We appreciate the efforts of the Referee #2, L. Flynn, and want to express our gratitude.

Referee #1 L. Flynn

The manuscript provides an overview of the operational products available from the METOP GOME-2 series of instruments and their performance. It gives references (articles and web sites) for more detailed algorithm descriptions and validation analyses as well as product access. The results are well-referenced and give a good introduction to the different products available from different program components and their validation. The paper is well-written and logically ordered with good thought to the selected figures.

Question:
There is not a section devoted to the Level 1 Product quality or validation. Some discussion should be included giving the status of the measurements – long-term stability, scan angle dependent errors, and polarization corrections. These should be related within the rest of the paper to individual product's sensitivities to such errors. For example, the ozone profile and surface reflectivity products are sensitive to drifts in the radiance/irradiance measurement ratios.

Answer:
We invited a L1 product expert to participate and write a L1b section. This is now added as Section 3. Furthermore, references to Munro et. al. 2015 were added too because that paper provide all details of L0 and L1b.

Due to this change as wella s other structrulal changes proposed by the Referee #2, the order of Sections and Figures has been changed.

Question:
Some additional details should be included for the total ozone column algorithm. Where does the algorithm get its cloud height and cloud faction information? (The ozone profile retrievals use Cloud pressures retrieved from O2 A-Band measurements. The Level 1 products contain some cloud information. If this is used in the product retrievals, then its theory and validation should be in this paper.) How is this information validated, and is its accuracy stable over time? What are the principle quality/error flag considerations (Aerosols? SO2? SAA?) in terms of frequency and severity for the total column ozone and ozone profile products? Does the product output file include the assumed profile and a one-dimensional averaging kernel to allow comparison and interpretation of the results with respect to other column ozone measurements or forecasts? If not, is this under consideration for future improvements?

Answer:
Please note that this Section is now 5.1

The cloud parameters used for total ozone algorithm are retrieved from GOME-2

measurements using OCRA and ROCINN algorithms. An explanation was added at the beginning of the Section describing Total Ozone product. This is valid also to other column products.

For the total ozone product, currently the flags for aerosol, SO2 and SAA are not included. Large SO2 values (from volcanic eruption) have a significant effect on O3 DOAS retrieval. Therefore, we plan to add the flags for volcanic SO2 eruption in the GOME-2 O3 product (refer to the answers for SO2 flags below). Also the SAA flags will be included.

A priori O3 profile used in the O3 retrieval has been included in the output file. We agree that the O3 averaging kernel information is very useful for the comparison with other O3 measurements and model simulations. In the future, we will include the DOAS averaging kernel for the retrieved total O3 columns in the GOME-2 O3 product.

Considering the quality/error flag for O3 profile:

Quality flag considerations for aerosols: The L2 product copies a number of flags over from the L1b product, such as the DegradedProcessor, DegradedInstrument and SouthAtlanticAnomaly flags. The L1b SAA flag is not used to select a different retrieval method. The L2 processor checks the contents of the L1b product and flags out of bound spectral measurements, which is the main driver for handling quality control inside the ozone profile retrieval algorithm. The processor uses the wavelengths it believes are reliable to use. In the output product there is a variable SpectralMask that indicates which spectral pixels have been used for the retrieval.

Question:
Given the variety of approaches used to separate stratospheric and tropospheric NO2 columns, another sentence or two should be added expanding on the sentence beginning on Page 7004, Line 24.

Answer:
Please note that this Section is now 5.3

The STS algorithm in the GDP 4.7 is based on a spatial filtering approach using total NO2 column measurements over “clean” regions, which allows for longitudinal variations. This is shown to be an improvement on the Pacific reference sector method. We have extended the description in the manuscript.

Question:
As for the total column, details on where cloud information is obtained should be

provided.

Answer:
Please note that this Section is now 5.3

The cloud information for the NO2 column retrieval (and for the other trace gas column products described in old Sect. 5.1) are obtained from the OCRA/ROCINN cloud algorithms for GOME-2 (see comment on the ozone columns above as well as the explanation in the updated description of the total ozone column product with cloud retrieval information).
Thus, we have added this information in the manuscript.

Question:
I do not understand the statement starting on Page 7005, Line 29. While on any given day the two instruments will make measurements at different times for a specific location, over the 29-day orbital cycle, I would expect that these variations would be similar. That is, that a given location would be seen at a variety of scan positions to the east or of the nadir tracks west (earlier or later than the local overpass time at nadir for that latitude which is the same for both Metop-A and -B) for both instruments over the course of a month. It was the case that cross-track biases in the Absorbing Aerosol Index for OMI were aliased into an orbital-track aligned set of features in monthly maps associated with the EOS-Aura's 16-day orbital track repeat cycle but these were very small.

Answer:
Please note that this Aerosol Section is now 5.3

We agree with the remark of the reviewer about elimination of the different times for a specific location of the 2 Metop when averaging. This sentence has been removed.

Question:
For the H2CO discussion on Page 7008, Line 15, clarify whether this is primarily due to increase noise caused by decreased throughput or some other factors affecting the measurements. This could be handled be referring to an expanded Level 1 section where the GOME-2A signal level changes would be discussed.

Answer:
Please note that this Section is now 5.4

In the paper, we have changed the sentence of line 15 " The effect of Metop-A degradation (increase of the DOAS fit RMS with time) is clearly visible." By "The effect of Metop-A degradation (increased noise caused by decreased throughput,

see Munro2015) is clearly visible in figure 9, with the DOAS fit RMS increase with time for both operational and scientific product."

Question:
Besides the use of BrO (Page 7008, Line 20) estimates used for the H2CO retrieval, are there any other combined retrievals in use or planned, e.g., for the new OClO and Tropospheric BrO Page 7019 Line 18? There could be complications from O3, SO2, NO2, and H2CO absorption features depending on the wavelength intervals selected.

Answer:
Please note that this Section is now 5.4

2-step fit method will also be used for CHOCHO retrieval to limit the impact of the interference from liquid water absorption over remote oceanic areas (see Page 7021 Line 5-10). The interference of other absorbers has been considered for the selection of BrO and OClO fitting window. Currently there is no plan to use combined retrieval for BrO and OClO.

Question:
Are the SO2 flagged pixels described on Page 7009, Line 20, used to set quality flags for any other products?

Answer:
Please note that this Section is now 5.5

The elevated SO2 flag has been specifically designed for the assimilation of the GOME-2 SO2 columns in forecast models (e.g. in the MACC-system). However, the SO2 flag would also be of interest as a quality indicator for the total ozone product. This will be further investigated for implementation in a future version of the GDP (see comment above).

We have added a note on this.

Question:
Can you provide an overlay contour (say at 2 DU) for the SO2 product for Figure 19? That is, show how the AAI and SO2 products see the dispersions of the volcanic ash and SO2 clouds, respectively?

Answer:
Please note that this Figure is now figure 18.

The referee asked for an SO2 contour overlay but since the horizontal variation of this trace gas is so irregular (noisy), the bundle of lines would be too much and not clear.

We have added the SO2 data set with a different color scale over the AAI. The SO2 values are in Dobson Units (DU).

Question:
Since the GOME-2 instruments measure both radiance and irradiance, you should provide more information on the source of the "Degradation" mentioned on Page 7015, Line 26. How large would the calibration drifts in radiance/irradiance ratios for GOME- 2A have to be to produce the trends seen in Figure 18?

Answer:
The multiplicative bias factors for 267 and 329nm (to be applied to the measured spectrum) are given in a new figure. Indicated is the ratio between measured reflectance and the simulated reflectance. Both ends of te used spectral range change with time, so the whole spectral range is affected by the degradation, with the short UV wavelengths degrading stronger than the longer wavelengths. At the start of 2013, the bias correction for 267nm was around 1.6 and the bias correction for 329nm was around 1.1.

The text of the manuscript has been updated.

Question:
Given the large differences between the two ozone profile products apparent in Figure 17, are there any on-going efforts to adjust the two GOME-2 Level 1 products to agree with each other?

Answer:
Please note that this is now figure 4.

There is an effort to establish a method that can handle the degradation correction as well as an initial bias of the GOME-2 instruments. The expectation is that this brings the two instruments close enough together that the values of both instruments can agree with each other.

Question:
Is similar degradation expected for the GOME-2B (Page 2016, Line3 "show this effect yet.") and if so what is its source?

Answer:
A similar degradation is expected for GOME-2b, but this still needs to be established properly.

Question:
Will reprocessing (Page 7019, Line 13) help to remove these drifts and differences?

Answer:
A reprocessing using a degradation correction for the ozone profiles will likely de-trend the slowly increasing differences. This has been shown for the tropospheric ozone in one of the last O3MSAF reviews.

Question:
Are there any adjustments made to the Level 1 measurements or Level 2 products to reduce cross-track biases in the AAI products? That is, are "clean" atmospheric regions examined to check for scan dependent biases in these products? I suspect not as Figure 19 shows a cross tack bias of approximately 1 unit at the Equator with middle scans lower than edges. Are these expected from some path-length dependent properties of the AAI retrieval? How do these cross-track biases relate to the target 0.5 index accuracy? (Note: the scale for Figure 19 does not identify a value for pink. The reader should not have to assume that the color is used for values > 4.) Are any time-dependent adjustments made to the AAI products? Are any adjustments made to bring the GOME-2A and GOME-2B products into agreement?

Answer:
This is now figure 18.

Before calculating the AAI from them, the Earth reflectances are first corrected for the effects of instrument degradation. For this we make use of the method introduced in (Tilstra et al., JGR 117, 2012) for the SCIAMACHY instrument and later applied to the GOME-2 instrument (Tilstra et al., EUMETSAT 2012 Conf. Proc., 2012).

This correction is in fact a pre-processing adjustment. The reflectances from which the AAI are derived are corrected; the AAI is not corrected afterwards.

The AAI shown in Figure 19 of the AMTD paper was not corrected for instrument

degradation, thank you for spotting this. We have created a new version of Figure 19 (with instrument degradation switched on). Also, we changed the colour bar. It now explains the meaning of the pink colour.

Both the GOME-2A and GOME-2B AAI are corrected separately for the impact of instrument degradation, as if the other instrument would not exist. No effort is made to bring the two instruments in agreement.

We have explained this more clearly in section 5.3 of the revised manuscript.

Question:
How are the changes in the GOME-2A instrument which produce trends in the radiance / irradiance ratios taken into account in the LER computations over the entire mission (Page 7017, Line 14 et seq.)? Do the dates when minima are found show any temporal patterns?

Answer:
The Earth reflectances are first corrected for the effects of instrument degradation using the method introduced in (Tilstra et al., JGR 117, 2012) before the surface LER is derived from them. This correction is dependent on time, wavelength, and scan mirror position. It is also different for the main science channels and the PMD bands.

The surface LER follows a seasonal pattern, which is why the database is provided for each month (January to December). When focussing on one particular month of the year (say, January), and analysing from which years the minima come from, we find that, at least globally, there is no real preference for a certain year in the time series.

There is, in other words, for a specific month not a specific year in the time series from which more monthly minima originate from than from the other years. (Exceptions are snow/ice regions where in one year there may be snow lying on the surface, while in another year the surface is snow-free.)

We have added a few lines to section 5.4 of the manuscript about the correction for instrument degradation.

Question:
Unless the photolysis product (Page 7018, Line 5) is not made for GOME-2A, it should be simple to fill in the gaps in Figure 21. If so, this option (as used for Figure 2) should be noted.

Answer:

Pelase note that this is currently figure 20.

Currently, the product coverage is determined by the total ozone input from a single GOME-2 instrument as shown in Fig. 21. We do not apply any image processing to fill in the gaps. But indeed, the product will be changed in the near future so that total ozone will be collected from all available GOME-2 measurements for a given day, and quality flags will indicate the source of total ozone for interested users. Then, for example, the gaps at low latitudes in Fig. 21 would be filled with GOME-2A operated in reduced swath mode (960km) in tandem to the operational GOME-2B in full swath mode (1920 km).

Question:
What is the purpose of the reprocessing (Page 7019, Line 13)? Are there newly reprocessed Level 1 data sets? How much do they change from the operational ones, what were the sources

Answer:
The main purpose is to have a data set that is processed with the same algorithm version for the whole GOME-2 period. This is useful for several applications like climatological studies, trend research as well as other reprocessing campaigns. Furthermore, implemented degradation corrections for sensitive products, like AAI, will be applied for the whole data period which improve the homogeneity of the data because early versions of algorithms doesn't include correction for degradation. This is the first reprocessing campaign for all archived GOME-2 products. The last L1b reprocessing has been done in 2012.

Editorial comments:

Question:
The paper introduces the shorthand GOME-2A and GOME-2B for Metop-A GOME-2 and Metop-B GOME-2 but sometimes has "Metop-A degradation" or "GOME-2 A and B" or "for Metop-B" or "GOME-2 on the Metop-A" instead (E.g., see Page 7008.). Make these consistent.

Answer:
The consistency has been checked and the needed corrections have been made.

Question:
Page 6995, Line 15, "quarantined" -> guaranteed"

Answer:
Corrected

Question:
Page 6999, Line 3, "build" -> "built"

Answer:
Corrected

Question:
Page 6999, Line 16 “The Metop satellites are flying on a sun-synchronous orbit with an equator crossing time of about 09:30 LT (descending node) and a repeat cycle of 29 days.” —> “The Metop satellites are flying in sun-synchronous orbits with equator crossing times of approximately 09:30 LT (descending node) and a repeat cycle of 29 days.”

Answer:
Updated

Question:
Page 7015, Line 27. The “tropospheric” and “at most altitudes” do not go together. Remove one of them. Also, 25-30 km is mid-stratosphere not higher.

Answer:
The word tropospheric has been removed. For the higher stratosphere the altitude range was changed to “above 30 km”.