Answer to referee #2 comment on "Validation of satellite SO2 observations in northern Finland during the Icelandic Holuhraun fissure eruption" by I. Ialongo et al.

The authors thank the referee for his useful comments. The following text includes the point-to-point answer to the referee's comments. The referee's comments are in italics while the authors' answers are in roman.

1) *This paper provides a study on the validation of satellite SO2 observations after the Holuhraun eruption. The paper is well written and structured but I find the conclusions weak. A large portion of the paper is to show that the emitted SO2 was located in the lower troposphere but this information is superfluous because this was shown many times by several groups and air quality stations. Overall I find the validation results quite limited (one Figure) and qualitative. There are many statements given without demonstration. Therefore, I think this study could be published in AMT, but after addressing the following points.*

We stress now that the focus of this paper is the evaluation of the operational SO2 products at high latitudes (also when the volcanic plume is located at low altitudes) and to show that the observed spatio-temporal link between high SO2 concentration values at surface and large total columns from satellite adds confidence in satellite-based observations for volcanic emission monitoring also at surface levels.
This is added in the revised manuscript.
We want to mention also that this paper includes several new aspects (e.g., the focus on high latitudes, the satellite data capability of detecting the volcanic plume position as compared to the surface observations, the direct broadcast products comparisons) and we believe it provides a significant added-value to the volcanic emission research using satellite data. This is also the first work in which these PBL products are presented to analyse volcanic emissions.

Main comments
2) *Such a validation paper should be a good opportunity to make sensitivity tests/alternative retrievals to solve the discrepancies, but these tests are not done. E.g. one parameter that really limits the accuracy of the satellite retrievals is the (incomplete) knowledge of the shape of the SO2 vertical profile. The discussion on the latter point is limited to the use of the different baseline products (PBL, TRL, TRM and STM) and the actual radiative transfer is not well modeled. In the text, it would also be good to say if there was snow over Sodankyla and how it could influence the satellite retrievals and validation results.*

We aim here at evaluating the performance of the operational SO2 products at high latitudes (including the direct broadcast product). In order to provide a more complete description of the accuracy of satellite retrievals, we now include a new section (3.3 in the revised manuscript) including the uncertainties of the operational retrievals and the errors expected from using the operational assumptions. We add also two figures (S4 and S5) to the supplementary material including the AMF calculation for different SZA values and the averaging kernels to account for the effect of the unknown SO2 profile.

There was not yet snow in Sodankylä in September.

3.1) *Figure 2 gives little information. If I understand correctly, only the pixels containing the Brewer station are shown. Therefore the comparison is statistically insignificant. To increase the statistics, It would be important to redo the analysis by considering all pixels with centers falling in a given area around Sodankyla.*

We now include all the overpasses within 60 km from Sodankylä in Fig. S2 and S3 in the supplementary material. Both OMI and OMPS PBL products are presented, separating large and small pixels and clear-sky and cloudy scenes.
Due to the narrow-structured volcanic plume, this dataset includes both in-plume and off-plume pixels and this should be kept in mind in the comparison. The new figures are discussed in section 3.2.

3.2) *It is also not mentioned whether the displayed values are above the OMI/OMPS detection limit and for several data points it is clear there are not.*

The detection limit is introduced in the new section 3.3. In section 3.3 we point out that the satellite retrievals are often close to the detection limit.

3.3) *The OMI BRD results for 5&27/09 are far away from the other PBL products and at the same time close to the Brewer. It is hard to know what it means (BRD are not supposed to be better than PCA product). No explanation is given.*

Because the PCA algorithm uses the entire spectrum in the SO2 fitting to reduce interferences from instrumental or geophysical effects in general it was found PCA SO2 results to be smaller than BRD, particularly for high latitudes.
This comment is added to the text.

3.4) *05/09 and 06/09: there is a clear bias between ground-based and satellite data (no matter the product selected) but this is not even mentioned by the authors. It is not enough to say that PBL are closer to the truth but the authors should discuss and understand the discrepancy.*

We mention in the text about this difference and we discuss the possible reasons in section 3.3 (for example the AMF value used in the algorithm).

4) *In the comparison with the Brewer data, the authors claim that some differences are due to different solar and viewing angles than assumed in the retrievals. For this paper to be useful, the authors shall demonstrate that.*

In section 3.3 of the revised manuscript we present the AMF calculated for different SZA values. The AMF decreases with increasing SZA, meaning that the vertical column will be underestimated for large SZAs. Also we discuss about the potential instrumental effects at high latitudes (such as stray lights and other spectral artifacts), which can produce very large biases.

5) *Page 6, on precision of satellite data: I don't understand how the STDs in the high latitude box can come any close to the estimates from the README file (equatorial pacific). There is a large difference in SZA and it should be reflected in the estimated values. Actually on Fig 1, it is clear that the noise increases at high latitudes.*

Indeed the noise increase at high latitudes. The values calculated for 1 September correspond to SZA values closer to the values calculated for equatorial pacific. When repeating the calculation in October the SZA values are much larger. We conclude indeed that the noise is larger for larger SZAs, in agreement with the referee comment.

6) Figure 3: it is not clear what the figure brings to the validation exercise.

Despite satellite vertical columns and ground-based surface concentrations are not quantitatively comparable, the observed spatio-temporal link between high SO2 concentration values at surface and large total columns from satellite adds confidence in satellite-based observations for volcanic emission monitoring also at surface levels. In particular, satellite instruments show their capability to detect the position of the volcanic plume as compared to independent ground-based observations.

We add this discussion in the text.

Minor comments
7) Figure 2 is difficult to read. I suggest to split the figures in two (one for OMI and one for OMPS). For a better readability, I also suggest to remove the STL data points (it is clear that this is not a stratospheric eruption).

Fig. S2 and S3 are added in the supplement including OMI and OMPS PBL retrievals separately. They include all overpasses within 60 km from Sodankylä. We keep all products in Fig. 2 in order to compare them to each other.

8) P1, l50-57: please specify that this if for UV sensors (there are also space infrared measurements of SO2 dating back to the mid-seventies). I suggest you chose an-other reference than Krueger et al. (2008) for the first SO2 measurement made in the 1980's. Also, please change the reference to Krotkov et al (2006) which is for OMI measurements, not TOMS.

These references are added:

Krueger, A. J.: Sighting of El Chichon sulfur dioxide clouds with the Nimbus 7 Total Ozone Mapping Spectrometer, Science, 220, 1377, doi:10.1126/science.220.4604.1377, 1983.

Gurevich, G. S., and Krueger, A. J.: Optimization of TOMS wavelength channels for ozone and sulfur dioxide retrievals, Geophys. Res. Lett., 24, 17, 2187–2190, doi:10.1029/97GL0209, 1997.

9) P2, l73-74: “Quality and timelines of. . .” It is not clear what ‘timelines’ means here.

This was a typo. We mean “timeliness”.

10) section 2.2: a detection limit of 1 DU for the Brewer data is given here but no information on possible offsets (bias) is provided. Please give details as it directly impacts the findings of the paper.

We add this sentence in section 2.2:
“The calibrations have been performed on regular basis. During the calibration the extraterrestrial constant is determined using the Langley extrapolation method as described by Redondas (2007). Since the measurements at short wavelengths are affected by stray light effects, the DS measurements corresponding to high air mass values (after 14:20 UT) are not provided. No significant bias has been estimated during the calibration.”

11) page 5, l300: It is written “the agreement is weaker because of the challenging retrieval conditions (e.g., high SZA and cloudy conditions)” but this is stated with no proof. L345: “This makes the retrieval from satellite more difficult.” but it is not explained why and what is the expected effect on the retrievals. Please clarify.

These two sentences are removed and replaced with a more complete discussion of uncertainties of the satellite retrieval reported in section 3.3 of the revised manuscript.