Response to referee #1's comments

The author would like to thank Anonymous referee #1for the constructive and helpful suggestions on this manuscript.
We replied to 9 main comments and 9 minor comments

Main Comment

1-3. this paper is about the testing of the GEMS retrieval algorithm on OMI radiances, and the subsequent comparison with ozonesondes. This is interesting and worthwhile study in the run-up to the launch of GEMS. I see the paper as having two main purposes with respect to GEMS. First, it is an exercise of the retrieval algorithm on real OMI data. The fact that this is successful gives confidence that the retrieval is ready to receive the first GEMS data after launch. However, quantitative verification of the retrieval performance is harder, and the discussion of the GEMS retrieval algorithm performance on OMI radiances against sondes, compared with the OMI-algorithm retrievals against the same sondes should be expanded. The second purpose is to identify those ozonesonde measurements that might be good for GEMS validation, in as much as the work here suggests that they are useful or not for OMI validation. If there is a cross-verification here, it is really about OMI validation between OMI and the radiosondes. The fact that the GEMS algorithm is used to process the OMI radiances does not change this, especially with comparisons that should adequately account for how a priori profile and smoothing error assumptions differ between the GEMS and usual OMI algorithms. As such, the title of the manuscript does not clearly describe what is done in the paper, and I suggest that the authors modify the title to better reflect the above two goals.

- As indicated by this reviewer, the simulated GEMS retrievals are similar to OMI retrievals (PROFOZ), but with different and better implementations. Therefore the cross-verification performed in this paper is actually close to OMI validation than GEMS validation. The smoothing errors between OMI and GEMS have been addressed in Bak et al. (2013), indicating that GEMS will provide the comparable ozone profile information below ~ 22 km and the reduced information up to ~ 40 km. The tropopause-based ozone profile climatology is implemented as a priori ozone information for GEMS. However, the goal of this paper is not to detail the difference between OMI and GEMS ozone profile retrievals, but to evaluate the simulated GEMS tropospheric ozone retrievals using the limited UV information (300-330 nm) and to prepare the good reference dataset for GEMS validation. The quality comparison of simulated GEMS retrievals with OMI retrievals are additionally performed to demonstrate the confidence of the presented GEMS ozone profile algorithm. According to reviewer's suggestion, the title is changed to "Evaluation of GEMS tropospheric ozone retrieval performance using OMI data and validation ozonesonde dataset Over Northeast Asia" to clearly reflect what we did through this work.

4. **Section 2.1** describes the retrieval algorithm applied to the OMI radiances. A discussion should be included as to how the retrieval algorithm characteristics are expected to change for the GEMS radiances.

- GEMS is a scanning Uv/Visible spectrometer with a single UV-enhanced CCD for the spectral measurements of 300-500 nm (FWHM: 0.6 nm, spectral sampling: 0.2 nm) with at least comparable radiometric/wavelength accuracy (4% including light source uncertainty/0.01 nm) as OMI. However, GEMS data processing is expected to be different to OMI mainly in two ways: 1) OMI use a depolarizer to scramble the polarization of light. However, GEMS requires polarization factor < 2% and adopts polarization correction using RTM-based look-up table and pre-flight measured polarization characterization at the level 1b data processing. The GEMS polarization correction is less accurate and hence additional fitting process might be required in the level 2 data processing, especially for ozone profiles which the retrieval sensitivity to the polarization error is expected to be more significant compared to other trace-gases. 2) GEMS has a capability to perform diurnal observation and hence the diurnal meteorological input data are required to account for the temperature dependent Huggins band ozone absorption. Hence the numerical weather prediction (NWP) model analysis data will be transferred to the GEMS science data processing center (SDPC). This part is clarified in the Section 2.1 of the revised manuscript.

5. **Section 3.1**. The discussion of the differences between satellite/sonde agreements at the different sites is interesting. In addition to the differences between the sonde characteristics and reliabilities, one might expect greater standard deviations at sites that are polluted and/or show greater variability in ozone loadings due to meteorology. Some further discussion would be useful about the chemical-transport environment before eliminating sites from potential GEMS validation based on instrumentation/ experimental method arguments alone. It would also help this reader if the current dense text were broken up into descriptions of the various reasons for good/bad agreement.

- We have added the figure 4 in which the seasonal mean and standard deviations of ozonesonde measurements are presented to see the stability and characteristics of ozonesonde measurements at each site. Instabilities of measurements are apparently observed from New Delhi ozonesondes. High surface ozone concentration at Trivandrum in summer is believed to be caused by measurement errors because low level of the pollutants has been reported at this site under the geolocation and meteorological effects (Lal et al. 2000). Besides Trivandrum, Naha could be regarded as background sites according to low surface ozone and its precursor concentrations compared to neighboring stations (Fig. 2 and Fig. 4), and previous studies (Oltmans et al., 2004; Liu et al., 2002). In the lower troposphere high ozone concentrations are captured at Pohang, Tsukuba, and Sapporo in the summer due to enhanced photochemical production of ozone in daytime, whereas tropical sites, Naha, Hanoi, and Hong Kong show the ozone enhancement in spring mainly due to the biomass burning in Southeast Asia, with low ozone concentrations in summer due to the Asian monsoon and in winter due to the tropical air intrusion (Liu et al., 2002; Ogino et al., 2013). Singapore and Kuala lump are supposed to be severely polluted area, but ozone pollution is not clearly captured over the seasons. It could be explained by the observation time carried on in the morning. In addition, instabilities of Singapore measurements are noticeable such as abnormally large variability and very low ozone concentration in the stratosphere. The effect of stratospheric intrusions on the ozone profile shape is dominant at the mid-latitudes (Pohang, Tsukuba, and Sapporo) during the spring and winter when the ozone pause goes down to 300 hPa, with the larger ozone variabilities in the lower stratosphere and upper troposphere, whereas ozonepause is placed around 100 hPa with much less variabilities of ozone over the pressure at other seasons. This discussion will be included in Section 3.

- Lal, S., Naja, M., and Subbaraya, B: Seasonal variations in surface ozone and its precursors over an urban site in India, Atmospheric Environment, Volume 34, Issue 17, 2000, Pages 2713-2724, 2000.
- Liu, H., D. J. Jacob, L. Y. Chan, S. J. Oltmans, I. Bey, R. M. Yantosca, J. M. Harris, B. N. Duncan, and R. V. Martin, Sources of tropospheric ozone along the Asian Pacific Rim: An analysis of ozonesonde observations, J. Geophys. Res., 107(D21), 4573, doi:10.1029/2001JD002005, 2002.
- Ogino, S.-Y., M. Fujiwara, M. Shiotani, F. Hasebe, J. Matsumoto, T. H. T. Hoang, and T. T. T. Nguyen (2013), Ozone variations over the northern subtropical region revealed by ozonesonde observations in Hanoi, J. Geophys. Res. Atmos., 118, 3245–3257, doi:10.1002/jgrd.50348.

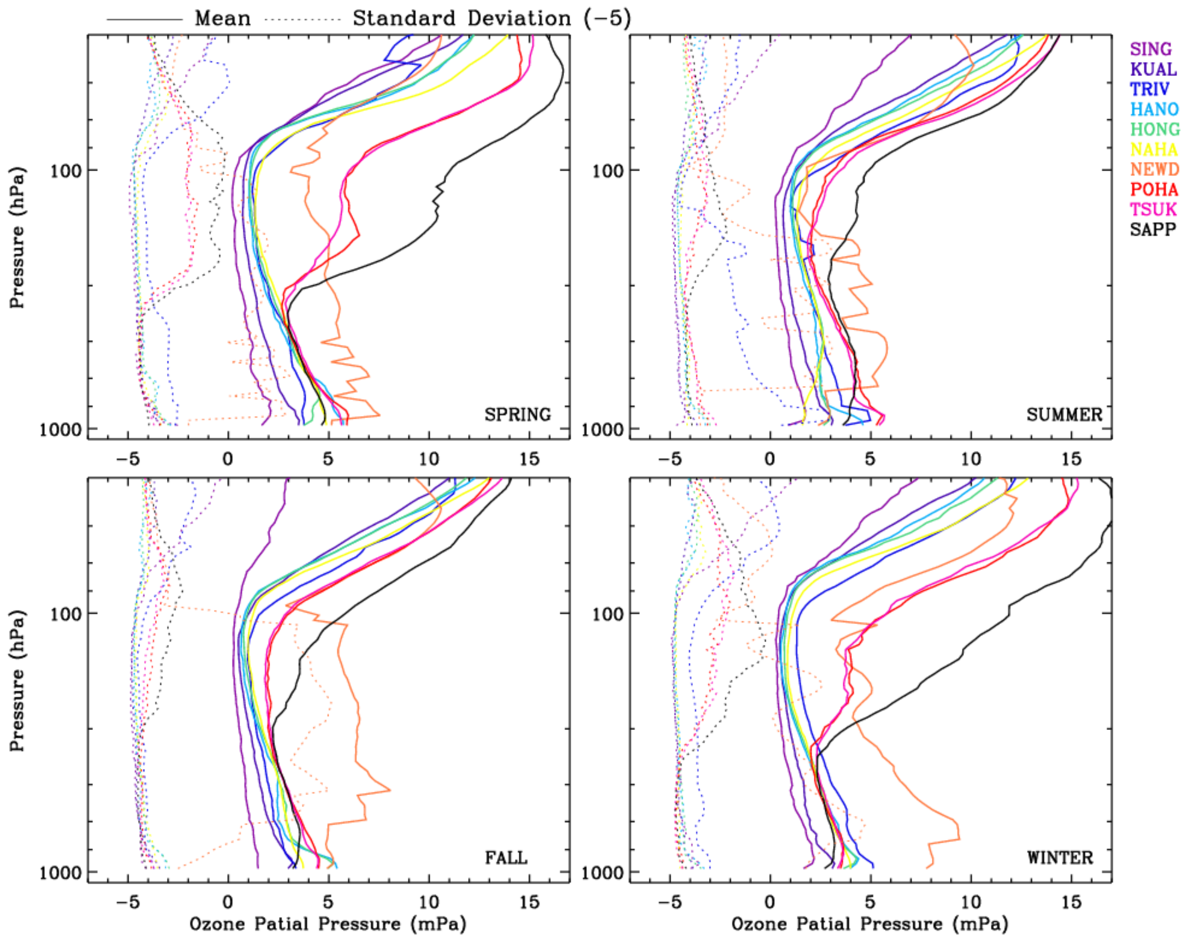


Fig. 4. Seasonal mean (solid) and standard deviation (dashed) of ozonesonde soundings from 2005 to 2015 at 10 sites. 5 mPa is subtracted to standard deviations to fit in the given x-axis.

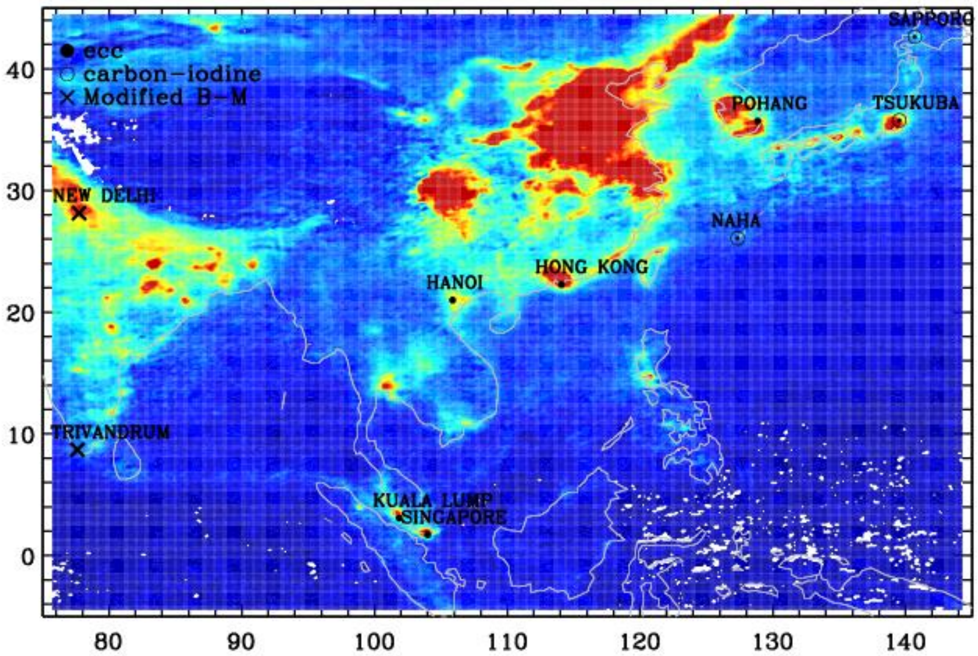
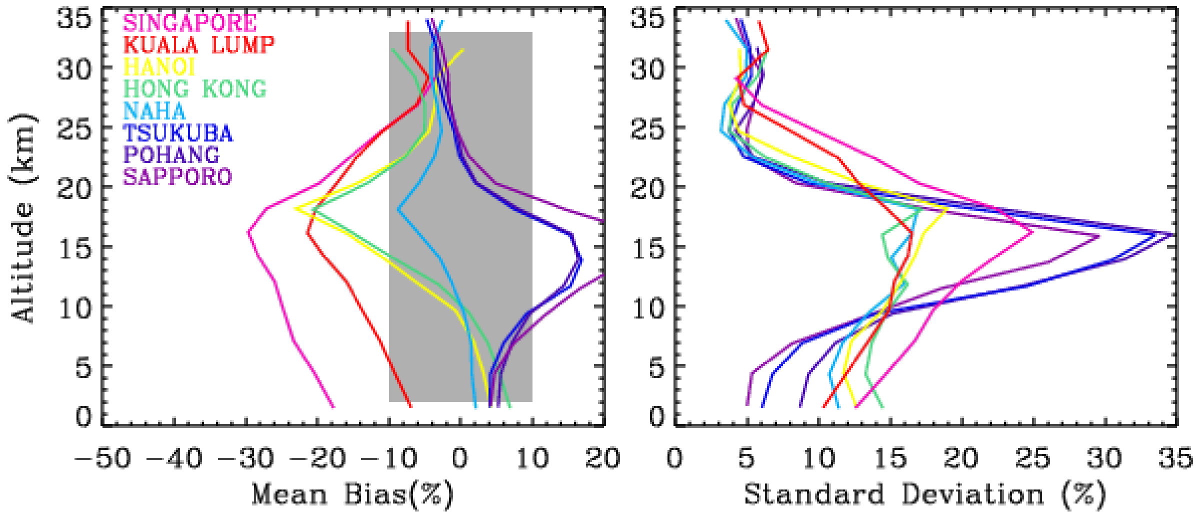


Figure 2. Geographic locations of the ozonesonde stations available since 2005 over the GEMS observation domain. The background map illustrates the OMI NO2 monthly mean in June 2015.

6. **Section 4.2 and Fig. 7**. The impact on correlation of smoothing or not smoothing the sonde profiles might be dependent on how close the GEMS retrieval a priori profile is to the sonde "truth". How do these compare and how do they vary between locations of good and poor comparison? How far does a priori profile go toward explaining the bias?

- To explain this impact, the comparison between GEMS a priori and ozonesondes is presented below, similarly to Fig.6 (profile comparison). This a priori information is taken from the tropopause-based ozone profile climatology (TB) which adjusts a monthly and zonal mean ozone profile with a daily tropopause height. Bak et al. (2013) demonstrated that TB based a priori better represents the ozone variabilities in the extra-tropical upper troposphere and lower stratosphere, especially during the winter and spring when the atmospheric status is strongly controlled by the dynamics. A priori information is very important to the quality of UV ozone profile retrievals, but the retrieved ozone profiles show much better agreement with ozonesondes than a priori ozone profiles, implying that the independent piece of information are available from these UV measurements. We can see that the biases seen in a priori are significantly reduced for most stations except for the Singapore station. Positive biases around 15 km in the a priori at the three mid-latitude sites still remain in the retrievals but with much smaller magnitude. Negative biases around ~17 km at other lower-latitude sites (except for Singapore) are almost eliminated in the retrievals.



S1. Same as Fig 6, but for comparison of ozonesondes and a TB-based priori ozone profiles.

7-8. Section 4.2 should be split into another sub section at Line 354 that starts the discussion of the evaluation of the GEMS algorithm against the OMI algorithm. This should be presented as one of the main results sections of this paper: a quantitative evaluation of the GEMS algorithm against other widely used algorithms based on the same OMI radiances. Given the previous discussion in the paper of the various limitations of some of the sondes, it might additionally be useful to directly compare the results of the different retrieval algorithms and explain differences in results on the basis of different features of the algorithms. This would help make the case that the GEMS algorithm is performing as expected.

- In this paper, the ozonesonde measurements available in the GEMS domain are characterized and validated to better evaluate the performance of the GEMS ozone profile algorithm. The accuracy and precision of simulated GEMS ozone profiles are established against the selected true reference. Additionally, the consistent evaluation is performed for the existing OMI ozone products to check the confidence of the GEMS retrieval algorithm, demonstrating that the comparable or better performance of GEMS ozone profile retrievals in the comparison with ozonesonde measurements. The different validation results between retrieval algorithms were discussed such as "GEMS algorithm is developed based on the heritages of the SAO ozone profile algorithm with several modifications. There are two main modifications: a priori ozone climatology was replaced with a tropopause-based ozone profile climatology to better represent the ozone variability in the tropopause. Irradiance spectra used to normalize radiance spectra and characterize instrument line shapes are prepared by taking 31-day moving average instead of climatological average to take into account for time-dependent instrument degradations. These modifications reduce somewhat spreads in deviations of satellite retrievals from sondes, especially in TCO comparison. KNMI retrievals systematically overestimate the tropospheric ozone by ~ 6 DU (Fig. 9.c), which corresponds to the positive biases of 2-4 % in the integrated total columns of KNMI profiles relative to Brewer observations (Bak et al., 2015). As mentioned in Bak et al. (2015), the systematic biases in ozone retrievals are less visible in SAO-based retrievals (GEMS simulation, OMPROFOZ) as systematic components of measured spectra are taken into account for using an empirical correction called "soft calibration". The GEMS algorithm is very similar to the SAO algorithm except for the use of TB climatology and the impact of TB on the retrievals was discuss in detail in Bak et al. (2013). Also the comparison of SAO and KNMI algorithms were discussed in detail in Bak et al. (2015). So we think that it is more efficient to place the discussion related to Figures 7-9 in the same section and the direct comparison of GEMS and other OMI product is beyond the scope of this paper.

9. The paper requires careful, and extensive, editing for English usage, and cut-paste typos, e.g. line 75, that should have been corrected before manuscript submission.

- This manuscript is going to be carefully revised though native English co-author before the submission of the revised manuscript.

Minor Comments

1. Several times "GEMS" measurements are described. The word "simulated" should be added each time to avoid confusion

- In this revised manuscript, "GEMS measurements" was edited to " simulated GEMS measurements"

2. Line 83: consistent perhaps, but not homogeneous as the authors point out in the text above.

- The indicated sentence was revised from "a homogenous, consistent ozonesonde" to "a consistent ozonesonde".

4. Line 105: Instrument errors certainly, but also instrument design sensitivity.

- The indicated word, "instrumental errors" was revised to "Instrument errors certainly, but also instrument design sensitivity"

5. Line 106: Common geophysical conditions can reduce sensitivity, not just extreme

- The associated sentence was edited from "The impact of a priori information on retrievals become important ~~ under extreme geophysical conditions to ~~ under certain geophysical conditions.

6. Line 123: Information may be limited but is a goal of the GEMS mission. This should be clarified.

- The GEMS mission was originally planned to develop the spectrometer for measuring the tropospheric pollutants, the spectral coverage of 300-500 nm satisfies to observe the tropospheric ozone as well as the lower/middle stratospheric ozone.

7. Line 157: Any more recent references to new measurement technique and instrumentation?

- More references (Thompson et al., 2017; Witte et al., 2017; 2018) are added in this sentence.

8. Line 200: How do these coincidence criteria for OMI and the sondes affect the results? What is the expected variation within the time and space windows? What is the representativeness uncertainty? How do these results here with the OMI comparison inform on the expected GEMS comparisons with hourly measurements at ∼7km resolution?

- As mentioned in Section 2.3, the coincidence criteria between satellite and ozonesonde are: ±1.0° in both longitude and latitude and ±12 hours in time and then the closest pixel is selected. The actual spatiotemporal difference is much smaller than this criteria (57.5 km to 66.6 km, ~3 hours). The close collocation can significantly minimize the effects of spatiotemporal variability on the comparison and therefore GEMS validation accuracy could be enhanced compared to OMI, which is newly included in the summary of this paper.

9. Line 231: Is “troposphere” written where is should be stratosphere?

- The indicated sentence was edited to “much coarser vertical resolution of 10-14 km in the troposphere and 7-11 km in the stratosphere”.