Comments on Anonymous Referee #1

We would like to thank the Anonymous Referee #1 for their comments that helped improving the paper. Our response is organised as follows. After each referee's comment (in italic black font) can be found the authors' response (in normal black font), and where needed, the changes made in the manuscript (in blue). In the revised version of the paper, only the significant changes have been coloured in blue to help identifying any new important improvement.

Also to improve the clarity of the paper and following the referees' comments, we have slightly changed the organisation of the paper by splitting section 5 into two. The new Section 5 is only devoted to the evaluation (ex-Sect. 5.1). Section 6 is on the impact of the inventory update on the species concentrations (ex-Sect. 5.2). Also, the purpose of the CARN simulation was not very clear for the referees. This simulation is only used to understand the effect of improving the altitude of injection. This is why CARN results are only used now for the analysis of the species concentrations in the new section 6 (ex-Sect. 5.2). The manuscript has been revised accordingly.

Please note that the revised manuscript has been read and corrected by an English native speaker and that we have added co-authors to the paper that contributed to the responses to the referees and to the revised version.

## Major comments:

This modelling study of the impact of non-eruptive volcanic  $SO_2$  emissions could have been interesting, given the importance of such emissions for the global sulfate aerosol budget. The authors implement a recently developed volcanic emission inventory (Carn et al.) which represents a significant improvement in comparison with the widely-used GEIA inventory. Those inventories are tested using the global model MOCAGE and evaluated against spaceborne  $SO_2$  columns. The correlation coefficient between the model and the data is negative (-0.17) on the global scale, and it is insensitive to choice of the emission inventory. I'm afraid that any conclusion drawn from comparisons with the model are probably useless.

1) Unfortunately, they use GOME-2 SO<sub>2</sub> columns from ACSAF, maybe the worst possible choice of SO<sub>2</sub> satellite data. OMI SO<sub>2</sub> data would have been much more appropriate. The SAF dataset is not even the best GOME-2 dataset. In fact, examination of Figure S1 shows two things: 1) the filtering applied to the columns has an disproportionate impact on the columns, and 2) the filtered GOME-2 columns (Figure S1b) have a completely unrealistic distribution. Hot spots are found in every very dry areas on Earth including South Africa, Mongolia, Tibet, Central Australia and Western U.S. This is a strong and obvious artefact. The North China Plain shows a weak enhancement, but much less pronounced than Tibet. This is not credible at all. As far as I know, this dataset has not been validated nor has it been used for any specific investigation.

We agree with the reviewer that the Metop-A GOME-2 SO<sub>2</sub> columns presented show unrealistic features in some regions. Not being experts on satellite observations, we had chosen for the model evaluation to use GOME-2 MetopA SO<sub>2</sub> columns from DLR provided by ACSAF (ex- O3F-SAF) because those data provide an independent measurement of SO<sub>2</sub> with respect to OMI (used in the volcanic emission inventory). Indeed, these data present artefacts and noise. Although we had applied filtering, this was not enough to remove all the unrealistic features. This is probably the reason why these data were mainly used in the literature not at the global scale but on case studies at the regional and local scales [Rix et al (2009,2012), Koukouli et al (2015)], and to detect very large emission sources [Fioletov et al (2013)]. Note that we also investigated the use of GOME-2 MetopB SO<sub>2</sub> columns from DLR by ACSAF (ex- O3F-SAF) but the results showed similar unrealistic features in some regions as in GOME-2 MetopB SO<sub>2</sub> columns.

All this has lead us to change our evaluation strategy to base it on OMI products as suggested by Referee #1.

2) The author should use a better  $SO_2$  dataset. I do not accept the argument that "only instruments different from those used to set the inventories can be selected for an independent evaluation". On the contrary, it seems imperative to confront the model with OMI  $SO_2$  data and check the overall performance of MOCAGE against those data. It would make the paper much interesting. Confronting the model with GOME-2 could be interesting as well, but a better dataset would have to be used.

As suggested, we choose in the revised version to use  $OMI SO_2$  columns data for the model evaluation. We

also changed the approach chosen for the statistical evaluation based on the analysis of the literature. Section 4.2 "Observations used for the evaluation of the simulations" and 5.1 "Evaluation of the simulations" were rewritten to explain our new model evaluation strategy and associated results. Here are the main modifications written in the revised paper:

As for all satellite derived products, the relative uncertainties on SO<sub>2</sub> columns are large where the SO<sub>2</sub> signal is low, in particular for background SO<sub>2</sub> conditions. This is why in the literature, the SO<sub>2</sub> satellite comparisons or the model evaluations focus on specific areas close to SO<sub>2</sub> sources [*e.g.* He et al. (2012), Fioletov et al. (2013), Wang and Wang (2020)]. Similarly to these studies, our new strategy is to perform the model evaluation only in the vicinity of the volcanic sources. For each volcano, we select 9 model grid points (representing a square of 3°longitude x 3°latitude) with the middle point being where the volcano is located. Altogether it corresponds to 633 points. The mask is applied on each daily OMI SO<sub>2</sub> total column measurements and then we perform an annual average for each of the 633 data points. Similarly to the above mentioned studies, the results are shown as scatter plots and the statistical metrics used are the correlation coefficient and the RMSE.

There are various products available in the OMI dataset since OMI instrument has a variable sensitivity depending on altitude and the retrieval of  $SO_2$  requires the use of an *a priori* profile. We choose the OMI total column density constrained by the *a priori* profiles from GEOS-5 global model. To test if the evaluation is sensitive to this choice, we use another approach which consists in an interpolation from the altitude where the volcanic emissions are injected in MOCAGE to OMI products for the boundary layer, the low troposphere and the middle troposphere. More precisely, the OMI products PBL, TRL and TRM are used. They correspond to  $SO_2$  vertical column density with an *a priori* profile assuming fixed mixing ratio within the planetary boundary layer (around 1 km), lower troposphere (around 3 km) and middle troposphere (around 8 km), respectively. Depending on the altitude of the emissions in MOCAGE, either PBL and TRL, or TRL and TRM, are used for the interpolation.

The interpolation that we made is simple. We could have used the product "Scattering Weight" similar to an averaging kernel (provides information on the vertical distribution of  $SO_2$ ) to made a better validation of our model total column with the observation. However, this method is more complicated to do since it is necessary to pre-process the observation data, adapt them into the validation process in MOCAGE and re-run the simulations.

The comparison between the model and OMI  $SO_2$  columns clearly shows an improvement of the model performances in the CARNALTI simulation (see Fig. 1).

3) The paper insists several times that "the contribution of volcanic emissions is argued as non-linear on the sulfur species burden". There seems to be quite a confusion regarding the nature of non-linearity. Yes, volcanic  $SO_2$  is longer-lived than  $SO_2$  from other sources, because it is emitted at higher altitudes and is therefore less subject to dry and wet deposition. But this does not make the contribution of volcanic emissions "non-linear". It would be non-linear if the  $SO_2$  emissions would significantly alter their own lifetime (as is the case e.g. fro NOx, due to the strong influence of NOx on hydroxyl radical concentrations). I don't think this is what the authors mean here. The emphasis on the role of non-linearity should be removed from the paper.

The meaning of this statement which was not clear is that, the contribution of volcanic emissions is argued as non-linear with respect to the volcanic sulfur emissions. In other word, by emitting 15 % of volcanic sulfur, we do not find 15 % of sulfur burden in the atmosphere due to volcanic emissions. The use of the word "non-linearity" in the paper was referring to the term used in Graf et al (1997): "The most striking feature is that the contributions of the different sources to the SO<sub>2</sub> as well as to the sulfate burden are not linear with respect to their source strengths." In the paper, we were not precise enough because we did not refer clearly to the non-linearity as the non-linearity with respect to the emissions. This has been made clear in the revised manuscript.

4) The paper also insists that non-eruptive volcanic emissions were injected at the first model level in previous studies. This is not correct. The altitude of the mouth of the volcanoes is of course well known, since a long time, and was taken into account already in the early global studies of the sulfur cycle, e.g. Spiro et al (1992), Pham et al (1995), Chin et al (2000). The crater lies generally much higher than the lowest level of the model. It seems that this point was not clear in the paper. We did not intend to emphasise that in previous studies non-eruptive emissions were injected on the first model level. We were only referring to the previous versions of the MOCAGE model, in which non-eruptive emissions were injected on the first model level. The information on the actual altitude of the volcano vent was not taken into account previously even if often much higher than the model orography (which is by definition a weighted average over the  $1^{\circ} \ge 1^{\circ}$  grid box). We knew this was a weakness of the model. We made this clearer in the revised version of the paper.

Minor comments:



Figure 1: Scatter plots of annual mean OMI  $SO_2$  versus MOCAGE simulations (left: REF, right: CARNALTI) (a) considering total columns and (b) interpolating at the model level where volcanic emissions are injected. Also shown on the scatter plot are 1:1 line (solid grey), linear regression line (black dash), linear regression formula, correlation coefficient (R), root mean squared error (RMSE), number of collocated pairs (N), OMI mean and standard deviation in DU (x), MOCAGE mean and standard deviation in DU (y), and density of collocated pairs (colorbar).

L-32) COSPEC: here, make reference to section 3.1 which explains what it is. The reference to section 3.1 for COSPEC description has been added.

L-33) TOMS: make clear that TOMS provided only crude measurements of  $SO_2$  columns.

TOMS was the first satellite instrument to measure  $SO_2$  total column from space, and at this time, the instrument specifications and the retrieval algorithms were not providing  $SO_2$  estimates as accurate as nowadays. We included in the revised version a piece of text on the TOMS early-days measurements of  $SO_2$  in the paper.

L-117) "first five levels: indicate the approximate altitude range. Why not injecting emissions at the first level only?

For numerical reasons, in particular linked to the use of a semi-Lagrangian scheme for the tracer advection, it is not recommended to inject strong and localised emissions on a single level in the model. Therefore, the injection is prescribed on the first five levels (from the model surface up to approximately 500 m), but with an exponential decrease. This leads to around 50 % injected on the first level, 25 % on the second level and the remaining mass above. The sentence "The injection profile of anthropogenic and biogenic emissions follows an exponential decrease from the surface level of the model:  $\delta_L = 0.5\delta_{L+1}$ , with  $\delta_L$  the injection fraction of the mass emitted at the level L of the model; meaning that the majority of pollutants are emitted at the surface and then quickly decrease in altitude." has been changed as follows in the revised version to make clear the reason why the emissions are not emitted on the first level only.

In MOCAGE, with the exception of the species emitted from biomass burning [Cussac et al. (2020)], lightning  $NO_X$  [Price et al. (1997)] and aircraft [Lamarque et al. (2010)], all of the chemical species sources are injected in the first five levels of the model (up to approximately 500 m). This configuration is necessary for the numerical stability in the lowest model levels. The injection profile implemented follows an exponential decrease from the

surface level of the model (including model orography):  $\delta_L = 0.5\delta_{L-1}$ , with  $\delta_L$  the injection fraction of the mass emitted at the level L of the model. It means that the majority of pollutants are emitted at the surface level and then quickly decrease with altitude. Hereafter, we will refer to "the model surface" when this configuration is used.

L-142/144) The description of SOA parameterization is very brief, and could be expanded. How well does it perform against organic aerosol observations?

The parameterization used in MOCAGE is simple. This is why its description is brief. Nevertheless, it was not clear enough. We have improved it in the revised paper.

Secondary organic aerosols are treated in MOCAGE similarly to primary aerosols with its emissions scaled on the primary anthropogenic organic carbon emissions. The scaling factor is derived from aerosol composition measurements [Castro et al. (1999)]. The implementation in MOCAGE was done by Descheemaecker et al. (2019) in the frame of a study on data assimilation for air quality applications.

The evaluation in Descheemaecker et al. (2019) was only done against  $PM_{10}$  and  $PM_{2.5}$  concentrations over Europe, not targeting specifically the secondary organic aerosols. But note that two general papers describing and extensively evaluating the latest version of the chemistry and aerosols in MOCAGE are in preparation. These papers will include comparisons with observations of different types of aerosols including organic aerosols.

Table 1) states that the Carn et al. inventory relies on TOMS and OMI, whereas the test mentioned 7 different satellite instruments.

The mention in the text, that 7 instruments are used in Carn et al (2016), is correct. The mistake in Table 1 has been corrected.

L-253/234) "One simulation takes into account only anthropogenic emissions": strange, no biomass burning or natural S emissions?  $\rightarrow$  replace by "The first run (NOVOLC) neglects volcanic emissions". Adapt also the rest of the paragraph.

The sentence has been replaced. We wanted to say that only non-volcanic emissions are injected in this simulation. In the revised version, the general description of the simulations have been improved.

L-291) "daily mean SO<sub>2</sub>: satellites do not provide daily means.

This statement has been corrected in the paper. We wanted to say that we used GOME-2 daily measurements.

L-295) "thanks to fitting AMF": unclear. As far as I know, the AMF is not fitted.

We agree that the AMF are not fitted. The sentence was unclear. It is DOAS slant columns which are fitted and then the AMF is applied to produce vertical columns. The description of GOME-2 MetopA dataset is no longer in the revised paper, since we changed our validation strategy.

Table 3) Table 3 does not bring much, since the global MNMB is given in the text, and the correlation coefficient is negative. Could be moved to the supplement. The global MNMB is not much interesting given the compensation between very high and very low values apparent in Figure 3.

As explain in the response to major comment 1), the validation strategy has been changed. We do not use anymore the GOME-2 data. Therefore this comments and those (below) regarding lines 360, 361 and 362 are not relevant anymore.

L-360) "We notice small changes in the vicinity of volcanoes where MNMB score is improved": there are many cases where the MNMB is worsened, including Hawaii and islands (Vanuatu?) in the Southern Pacific.

L-361) "FGE is better"  $\rightarrow$  "The FGE is slightly improved"

L-362) Some comments are needed concerning the negative value of the correlation coefficient.

Minor (language) comments:

L-1) Why "Thus"? The sentence remains true even in the absence of non-linear behaviour.  $\longrightarrow$  deleted L-3) at the global surface  $\rightarrow$  at the global scale (?)  $\longrightarrow$  corrected

L-4/5) I would rephrase as "the changes induced by the update of the volcanic emissions inventory are studied using the  $\dots$ "  $\longrightarrow$  rephrased

L-7) "degassing"  $\rightarrow$  "degassing emissions"  $\rightarrow$  rephrased

L-8) "uncertainties by volcances": what does that mean?  $\longrightarrow$  The sentence was clarified as follows:

Eruptions are provided as daily total amounts of sulfur dioxide  $(SO_2)$  emitted by volcanoes. Degassing emissions are provided as annual averages with the related mean annual uncertainties of those emissions by volcano.

L-9) "negligible"  $\longrightarrow$  corrected

L-12) and elsewhere: remove the dot between Tg and  $yr \longrightarrow$  corrected

L-17) "necessity of estimates accurate volcanic volcanic sources"  $\rightarrow$  "need for accurate estimates of volcanic sources"  $\rightarrow$  rephrased

L-20) delete "naturel" before "volcanic"  $\longrightarrow$  deleted

L-25) "Plus"  $\rightarrow$  "Moreover"  $\rightarrow$  corrected

L-30) "to well constrain": ??  $\longrightarrow$  In this sentence, "constrain" means "define". This sentence has been removed because not necessary.

L-43) "were not very accurate in quantitative, spatial and temporal detection": weird wording, please rephrase  $\rightarrow$  Unnecessary details were deleted and rephrased as follow:

But at the time these inventories were built, techniques for measuring emission fluxes were not very accurate for the determination of volcanic sources.

L-44) "used on":  $?? \longrightarrow$  rephrased as "deployed at"

L-46) Andres and Kasgnoc (1998) cork  $\rightarrow$  The study of Andres and Kasgnoc (1998)  $\rightarrow$  replaced

L-51) "As well": ??  $\longrightarrow$  Unnecessary, removed.

L-55) "in its work": ??. "more numerous and qualitative data".  $\rightarrow$  Rephrased as follows:

Carn et al. (2016,2017) sought to compile all those new higher quality data, compared to Andres and Kasgnoc (1998), in order to provide a more representative inventory of volcanic  $SO_2$  emissions.

L-59) "for passive source strength":: ??  $\longrightarrow$  replaced as "for passive emissions"

L-61) "huge change"  $\rightarrow$  "stark improvement"  $\rightarrow$  replaced

L-61) "last decades studies"  $\rightarrow$  "studies of the last decades"  $\rightarrow$  replaced

L-63) "the radiative forcing induced"  $\rightarrow$  "the subsequent radiative forcing"  $\longrightarrow$  replaced

L-67) "on surface species concentration and deposition"  $\rightarrow$  "on the surface concentration and deposition of sulfur species"  $\rightarrow$  replaced

L-67) "We want":  $?? \longrightarrow$  corrected to "We aim"

L-74) the configuration of simulations with  $MOCAGE \longrightarrow$  rephrased

L-76) "updating inventory":  $?? \longrightarrow$  mistake, corrected as "updated"

L-76) "the comparison for"  $\longrightarrow$  replaced

L-77) "Then"  $\rightarrow$  "Next"  $\rightarrow$  replaced

L-83) "Its use is applied":  $?? \longrightarrow$  cleared up to "It is applied"

L-94) "the duration emissions"  $\rightarrow$  "the duration of the emission"  $\rightarrow$  replaced

L-115) "from biomass burning process"  $\rightarrow$  "emitted from biomass burning"  $\rightarrow$  replaced

L-124) "completed":  $?? \longrightarrow$  corrected

L-158) "It was carried out over a period of about 25 years": I suppose you mean the measurements span 25 years. Please rephrase.  $\rightarrow$  Rephrased as It ranged over a period of about 25 years.

L-167) "thanks to the similar molecular structure  $SO_2$  and ozone": misleading, rephrase or omit; "Thus": ?? The following sentence is unclear. This could be simplified, as not really necessary.

Indeed, the reference to the similar molecular structure of ozone and sulfur dioxide was too straight forward. The two species have overlapping UV absorption bands (between 300-340 nm). Therefore, TOMS measurement of  $SO_2$  is tangled to  $O_3$ . Krueger et al (1995) explained this phenomenon as follow. "Typically, the amount of sulfur dioxide in the region of the atmosphere that affects TOMS-measured radiances (above the boundary layer) is too small to cause significant absorption. However, a volcanic eruption can produce enough  $SO_2$  in a localized region to produce UV absorption comparable to or even exceeding the ozone absorption at the shortest two TOMS wavelengths. In such cases the present TOMS algorithm incorrectly interprets  $SO_2$  as enhanced ozone. The problem is to discriminate between sulfur dioxide and ozone.". Thus, an algorithm is needed to discriminate ozone from sulfur dioxide measurements. This level of details is not necessary. The sentence has been deleted.

L-178) "constancy"  $\rightarrow$  stability  $\rightarrow$  replaced

L-178) "Thus" could be omitted  $\longrightarrow$  deleted

L-178/179) "in order to incorporate natural variations due to temporal and even chemical inhomogeneities": confusing. Could be omitted.  $\longrightarrow$  deleted

L-181) "as the one ...": replace by "as being among the largest..."; "passive": ??  $\longrightarrow$  replaced and deleted

L-181) "For them..."  $\rightarrow$  "For those volcanoes, fluxes (...) supersede the averages"  $\rightarrow$  replaced

L-185) "Knowing that"  $\rightarrow$  "Since"  $\longrightarrow$  replaced

L-189) "lowest levels" or "lowest level"?  $\longrightarrow$  As explained in major comment 4), due to numerical issues, it is not possible near the surface to inject emission on a single model level. Therefore, volcanic emissions were previously emitted on the first five levels of MOCAGE, such as anthropogenic and biogenic emissions. It was rephrased in the revised paper.

Since no configuration was developed in MOCAGE to inject volcanic emissions aloft until this study, they were implemented similarly as the other pollution sources. Volcanic  $SO_2$  were thus emitted at the model surface (see Sect. ??). However, the surface elevation of the model (orography) is mainly below the actual elevation of the volcances.

L-191) "technological improvements in satellite technology": awkward  $\rightarrow$  clumsy repetition, rephrased as "With the improvements in satellite technology"

L-197) "The work of Carn et al. (...) updates and completes the study of Andres and Kasgnoc (1998).  $\rightarrow$  replaced

L-203) "given is"  $\rightarrow$  "given includes"  $\rightarrow$  corrected

L-204) "measured"  $\rightarrow$  "estimated"  $\rightarrow$  replaced

L-205) "We will...": is somewhat ambiguous. Within this study or later on?  $\longrightarrow$  clarified by "Within this study"

L-207) "the daily frequency allows to take into account the eruptions in simulations for the period...": weird statement  $\rightarrow$  unnecessary, deleted

L-211) "could distinguish"  $\rightarrow$  "made possible to distinguish"  $\rightarrow$  corrected

L-222) "every day of the year"  $\rightarrow$  "throughout the year"  $\rightarrow$  replaced

L-236) "the update of the ... "  $\rightarrow$  "the updated"; "compiles"  $\rightarrow$  "includes"  $\rightarrow$  corrected

L-237) "spread over the globe":  $?? \longrightarrow$  unclear, it means "worldwide", not necessary so deleted

L-244) Delete words "lat" and "lon"  $\longrightarrow$  deleted

L-246) "The same global annual sulfur emissions are computed for all other sources": of course since the same inventories are used  $! \longrightarrow L-245/246$  deleted

L-248) "emissions are"  $\rightarrow$  "emissions amount to". You don't need two significant digits after the decimal point, one is enough.  $\rightarrow$  replaced

Table 1) legend "Summary information on"  $\rightarrow$  "Summary of"; "Nb of volcano"  $\rightarrow$  "Numbers of volcanoes"  $\rightarrow$  corrected

L-253) "characteristics"  $\rightarrow$  "main features"  $\rightarrow$  replaced

L-256) "However, one injects the volcanic  $SO_2$  emissions"  $\rightarrow$  "In simulation CARN, volcanic emissions are injected". Adapt also the rest of the sentence.  $\rightarrow$  The paragraph was rewritten to make it clearer, as follows: The first simulation, named REF, takes into account the previous volcanic inventory [from Andres and Kasgnoc (1998)] with the injection at the model surface. The second simulation, named CARNALTI, uses the updated volcanic inventory [from Carn et al. (2016, 2017)] and the new configuration to inject volcanic emissions from the volcano altitude as described in Section 3.2. By comparing REF and CARNALTI runs, we can analyse the changes brought by the updated volcanic emission inventory with respect to the previous one. These two simulations are evaluated in Section 5 and the associated global distribution of sulfur species is compared in Section 6.

In order to distinguish between the impact of the height of emission and of the quantity of  $SO_2$  emitted, another simulation, named CARN is run and used for the analysis of the differences between REF and CARNALTI global distribution of sulfur species. Volcanic emissions are from Carn et al. (2016, 2017), like in CARNALTI but they are injected at the model surface, like in REF.

L-261) "in altitude"  $\rightarrow$  "in the vertical"

L-262) "Then": ?? The entire sentence is weird. You could drop it since you explain what you do in the following sentence.  $\longrightarrow$  deleted

L-263) "The CARNALTI run is expected to provide the best..."  $\longrightarrow$  rephrased

Figure 1) legend: drop "annual" (since monthly values are shown). "anthropogenic" or "other emissions"?  $\rightarrow$  corrected to "non-volcanic emissions".

L-269) lowest eruptive emission flux (Carn et al., 2016)  $\rightarrow$  replaced

L-269) "is negligible". This sentence could be dropped.  $\rightarrow$  deleted

L-270) Why the upper-case AND?; This sentence is weird, not really useful.  $\rightarrow$  unnecessary, deleted

L-273) "adds": ??  $\longrightarrow$  This sentence aims to explain that the blue line in Fig. 1 is the addition of non-volcanic emissions (represented is green in Fig. 1) and volcanic emissions from Andres and Kasgnoc (1998). We cleared up the revised paper as follows:

We notice the monthly variation due to non-volcanic emissions (NOVOLC run in green), with less emissions during the northern hemisphere summer and the highest values in the northern hemisphere winter. Volcanic emissions from Andres and Kasgnoc (1998) are stable throughout the year, as we can see in REF run (in blue). They are lower than the volcanic emissions of CARNALTI and CARN runs (in red), with strong constant passive degassing throughout the year and a few sporadically eruptive events.

L-276/277) "counts": ??; "into": ??, "are"  $\rightarrow$  "amount to". Use only one significant digit for the totals  $\rightarrow$  replaced

L-279) "current": the use of this word for the previously used inventory is weir. Replace maybe (here and elsewhere) by "previous"  $\rightarrow$  corrected

L-280) "inventory against"  $\rightarrow$  "not accounted for by". Delete "one"  $\longrightarrow$  corrected

Figure 2) legend: "round"s  $\rightarrow$  "circles"  $\rightarrow$  corrected

L-284/285) Weird sentence, provide more direct formulation.  $\rightarrow$  clarified by "The target chemical species that we evaluate are SO<sub>2</sub> and aerosols, since SO<sub>2</sub> is the precursor of sulfate aerosols."

L-285) "benefit". The sentence is true but too obvious.  $\longrightarrow$  deleted

L-289) "indirectly correlated to  $SO_2 \rightarrow AOD$  depends on the quantity of all aerosol species, including sulfate aerosols. And sulfate aerosols are notably formed by sulfur dioxide, therefore  $SO_2$  can indirectly impacts the AOD. Nevertheless this statement is unnecessary, deleted

Comments between L-290 to L-306 and between L-356 to L-389 will was ignored since GOME-2 dataset is not considerated in the revised paper anymore.

L-299) "Plus"  $\rightarrow$  "In addition"; "presence of offsets"  $\rightarrow$  "offsets" (?); "lead"  $\rightarrow$  "leads"; "criteria" is plural, replace by "criterion" (if meant as singular)

L-301) "subtracted at"  $\rightarrow$  "subtracted from"

L-358) "higher"  $\rightarrow$  "less negative"

L-359) "againts"  $\rightarrow$  "against"

Table 4) "Coorelation"  $\rightarrow$  "correlation"; "specifics"  $\rightarrow$  "specific"

L-311) "low confident"  $\rightarrow$  "low-confidence"  $\rightarrow$  replaced

L-311) "filtered"  $\rightarrow$  filtered out"  $\rightarrow$  replaced

L-318) "we can use several statistical metrics": delete, and merge with next sentence "we use the fractional bias..."  $\rightarrow$  rephrased

L-335/337) This paragraph could be omitted. Delete "Therefore" from the next paragraph.  $\rightarrow$  rewritten with a brief description of the new validation strategy.

L-340) "Plus"  $\rightarrow$  "Furthermore". u I don't understand well the rest of the sentence. Rephrase.  $\rightarrow$  The paragraph has been rewritten and this sentence has been removed.

L-342) Drop "The" before Zone 1. Same elsewhere.  $\longrightarrow$  deleted

L-343) You might drop the word "inventory" after the reference. Same remark applies elsewhere in the text.  $\rightarrow$  deleted

L-346) "are"  $\rightarrow$  "amount to"  $\rightarrow$  replaced

L-354) "counting": ??  $\longrightarrow$  replaced by "totalling"

L-698) the link does not work  $\rightarrow$  corrected by https://doi.org/10.1016/j.jvolgeores.2016.01.002

Fioletov, V. E., C. A. McLinden, N. Krotkov, K. Yang, D. G. Loyola, P. Valks, N. Theys, M. Van Roozendael, C. R. Nowlan, K. Chance, X. Liu, C. Lee, R. V. Martin, Application of OMI, SCIAMACHY, and GOME-2 satellite SO<sub>2</sub> retrievals for detection of large emission sources, JGR Atmospheres, (2013), doi: https://doi.org/10.1002/jgrd.50826.

He, H. and Li, C. and Loughner, C. P. and Li, Z. and Krotkov, N. A. and Yang, K. and Wang, L. and Zheng, Y. and Bao, X. and Zhao, G. and Dickerson, R. R., SO<sub>2</sub> over central China: Measurements, numerical simulations and the tropospheric sulfur budget, JGR, (2012), doi: https://doi.org/10.1029/2011JD016473.

Koukouli, M. E. and Balis, D. S. and Theys, N. and Brenot, H. and van Gent, J. and Hendrick, F. and Wang, T. and Valks, P. and Hedelt, P. and Lichtenberg, G. and Richter, A. and Krotkov, N. and Li, C. and van der A, R., OMI/AURA,

SCHIMACHY/ENVISAT and GOME2/MetopA sulphur dioxide estimates; the cas of Eastern Asia, 'ATMOS 2015, Advances in Atmospheric Science and Application', Heraklion, Greece, June 2015 (ESA SP-375, November 2015)

Rix, M. and Valks, P. and Hao, N. and van Geffen, J. and Clerbaux, C. and Clarisse, L. and Coheur, P-F. and Loyola, D. and Erbertseder, T. and Zimmer, W. and Emmadi, S., Satellite Monitoring of Volcanic Sulfur Dioxide Emissions for Early Warning of Volcanic Hazards, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, (2009), doi: https://doi/org/10.1109/JSTARS.2009.2031120.

Rix, M. and Valks, P. and Hao, N. and Loyola, D. and Schlager, H. and Huntrieser, H. and Flemming, J. and Koehler, U. and Schumann, U. and Inness, A., Volcanic SO<sub>2</sub>, BrO and plume height estimations using GOME-2 satellite measurements during the eruption of Eyjafjallajökull in May 2010, JGR, (2012), doi : https://doi.org/10.1029/2011JD016718.

Wang, Y., Wang, J., Tropospheric SO<sub>2</sub> and NO<sub>2</sub> in 2012–2018: Contrasting views of two sensors (OMI and OMPS) from space, Atmospheric Environment (2020), doi: https://doi.org/10.1016/j.atmosenv.2019.117214.