Reply to Anonymous Reviewer #1

We highlight the Reviewer comments in red (R) and our answers in black (A).

General comments:

R1. Does the paper address relevant scientific questions within the scope of SE? The topic focus of the paper lies well within the scope of SE, as it explores new techniques for the detection of volatile emissions from volcanoes and the implications for subsurface magmatic processes.

A1. Thanks for this comment.

R2. Does the paper present novel concepts, ideas, tools, or data? The methodology presented in the paper is novel, and offers a useful approach for maximising the data yield from satellite images of SO_2 detection. It offers a significant advance to the subject that I have not seen before in the literature.

A2. Thanks for this comment.

R3. Are substantial conclusions reached? The conclusions are physically reasonable, and well-justified. Although the conclusion itself not particularly novel (as the presence of pre-eruptive vapour phase is now a well-established concept), the method by which this conclusion was reached is of substantial value.

A3. We agree with the reviewer. In general, the presence of pre-eruptive volatile phase is now a well-established concept, but this is important for the Calbuco eruption. Our data provide us new insights to better understand which conditions and processes have led the triggering of the two sub-Plinian eruptions.

R4. Are the scientific methods and assumptions valid and clearly outlined? The methods and assumptions related to the trajectory modelling and numerical calculations are very clearly explained. However, discussion of the sources of uncertainty related to the petrological technique are lacking; the overall discussion of the petrological methodology is unsatisfactorily brief. For example, are any post-entrapment crystallisation corrections applied to MI compositions? Uncertainties related to the acquisition of the initial SO_2 satellite images is also lacking.

A4. We revised the text according to this comment. In particular, the petrological methodology has been better explained and inserted in the main text as stand-alone section (Section 3.2 – Petrological Method to Estimate Sulfur in Magma). Also, the application of this method for the study of the Calbuco eruptions is now better described in Section 4.3 – Petrological Results. We also added a more detailed explanation of the uncertainties related to satellite SO_2 acquisition in Section 4.1, adding the following text:

"The Differential Optical Absorption Spectroscopy (DOAS) (Platt and Stutz, 2008) technique is applied to retrieve SO_2 vertical column amount by measuring the portion of the sunlight backscattered in the

atmosphere. For GOME-2 retrievals, the overall error in SO_2 vertical column estimates is in the range 20-70% (Rix et al., 2012). This range of uncertainty accounts for both random and systematic errors. Random errors are mainly due to instrument noise and they are typically of 5-20%. The main contribution to systematic errors comes from the difficulty in assessing the plume height at measurement time and it is estimated to be in the range 10-60% (Rix et al., 2012). Plume height is a central parameter when converting SO_2 slant column density (i.e. the gas concentration along the entire light path) into vertical column density (i.e. the gas concentration right above the satellite footprint). When using an SO_2 retrieval done assuming the plume located at a certain height, errors up to 50% on vertical column amount can rise if the actual plume height is not the one used for the SO_2 retrieval. In order to deal with the missing information on plume height at measurement time, for GOME-2 retrievals, three different SO_2 estimates are given for three hypothetical plume altitudes equal to 2.5 km, 6 km, 15 km. "

R5. Are the results sufficient to support the interpretations and conclusions? The results are convincing, and are clearly in line with those of other independent studies. Although this is unlikely to change the overall conclusion, I would like to see an expanded discussion of the petrological analyses (e.g., the overall variability in both MI and glass sulfur concentrations, relationship to major element compositions, potential for postentrapment crystallisation of MIs, potential to volatile losses from compromised MIs etc.) which would provide stronger support to the quantification of the 'excess' sulfur.

A5. Thanks for the suggestion; the petrological method has been revised as shown in comment A4.

R6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? The description of the modelling and numerical calculations were very clear, and well-illustrated. However, this is largely on a qualitative (at best semi-quantitative) level and the analysis could not be fully reproduced from this description. The authors might consider publishing their algorithm as a supplement to the paper, if indeed their intention is for this to be a useful tool to the wider community.

A6. Our plan is to publish our numerical tool. However, for the present paper we preferred to focus on the application of the numerical method to a recent eruption. The data can be shared under request.

R7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution? Yes

A7. Thanks.

R8. Does the title clearly reflect the contents of the paper? Yes, the title is relevant to the subject matter.

A8. Thanks.

R9. Does the abstract provide a concise and complete summary? The abstract is well written and provides a good summary of the paper. However, it is too long and could be condensed more effectively to have more of an impact: a brief summary abstract is more informative to the reader than a lengthy detailed abstract that repeats sections of the introduction and conclusions.

A9. The abstract has been revised according to this suggestion.

R10. Is the overall presentation well-structured and clear? Is the language fluent and precise? Are mathematical formulae, symbols, abbreviations, and units correctly defined and used? Yes, the paper is well-structured and well-written overall and the organisation of ideas flows well. There are a couple of typo errors in the units, which I have highlighted on the line by line comments below.

A10. Thanks, we corrected them.

R11. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? The figures are very good. I particularly like how the data are presented in figure 8. Panels a,b,c,d need to be labelled in figure 6. I feel that the petrological methods should be described up front in the methods section, rather than being in the supplementary information. The petrological data are central to the discussion of the magmatic processes involved, and so should be set up in the main paper.

A11. We revised the paper according to this suggestion. The petrological method is now fully inserted and described in the main text as shown in comment A4.

R12. Are the number and quality of references appropriate? Mostly yes. There are a few additional papers that are relevant and I would suggest considering, particularly those relating to the fidelity of melts inclusions as a volatile record, for example:

Wallace, P.J. and Edmonds, M., 2011. The sulfur budget in magmas: evidence from melt inclusions, submarine glasses, and volcanic gas emissions. Reviews in Mineralogy and Geochemistry, 73(1), pp.215-246.

Andres, R.J., Rose, W.I., Kyle, P.R., DeSilva, S., Francis, P., Gardeweg, M. and Roa, H.M., 1991. Excessive sulfur dioxide emissions from Chilean volcanoes. Journal of Volcanology and Geothermal Research, 46(3-4), pp.323-329.

Wallace, P.J., 2005. Volatiles in subduction zone magmas: concentrations and fluxes based on melt inclusion and volcanic gas data. Journal of Volcanology and Geothermal Research, 140(1), pp.217-240.

Danyushevsky, L.V., McNeill, A.W. and Sobolev, A.V., 2002. Experimental and petrological studies of melt inclusions in phenocrysts from mantle-derived magmas: an overview of techniques, advantages and complications. Chemical Geology, 183(1), pp.5-24.

A12. All the references have been added.

R13. Is the amount and quality of supplementary material appropriate? Yes the SI is fine and provides useful extra detail. However, the authors should consider including the description of the petrological methods up front in the main methods section of the paper.

A13. This has been done as shown in comment A4.

Line by line comments:

R14. P2 Line 11: replace 'which' with 'that'

A14. Done

R15. P2 Line 16: replace Westrich et al., 1992 with Westrich and Gerlach, 1992 (also consider citing some additional references here)

A15. The reference has been corrected and three more have been added:

Campion, R. (2014), New lava lake at Nyamuragira volcano revealed by combined ASTER and OMI SO2 measurements, Geophys. Res. Lett., 41, 7485–7492, doi:10.1002/2014GL061808.

Carn, S. A., Krueger, A. J., Arellano, S., Krotkov, N. A. and Yang, K.: Daily monitoring of Ecuadorian volcanic degassing from space, Journal of Volcanology and Geothermal Research, 176(1), 141–150, doi:10.1016/j.jvolgeores.2008.01.029, 2008.

Carn, S. A., and F. J. Prata (2010), Satellite-based constraints on explosive SO2 release from Soufrière Hills Volcano, Montserrat, Geophys. Res. Lett., 37, L00E22, doi:10.1029/2010GL044971.

R16. P2 Line 32: Should 'box method' be 'delta method' on this line?

A16. It is, thanks.

R17. P3 Line 20: Consider adding some additional explanation to this statement such as '. . .allow us to infer the presence of excess SO_2 at depth before the eruption, when combined with petrological and deposit volume constraints.'

A17. The sentence has been changed in:

''Our retrieved SO_2 injection height and flux time-series, together with estimates of masses of erupted material, allow us to infer the presence of excess SO_2 at depth before the eruptions.''

R18. P4 line 4: 'appeared to be more violent' – in what way? Considering adding additional description.

A18. The sentence has been changed in:

"The intensity of the second eruption appeared to be higher than the first one and with a ~50 % greater mass eruption rate (Van Eaton et al., 2016)."

R19. P4 line 13, 14: Why were such deposit densities used by the cited studies? (1000 kg m-3 vs. 2500 kg m-3)

A19. Because this density has been measured by the cited studies. However, we re-phrased the sentence in order to clarify it. In the paper by Romero et al., [2016], the authors evaluate a density of the deposit equal to 997.3 kg m^{-3}, while they consider a magma density of $\sim$ 2500 kg m^{-3}. In the paper by Castruccio et al.,[2016] a density of 1000 kg m^{-3} is calculated for the deposit. The sentence is now:

"Considering both the tephra fall and PDC deposits, the deposit volume estimated by Castruccio et al., (2016) is 0.38 km^3 assuming a deposit density of 1000 kg m^{-3} (0.15 km^3 dense rock equivalent DRE), while Romero et al., (2016] report a tephra fall deposit volume of 0.28 km^3 considering a deposit density of 997.3 kg m^{-3} (0.11-0.13 km^3 DRE). These values are both within the 0.56$\pm$0.28 km^3 volume calculated by Van Eaton et al., (2016), which presents a DRE of 0.18$\pm$0.09 km^3 assuming a magma density of 2500 kg m^{-3}."

R20. P4 line 14: units typo – replace '2450 km m-3' with '2450 kg m-3'

A20. This value is not present anymore in the text.

R21. P6 line 12-14: The justification for using the 2.5 km plume is not clear to me, please add some additional explanation or re-phrase. How is overestimation of the SO$_2$ plume a good thing?

A21. This point has been clarified and re-phrased as follows:

"As previously discussed, plume altitude is one of the main parameters influencing the retrieval of SO$_2$ vertical column amount. We use as an input image for our numerical procedure the SO$_2$ column amount image calculated assuming a plume height of 2.5 km. This means that the maximum accuracy in SO$_2$ estimation is achieved if the actual plume is located at 2.5 km. In case of different plume height (higher than 2.5 km), SO$_2$ column amount for a single pixel can be overestimated up to 70-80% (Carn et al., 2013). On the contrary, when using the 6 and 15 km retrievals for a plume which is actually located at lower heights, the SO$_2$ column amount results to be underestimated and thus information on plume spatial distribution can be lost. Since we do not want to make assumptions on both plume height and SO$_2$ spatial distribution, we use as input data for our numerical model the SO$_2$ image at 2.5 km. In this way we operate our numerical procedure on a plume which is eventually broader than the actual one and we let the model retrieve the actual plume spatial distribution in term of SO$_2$ vertical column corrected for plume height."

R22. P7 line 3, 4: 'This is due to several uncertainties given by wind data, trajectory calculation and SO$_2$ spatial distribution' – give more specific details of what exactly these uncertainty sources are, and how significant.

A22. This sentence has been added to the main text:

"An overall error in the range of 15-30% of the travel distance can be estimated for the trajectories computation. This error is due to computational inaccuracy, interpolation errors, starting position errors and wind field errors (Stohl 1998)."

R23. P8 Line 19: should Figure 5(d) be Figure 5(b)? (there are only two panels in figure 5. ..)

A23. Figure 5 has been modified as shown in the text. The old one has been replaced with a new one showing the comparison between the Calbuco plume as seen by GOME-2 and as retrieved with our model.

R24. P9 line 18, 19: Uncertainties should be given along with calculated MER values.

A24. We added MER and uncertainties on masses of erupted material and volume estimates. The new values are presented in Section 4.2.1. All the calculations about the SO$_2$ loading and the "excess" SO$_2$, have been revised considering these new values. According to the new calculations, the paragraph has been re-phrased as:

"We use our mean injection height time-series to calculate a mean mass eruption rate (MER) and we use it to evaluate the mass of erupted solid material. From this calculation we compute a mean MER of $1.14\pm0.42\cdot10^7$ kg s^{-1} for Eruption 1 and of $1.09\pm0.38\cdot10^7$ kg s^{-1} for Eruption 2 and we infer $6.2\pm2.2\cdot10^4$ kt emitted during Eruption 1 and $24\pm8.2\cdot10^4$ kt during Eruption 2."

R25. P9 Line 21: should 'first or the four layers' be 'first of the four layers'?

A25. Yes, thanks.

R26. P9 Line 23: should 'despite the two authors agree' be 'despite this, the two authors agree'

A26. It is, thanks.

R27. P9 Line 23: What is the basis for the disagreement between the origin of layer 2? Why is it ambiguous? How much does this affect you conclusion if layer 2 is attributed instead to eruption 1?

A27. This disagreement comes from the different way the two authors, Romero and Castruccio, interpreted the stratigraphy of the deposit. However, this has no implications on our numerical results since our MER and masses estimates come purely from space and from our retrieved injection height time series. According to the cited studies, the deposit present four layers: A, B, C and D. The attribution of layer B to the first or the second eruption can have an impact on the results of the petrological analysis. This is because it can result in a different SO$_2$ loading calculated for the two eruptions. However, even if Layer B (layer 2) is attributed to Eruption 1 (as in Romero et al., [2016]), no significant changes in our final results are obtained. This is

due to the fact that the sulfur concentration measured in the samples of layers A, B and C do no present significant differences. Thus, attributing layer B to the first or to the second eruption does not make any difference in terms of our results. To show this, we report in Table 1 the results for Eruption 1 considering both the interpretations, while the same has been done for Eruption 2 as shown in Table2.

	S_{MI}	S_{gm}	$m(SO_2)_{PETR}$	$m(SO_2)_{ex}$
E 1: A	0.035±0.01wt%	0.009±0.0006wt%	16±7 kt	144±26 kt
E 1: A+B	0.035±0.008wt%	0.011±0.002wt%	15±6.5 kt	145±25.5 kt

Table1: the first line shows results as presented in the paper attributing layer A to Eruption 1 as in Castruccio et al.,[2016]. In the second line we report the same calculations done attributing layer A and B to Eruption 1 as in Romero et al., [2016].

	S_{MI}	S_{gm}	$m(SO_2)_{PETR}$	$m(SO_2)_{ex}$
E2: B+C+D	0.04±0.007wt%	0.01±0.003wt%	71±26 kt	69±39 kt
E2: C+D	0.045±0.009wt%	0.008±0.004wt%	79±19 kt	61±40 kt

Table2: the first line shows results as presented in the paper attributing layer B,C and D to Eruption 1 as in Castruccio et al.,[2016]. In the second line we report the same calculations done attributing layer C and D to Eruption 2 as in Romero et al., [2016].

R28. P9 line 32: 'well-correlated' – looking at figure 8 I would say this is a slight overstatement. I agree there is a correlation, but there is still quite a bit of scatter in the data.

A28. We corrected the sentence with:

"SO₂ flux appears to be correlated with mean injection heights for both eruptions (Figure 8(b))."

R29. P10 line 23: 'bubbles migrated to the top of the. . .'

A29. This sentence is not present anymore in the text.

R30. P10 line 29: methodology should not be in the SI (see earlier comments) Also, just a consideration - potentially pyrite is not the optimal standard for S in this case. I anticipate that the Calbuco magma is quite oxidising, such that much of the S will be in the S6+ phase. Use

A30. Yes, we agree with the reviewer that Barite may have been another good standard and probably a more appropriate standard in this case. However, at the beginning of our analyses we tested also Barite as a standard and we compared the results obtained with Pyrite. At the end the results were really similar. Furthermore, we used also a secondary standard (Standard Glass Basalt (A-99)) to have a double check on the quality of the analyses. Analyses on the Standard Glass Basalt (A-99) are reported below:

Oxide (wt.%)	Standard Galss Basalt (A-99)	Oxide (wt.%)	Analysis_1	Analysis_2	Analysis_3	Analysis_4	Analysis_5
		Microprobe analysis - Jeol JXA 8530F					
SiO2	50.94	SiO2	50.92	50.88	50.76	51.01	50.89
TiO2	4.06	TiO2	4.01	4.02	4.16	4.09	3.92
Al2O3	12.49	Al2O3	12.6	12.53	12.55	12.46	12.42
Fe2O3	1.87	Fe2O3	1.79	1.81	1.93	1.85	1.94
FeO	11.62	FeO	11.66	11.64	11.69	11.71	11.58
MnO	0.15	MnO	0.16	0.15	0.14	0.16	0.17
MgO	5.08	MgO	5.11	5.15	5.09	4.99	5.04
CaO	9.3	CaO	9.41	9.39	9.26	9.73	9.38
Na2O	2.66	Na2O	2.58	2.69	2.71	2.66	2.79
K2O	0.82	K2O	0.81	0.79	0.88	0.78	0.86
P2O5	0.38	P2O5	0.34	0.37	0.38	0.32	0.39
S	0.0143 ±0.001	S	0.0138	0.0155	0.0148	0.0135	0.0164
Total	99.38	Total	99.40	99.47	99.56	99.78	99.40

The results obtained analysing Basalt (A-99) support that Pyrite has been a reliable standard, therefore, we may have only slightly underestimated the total dissolved S in the glass. More comments on the uncertainty of these measurements are reported in the new version of the manuscript (Section 3.2 and 4.3).

R31. P10 line 32: Have you considered that MI hosted in plagioclase may not represent that initial S concentration in the melt? Late crystallisation of plagioclase may yield MIs of slightly more evolved composition. Have you performed any post-entrapment crystallisation corrections on your MI compositions? Some discussion of the major element systematics of your MI and matrix glass data would help to shed light on these points.

A31. We agree with the reviewer. MI hosted in plagioclase may not represent the initial S concentration in the melt before the crystallization of the magma. The bulk composition of Calbuco rocks is basaltic-andesite, whereas MIs have andesitic compositions. We think that fractional crystallization produced exsolution of volatiles and an evolution of the melt (the composition of MIs is the evidence). This means that the total initial S estimated may be a little bit higher because a part of it may have been exsolved during the initial crystallization in the reservoir (crystallization of olivine, pyroxene and plagioclase). It was difficult to analyse MIs in olivine, orthopyroxene and clinopyroxene because they usually are smaller than 30 μm in these samples. These minerals may have yielded MIs of slightly less evolved composition.

Post-entrapment crystallization is evident in some MIs, but we have analysed only glassy MIs avoiding crystallized MIs. We are preparing another paper to study in detail the triggering mechanism of the 2015 calbuco sub-Plinian eruption, where we have also analysed MIs in detail with the Raman spectrometer. This analysis shows that MIs chosen are crystal-free glasses. Furthermore, MIs and the residual melt have similar K_2O concentration, whereas S concentration decreases in the residual melt indicating that degassing and sulfide exsolution occurred in syn-eruptive conditions (see Figure9(a)) Whereas, MIs have a higher MgO than matrix glass, this is due to the crystallization of microlites of orthopyroxene in the groundmass rather than post-entrapment crystallization in MIs (see Figure 9(b)). Therefore, we have not performed any post-entrapment crystallisation corrections on our MI compositions. These discussions are reported in the new version of the manuscript.

R32. P11 line 3: Do the errors on your petrological S yields include the errors attached to the deposit volumes?

A32. Yes. We have considered errors in the calculation in order to estimate the potential minimum and maximum value of sulfur emitted during the first and second pulse of the eruption.

R33. P11 line 4: 'it is'

A33. Corrected

R34. Table 1: Matrix glass S values are close to the limit of detection, I am surprised to see such low uncertainties on these measurements.

A34. Yes, the reviewer is right. The uncertainty of some matrix glass S values is higher than that reported in the previous Supplementary Table. It was just an error in our excel spreadsheet for the calculation of 2σ. Now the right values are reported in the new Supplementary Table 1.

R35. '0.8 is a coefficient accounting for 20 vol% of crystallisation' – this is thrown in here as a footnote without any mention or explanation in the text. Please clarify this in the main paper.

A35. This statement has been clarified in the main text as:

"Thus, according to the petrological method, the SO_2 loading released into the atmosphere during an individual eruption ($m(SO_2)_{PETR}$) can be expressed as:

$$m(SO_2)_{PETR} = M \cdot \left(S_{MI} - S_{gm}\right) \cdot \frac{MW(SO_2)}{MW(S)} \cdot (1 - CF), \tag{8}$$

where M is the mass of erupted material, S_{MI} and S_{gm} are the sulfur concentrations measured in melt inclusions and glassy matrix, $MW(SO_2)$ and $MW(S)$ are the molecular weights of SO_2 and S and CF is a coefficient accounting for the volume of syn- or post-eruptive crystals in the melt."

And:

''In order to evaluate the atmospheric SO$_2$ yield as shown in Eq. (8), we consider a crystal fraction of 50 vol% (i.e. CF equal to 0.5 (Arzilli et al., (2017)).''