

General comments:

The manuscript presents the results of a two-year comparison of total ozone measurements from Pandora, Brewer and Dobson spectrophotometers and satellite estimates (OMI, two algorithms). Results are given in terms of a least-squares analysis and some statistical indicators are provided (max, min, average, standard deviation, relative differences, Pearson's correlation index, R²) together with the analysis of variance (ANOVA). The manuscript addresses relevant questions and could potentially provide interesting results, since a complete set of instrument operates at the measuring station of Seoul.

→ We really appreciate the reviewer's all comments and suggestions about this manuscript. They are really helpful, thus helped improve the quality of our manuscript. The following is our responses to the reviewer's comments and the revised manuscript is attached in supplement.

Specific comments:

1. No mention is made in the manuscript of the processing/algorithms of the measurements. Do they differ for different instruments? For example, how much can the used spectroscopic sets of cross-sections (and their relative dependence on effective ozone C₂ AMTD Interactive comment Printer-friendly version Discussion paper temperature) impact on the comparison?

→ General manuscripts of the processing/algorithms of four measurements are mentioned in Section 1. And the specific mention for each instrument in Seoul is summarized in Section 2.

2. Table 2: do the authors have an idea why some instruments have a negative intercept compared to Pandora and others a positive intercept? How do the authors cope with straylight, which affects all data from the presented ground-based instruments?

→ The intercept represents bias of a measurement, associated with the slope of regression line. Pandora measurements tend to underestimate total column ozone(TCO) compared to other measurements. Therefore, it is possible to find the negative value of the intercept if the difference is somewhat large at the small TCO. However in Table 2, all absolute values of t-ratio are in 0.27~1.14, so it is hard to say that the intercept is significantly negative.

As for stray light, it affects all data especially in shorter UV band. Pandora corrects straylight empirically as a function of air mass factor(AMF), where the retrieved TCO decreases rapidly at very large air masses(sunrise and sunset) (Herman et al., AMT 2015). Sentences were added in Section 2.4.

3. Lines 179-180 explain that measurements from Pandora at high SZAs are removed from the analysis, but what about the other instruments? The comparison is performed in terms of monthly and daily averages. Then, a section (3.3) explains that large inconsistencies may originate due to the ozone variability when daily averages are performed on different datasets (because of different sampling frequencies). Could the authors additionally perform a comparison of nearly simultaneous measurements? It is stated that the Pandora temporal resolution is about 2 minutes, therefore those measurements could be interpolated to the nearest Brewer/Dobson/OMI estimate, thus avoiding the bias illustrated in Sect. 3.3;

→ Basically, Dobson measures TCO under condition of $\mu < 2.5$ as mentioned in Section 2.1

(noon/close to noon) and Brewer instrument is also set up to $\mu < 3.0$. As mentioned in Section 3.3, large inconsistencies may generate due to the ozone variability, but the main purpose of this study is focused on the comparison of daily TCO by Pandora, with other measurements as stated. Section 3.3 is the additional part in order to explain diurnal variation of TCO and Pandora data clearly shows this variation with its high temporal resolution as mentioned in line 321. Validation of real-time data from each instrument leading to minimize bias has been done in previous studies and is not the main focus of current study.

4. Line 196: if the only condition for the comparison is that the number of daily observations must be greater than 10 days, it is likely that some differences are due to the sampled subset. Since the results orient the following comparison of the instruments, I think that the authors should revise their criteria or state the uncertainty of their results due to the day-to-day variability and different days sampled for each month in the datasets (as done in Sect. 3.3 for the daily averages);

→ As mentioned above, the main objective of this study is to compare Pandora data with other instruments in terms of daily values. Thus, despite of some differences in sampling issue, daily Pandora TCO was calculated and the comparison result actually shows good agreement with Brewer, Dobson and OMI measurements, respectively. Also, monthly TCO values are relatively less affected by daily variation and they are calculated when all data sets are available. The number of daily data is considerably small especially in summer season due to lack of clear days. Considering this, the comparison condition (daily observations > 10 per month) is set and monthly TCO values from all instruments are in good agreement with each other. The result of small mean relative difference stands for small day-to-day variability.

5. The dependence on solar zenith angle is listed as an important factor impacting the comparison (line 365). Could the authors present some plots of the differences between instruments as a function of the solar zenith angle or the airmass? What is the expected magnitude of the SO₂ effect in Seoul (line 365)?

→ Unfortunately, other instruments except for Brewer have taken measurements one to maximum three times per day, which limits the direct comparison of SZA effect. Two graphs in Figure 1 are generated from the comparison of Brewer and Dobson measurements in Seoul, Korea. The left figure shows the comparison of Brewer data and Dobson data after SO₂ correction from 1999-2005 except for year 2001 in Seoul and the right figure shows the same comparison but from January 2007 to March 2007. From both figures, the agreements between the Dobson and Brewer dataset are reasonably good after SO₂ correction with slopes of 0.996 (left) and 1.004 (right) respectively and R² values of 0.986 (left) and 0.985 (right) respectively. Typical concentration of SO₂ in Seoul is 0.02 ppm (annual average : 0.005 ppm, ~1 DU when SO₂ is assumed to distribute constantly up to 1 km in altitude), which affects O₃ measurements but not significantly. Although there are some differences in data sampling, it can be known that slopes and R² values in these figures are closer than those in Figure 4(e).

Technical corrections:

1. Line 5 (and 209): it should be clarified why Pandora is taken as a reference for the comparison. Is it because of its high temporal resolution? In this case, this should be C3 AMTD Interactive comment Printer-friendly version Discussion paper explicitly stated;

→ Sorry for the confusion. Pandora is not a reference in this study. Rather, it meant to compare the recent Pandora measurement with other collocated instruments such as Dobson and Brewer in Seoul, in addition to OMI satellite data. So this study attempts to assess the reliability of Pandora data in terms of daily representative values through the result of inter-comparison with other instruments.

2. Line 10: reporting the slope and R^2 for the comparison with the Brewer is redundant, since these data are already provided few lines above;

→ The redundant phrase is deleted. “In particular, they show a close agreement with the Brewer TCO measurements...”

3. Line 12-13: are both instruments affected by these factors in a similar way?

→ Line 12-13 lists the possible explanation for the difference between the Pandora and Dobson data is affected by these factors. The way the two instruments are affected by those factors except for SO_2 are similar in qualitatively. Pandora retrieves ozone amounts from spectral fitting to cover the entire 310 to 330 nm range, while Dobson retrieves ozone amounts from the difference between the intensity of selected wavelength pairs in the range 300-340 nm. Sentences were added in Section 2.4.

4. Line 13: does "temperature" mean effective ozone temperature or instrumental temperature? Or both?

→ Yes, it means “Effective Ozone temperature”. Thanks. Corrected.

5. Line 25-27: please, add some bibliographic references;

→ Bibliographic references are added in manuscript (Liou, 2002; Schott, 2007).

6. Line 31: the recovery of the "ozone hole" is still an open question (e.g., Solomon et al., Science, 2016), and the cited articles (1997-2003) do not pretend to report the recovery of the ozone hole, contrary to what the authors state. Please, notice that slowdown of depletion does not necessarily mean recovery. Moreover, a quite confusing explanation of the "ozone hole" is offered to the reader (without even specifying where it occurs) and no mention to the decrease of ozone at midlatitudes is made. Please, rewrite this part;

→ This part is rewritten.

7. Line 94: why "An" OMI?

→ It is corrected to OMI

8. Line 216: how can the authors state that Fig. 3 shows a "generally gaussian distribution" based on only 6 bins? Can they support their sentence on the basis of a normality test?

→ Lin3 216 is rewritten.

9. Line 221: "Dobson unit" or "Dobson instrument"?

→ It means Dobson instrument

10. Line 224: are you comparing Pandora, Brewer and OMI from 2012-2014 to the Dobson in the period 1985-2000? Please rewrite this sentence, since it is very confusing. Furthermore, can you

assess an increase/decrease by comparing datasets from different instruments?

→ Sorry for the confusion. It meant to compare the current study period of 2012-2014 to previous data record. Sentences were rewritten and deleted to avoid confusion.

11. Lines 312: "fixed temperature". Do you mean effective ozone temperature? Does humidity refer to instrumental humidity or atmospheric humidity?

→ Effective temperature used in TCO retrieval. Humidity was meant to be the instrumental humidity due to humidity outside.

12. Line 365: "temperature". Instrumental or ozone effective?

→ It refers to effective temperature for TCO retrieval.

13. Table 1: please, write more clearly the year (2012-2013) and the day;

→ The title in Table 1 states the period of measurements clearly..

14. Fig. 4: define the used acronyms (OMT, OMD, DBS, etc.). Also, Fig. 4d has no dashed lines. Explain in the caption what the dashed lines indicate;

→ Definition of acronyms and corrected line are shown in fig. 4 and its title.

15. Fig. 5, caption: "blue lines". There are no blue lines in the figures;

→ It now changed to dashed lines in Fig. 5

16. Figs. 4-5: why n=115 for all plots? Only days with all instruments measuring were chosen? Is it necessary to report the same number in all plots?

→ In this comparison, only days with all instruments measuring were chosen to have fair and reliable intercomparison as mentioned in line 229-231.