

Dear ACP Editor and Anonymous Referees,

Please find below our answers to the 2 Anonymous Referees. In blue, the referee's comments, in black our responses.

The manuscript has been improved following the reviewers' requests, in particular note that the effect of clouds above the volcanic plume could significantly change the calculated SO₂ total emissions by 50% (from 4.4 Tg to 6.7 Tg).

Anonymous Referee #1

In the manuscript entitled "Satellite-derived sulphur dioxide (SO₂) emissions from the 2014-2015 Holuhraun eruption (Iceland)", the authors derive the first timeseries of the SO₂ emissions for the entire Holuhraun eruption. In a first stage, using a retrieval scheme previously developed, the authors retrieve the SO₂ amount and altitude of the Holuhraun plume from IASI observations. Based on these, the authors then determine SO₂ total masses every 12 hours in a large box (30_N-90_N) covering the Northern Hemisphere. They finally retrieve the SO₂ fluxes for 12-hour periods using an optimal estimation scheme, considering the retrieved total masses as the measurement vector.

To assess their results, the authors compare the retrieved SO₂ columns, plume altitude and SO₂ emissions with different type of ground-based measurements. While the SO₂ emissions determined in this paper are important for different applications, a few main issues and several specific comments should be addressed before publication.

Major comments

1) Page 2, lines 27-30, the authors mention that because of some geophysical conditions, part of the SO₂ plume can be missed by IASI, and thus, the derived SO₂ masses should be considered as minima. This is totally true, the presence of clouds and/or low thermal contrast can hamper the detection of the SO₂, and this is a complicated problem to deal with when estimating the SO₂ masses. However, the authors stay very qualitative on this problem and more particularly do not mention this problem anymore in the rest of the paper (e.g. in the comparisons). It seems that the SO₂ total mass derived by the authors for the entire eruption is lower than those previously estimated (Gauthier et al, 2016; Pfeffer et al., 2018; Gíslason et al., 2015; Thordarson and Hartley, 2015), but this is not discussed in terms of the underestimation of the SO₂ masses and its effect on the estimated fluxes. Since the latter will be available for future comparisons and for model simulations, the authors should discuss this deeper and try to evaluate (and I realize that it is a complicated problem) how large can be the underestimation of the retrieved SO₂ masses and how this underestimation affects their SO₂ fluxes.

Thank you for this comment. Indeed it is not trivial to estimate the underestimation.

We have now included in the paper a way to estimate how much the SO₂ mass could be underestimated due to meteorological cloud above the SO₂ plume. This correction can be applied to other datasets that include the altitude of the plume, and is based on monthly cloud statistics from the ESA CCI project.

We also included an estimate of SO₂ plume missed due to low thermal contrast using the OMI BIRA SO₂ dataset.

Our approach is summarised as follows:

- We estimated the percent of SO₂ missing due to cloud above the plume, as a function of cloud optical depth and the altitude of the meteorological cloud above the SO₂ plume, using simulations with a standard atmosphere as done in Fig 6 of Carboni et al 2012: (<https://www.atmos-chem-phys.net/12/11417/2012/acp-12-11417-2012.pdf>)

- Using the ESA cloud CCI dataset of AVHRR (carried on the same platform as IASI and so having the same overpass time) L3 monthly mean statistic, we computed:

1) Monthly mean histograms of frequency of cloud optical depth (COD) at 550 nm, τ , averaged over the globe. Cloud optical depth is not present in the cloud L3 database for locations without daylight (e.g. visible channels) and most of the Icelandic plume in the winter months is without daylight, as a consequence here we are assuming the global histogram of frequency of COD is valid over the plume region.

2) Monthly mean histogram of frequency of cloud altitude, averaged on the plume region (30_N-90_N). Cloud altitude is available for all locations and during winter months.

We consider the measured mass M_{meas} to be the difference between true mass M_{corr} and the missing one M_{miss} :

$$M_{corr} - M_{miss} = M_{meas}$$

$$M_{corr} \left(1 - \frac{M_{miss}}{M_{corr}} \right) = M_{meas}$$

$$M_{corr} = M_{meas} \left(\frac{1}{1 - \frac{M_{miss}}{M_{corr}}} \right)$$

We compute the correction factor, C , for every month of the eruption as a function of altitude, and applied to the vertical distribution dataset.

$$M_{corr}(h) = M_{meas}(h) \cdot C(h)$$

With:

$$C(h) = \frac{1}{(1 - Z(h))}$$

Where $Z(h)$ is the SO₂ mass fraction 'missed' in the measurements due to cloud above the plume. $Z(h)$ is estimated as the product of probability of having cloud above altitude h , $F(h)$, times the attenuation due to cloud, A ,

$$Z(h) = F(h) \cdot A$$

The probability of having cloud above h has been estimate from CCI data for the region considered for the volcanic plume (latitude > 30 N) as the number of cloud retrievals above altitude h divided by number of observations.

Attenuation due to cloud (A) is the sum of the frequency of having a cloud with a cloud optical depth $f(\tau)$ times the attenuation due to a cloud with the same optical depth $a(\tau)$.

$$A = \sum_{\tau=0}^n f(\tau) a(\tau)$$

$f(\tau)$ has been estimated using the monthly mean histogram of frequency of cloud optical depth, estimated over the globe.

$a(\tau)$ has been estimated by running the SO₂ retrieval using, as IASI measurements, simulated spectra with water cloud above the plume, using the default atmosphere, and different optical depths at 550 nm. For optical depths bigger than 10 the attenuation is 1 (cloud is opaque and completely mask the SO₂ signal).

The figures below show:

- Correction factor.
- SO₂ vertical distribution obtained from IASI retrieval (was already in the paper).
- SO₂ vertical distribution corrected (for underestimation due to cloud cover).

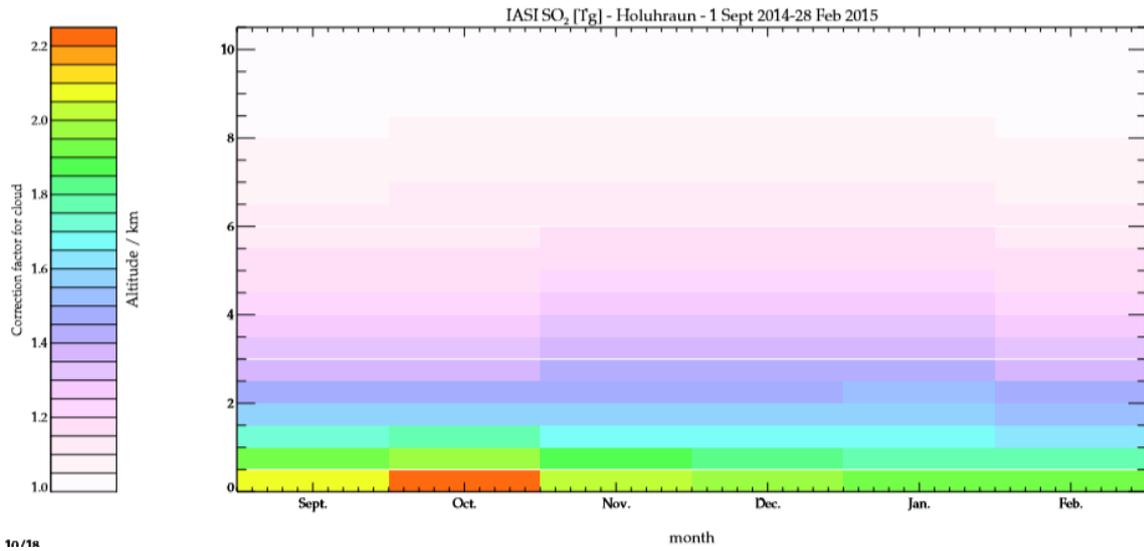


Figure 1. Correction factor for the SO₂ masses to estimate to correct for the presence of cloud above the plume.

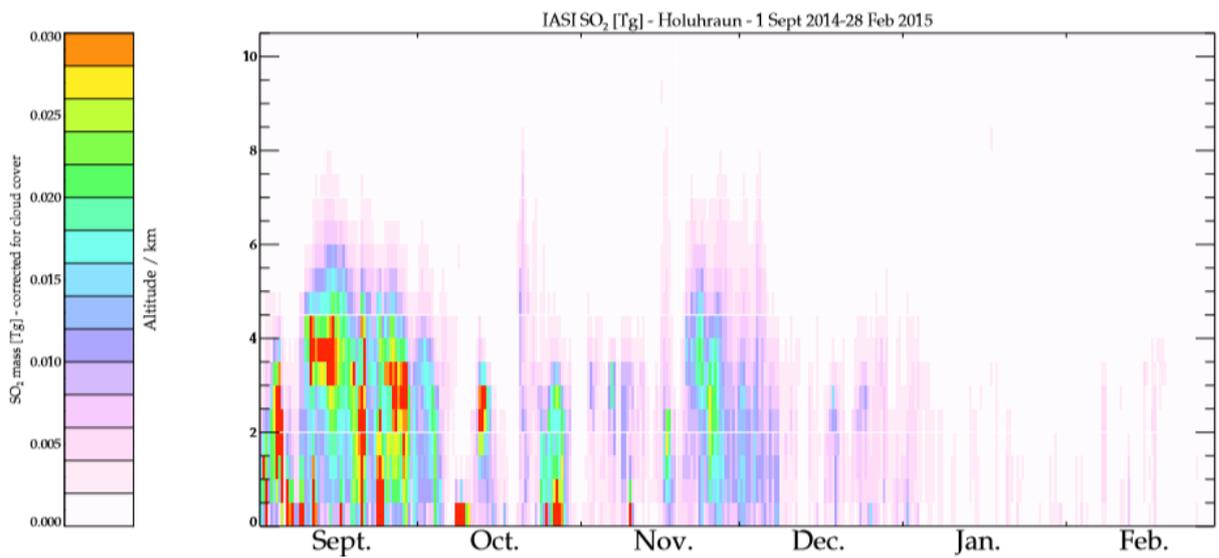
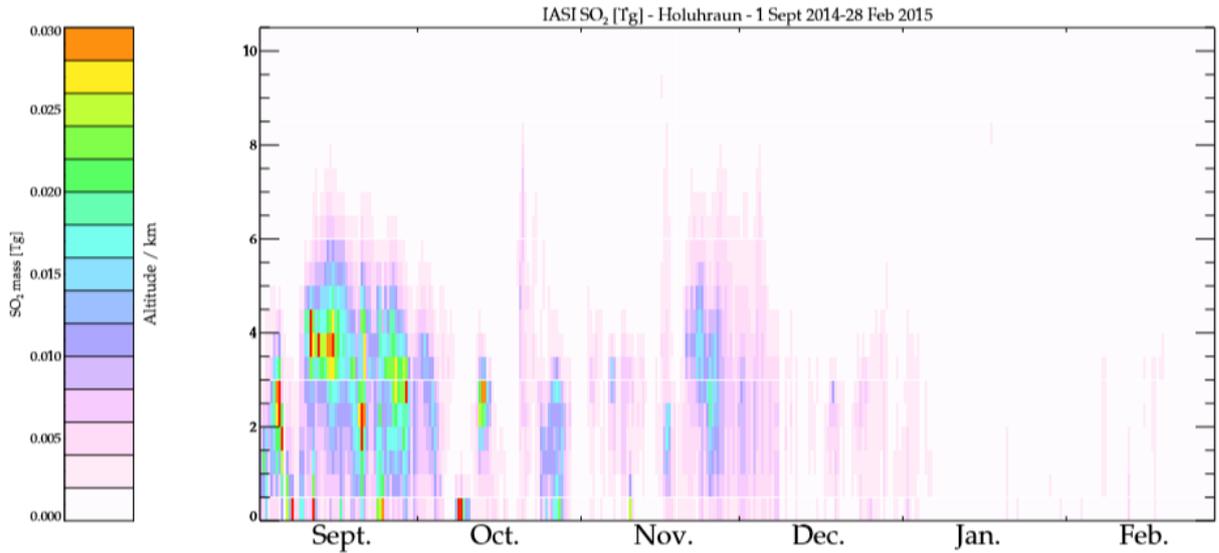


Figure 2. SO₂ vertical distribution in km above sea level. The colour represents the mass of SO₂, dark-red represents values higher than the colour-bar. Every column of the plot is generated from an IASI map (one every 12 hrs). First plot show the data obtained from IASI maps, second plot is the first plot times the correction factor (to include SO₂ that statistically has been missed by the IASI measurements due to cloud above the plume)

The emission fluxes have been estimated with both the original SO₂ masses (from IASI retrieval) and the masses corrected by cloud cover.

The total emission estimated with a cloud correction is 6.7 Tg (without the correction it is 4.4 Tg).

These results have been added to the manuscript.

We estimate the missing SO₂ due to thermal contrast by comparison with OMI SO₂ (UV dataset) for the month of September and October 2014, as the OMI dataset doesn't fully cover the eruption time period due to the lack of solar radiance during the winter.

We visually inspected the daily maps of IASI and OMI and identified the parts of plume missing from the IASI detection (and consequently missing in the IASI retrieval).

Here is the list of all areas identified and the SO₂ estimate from OMI (BIRA-IASB)

Date	Max latitude	Min latitude	Max longitude.	Min longitude.	OMI SO ₂ for 7km height [kT]	OMI SO ₂ for 0-1 km a.g.l. [kT]
20140901	70	60	-20	-36	7.25	9.6
20140915	70	75	-10	-20	9.5	11.1
20140915	75	70	30	0	21.5	33.7
20140929	70	65	-15	-30	2	3
20140929	77	63	-20	-40	12	26

During the first 2 months we miss part of the plume corresponding to (summing them all) 83.4 kt = 0.08 Tg (1 [Tg] = 1000 [kt])

The emission estimate (sum of fluxes * interval of time between 2 maps) from IASI for the first 2 months is 2.71 Tg, the missing mass of SO₂ (estimate with OMI 0-1 km) summed over the first 2 months is 0.08. Then the estimate of SO₂ missing due to thermal contrast is around 3%. (0.08/2.71 = 0.03).

The total mass of SO₂ missed due to thermal contrast is estimated to be few percent of the emission estimate by IASI. In particular the missing plume for the first 2 months has a total mass of 0.08 Tg of SO₂ that corresponds to 3% percent of the emission estimate by IASI.

Low SO₂ cloud also be a problem for OMI. If the OMI values are wrong by a factor 2-3, the underestimation will change to 6-9% (instead of 3%).

This estimate has been added to the text.

2) Page 3, lines 10-18, it is explained that the estimation of the SO₂ masses is performed for a box going from 30_N to 90_N, considering that the SO₂ detected comes from the Holuhraun eruption only. However, in this large box, other SO₂ sources (China, Norilsk, volcanoes) are located and contribute to the total SO₂ in this box.

While the SO₂ amount emitted by these sources is probably negligible compared to the one emitted at the beginning of the eruption (masses of 0.1-0.3 Tg), I am afraid these sources could contribute to a larger percentage the days where SO₂ masses lower than 0.1 Tg are estimated. For instance, the annual SO₂ emissions of Norilsk are estimated to be around 2 Tg (Fioletov et al., 2016). On a daily basis, this can lead to SO₂ masses around 0.001-0.01 Tg, and this can represent a large percentage of the estimated SO₂ masses. This is especially the case from December, when the SO₂ masses are mostly lower than 0.1 Tg. Moreover, in this period, the thermal contrast values and humidity conditions in the Norilsk area, but also in China, were shown to favour the measurement of near-surface SO₂ (Boynard et al., 2014; Bauduin et al., 2014; 2016).

How did the authors take into account these extra sources? Did they remove a background of SO₂ from their masses? How large do they estimate the contribution of the other sources and how does it affect the estimated SO₂ masses and fluxes? This issue deserves more explanations and investigations.

We expanded the algorithm description (as also requested by referee 2) in section 2, and we hope that it is clearer now.

The IASI SO₂ algorithm is based on

1) A detection scheme that only uses the v3 band.

And for all the pixel identified by the detection

2) An iterative retrieval scheme that includes both the v1 and v3 bands.

Our detection scheme, whose theory is explained in Walker et al. (2011, 2012), is a linear retrieval with one free parameter - the column amount of SO₂. In particular we assume the vertical distribution of SO₂ and the atmospheric vertical profiles (temperature and trace gases). We don't take into account negative thermal contrast so that regions with negative thermal contrast (such as Norilsk) often give negative values of SO₂ column amount. You can see this artefact for the month November 2013 in the following plot where negative monthly means around Norilsk are white.

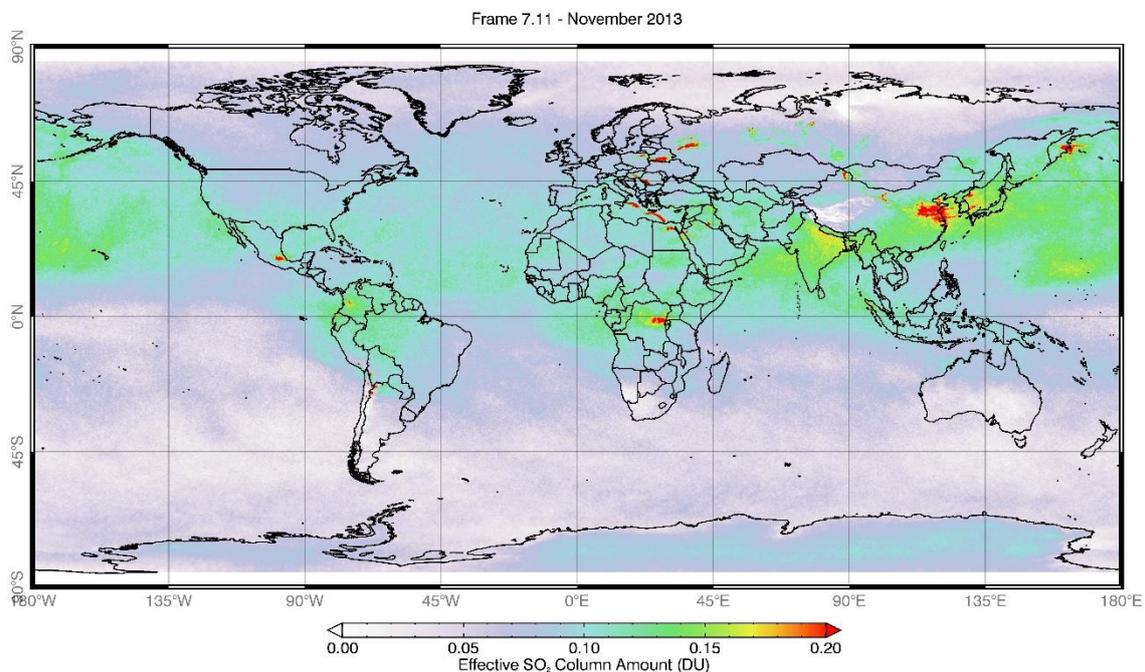


Figure 3: Global IASI SO₂ linear retrieval output averaged for November 2013.

More monthly mean plots are available here (as supplementary information of the paper Taylor et al 2018):

<https://agupubs.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2F2017JD027109&file=jgrd54568-sup-0002-supinfo.gif>

In our scheme we consider detection 'positive' if the output of the linear retrieval is greater than a defined positive threshold.

The full IASI dataset from 2007 to 2014 has been analysed with the same linear retrieval and results are presented in Taylor et al (2018).

Taylor, I.A., J. Preston, E. Carboni, T.A. Mather, R.G. Grainger, N. Theys, S. Hidalgo and B. McCormick Kilbride, Exploring the utility of IASI for monitoring volcanic SO₂ emissions, *Journal of Geophysical Research: Atmospheres*, **123**, 5588–5606, 2018. (doi:10.1002/2017JD027109)

The threshold for positive detection is only exceeded in the area of Norilsk when this region is affected by a volcanic plume (such as from Kasatocy, Sarakyev...). This is a limitation of our detection scheme, and improvements are under development, but for the purpose of this paper this means that Norilsk's emissions are not included as part of a volcanic plume.

The movie of the IASI SO₂ plume (present in supplementary data) shows the absence of the Norilsk contribution to the plume: there is only SO₂ detection and retrieval over the Norilsk area when a volcanic plume overpasses the area.

3) I have a few comments on the method used to estimate the SO₂ fluxes. First of all, the authors should explain the advantages of the method they use compared to others ones (Theys et al., 2013). Then, I am concerned about the assumption of an averaged constant lifetime for the entire eruption. As mentioned by the authors, the lifetime of SO₂ is very variable and depends on humidity, solar irradiation and altitude of the plume. Because the eruption lasted 6 months and the plume travelled very far from the source, these conditions significantly varied during the eruption and according to the location of the plume. Therefore, I am wondering why the authors have made this choice of method and why they did not consider a more sophisticated method, using a dispersion model to estimate the SO₂ fluxes (Theys et al., 2013). At page 7, line 26, the authors mention that the flux uncertainties include the possible variation of the e-folding time. This is not clear how this is done. In conclusion, the authors should justify their choice of method and provide a clear explanation of how they assess the impact of a constant lifetime on the retrieved SO₂ fluxes.

We agree that the best way to estimate the fluxes will be to combine satellite measurements with a dispersion model, possibly using a scheme with a variational assimilation of the SO₂ plume height and column retrievals as we have done for Eyjafjallajökull (Vira et al 2017).

Vira, J., E. Carboni, R.G. Grainger and M. Sofiev, Variational assimilation of IASI SO₂ plume height and total column retrievals in the 2010 eruption of Eyjafjallajökull using the SILAM v5.3 chemistry transport model, *Geoscientific Model Development*, **10**, 1985–2008, 2017. (doi:10.5194/gmd-10-1985-2017)

This manuscript presents a way to derive the emission flux from a time series of total mass loading from satellite data only, and doesn't require a dispersion model. Adding this would be an enormous effort outside the scope of the paper.

We have added text to the paper to suggest further work on the use of an assimilation scheme to give the emission fluxes, the vertical distribution of emissions and associated errors.

For the first month of the eruption we compared satellite datasets (IASI and OMI) with a dispersion model (NAME) simulations (Schmidt et al 2015). Different fluxes and emission altitudes were tested to estimate the values that better match model and satellite, The fluxes found with this comparison are consistent with the values estimated here and the following figure has been now added to the manuscript.

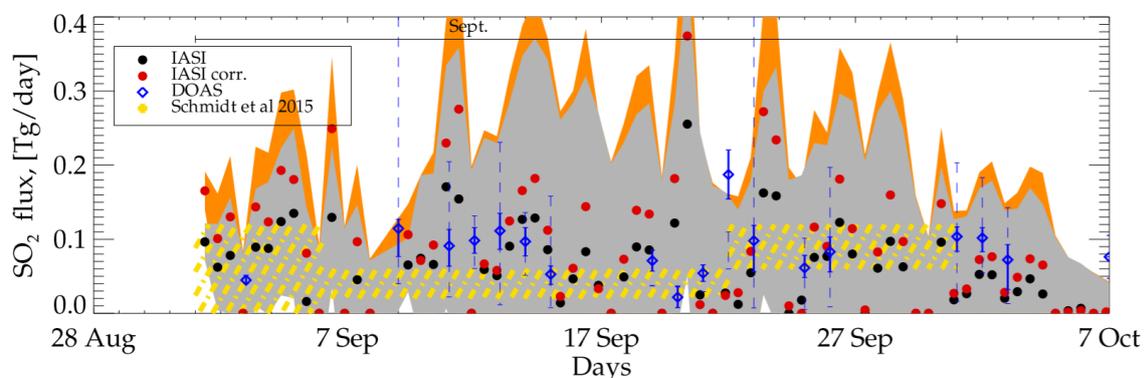


Figure 3. SO₂ flux time-series from IASI/IASI corrected by cloud cover in black/red with grey/orange error-bars. Ground-based DOAS measurements in blue (blue bars show the errors, dotted bars show the maximum and minimum values measured that day) and range of fluxes from Schmidt et al 2015 in yellow.

The optimal estimation scheme gives a vector of parameters that we wish to estimate (the state vector) and associated errors. It is a Bayesian scheme that fits the measurements and a priori knowledge of the state vector. In particular we minimize a cost function:

$$\chi^2 = [\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{b})]^T \mathbf{S}_e^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}, \mathbf{b})] + [\mathbf{x} - \mathbf{x}_a]^T \mathbf{S}_a^{-1} [\mathbf{x} - \mathbf{x}_a],$$

where \mathbf{x} is the state vector, \mathbf{y} is the vector of measurements, \mathbf{F} is the forward model (function of \mathbf{x} and auxiliary data \mathbf{b}), \mathbf{x}_a is the *a priori value of the state vector* and \mathbf{S}_e and \mathbf{S}_a are the *measurement and a priori error covariance matrices* respectively. The output is the more probable state vector together with the a posteriori covariance matrix of the state vector \mathbf{S}_x . The square root of the diagonal elements of \mathbf{S}_x are the uncertainties given for the retrieved parameters. In this case the resulting uncertainties in the fluxes are a function of the errors in total mass, a priori errors and information content of the measurements.

4) Regarding among other things the previous comments, the present paper lack references, especially related to the estimation of volcanic SO₂ emissions and the SO₂ lifetime. The following references should be at least added:

McCoy and Hartmann, GRL 42, 10409-10414, 2015, doi:10.1002/2015GL067070;

Malavelle et al., Nature 546, 485-491, 2017, doi:10.1038/nature22974;

Lee et al., JGR 116, 2011, doi :10.1029/2010JD014758;

Theys et al., ACP 13, 5945-5968, 2013, doi :10.5194/acp-13-5945-2013;

Carn et al., J. Volc. Geoth. Res. 311, 99-134, 2016, doi:10.1016/j.jvolgeores.2016.01.002.

Thanks for these suggestions. References has been added to the manuscript.

5) Some parts of the manuscript are difficult to read and are not clear. For instance, the following paragraphs could be improved (see also specific and technical comments):
Page 3, lines 3-6; Page 7, lines 24-31; Page 8, lines 1-12.

Manuscript has been expanded and rewritten in these parts.

Specific comments

-Abstract, line 5: the use of the optimal estimation to infer the SO₂ emissions is not something new (Theys et al., 2013).

We don't agree. Optimal estimation (OE) has been used previously to estimate the mass loading of SO₂ (Carboni et al 2012, Clarisse et al 2008, 2012) not to estimate fluxes.

Theys et al (2013) use optimal estimation (in section 3.4 'inversion modelling method') to fit a dispersion model to observations.

Here we use optimal estimation to fit the measured time series of total mass loading with a forward model that reproduces the time series of total mass loading (a function of emission flux and SO₂ e-folding time)

This forward model (eq 4 in the manuscript) is the inverse of equation 6 in Theys et al 2013 (section 'Delta-M method'). Both equations derive from the solution of the same differential equation:

$$dM(t) / dt = F(t) - k \cdot M(t).$$

In Theys et al (2013) the fluxes are obtained assuming the e-folding time, here the fluxes and the averaged e-folding time and their uncertainties, are estimated simultaneously based on OE.

-Abstract, line 6: "The algorithm is used to estimate SO₂ fluxes of up to 200 kt : :". This sentence sounds weird to me. I understand that the algorithm cannot estimate fluxes larger than 200 kt.

'The algorithm is used to estimate SO₂ fluxes of up to 200 kt per day and a minimum total SO₂ erupted mass of 4.4±0.8 Tg'

Has been replaced with

'For the six months eruption studied, the SO₂ flux was observed to be up to 200 kt per day and the minimum total SO₂ erupted mass was 4.4 ± 0.8 Tg'

-Abstract, line 8: you say that you compared your results to model simulations. Do you refer to the comparison with the work of Schmidt et al. (2015)? You should rephrase the sentence because, the way it is written, the reader understands that you do model simulations.

'Where comparisons are possible, these results broadly agree with ground-based near-source measurements, independent remote-sensing data and model simulations of the eruption.'

Changed in:

'Where comparisons are possible, these results broadly agree with ground-based near-source measurements, independent remote-sensing data and values obtained from model simulations (Schmidt et al 2015).'

-Page 2, line 12: can you specify what coverage? Is it temporal, spatial or both?

Both.

'Retrievals of SO₂ amount from Metop-A satellite are binned and averaged for successive 12 hour periods to give coverage for the entire period of the eruption.'

Changed in:

'Retrievals of SO₂ amount from Metop-A satellite are binned and averaged for successive 12 hour periods to give global coverage twice a day for the entire period of the eruption.'

-Page 2, line 22: did you use IASI data of level 1b (not apodized)? Or 1c (obtained after apodization)?

1c Apodized, this has been added to the text.

-Page 2, line 24: Can you briefly remind what is a positive result in the SO₂ detection scheme?

An output values (of the linear retrieval) higher than 0.45 effective DU.
Description of the scheme has been expanded including this.

-Page 2, line 28: I would specify that you miss part of low-altitude SO₂ plumes in case of low thermal contrast. On the same line, you say that IASI can miss part of the SO₂ plume in case of negative thermal contrast conditions. Why? It has been shown by Bauduin et al. (2014, 2016) and Boynard et al. (2014) that negative thermal contrasts are favourable conditions to measure SO₂ close to the surface.

The referee is right, it is not IASI, but the IASI detection scheme used in this manuscript, that can miss part of the plume.

Session 2 has been expanded including this.

-Page 3, line 2: Why cannot you use the ν_3 band to measure SO₂ down to the surface?

Is the band saturated? Or is it because of large humidity close to the surface? Can you please explain? If none of these two reasons is true, you should be able to measure SO₂ close to the surface using the ν_3 band (Bauduin et al., 2014; 2016).

Sentence in the manuscript:

'All the channels in the ranges 1000-1200 and 1300-1410 cm⁻¹ (the 7.3 and 8.7 μm SO₂ bands) are simultaneously used in the iterative optimal estimation retrieval scheme to obtain the SO₂ amount, the altitude of the plume and the surface temperature. The SO₂ band around 8.7 μm (1000 to 1200 cm⁻¹) is within an atmospheric window. This allows the radiation from the surface to reach the satellite from deep within the atmosphere enabling the retrieval of SO₂ amount down to the surface.'

You can use only the ν_3 band but, for standard atmosphere conditions, your measurements will not be affected by SO₂ close to the surface, due to strong water vapour absorption. The ν_3 band will only allow SO₂ retrieval close to the surface in dry condition as stated in Bauduin et al. (2014; 2016).

Using ν_1 band together with the ν_3 band increases the information content of the measurements so that SO₂ is measured in all water vapour conditions.

-Page 3, line 3: How did you build the error covariance matrix? Is it a global one or did you build one more specific for the eruption? This should be explained.

We used the global covariance matrix, this has been added to section 2.

-Page 3, line 6: Can you specify what is your quality control?

Quality control (now added to section 2) is:

Cost function < 10, retrieved pressure between 0 and 1100 mb, positive SO₂ column amount, plus convergence of the iteration algorithm.

-Page 3, line 7: You say that the SO₂ retrieval is not affected by an underlying cloud. However, this cloud has to be taken in the retrieval, at least in the radiative transfer. How did you take into account the underlying clouds? How did you detect underlying clouds?

We don't detect cloud, the variability of the spectra due to cloud presence is included in the measurement error covariance matrix.

We build up the measurement error covariance matrix with the differences between our forward model (radiative transfer with no cloud, driven by ECMWF profiles) and the IASI measurements. In this way the measurement error covariance matrix include the variability of all the parameters that are not retrieved and not well represent by the forward model. The biggest contribution to this covariance matrix is the presence of cloud.

More detail on the retrieval scheme in Carboni et al (2012).

Moreover comparisons with CALIPSO measurements (Carboni et al. 2016) for cases strongly affected by underling cloud show consistency between the IASI retrieved altitude and CALIPSO backscattering profiles.

-Page 3, line 11: it is not clear to me what you combined. I suppose you created AM and PM maps each day?

Yes.

In the manuscript:

'The retrieval results from the Metop-A orbits during the period from September 2014 to February 2015 and from 30° N to 10 90°N are combined to produce two maps per day of retrieved SO₂ amount and altitude.'

-Page 3, lines 15-18: you say that you cannot make the distinction between the Holuhraun plume and the other SO₂ sources, but then you make the distinction for the 21st and the 31st December. How did you do this distinction? How did you take this into account in the evaluation of the SO₂ masses and fluxes? (major comment 2)

'Satellite observations at the pixel level do not provide sufficient information to distinguish between SO₂ from Holuhraun and SO₂ from other sources.'

The following comments in the manuscript (below) describing other sources come from the visual inspection of the sequence of daily maps (show as a movie in the supplemental material).

'For example, the elevated SO₂ near Beijing on 21st December 2014 appears to be from an anthropogenic source but the elevated SO₂ in the same area on 31st December 2014 is from the Holuhraun eruption. '

As is possible to see in the supplement material movie, in the maps before 21 December there is no presence of volcanic plume moving toward China, while in the maps before 31 December we can follow the evolution of the Icelandic plume over Asia and reaching Beijing.

-Page 3, line 24: the reference Boichu et al. (2015) should be added.

Done

-Page 3, line 30: can you specify how you calculate the errors on the total masses?

'The SO₂ mass present in the atmosphere for each IASI overpass was found by regridding the observations of column amount and plume altitude into a 0.125° latitude/longitude boxes following Carboni et al. (2016). The SO₂ mass time-series is obtained by summing the mass values of the regularly gridded map for each 12 hour period. The time-series of SO₂ mass, together with the errors, are presented in the top plot of Figure 2'

Changed into

'The SO₂ mass present in the atmosphere for each IASI overpass was found by regridding the observations of column amount and plume altitude into a 0.125° latitude/longitude boxes following Carboni et al. (2016). The SO₂ mass time-series is obtained by summing the mass values of the regularly gridded map for each 12 hour period. The same procedure has been used for the errors, this means that the sums of errors of grid boxes are considered as errors on the total masses (this could be an error overestimation but we cannot consider the usual errors in quadrature due to the possible presence of systematic error, e.g. the errors are not independent). The time-series of SO₂ mass, together with the errors, are presented in the top plot of Figure 2'

-Page 6, lines 10-11: why did you choose these a priori values? Did you rely on the previous literature?

'The a priori values used were 0.2 ± 0.2 Tg/day for flux and 2 ± 2 day for the e-folding time.'
For the flux we choose 0.2 ± 0.2 Tg a day for a priori flux because this allows the retrieval to move easily (if there is enough information in the measurements) between 0 and 0.4 Tg/day and 0.4 is greater than the maximum value of total mass. Only a few values of total mass in September exceed 0.2 Tg and in particular none of them show a total mass greater the 0.3.
We think that the a priori e-folding time between 0 and 4 will cover the real lifetime of SO₂ for low tropospheric plumes. We experimented with longer e-folding time but the resulting fitting is worse. Assuming shorter e-folding time results as good a fit as the one presented in the manuscript, this is why we have written this at line 28 page 7: 'Also note that any e-folding time shorter than the retrieved one can fit the measurements and give higher fluxes.'

-Page 7, line 7: you did not explain what Se you considered. Can you specify it?

Not sure I understand the comment here.

Page 7 line 7 states:

'where λ (with $\lambda = 1/k$) is the average e-folding time. Equation 4 is the forward model $F(x)$.'

We define Se in pag 6, line 9-10: 'Se and Sa are diagonal matrixes with the variances (square of errors) of y and xa respectively as their diagonal elements.'

-Page 7, line 13: The averaged fluxes reported for December, January and February are very low. The impact of other sources is in this case non negligible (major comment 2). Did you take this into account?

We did consider all the SO₂ as Icelandic source. See answer to major comment.

-Page 7, lines 14-15: You did not give a tentative explanation for the fact that 1) the monthly averages of IASI fluxes are lower than those calculated from ground-based observations, and 2) the maximum values are larger for IASI than for ground-based observations. Is this because of the underestimation of the masses? Or variations in the lifetime? Or the inclusion of other sources? I think you can extend the discussion.

Here is the paragraph from the manuscript: 'The estimates for December, January and February show decreasing flux with monthly averages of 0.016, 0.006 and 0.005 Tg/day respectively (0.026, 0.028, 0.016). The monthly averages are lower than those measured by the ground-based measurements while the maximum daily averages for each month are generally higher.'

We added this to the text:

'The UV ground-based measurements for the dark months of December, January and February are sparse, with only 10 measurements over these 3 months. There was only one day with measurements at the beginning of December, and then 6 days with measurements in the second half of January and three days with measurements in February. The extrapolated flux from the ground-based measurements through December to the first half of January is consistent with the error bars from the IASI estimates. The differences in monthly means between the ground-based measurements and the IASI flux estimates in the sunny months are explainable by low values with large error.'

-Page 7, line 16: You compare the SO₂ fluxes you determined with the modelled fluxes of Schmidt et al. (2015). Why didn't you also compare your emissions with the fluxes they determined from OMI and IASI observations?

In Schmidt et al 2015 we did not estimate fluxes from IASI and OMI, we estimated fluxes comparing maps of SO₂ column amount from IASI, OMI and NAME. The NAME simulation that best matched the satellite was used to estimate the flux range reported in Schmidt et al (2015). We added the Schmidt et al 2015 fluxes estimate in fig 2.

-Page 7: I would add the errors of the fluxes in the text.

I'm not sure what this comment refers to but I guess the referee refers to these lines 15-17 at page 7: 'The fluxes calculated for September 2014 are consistent with Schmidt et al. (2015) (e.g. up to 0.120 Tg/d during early September, 0.02-0.6 Tg/day between the 6th and 22nd of September, 0.06-0.120 Tg/day until the end of September).'

We now added a 'zoomed' plot with fluxes for September (only) including IASI estimate of fluxes, IASI error-bars and fluxes estimate from Schmidt et al. (2015) that show the consistency of IASI fluxes with Schmidt et al. (2015).

-Page 7, line 18: you calculate the total mass of the eruptions from the SO₂ emissions you derived. These emissions are strongly affected by the fact that you use a constant averaged SO₂ lifetime. Wasn't it more accurate to use the daily masses you estimated (even though they are underestimated)?

Summing all the daily masses, and considering this sum as the total emission, would assume that the plume in one retrieval is completely gone in the following retrieval after 12 hours. This is not true as we can follow the plume evolving in time and reaching different parts of the northern hemisphere and we can visually track the same part of plume through consecutive retrievals for multiple days. We need an estimate of flux that takes into account an e-folding time. In this manuscript we chose to take into account one e-folding time with a big a priori error that includes all the e-folding time variability.

-Page 7, lines 32-33: the "spikes" you see in the SO₂ fluxes, are they real? Or do they come from the forward model you used? The Delta-M method is known to produce spikes in time-dependent fluxes (Theys et al., 2013).

(Theys et al., 2013) state that (in the Delta-M section):

'As these series (referring to total mass series) do contain some uncertainty, the resulting flux curves often display spikes that are likely not related to real source variation.'

Delta M spikes are mostly coming from incomplete coverage of the plume, e.g. due to orbital position. In the case of the Holuhraun eruption we have complete coverage of the plume using IASI data.

Nevertheless it is true that the fluxes estimated in this manuscript show results with some spikes but the fluxes obtained here have to be considered together with their error estimates. Here we consider the time series of total masses and associate errors in a comprehensive optimal estimation scheme.

Note that the errors in the resulting fluxes are often of the order of 100%. In figure 2 of the manuscript the grey colour band represents the flux errors and this grey band rarely is detached by the horizontal line of 'zero' flux line, even in correspondence of the spikes. It is plausible that spikes in the SO₂ fluxes are genuine.

-Page 8, lines 8-10: How do you explain that Gauthier et al. (2016) estimated lower SO₂ daily masses but a higher total SO₂ mass? Is this related to their choice of lifetime? This comparison should be discussed deeper.

Although is not 100% clear, from the Gauthier et al. (2016) paper, how they estimate the total emitted mass (from sparse fluxes only in correspondence of 'positive gradient' between 2 successive SEVIRI masses), there are 2 possible explanations for disagreement.

One explanation could indeed be the assumed SO₂ lifetime, but Gauthier et al. (2016) estimate that this could only affect their masses by 1.3–5.2% percent (Gauthier et al 2016, section 3.1).

The other explanation is mainly related with the fact that they use a less sensitive instrument (SEVIRI) and an algorithm that, in case of no valid measurements, interpolates between valid flux estimates.

This paragraph, that was in the manuscript:

'Had a less sensitive instrument been used that only produced 'valid' measurements in correspondence with higher flux values (e.g. > 0.05Tg/day) and had considered these fluxes as representative of the period without valid measurements (i.e. period between two 'valid' measurements), this would result in very different (and higher/overestimated) estimated values of total SO₂ emitted. An example of this is Gauthier et al. (2016) where they used TIR data from SEVIRI, on board the geostationary satellite Meteosat Second Generation (MSG), to retrieve an SO₂ mass time-series from 1 September 2014 to 25 November 2014. Their retrieved mass values are lower compared to the IASI values here (due to the smaller geographic area considered and possibly due to a smaller sensitivity or detection threshold of SEVIRI), nevertheless they estimate a total SO₂ emitted mass of 8.9±0.3 Tg for the period September 2014 to November 2014, which is a factor of two higher than calculated here.'

Have been substitute with this one:

'Gauthier et al. (2016) used TIR data from SEVIRI, on-board the geostationary satellite Meteosat Second Generation (MSG), to retrieve an SO₂ mass time-series from 1 September 2014 to 25 November 2014. Their daily retrieved mass values are lower compared to the IASI values reported here due to the smaller geographic area considered and possibly due to the lower sensitivity of SEVIRI. Nevertheless they estimate a total SO₂ emitted mass of 8.9±0.3 Tg for the period September 2014 to November 2014, which is a factor of two higher than IASI. Our understanding is that the Gauthier et al. (2016) scheme produced 'valid' measurements when the SO₂ loading increased. The resulting data gaps were filled by linearly interpolating between two 'valid' measurements. This has a tendency to bias flux estimates in favour of increasing SO₂ loading. The dataset of Gauthier et al. (2016) contains several days with no valid fluxes estimates and for these data gaps the interpolation of valid measurements into data gap could account for their discrepancies with our dataset.'

-Page 8, lines 11-12: Following the previous comment, I think the comparison could be extended. You did not compare the total mass you estimated with those reported by Schmidt et al. (2015) (2 Tg for September 2014) and by Gíslason et al. (2015) (11 Tg). Moreover, you should compare the total masses calculated for a same period.

The comparison should mention the difference in sensitivity, in lifetime, : : :

The total masses reported by Schmidt et al. (2015) are relative to a smaller area that the one considered here (summed from a lat-lon box of 60°W-40°E and 75°N-45°N) and in particular the IASI total masses are coming from the same IASI dataset. We did add the comparison of the fluxes estimate from Schmidt et al. (2015) in a new figure of the paper (reported as fig 3 in this document)

-Page 8, lines 17-18: As already mentioned above, negative thermal contrasts have been shown to increase the sensitivity to near-surface SO₂ (Bauduin et al, 2014; 2016; Boynard et al., 2014).

See discussion above and rewritten section 2.

-Page 8, line 34: Can you explain why you compare ground-based measurements with the average of all IASI pixels located within 200 km? Why not taking the closest IASI pixel?

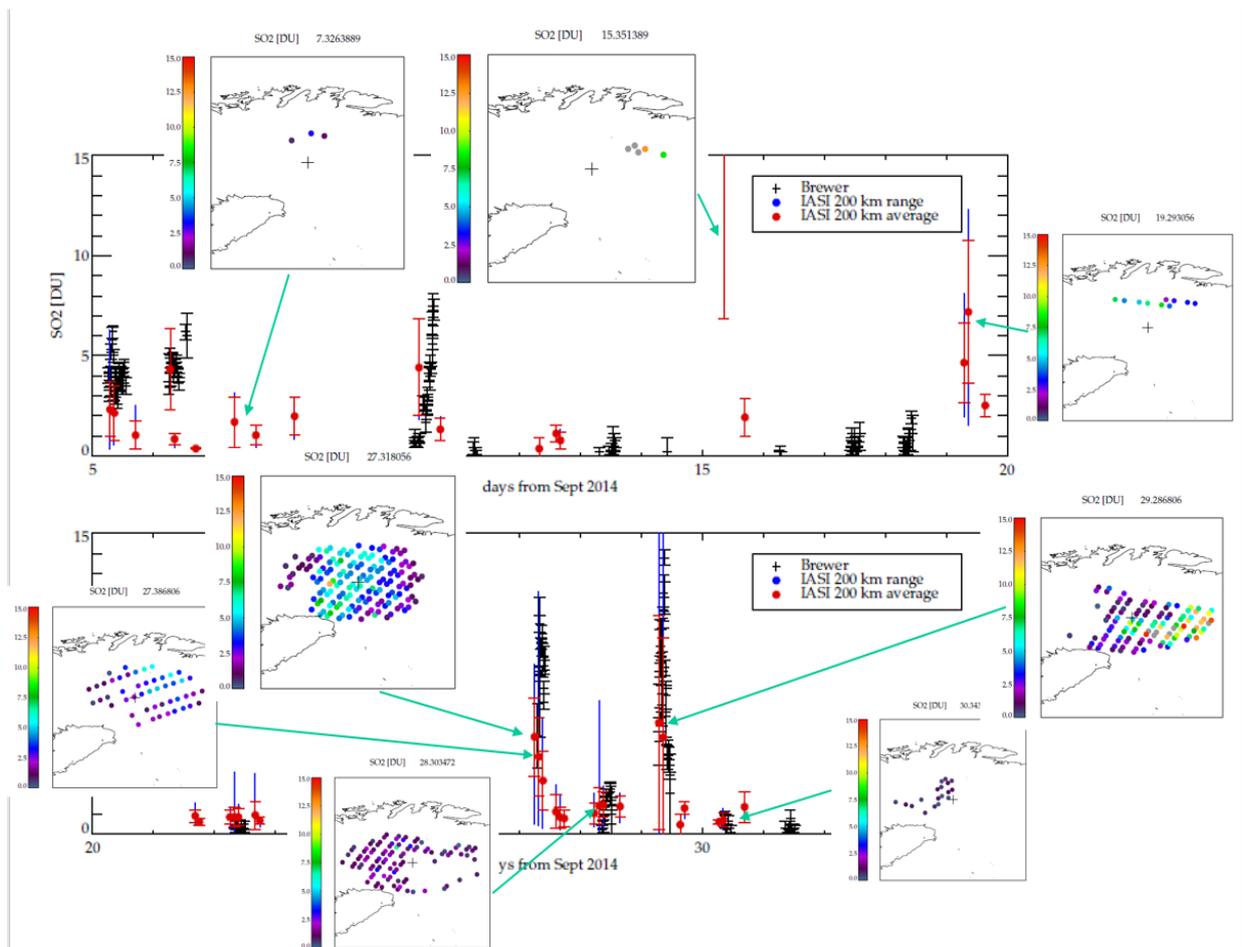
Due to variability of the volcanic plume we can have strong variation in space (in satellite maps) that are seen as variation in time (in ground measurement when the plume with different loading overpass the ground location at different time). It is common procedure to average satellite data over some distance from the ground location and to compare this with ground measurements taken during a time period. It is essentially assumed that variability in time is related to variability in space.

-Page 9, line 9: You say that for some days, SO₂ is detected by only one instrument because the limit of detection of the other instrument is not exceeded. Is it really true? Did you check that the fact that you consider a circle of 200 km radius around the ground-based station for calculating the IASI average does not play a role (i.e. IASI can detect SO₂ in a part of the circle far away from the Brewer)?

Yes we did look into everyone of the 200 km circle IASI data, and this is why we describe this in the manuscript:

'All the 'plume' episodes (with SO₂ amount larger than 2 DU) are consistent between the two datasets with the exception of 15th and 19th September where the plume only passes over the northern part of the 200 km circle in the IASI data and does not pass over the ground measurement station.'

In attach here a slide with fig 4 of the manuscript together with some of the IASI map (of column amount [DU] in colour, Day from 1st Sept of the IASI measurements in the titles) that visualized the IASI plume pixel within 200km. This shows the case of 15 and 19th September IASI overpass where the plume detect by IASI doesn't overpass the Brewer location.



-Page 10, line 2: Can you define what is IMO (indicated in Figures 2 and 3)?

IMO = Icelandic Met Office. IMO dataset were referring to the ground-based measurements described in (Pfeffer et al., 2018). We now removed 'IMO' and replaced with DOAS for the figure with fluxes and 'observation' for the altitudes.

-Page 10: Could you specify what are the errors on altitude and SO₂ fluxes calculated from ground-based measurements? How were they calculated?

Ground base measurements are described in (Pfeffer et al., 2018).

-Page 10, line 15: you say that IASI values can be underestimated. Is it because of low thermal contrast? Did you check the values?

Yes. The plume close to the surface could be underestimate due to thermal contrast.

-Page 10, lines 16-18: Since you linearly interpolated the fluxes, you overestimate the total mass when integrating below the red line (especially if the variations of IASI are real). Calculating the total mass on periods where you do not need to interpolate might improve the comparison (when the mass is compared with the one of IASI calculated for the same period).

The reviews are right but, and the interpolation (or extrapolation of fluxes where there is no measured data) is a key factor that could produce discrepancies. The problem is that we mainly need to interpolate at any time, at the best the ground measurement are performed during daylight and IASI measurements cover all the 24 h. To show the difference in time sample see fig 3 of this document that show a plot of different flux estimates for September only (the month with more ground measurements). This plot has been added to the paper (to include the comparison with Schmidt fluxes estimate)

-Page 10, conclusions: I find the conclusions a bit too short. I would say a word about some of your limitations (underestimation of the masses), about the comparison of the total mass.: : :

We added this paragraph into the conclusion:

'By comparison with OMI dataset we estimate that the SO₂ masses missed due to low thermal contrast is of the order of few percent (3%) of the total emission.

We did estimate the SO₂ mass missed, due to cloud above the SO₂ plume that is masking the signal, using AVHRR cloud CCI dataset monthly mean statistic. Results show a correction factor increasing with decreasing altitude, from one (no underestimation of SO₂ masses) up to a factor two (we measure 50% of the 'true' mass) for plume between 0-1 km. Applying this correction result in a total mass, emitted during the 6 month of eruption, of 6.7 ± 0.4 Tg and little change in the average e-folding time (2.5 ± 0.7). The IASI fluxes data reported here are representative of ~12 hours and with no data-gaps but when comparing with different source of emission estimate the interpolation (or extrapolation of fluxes) where there is no measured data, is one of the key point that could produce discrepancies.'

The following comments have been taken into account in the new version of the manuscript:

Technical comments

-Abstract, line 4: remove the comma after "data".

-Abstract, lines 9-11: The last sentence is very long and hard to read. You should rephrase it.

-Page 2, lines 2-5: I think you should rephrase the end of the paragraph, it does not read well.

-Page 2, line 11: I would replace "the first time series of the Holuhraun SO₂ plume" by "the first time series of the Holuhraun SO₂ emissions".

-Page 2, lines 16-18: I would rephrase the last sentence, it is a little bit too long.

-Page 2, line 21: add "a" before "sampling" and "almost" before "global".

-Page 2, line 26: "are estimate" ! "are estimated".

-Page 3, line 30: Replace "The SO₂ mass is highest" by "The largest SO₂ mass is found".

-Page 6, line 6: "y" should be bold.

-Page 6, line 9: matrixes ! matrices.

-Page 7, line 28: "then" ! "than".

-Page 8, line 6: Gauthier et al. (2016) is not included in the references at the end of the manuscript.

-Page 8, lines 21-24: I found this sentence very long and difficult to read. I would rephrase it.

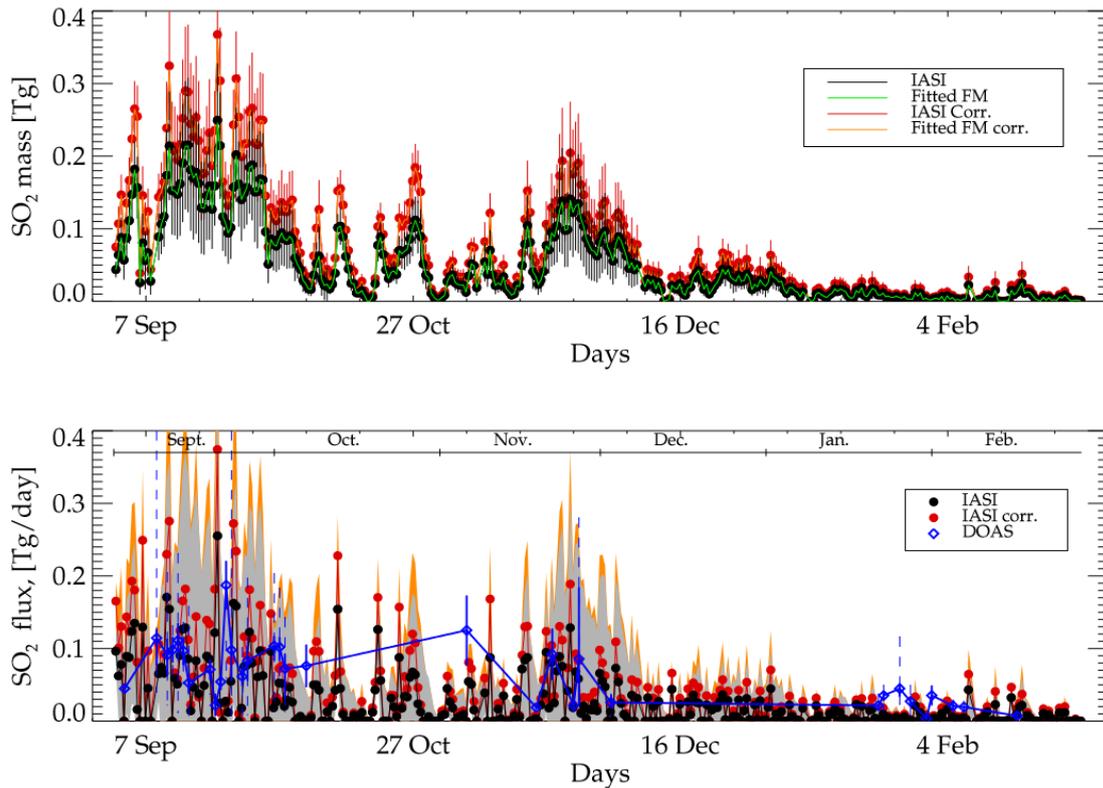
-Page 10, line 1: "collated" ! "collected"?

-Page 10, line 6: "groud" ! "ground".

-Page 10, line 15: "underestimates" ! "underestimated".

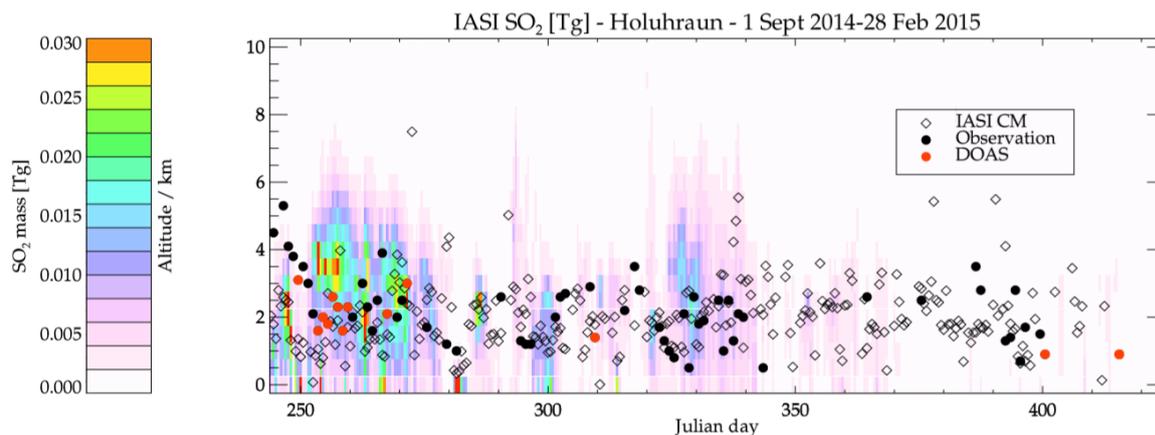
-Figure 2, top: The blue line is difficult to distinguish from the black dots. Maybe change the colour?

Blue changed with light green. Following addition of estimate of missing SO₂ we now added both estimate of total masses and fluxes with uncorrected and corrected data. New fig 2 below.



-Figure 3: In the text, you say that some of the ground based measurements provide the altitude of the plume center-of-mass, and the others the altitude of the top of the plume. Maybe you could make the distinction between the two cases in the Figure (circle and triangle, or something else). It would be easier to see which ground-based observations provide the same information than IASI.

Fig 3 in the manuscript has been change with this one where IMO data with DOAS are plotted in red.



References

- Bauduin et al., J. Geophys. Res. Atmos. 119, 4253-4263, 2014, doi:10.1002/2013JD021405;
 Bauduin et al., Atmos. Meas. Tech. 9, 721-740, 2016, doi :10.5194/amt-9-721-2016;
 Boynard et al., Geophys. Res. Lett. 41, 645-651, 2014, doi :10.1002/2013GL058333;
 Fioletov et al., Atmos. Chem. Phys. 16, 11497-11519, 2016, doi :10.5194/acp-16-11497-2016;
 Thordarson and Hartley et al. Geophys. Res. Abstracts, 17 (EGU2015-10708).

Anonymous Referee #2

This paper developed a new scheme to calculate daily SO₂ fluxes and average e-folding time for volcanic SO₂ emissions in Iceland. In order to overcome the difficulties in latitude and time, the authors propose to use satellite-based thermal infrared spectrometers instead of UV bands to study the volcanic SO₂. The results look sound and interesting. I recommend publishing the paper after addressing the comments below.

General comments:

1. Page 3, line 18. In this study all the SO₂ measured from 30N to 90N between September 2014 and February 2015 is referred to as Holuhraun SO₂. What is the uncertainty of this assumption?

We estimate the atmospheric loading of non-Holuhraun source as no higher than 0.01 Tg. This estimate comes from the SO₂ total mass loading during the second half of February where there is no presence of plume from Iceland. The SO₂ in these two weeks mainly comes from China and from some volcanic activities in Kamchatka.

2. This paper is based on the previous work performed by the same author. I understand the authors would like to keep the text simple and avoid repeating contents mentioned by their previous work. However, sometime the text seems to be too brief to keep all important information. For example, Page 3, line 27-28. "regridding the observations of column amount and plume altitude into a 0.125 latitude/longitude boxes following Carboni et al. (2016)." What is special of the regridding approach in Carboni et al. (2016)? I have the similar concern for Section 2.

Section 2 with the algorithm description have been expanded including more detailed explanation of the regridding

Specific comments:

1. Page 2, line 22. The exact location of the IASI data should be added.

Added.

2. Page 2, line 30. Putting a rough quantification of the uncertainty of the “minimum” here would be appreciated.

We added into the manuscript an estimate of the underestimation due to cloud cover and thermal contrast, using respectively AVHRR Cloud CCI dataset and OMI SO₂ dataset, See answer to review1.

3. Page 6, Line 10. The a priori values used were 0.2 ± 0.2 Tg/day for flux and 2 ± 2 day for the e-folding time. Is there any sources for the priori values? If not, will the fitting results be sensitive to the choices of the priori values?

As for answer to review 2:

‘The a priori values used were 0.2±0.2 Tg/day for flux and 2±2 day for the e-folding time.’

For the flux we choose 0.2± 0.2 Tg a day for a priori flux because this correspond to allow the retrieval to move easily (if there is enough information in the measurements) between 0 and 0.4 Tg/day and 0.4 is greater than the maximum values of total masses. Only few values of total masses in September exceed 0.2 Tg and in particular none of them show total mass greater the 0.3.

We think that the a priori e-folding time variation between 0 and 4 will cover the real lifetime of SO₂ for low tropospheric plume, we experiment assuming longer e-folding time to the analysis and the resulting fitting is worse. Assuming shorter e-folding time instead result in good fit as well as the one presented in the manuscript, this is why we have written this at line 28 pag 7: ‘Also note that any e-folding time shorter than the retrieved one can fit the measurements and give higher fluxes.’

4. Figure 2. The color of blue is difficult to see.

Changed with light green. See new plot in answer to referee 1.