Response to Reviewer #2

We thank Reviewer #2 for her/his very detailed and helpful comments. Please find below the reviewer's comments (black) and our responses (blue) which also indicate the changes made in the manuscript.

1) Scientific Significance: The manuscript provides a good description of the methods and results of a study to evaluate and improve the stability of the GOME Level 1 record.

2) Scientific Quality: The results are well-structured and well-referenced and use good statistical analysis methods. There are good references to detailed reports for interested readers.

3) Presentation Quality: The paper is well-written and the figures and tables are good in both content and structure.

Editorial Comments and Suggestions:

Make Figure 2 larger. It should at least be full page width.
Done as suggested.

In Section 2.3.2, rewrite and clarify the last line. Is this 10% the accuracy of the stray light estimates relative to the true stray light? That is, is if the stray light error is 20 units, then the correction will be between 18 and 22 units and the final result will have an error of ±2 units?
We added: "… not more accurate than ∼10%, i.e. processing errors of 10% of true straylight."

Page 19, Line 14, "raise" should be "rise".
Corrected.

The value of 1100 for the SNR for Channel 1 in the Table 1 seems high even for the 305 nm wavelength. What is the corresponding integration time and the size of the FOV? I believe there was a change in the Channel 1A/1B wavelength boundary during the mission. Is this before or after that change? Also, provide an SNR value for a shorter wavelength in the table, say 290 nm.
We added to Table 1:

- the information on the integration time for channel 1 (6s);
- values for 290nm;
- values for middle (2001) and end (2010) of the mission.

The change in the Channel 1A/1B wavelength boundary was implemented in June 1998.

While the views of the Moon are complicated by scan mirror differences with angle and the phases of the Moon, more accurate lunar models are now available. For example, Eumetsat's GSICS Implementation of the ROLO model (GIRO) and the GSICS Lunar Observation Dataset (GLOD) introduced at
https://www.eumetsat.int/website/home/News/DAT_3460357.html?lang=EN&pState=1
could be explored to allow the lunar measurements to be used to monitor instrument changes.
Thank you very much for the reference to these data sets! For the GOME instrument one problem is that the Moon measurements do not fill the entire slit so that the calibration key data cannot be applied just like that. Moreover, the Moon is always observed at a scan angle in the western Limb, whereas scan angle dependency observations corresponding to East or Nadir pixels would be needed to significantly improve on the degradation correction using the Sun. And since the Moon merely reflects Sun light, it cannot be used as calibration source independent of solar activity.

Questions on Science:

Section 3.1

Page 11 line 8 et seq. While arguments can be made for estimating degradation by avoiding lines with high solar activity, this will not work well for Channel 1. See

V. Marchenko, Sergey & Deland, Matthew & Lean, Judith. (2016). Solar Spectral Irradiance Variability in Cycle 24: Observations and Models. Journal of Space Weather and Space Climate. 6. 10.1051/swsc/2016036.

for estimates of solar variability for 270 nm to 500 nm over a solar cycle. After estimating the changes in the instrument throughput, the final time-dependent solar provided in Level 1 should be constructed with realistic solar activity variations.

We agree with the reviewer that in principle realistic solar variations should be taken into account. Thank you for pointing us to this reference. Nonetheless, providing an optimum solar irradiance product is not the main focus of this study.

Also, how large are the Etalon Effects in Figure 2? What errors would they be expected to produce in the radiance/irradiance ratios? Why wasn't a correction applied? It appears that the authors have access to estimates of these corrections from other analysis:

https://wdc.dlr.de/sensors/gome/degradation_files/degradation.php

The etalon amplitudes may be estimated from Figure 4 (former Fig. 2) as the amplitude of the semi-regular wiggles (~10 wiggles in channel 1 to ~5 in channel 4). Since GOME does not have a flat field mode (e.g. using a white light source) etalon cannot be directly derived. Correction using the Sun would be possible but only relative to a certain reference date, not in an absolute sense. The main focus of the Level 1 product has been to function as input for Level 2 retrievals. For Level 2, etalon is irrelevant as long as the structure is identical for solar and for earth-shine measurements. There are in fact some indications that this may not completely be the case, depending on the solar azimuth, but attempts to characterize solar-azimuth dependent etalon-like structures of the diffuser BSDF have not been deemed reliable enough to be applied in the GOME calibration. The errors are shown in Fig.8 of reference (Slijkhuis 2004) on

https://wdc.dlr.de/sensors/gome/degradation_files/degradation.php

And there are the earlier results in

Weber, Mark & Burrows, John & Cebula, R. (1998). GOME Solar UV/VIS Irradiance Measurements between 1995 and 1997 – First Results on Proxy Solar Activity Studies. Solar Physics. 177. 63-77. 10.1023/A:1005030909779.

Section 3.3

From Section 2.3.2, the angle for the mirror for Solar measurements is 41◦ and those for the Earth measurements range from 49◦ ±15◦. What are the results for Figure 6 for the ground pixels at this matching angle? If they are not equal to 1.0 what are the likely instrument changes that produce time-dependent differences in the radiance/irradiance ratios?

In Fig. 8 (former Fig. 6) the results for the west pixels (incidence angle 44°-34°), cyan curves, match the angle for the solar measurements. In general, west pixels show the minimal degradation in reflectance compared to the other ground pixels types. Explicit characterization and indication of the instrument changes that produce the reflectance degradation is difficult. As mentioned on page 10: “The main degradation as a consequence of extensive exposure to the space environment can be attributed to deposits on the scan mirror (which is coated with a MgF 2 layer) thereby changing its reflective properties”. This change in mirror coating also changes the scan-angle dependent polarization properties of the instrument (Snel, 2001) .

Is it correct that the analysis in the section is just an evaluation of errors in the Level 1 product and that no

corrections based on the PICS results have been applied?
Yes, this is correct. We added this to the summary.

If so, degradation is only shown for 325 nm and 335 nm measurements and the changes are over 20% and differ by over 5%. This does not suggest that the shorter channel are well characterized for absolute radiance / irradiance calibration. All algorithms are sensitive to the reflectance if they need parameters associated with cloud cover. What are the effects of a +10% error in the UV cloud fraction on GODFIT ozone retrievals?
A 10% error in cloud fraction is expected to have an impact of <1% on total column ozone retrieval (Christophe Lerot, personal communication, July 2018).

Do the authors recommend that Channel 1 data in this product be used for ozone profile retrievals? What about the use of data from 300-310 for tropospheric retrievals requiring radiance / irradiance calibration?
We agree with the reviewer that the retrieval of ozone profiles and tropospheric columns from GOME requires a very careful handling of the measured spectra and additional corrections to account for degradation. However, several studies successfully demonstrated the feasibility (Liu et al., 2005, Cai et al., 2012, Miles et al., 2015). Moreover, Keppens et al. (2018) have shown that decadal drift values for GOME level-2 ozone profiles are overall insignificant.

Section 4.2

How large are the variations in the wavelength scales along an orbit from measurement-based estimates? Do they match with the variations predicted from the effects of the measured pre-disperser prism temperature changes combined with the laboratory sensitivity characterization or are there other complicating factors?
On average 9 different wavelength scales are used along one orbit. The variation in the wavelengths depends on the spectral region. In general, the variation is <0.002 nm, except for the beginning of channel 3 and the end of channel 4, where the variation is 0.004-0.005 nm along one orbit. This analysis is based on ~2000 randomly selected orbits.
The use of spectral calibration as function of pre-disperser temperature was a recommendation based on on-ground measurements of instrument performance. However, we cannot retrieve the original data.

DOAS- based retrievals often generate internal estimates of the wavelength scale shifts as part of the fitting process. Have any of these been compare to this bottom-up analysis based on the prism temperatures?
No, unfortunately, these comparisons have not been performed.

References:

Cai, Z., Y. Liu, X. Liu, K. Chance, C. R. Nowlan, R. Lang, R. Munro, and R. Suleiman (2012), Characterization and correction of Global Ozone Monitoring Experiment 2 ultraviolet measurements and application to ozone profile retrievals, J. Geophys. Res., 117, D07305, doi: 10.1029/2011JD017096.

Keppens, A., Lambert, J.-C., Granville, J., Hubert, D., Verhoelst, T., Compernolle, S., Latter, B., Kerridge, B., Siddans, R., Boynard, A., Hadji-Lazaro, J., Clerbaux, C., Wespes, C., Hurtmans, D. R., Coheur, P.-F., van Peet, J. C. A., van der A, R. J., Garane, K., Koukouli, M. E., Balis, D. S., Delcloo, A., Kivi, R., Stübi, R., Godin-Beekmann, S., Van Roozendael, M., and Zehner, C.: Quality assessment of the Ozone_cci Climate Research Data Package (release 2017) – Part 2: Ground-based validation of nadir ozone profile data products, Atmos. Meas. Tech., 11, 3769-3800, https://doi.org/10.5194/amt-11-3769-2018, 2018.

Liu, X., K. Chance, C. E. Sioris, R. J. D. Spurr, T. P. Kurosu, R. V. Martin, and M. J. Newchurch (2005), Ozone profile and tropospheric ozone retrievals from the Global Ozone Monitoring Experiment: Algorithm description and validation, J. Geophys. Res., 110, D20307, doi: 10.1029/2005JD006240.

Miles, G. M., Siddans, R., Kerridge, B. J., Latter, B. G., and Richards, N. A. D.: Tropospheric ozone and ozone profiles retrieved from GOME-2 and their validation, Atmos. Meas. Tech., 8, 385-398, https://doi.org/10.5194/amt-8-385-2015, 2015.