the standard deviation).

O3Brewer *vs*	N	RHO	MB (DU)	MPE (%)	RMSE (DU)

EUBREWNET					
Rome					
Individual values	38227	0.996	-0.2±3.8	- 0.05±1.00	3.80
Daily averages	2972	0.996	-0.1±4.6	- 0.02±1.20	4.60
Aosta					
Individual values	35746	0.997	0.3±5.3	0.2±2.4	5.33
Daily averages	2186	0.994	0.5±7.6	0.2±3.2	7.76
BPS vs EUBREWNET					
Rome					
Individual values	38227	0.995	1.0±4.1	0.3±1.1	4.27
Daily averages	2972	0.995	1.2±5.0	0.4±1.3	5.11
Aosta					
Individual values	35746	0.997	0.2±5.3	0.1±2.4	5.34
Daily averages	2186	0.994	0.5±7.6	0.2±3.2	7.59

However looking at Figs. 13-14 the differences between the individual ozone values calculated by BPS and EUBREWNET (Fig.13) and, by O3Brewer and EUBREWNET (Fig.14) are in some cases relevant. It seems that problems of the standard lamp values not properly filtered by the currently applied 7-days window smoothing, generate results less reliable (see the temporal behaviour of $R6_{EUBREWNET}$ *in Fig.15). This problem could be solved in the level 2 data, in which the setting a filter in the R6 values and smoothing the R6 time series is planned to be taken into account in the EUBREWNET algorithm (Fountoulakis, personal communication 2018). However, although these options exist in the configuration form they are still inactive.*

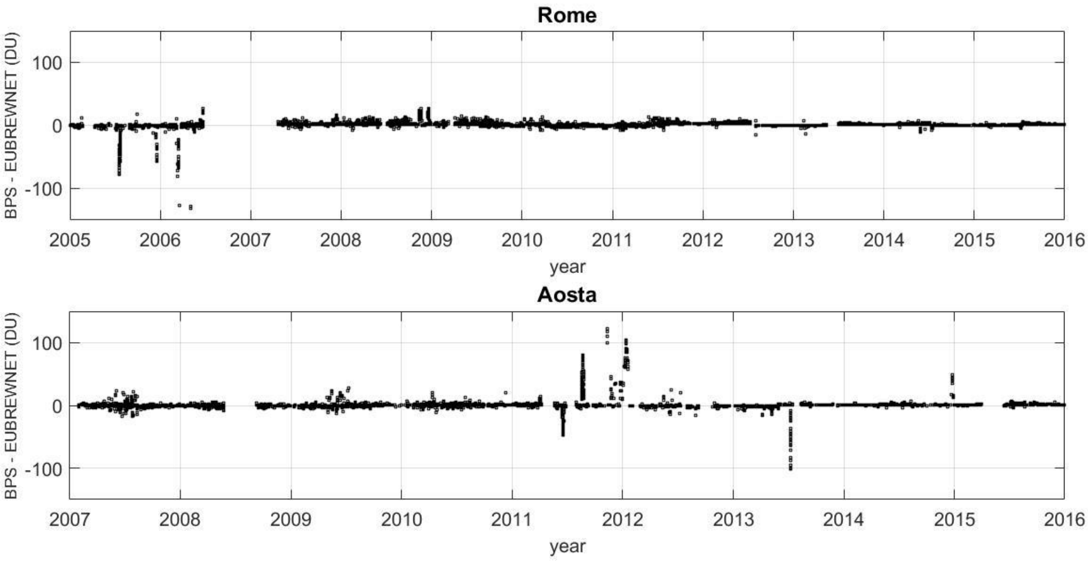

Figure 13. *Difference between individual TOC values generated by BPS and EUBREWNET (Rome upper panel and Aosta lower panel).*

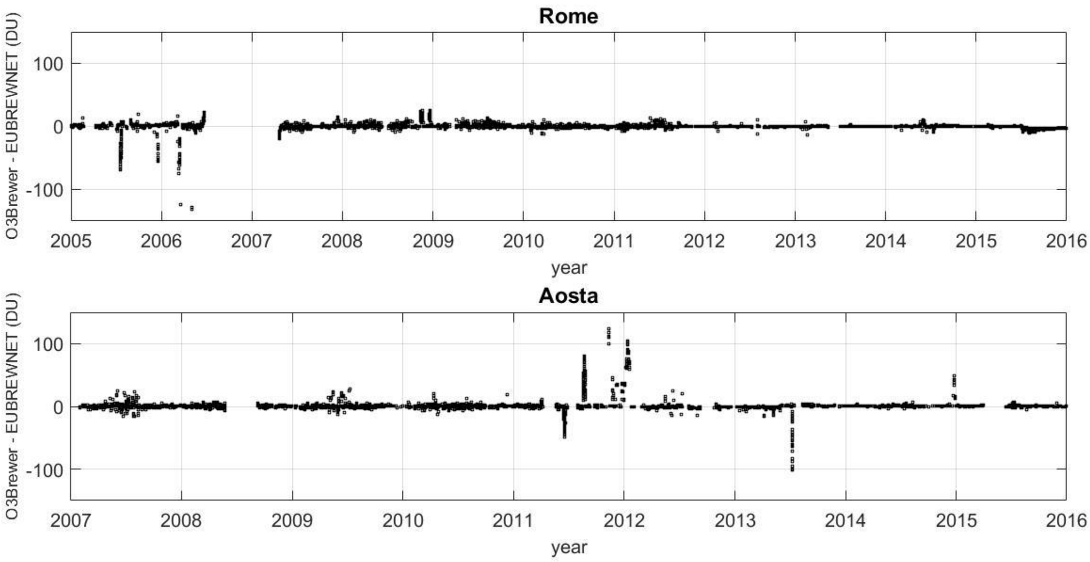

Figure 14. *Difference between individual TOC values generated by O3Brewer and EUBREWNET (Rome upper panel and Aosta lower panels).*

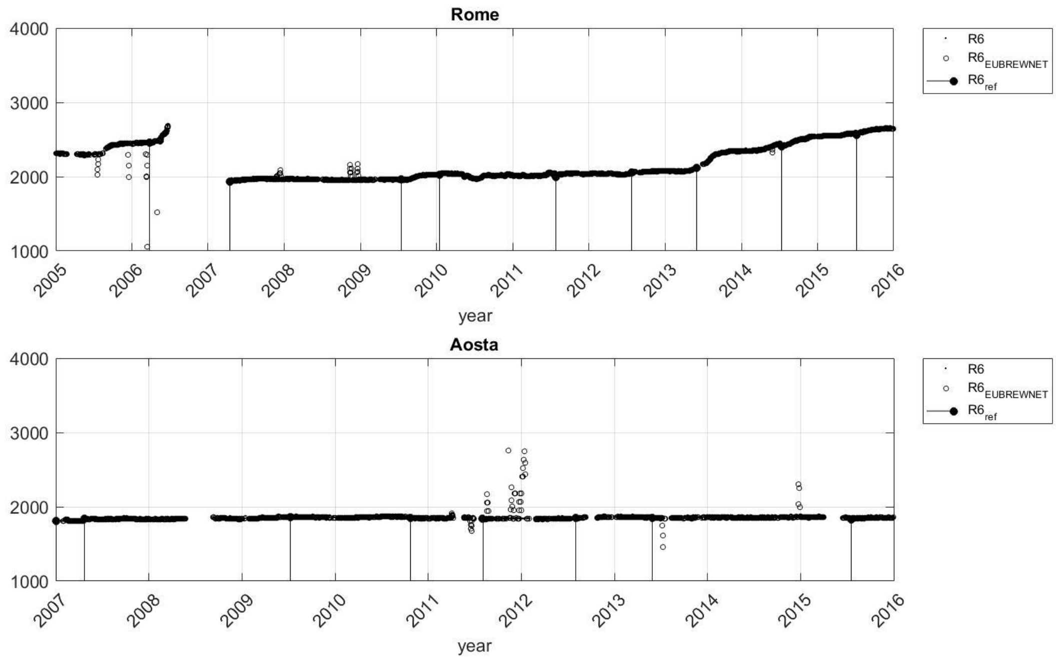

***Figure 15.** Daily series of the ratios R6, $R6_{EUBREWNET}$ at Rome (upper panel) and at Aosta (lower panel). Periods with the R6 drift or spikes were removed. Vertical lines represent R6ref established during each calibration campaign.*

3.3 Comparison of BPS, O3Brewerand EUBREWNET TOC retrievals with OMI data

OMI overpasses were also compared with the processed Brewer TOC retrievals. The comparison was performed taking into account the same design criteria described in the previous section. The scatterplots of OMI *vs* Brewer data are shown in Fig. 16. However depending on the Brewer processing software a different behaviour is visible, even when only "good" data are considered. *It can be observed that EUBREWNET data show larger deviations from the bisectrix with respect to the other retrievals.*

~~In the latter case a high degree of proportionality between OMI and the ground-based total ozone column data can be noticed. However depending on the Brewer processing software a different behaviour is visible, even when only "good" data are considered.~~

The results of the statistical analysis are summarized in Table 5. The results of the statistical analysis are summarized in Table 5. ~~In general in both sites, the TOCs retrieved by the three processing software show an excellent agreement with OMI products (the Spearman coefficient is very high).~~ *In general, the scaled correlation is, for both sites, on average RHOs= 0.8 which represents how the series are well connected in the short term.*

OMI products show a systematic underestimation with respect to ground-based data. *At Rome satellite data are less than 1 % for both O3Brewer and EUBREWNET whereas at Aosta about 2.5%;* 1.2% (Rome) and 2.5% (Aosta) in the case of BPS data. These results are in agreement with previous studies on validation of the OMI total ozone column by Brewer spectrophotometry conducted at the same latitudes (Ialongo et al., 2008; Anton et al., 2009).

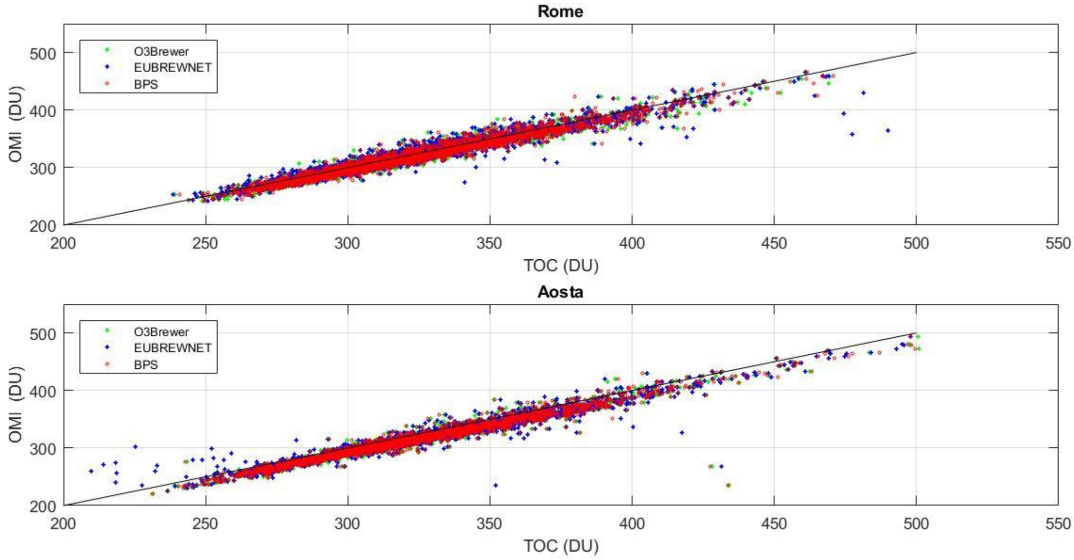

Figure 16. *Scatterplots OMI versus Brewer total ozone column at Rome (upper panel) and Aosta (lower panel). The solid line represents the bisectrix, The comparison is carried out with O3Brewer (green), EUBREWNET (blue) and BPS (red) data.*

Table 5. Summary of the statistics of the comparison between OMI versus BPS, O3Brewer and EUBREWNET (N= number of data; RHOs= Spearman scaled correlation; MB =Mean Bias, MPE=Mean Percentage Error, RMSE =Root Mean Square Error, the uncertainty of MB and MPE is characterized by the standard deviation).

Rome	N	RHOs	MB (DU)	MPE (%)	RMSE (DU)	**Aosta**	N	RHOs	MB (DU))	MPE (%)	RMSE (DU)
	OMI vs BPS										
	2622	0.841	-4.0±7.8	-1.2±2.3	8.63		2022	0.9	-8.6±10.4	-2.5±4.4	13.45
	OMI vs O3Brewer										
	2622	0.843	-2.8±8.4	-0.8±2.5	8.85		2022	0.882	-8.6±10.7	-2.5±4.8	13.74
	OMI vs EUBREWNET										
	2522	0.814	-2.8±9.6	-0.8±-2.7	9.99		1849	0.835	-8.2±10.5	-2.4±3.5	13.30

~~When comparing RMSE values it can be noticed that RMSE changes at Rome from 8.39 DU to 37.63 DU, at Aosta from 11.12 19.71 DU (higher in the case of all data reprocessed by O3Brewer) which supports the observed scatter plot shown in Fig. 5.~~

~~The slight differences among the statistical parameters used in the comparison of "good" cases are observable. A possible explanation is that the comparison was performed using Brewer data~~

~~averaged on daily basis which includes local and temporal fluctuations that cannot be detected by overpasses and from approaches of the standard lamp correction in the software in use~~. *When comparing RMSE values it can be noticed that RMSE at Rome is lower than that found at Aosta, which supports the observed scatter plot shown in Fig. 16.*

Besides, systematic differences between ozone estimated from OMI and from Brewer at Aosta could be related to the ground pixel size which can affect ozone amounts probed by the satellite, due to the complex orography of the valley.

3.4 Comparison among the trends estimated by the three processing software ozone retrievals

The detected trends in ozone series calculated by using the three processing software are reported in Table 6. The trends were quantified over the period 2005-2015 for Rome to be consistent with the EUBREWNET ozone data coverage, and 2007 -2015 for Aosta. Ozone data showing large differences among the codes, were not included in the trend analysis.

The QBO and solar cycle effects were not filtered in the ozone series. The former was found small at mid-latitude stations (Fountoulakiset al., 2016), whereas the latter was not taken into account due the short length of the analysed ozone series (< 11 years). All trends were found to be not statistically significant (p-value is 0.05).

It is clear from Table 6 that there are not significant differences in the trends expressed in terms of percentage variation per decade among the three codes, when data affected by rapidly changes in R6 or the spectral response of the instrument shows a persistent drift, were removed.

***Table 6**. The total ozone linear trends derived by the processed ozone values using three different processing codes*

	period	*BPS (% per decade)*	*O3Brewer (% per decade)*	*EUBREWNET (% per decade)*
Rome	*2005-2015*	*-0.23 ± 0.18*	*-0.32 ± 0.20*	*-0.34 ± 0.21*
Aosta	*2007-2015*	*0.07 ± 0.35*	*0.04 ± 0.34*	*0.00 ± 0.38*

4.Conclusions

This study analyzed the total ozone column recalculations at Rome and Aosta using three different software packages. We found that large differences in total ozone column retrievals can be experienced when the instrumental sensitivity exhibits ~~a long-term drift~~ *a fast and dramatic drift between two consecutive calibration*s or spikes. These conditions can affect TOCs retrievals due to the algorithm of the standard lamp correction applied. ~~When anomalous R6 values occur, the correction applied by O3Brewer software is a constant value producing anomalous TOCs. Similarly, the current Level 1.5 in the EUBREWNET can produce erroneous ozone recalculations when anomalous R6 values are experienced. This can be avoided if days with R6 outliers are removed manually The issue is expected to be solved in Level 2.0 products, when they will be released. The BPS ozone recalculations are less affected by abrupt changes in the sensitivity, even in case of R6 drifts. After discarding the periods with drifts or occasional abrupt changes in R6, a good overall agreement is found between BPS, O3Brewer and EUBREWNET (MPE about <0.3%). However a spread among the processing software was still found.~~

When R6 exceed the default value of the cut off (500 units) set in the configuration of the O3Brewer software during an occasional spike, the correction is not applied, whereas the BPS correction does. This could generate false high/low ozone values. In latest version of O3Brewer it is possible to set the cut off to higher value that is useful when there a large R6 drift is experienced. However, anomalous ozone values can be still observed, since in O3Brewer there are no filter conditions on the minimum and the maximum ozone values. Similarly, the current Level 1.5 in the EUBREWNET can produce erroneous ozone recalculations when anomalous R6 values are experienced. The issue is expected to be solved in Level 2.0 products, when they will be released. The BPS ozone recalculations seem to be less affected in the case of R6 drift. However when serious changes in the spectral sensitivity of instrument is experienced, a solution consists in dividing the periods of R6 drifts into shorter time intervals and for that period a new set of constants ($R6_{ref}$ and ETC) could be established by the user as the averages of R6 ratios in that time interval. This process ("synthetic calibration") allows the user to introduce standard lamp corrections larger than the software hardcoded thresholds. In any case the synthetic constants in use must be confirmed at the next calibration with the reference instrument.

Here we decided to discard the periods with drifts or occasional abrupt changes in R6, and a good overall agreement is found between BPS, O3Brewer and EUBREWNET (MPE about <1%). However a spread among the EUBREWNET individual ozone values and those retrieved by the other twos codes is still found, probably due to the standard lamp values not filtered properly by the currently applied 7-day window smoothing, generating results less reliable.

The analysis of the differences between recalculated TOCs and OMI overpasses showed that the latter dataset underestimate less than 2% ground –based total ozone columns at Rome and less than 3% at Aosta (using "good" cases). Yet, the estimate of the trends using the retrievals from the three different codes resulted not be affected when ozone data with anomalous R6 values are removed.

The operators should constantly monitor the sensitivity of the instrument and know carefully the processing software used to recalculate the total ozone. This means that the quality-controlled data cannot be assured only by automatic data rejection rules of the adopted software, but a rigorous manual data inspection is always necessary to prevent inconsistent data produced by the processing software package in use.

As a final remark, it is important to underline that for sake of consistency and comparability between the results from different stations which send ozone products to international data centres such as WOUDC or others, it is important to know the processing software used to generate individual ozone values, the time behaviour of the instrumental stability, the method applied for the standard lamp correction as well as the adopted rejection criteria to determine the daily means.

Data availability. The data used for the present study can be asked to the authors of the present paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgments: We thank the European Brewer Network (http://rbcce.aemet.es/eubrewnet/) for providing access to the data used in this investigation, and the COST Action ES1207 "A European Brewer Network (EUBREWNET)", supported by COST (European Cooperation in

Science and Technology). We also thank NASA Goddard Space Flight Center for OMI data available (https://avdc.gsfc.nasa.gov/).

The authors are grateful to M. Stanek, V. Fioletov, A. Ogyu and I. Fountoulakis for their helpful clarifications on the processing software.

The authors thank the anonymous reviewers for their valuable suggestions and comments.

This paper is dedicated in memory of Ken Lamb, founder of International Ozone Services Inc. (IOS), who zealously delivered accurate ozone and UV calibrations to the worldwide Brewer community.

Author Contributions: All authors have helped to develop the paper. A.M. S played the major role supervising and coordinating the whole work; G.R. C. and H. D. have equally provided helpful comments on the draft. F.F, F. S. and A. R. have contributed in the elaboration of the Brewer and satellite data. A.M. S and G.R. C. are responsible for establishing and maintaining Brewer 067; H. D. has contributed with data of Brewer 066 and in establishing and maintaining the site; F.F and M. P. have given Matlab support; V. S. has given support with the Brewer processing software.

References

Antón, M., López, M., Vilaplana, J. M., Kroon, M., McPeters, R., Bañón, M., and Serrano, A.: Validation of OMI-TOMS and OMI-DOAS total ozone column using five Brewer spectroradiometers at the Iberian Peninsula, J. Geophys. Res.-Atmos., 114, D14307, doi:10.1029/2009JD012003, 2009.

Balis, D. Kroon, M., Koukouli, M. E., Brinksma, E. J., Labow, G., Veefkind, J. P., and McPeters, R. D.: Validation of Ozone Monitoring Instrument total ozone column measurements using Brewer and Dobson spectrophotometer ground-based observations, J. Geophys. Res., 112, D24S46, doi:10.1029/2007JD008796, 2007.

Bass, A.M., and Paur, R.J.: The ultraviolet cross-sections of ozone, I, The measurements, in atmospheric ozone. In: Zerefos CS, Ghazi A, Reidel D (eds) Proceedings of the Quadrennial Ozone Symposium, Halkidiki, Greece, 1984. Dordrecht, Holland,606- 610,1985.

Bhartia, K. and Wellemeyer, C.:Toms-v8 total O3 algorithm, OMI Algorithm Theoretical Basis Document, 2, 15–31, 2002.

Bordi, I., Fraedrich K., Sutera A., Zhu X.,: On the climate response to zero ozone. Theor. Appl. Climatol. 109, 253 – 259, 2012.

Brewer, A.W.: A replacement for the Dobson spectrophotometer? Pure App.Geophys.106-108, 919 -927, 1973.

Brewer, A.W. and Kerr, J. B.: Total ozone measurements in cloudy weather, Pure appl. Geophys. 106-108, 928-937, 1973.

EUBREWNET (A European Brewer Network) COST Action ES 1207 (http://rbcce.aemet.es/eubrewnet, http://www.eubrewnet.org/cost1207, last accessed on June 2017; http://www.eubrewnet.org/cost1207/ last accessed on November 2017).

Dessler, A.: Chemistry and Physics of Stratospheric Ozone, International Geophysics Series 74. Academic press, London U.K., 2000.

Diémoz, H., Eleftheratos, K., Kazadzis, S., Amiridis, V., Zerefos, C. S.: Retrieval of aerosol optical depth in the visible range with a Brewer spectrophotometer in Athens, Atmos. Meas. Tech., 9, 1871–1888, 2016

Dobson G.M.B. and Harrison D.N.: Measurements of the amount of ozone in the Earth's atmosphere and its relation to other geophysical conditions. Proc. R. Soc. London Ser. A,110, 660-692, 1926

Fioletov, V. E., Kerr, J. B., McElroy, C. T., Wardle, D. I., Savastiouk, V., and Grajnar, T. S.: The Brewer reference triad, Geophys. Res. Lett., 32, doi:10.1029/2005GL024244, 2005.

Fioletov V.E. and Ogyu A.: Brewer Processing Software, http://exp-studies.tor.ec.gc.ca/pub/Brewer Processing Software/brewer processing software.pdf, last accessed on 2007.

Fountoulakis, I., , Bais, A. F., Fragkos, K., Meleti, C., Tourpali, K., Zempila, M. M.: Short- and long-term variability of spectral solar UV irradiance at Thessaloniki, Greece: effects of changes in aerosols, total ozone and clouds, Atmos. Chem. Phys., 16, 2493–2505, 2016

Ialongo, I., Casale G.R., and Siani A.M.: Comparison of total ozone and erythemal UV data from OMI with ground-based measurements at Rome station, *Atmos. Chem. Phys.*, 8, 3283–3289, 2008.

International Ozone Service (IOS) http://www.io3.ca/, last accessed on September 2017

Josefsson, W. A. P.: Focused sun observations using a Brewer ozone spectrophotometer, J. Geophys. Res.,97(D14), 15813–15817, doi:10.1029/92JD01030,1992.

Karppinen, T., Lakkala, K., Karhu, J.M., Heikkinen, P., Kivi, R., Kyrö, E.: Brewer spectrometer total ozone column measurements in Sodankylä. Geosci. Instrum. Method. Data Syst., 5, 229–239, 2016.

Karppinen T., Redondas, A., García, R.D., Lakkala, K., McElroy, C. T., Esko K.: Compensating for the Effects of Stray Light in Single-Monochromator Brewer Spectrophotometer Ozone Retrieval, Atmosphere-Ocean 53 (1), 66 – 73, 2015.

Kerr, J. B., McElroy, C. T., and Olafson, R. A.: Measurements of total ozone with the Brewer spectrophotometer, in Procs. of the Quadrennial Ozone Symposium, edited by J. London, 74–79, Natl. Cent. for Atmos. Res., Boulder, Colo.11., 1981.

Kerr, J.B., McElroy C.T., Wardle D.I. and Dorokhov V.: Measurements of arctic total ozone during the polar winter, Atmosphere-Ocean, 28:4, 383-392, 1990.

Kerr, J. B. and Davis, J. M.: New methodology applied to deriving total ozone and other atmospheric variables from global irradiance spectra, J. Geophys. Res., 112, D21301, doi:10.1029/2007JD008708, 2007.

Kerr, J. B.: The Brewer Spectrophotometer, in UV Radiation in Global Climate Change, edited by W. Gao, D. Schmoldt, and J. Slusser, Springer, Berlin, 160–191, 2010.

Muthama N.J., Scimia U., Siani A.M., Palmieri S.: Toward optimizing Brewer zenith sky total ozone measurements at the Italian stations of Rome and Ispra, J. Geophys. Res.,100, 3017-3022, 1995.

Ramanathan, V., Dickinson R.E.: The role of stratospheric ozone in the zonal and seasonal radiative energy balance of the earth-troposphere system, J. Atmos. Sci. 36:1084 – 1104, 1979.

Redondas A., Franco, J.R., Lopez-Solano, J., Carreno, V., Leòn-Luis, S.F., Hernàndez-Cruz, B.:The Regional Brewer Calibration Center - Europe: an overview of the X Brewer Intercomparison Campaign, WMO Commission for Instruments and Methods of Observation, TECO 2016.

Siani A.M, Casale, G.R., Galliani A.: Investigation on a low ozone episode at the end of November 2000 and its effect on ultraviolet radiation, Opt. Eng. 41(12), 3082-3089, 2002.

Siani, A. M., Modesti, S., Casale, G.R., Diemoz, H., and A. Colosimo: Biologically effective surface UV climatology at Rome and Aosta. AIP Conf. Proc. 1531, 903-906, 2013.

Stanek, M.: O3Brewer, http://www.o3soft.eu/o3brewer.html, last accessed on January, 2016.

Stübi, R., Schill, H., Klausen, J., Vuilleumier, L., and Ruffieux D.: Reproducibility of total ozone column monitoring by the Arosa Brewer spectrophotometer triad, Geophys. Res. Atmos., 122, 4735–4745, 2017.

Tzortziou, M., Herman, J. R., Cede, A., and Abuhassan, N.: High precision, absolute total column ozone measurements from the Pandora spectrometer system: Comparisons with data from a Brewer double monochromator and Aura OMI, J. Geophys.,Res., 117, D16303, doi:10.1029/2012JD017814, 2012.

Vanicek, K.: Differences between ground Dobson, Brewer and satellite TOMS-8, GOME-WFDOAS total ozone observations at Hradec Kralove, Czech. Atmos. Chem. Phys., 6, 5163–5171, 2006.

Vaz Peres, L., Bencherif, H., Mbatha, N., Passaglia Schuch, A., Toihir, A. M., Bègue, N.. Portafaix, T., Anabor, V., Pinheiro, D. K., Paes Leme, N. M., Bageston, J. V., Schuch, N. J.: Measurements of the total ozone column using a Brewer spectrophotometer and TOMS and OMI satellite instruments over the Southern Space Observatory in Brazil, Ann. Geophys., 35, 25–37, 2017.

WMO (World Meteorological Organization), Scientific Assessment of Ozone Depletion: 2014, Global Ozone Research and Monitoring Project, Report No. 55, Geneva, Switzerland, 2015.

WMO-GAW Seventh Intercomparison Campaign of the Regional Brewer Calibration Center Europe (RBCC-E) Lichtklimatisches Observatorium, Arosa, Switzerland, 16-27 July 2012, Report n.216, 2015.

WOUDC (World Ozone and Ultraviolet Data Centre), http://woudc.org, last accessed on June 2017.