Response to reviewer comments

J. Vira and co-authors

We thank both referees for the extensive reviews. In the following we will address the comments of reviewer #1 and the specific comments of reviewer #2. The general comment of reviewer #2 has been addressed in our response on 13 December 2016.

The main changes in the revised manuscript is extending experiments with synthetic data in Section 4 to cover assimilation of simulated plume height retrievals, and presenting a comparison between the simulated SO2 profiles and the lidar data from the CALIOP instrument. These major revisions are discussed in more detail in our response to reviewer #1.

Due to the volume of new material, we have removed the discussions of two minor items present in the original submission. The omissions are Figure 5 and the related discussion (additional experiments with randomly generated source terms), and Figure 10 and the related discussion (inversion with a simplified observation error covariance matrix).

The figures for Eyjafjallajökull have been rearranged as part of the comparisons for total columns are now presented together with the CALIOP data. Figures comparing the simulated total columns and plume heights with the IASI data for all days are included in a supplement.

The manuscript with changes highlighted is attached to this pdf.

The reviewer comments are presented with a blue italic font.

Response to comments by reviewer #1

We thank the referee for a thorough and detailed review. In order to address the main concerns raised by the referee, we have introduced the following major revisions to the manuscript:

1. The experiments with synthetic observations are extended to cover the case with assimilation of simulated plume height retrievals. The results are consistent with those obtained for the Eyjafjallajökull eruption. Assimilation of plume height retrievals results in a more accurate source term (in the sense of RMS error), although the difference is small for some of the assumed source terms. However, the experiment also confirms that assimilation of plume heights may have negative impact on the estimated total emission.
2. The results for Eyjafjallajökull are evaluated using profile data from the CALIOP instrument, and a detailed comparisons between the estimated injection height, the retrieved SO2 plume heights, and the radar observations of plume top are presented. The comparisons show consistently that the results with and without assimilation of plume heights are largely similar but the vertical distribution of the emissions especially on 6-9 May is improved by assimilation of plume height retrievals. For the other times, the injection height is mostly consistent with the data regardless whether the plume heights are assimilated or not.

Ideally the results for total columns should be compared with independent satellite data. However, as noted in earlier studies about Eyjafjallajökull (Boichu et al., 2013; Flemming and Inness, 2013), this is unlikely to lead into useful conclusions due to the large differences between satellite products. For this reason, we have instead focused on the vertical distribution which we consider to be the most interesting of aspect of the present study.

In the current version of the manuscript, it remains unclear what the benefit of the assimilation of the plume height is. Given the overall uncertainty and judging from the pictures, it seems that the plume height assimilation does not lead to an improvement. If this is the case it should be mentioned more clearly.

We believe that the revised manuscript gives a more accurate and detailed picture of impact of assimilating the plume height. However, we hesitate to declare one inversion better than other, because drawing such a conclusion would require three-dimensional observational data that do not exist. The possible benefits of assimilating the plume height in addition to total column cannot be evaluated based on their impact on the simulated total columns or the total emission, because these quantities are fit optimally when only total column data are assimilated.

Based on the experiments with synthetic data and the comparison to vertically resolved observations for Eyjafjallajökull, it seems reasonable to conclude that assimilating the plume height retrieval has a positive impact on the vertical distribution, even if the effect is small when taken over the whole eruption. Similarly, the experiments indicate that the impact on total columns and total emission is negative. Whether one of these effects outweighs the other depends on the application.

The presented 4D-VAR approach does not take into account correctly the error statistics of the assimilating model as no model error co-variances are considered. This seems a simplification which should be better justified.

We agree that the treatment of model errors is not ideal, however, as pointed out in the initial response to reviewer #2, the problem is challenging and we are not aware of any inversion study that would have addressed this aspect without simplifications.

Inclusion of the model errors into the R-matrix is necessary because the inversion is performed in a single, long assimilation window. As discussed in the revised manuscript, it would be in principle possible to set up the inversion as a sequence of shorter assimilation windows as done by Elbern et al. (2007) in context of air quality forecasting. However, when the primary interest is to estimate an emission source, this approach becomes more complicated because of two opposing requirements. The first requirement is that the assimilation window needs to be short enough so that the model errors arising within the assimilation window are not significant – this is the basic assumption of the currently used strong-constraint assimilation methods. The second requirement is that the assimilation window should be long enough to effectively constrain the emission. Shorter assimilation windows may be useful for constraining constant or slowly varying emissions, but such assumptions would not hold for volcanic emissions.

So far, the majority of studies on inverse modelling of atmospheric emissions have emphasized the second aspect and used long assimilation windows. This includes the inversions for volcanic emissions (Boichu et al., 2013; Kristiansen et al., 2010; Lu et al., 2016; Stohl et al., 2011), where the inversion is done for the entire eruption at once, but also inversions for other trace gas emissions (e.g. Meirink et al. 2008; Müller and Stavrakou, 2005) or for inverse modelling of accidental radioactive releases (e.g. Winiarek et al., 2014). An advantage of the long assimilation window without explicit model errors is that the no unphysical sources or sinks appear due to the assimilation, and the mass budget of the simulation remains closed. While this does not imply that the long-window approach is superior to the other options, a comparison of its advantages and disadvantages is outside the scope of the current paper.

Also, it seems that the SILAM model did not consider a chemical loss for SO2, which – if this was the case – would be an unnecessary simplification of the model.

As noted on L76 in Section 2.1 (Dispersion model), the model does include chemical removal of SO2.

L 17: Clarify "vertical centre of mass" and "first moment"

Done.

Done.

A discussion has been added onwards of L65.

Additional discussion has been added on lines 70-76.

"The algorithm and the dataset are explained in more detail by Carboni et al. (2012)" has been changed into "The algorithm and the IASI SO2 dataset (column amount and altitude) are explained in more detail by Carboni et al. (2012)".

The effective altitude is the altitude of the mean of the Gaussian distribution that fits the measurements. We called it "effective" because we assume one SO2 layer with a Gaussian profile. To clarify we have changed the text: "The scheme determines the column amount and effective altitude of the SO2 plume" is changed into: "The scheme determines the column amount and the altitude (mean of Gaussian distribution) of the SO2 plume".

This is the altitude of the SO2 plume, not the ash plume; to avoid misunderstanding we have changed the text:

"The altitude retrieved for the Eyjafjallajökull eruption plume"

into

"The SO2 altitude retrieved for the Eyjafjallajökull eruption".

The present study is best described as a (source term) inversion similarly to the studies cited in our response to the comment regarding the 4D-Var method. We nevertheless use the word "assimilate" in reference to the action of an observation operator (as in "to assimilate plume height retrievals"), since this is both intuitive and consistent with existing literature on inverse modelling.

Done.

Yes, as indicated above. A remark has been added also here.

L 140: Please discuss that the standard 4D-VAR approach would include the model error in form of the background error covariance matrix.

We have included a discussion on model errors and possible ways to handle them.

L 146: Spell out L-BFGS-B

Done.

L 156: Define also y and m_ij

Done.

L 160: Provide the formulae for centre of mass and 1st moment of mass

Done.

L 175: It is not clear if R also contains the error of the plume height retrieval. Is this error also provided and used?

The plume height retrieval error is provided and used. The corresponding text has been added.

L 204: This is not surprising given that the model/background error is ignored in equation (2)

L 209: Having only a diagonal model error covariance ignores the fact that the model advects the tracer.

We refer to our response above regarding the 4D-Var formulation.

L 219: Provide reference or explanation for the Tikhonov regularisation

References have been added.

L 252: Please describe the setting of the synthetic experiment better. What was period, region and meteorological data? Is only the emission term synthetic? What are the synthetic observations?

The section has been rewritten for more clarity. With synthetic observations we refer to simulated observations and observation errors. In addition, we have simplified the section by omitting the additional experiment (Fig. 5) with randomly generated source terms.

L 315: What is meant by "overall need"? Why is there an assumption about the source term? I thought regularizations is introduced to avoid an prior assumption for the source term? From which of the above are these generalisations deduced?

By the need for regularisation we meant that the sensitivity to under-regularisation varied between the assumed sources. However, we agree that this was formulated unclearly, and the paragraph has been rephrased.

L 316: In which of the experiments there was no model error? Please clarify. Perhaps a table of all the synthetic experiments would be useful.

We have clarified that model error was present in all experiments except for creating Fig. 2.

L 320: Say exactly which type of experiment you refer to

Done.

The iterates were chosen according to the algorithm described on lines 364-366 in the revised manuscript. We also examined the L-curves visually but kept the points chosen by the algorithm. The L-curves are provided in supplementary information.

As stated in Arason et al. (2011), the radar time series actually represent the upper border of the ash plume. While we agree that comparing the radar data with the inversion results is somewhat ambiguous, we have decided to keep the comparison, since the radar data provide an independent dataset evaluating the estimated injection height. However, we have also added a chart showing the comparison of the emission centre of mass with plume height retrievals averaged within 50 and 500 km radiuses from the volcano. We hope that Fig. 8 also exposes the differences between the inversion results in Fig. 7.

Please see our response to the general remark.

We have added (L562 onwards) a discussion regarding effects of the plume height assimilation on the total emission.

The total emission changes because when only total columns are assimilated, the inversion has more degrees of freedom to match the observed horizontal distribution. When a vertical constraint is added, the simulated total columns must change unless the model describes the real transport perfectly, which is not a realistic assumption. To give a concrete example, in Fig. 13 in the revised manuscript, the TC-only run appears to fit better the total column data between 50° and 54° N, shown in the lower panel with the line plots. However, the comparison with IASI and CALIOP data shows that this corresponds to an unrealistic vertical distribution. As a consequence, the TC+CM run agrees worse with the total column, but is quite likely to agree better with the true, three-dimensional distribution.

The synthetic experiments indicate that the TC+CM inversions tend to have a low bias for total mass, which is consistent with the results for Eyjafjallajökull. The likely reason is the regularisation, which among different emissions with similar likelihood prefers the one with lowest squared sum. Whether a different cost function would avoid the negative effect on total emission is a question for a future study.

Due to the volume of new material added, and since this aspect is peripheral to the main discussion, we have omitted Fig. 10 and the associated text from the revised manuscript.

L 349: In line 337 you say it is about 10%. Why is this the case? Were the observations errors different for the assimilated TC in the two cases? The TC and the plume height retrievals are not independent

We rewritten the text for better self-consistency. The same observation errors were used in both experiments. The correlation between TC and plume height retrieval errors is taken into account as indicated now more clearly in Section 3.2.

L 350: Please provide more evidence and discussion.

More evidence (especially Fig. 8) and discussion has been added.

L 355: I find it actually quite interesting that the additional assimilation of the plume height has so little influence.

L 357: Again, this is no proof that one inversion is better

A more detailed discussion about the changes and similarities between the results has been added.

L 366: The differences between GOME-2 and IASI plume height retrievals should be discussed in more detail. The explanation is too short.

We mention the GOME-2 retrievals, since they are one of the few published observations on SO2 plume height. As indicated, neither our data nor inverse modelling reproduces the GOME-2 plume height, but investigating the reasons for this would require a detailed comparison of the satellite products and the retrieval algorithms.

L 366 and L 370: Compare your results for both the injection height and the SO2 mass with the results published in the literature, for example Boichu et al. 2013, Flemming and Inness, 2013. This would be an important result and should be mentioned in the conclusion or abstract.

Comparisons with Boichu et al. (2013), Flemming and Inness (2013) and Stohl et al. (2011) have been added.

L 376: Which other studies? Provide references.

References (Boichu et al., 2013; Seibert et al., 2011) have been added.

L 380: This is all a bit unspecific. Please quantify the identified noise and how this could relate to the results from the 2010 case study.

The section has been revised based on the new results using synthetic plume height observations in Section 4.

L 390: 1000 days wall clock time or simulated days? How long does it need on a typical high performance computer architecture? What are the options for the parallelism of the application.

The 1000 days are simulated days. We have added a discussion about the required wall clock time and the options for parallelisation in each case.

L 400: It think it is fair to say that the assimilation of the plume height only had a small influence on the results. Also, the paper provides not enough evidence that one option is better than the other.

We agree that the difference is mostly minor. We have expanded the Conclusions section to cover both similarities and differences between the inversions as well as the impacts on both vertical distribution and total mass.

Response to specific comments by reviewer #2

This response extends our earlier response posted on 13 December 2016 and addresses the specific remarks by reviewer #2.

The reviewer comments are presented with a blue italic font.

There is a systematic confusion between "plume height" and (height of) "centre of mass". Both quantities seem interchangeable, so I would expect that this was described in a previous paper about the retrieval scheme. This should still be explained with proper references in section 2.2, with much more attention paid to the exact term used throughout the text (including title and abstract). Even if this was published already, consider adding a line plot in section 2.2 to compare the retrieved centre of mass with the plume height observed by radar and cameras (i.e. white dots and barely visible grey dots in figure 7). Note also that such a comparison directly above Eyjafjallajökull would still not be convincing for the downstream plume, where a multi-layered distribution probably happened.

The term "plume height retrieval" is used previous literature (Carboni et al., 2016, 2012; Rix et al., 2012), where it is interpreted as the midpoint of some parametric representation of the vertical profile. Since our observation operator does not assume a specific shape of the profile, it is more accurately described as an observation operator for the centre of mass (or rather, the vertical first moment of mass). However, since this definition is rather technical we prefer to use the term plume height retrieval except in cases where the distinction is essential.

We have added to the introduction a remark regarding the usage of "plume height" and "centre of mass" in the revised manuscript. The manuscript now uses "plume height" in all non-technical contexts.

The revised manuscript includes a comparison between the simulated emission, the IASI retrievals and the plume top observations, and also comparisons of the simulated SO2 concentration with the CALIOP profiles and collocated IASI retrievals.

What is the width of the Gaussian distribution assumed by the retrieval algorithm? The observation operator does not take into account this Gaussian shape. We are in a case where the vertical resolution leading to the simulated observation is much, much finer than the "vertical resolution" of the observation itself so this looks like a major oversight - especially in a context where Averaging Kernels were not taken into account. In order to simulate the observations correctly, the observation operator should first fit the modelled profile with a Gaussian shape before applying equation (4). Of course this should also be included in the adjoint of the observation operator...

The width of the Gaussian used in IASI forward model is 100 mb. The retrievals were also tested assuming different widths (50 and 10 mb), which gave results consistent with each other within the error bars.

The observation operator is based on the assumption that observations of the first moment of mass and the total column and the corresponding error covariance matrix are available. The retrievals (after application of the transformation described in Section 3.2) are consistent with this assumption, and the variability which is taken into account in the error estimates should not be included into the observation operator. Evaluating the centre of mass does not depend strongly on the model resolution.

The sentence on lines 134-135 is extremely problematic as it seems to show a fundamental misunderstanding about assimilation theory: "Finally, the vector y of observations is given by the possibly

non-linear observation operator H as y=H(x)+ε where ε denotes the observation error." H(x) is the model state in observation space. y-H(x) is the observation departure. Here ε does not denote only the observation error but all possible errors, including the model and representatitvity errors. Most importantly, y is not "given" by H, it is simulated by H! This awful confusion continues in equation (3) which does not provide y but the (total component of) H(x).

The sentence has been rephrased.

It appears from lines 307-312 that the assimilation algorithm can use both the algebraic solution with explicit computation of HM matrix, and the 4D-Var approach. This should be clearly explained in section 3.1. I assume that Vira and Sofiev (2012) provided all necessary details about this 4D-Var implementation and its adjoint model (e.g., was it generated automatically or manually? Is it as detailed as the forward model? How was it verified?). If that is the case, an additional reference and a few words may suffice. But if that was not the case, or if the 4D-Var implementation changed a lot (beyond the developments described in sections 3.2 and 3.3), then this should be fully described (since appropriate for the GMD journal).

The adjoint model is same as in Vira and Sofiev, (2012). The difference is that in the 2012 study the emission adjustment was multiplicative, horizontally variable but vertically constant, while in the current study, the adjustment is additive and vertically variable but confined to a single column. This corresponds to a different summation of the adjoint variable when evaluating the gradient, but the adjoint code does not need to be changed. A remark has been added to the revised manuscript.

References

Arason, P., Petersen, G.N., Bjornsson, H., 2011. Observations of the altitude of the volcanic plume during the eruption of Eyjafjallajökull, April–May 2010. Earth Syst. Sci. Data 3, 9–17. doi:10.5194/essd-3-9-2011

Boichu, M., Menut, L., Khvorostyanov, D., Clarisse, L., Clerbaux, C., Turquety, S., Coheur, P.-F., 2013. Inverting for volcanic SO2 flux at high temporal resolution using spaceborne plume imagery and chemistry-transport modelling: the 2010 Eyjafjallajökull eruption case study. Atmos. Chem. Phys. 13, 8569–8584. doi:10.5194/acp-13-8569-2013

Carboni, E., Grainger, R., Walker, J., Dudhia, a., Siddans, R., 2012. A new scheme for sulphur dioxide retrieval from IASI measurements: application to the Eyjafjallajökull eruption of April and May 2010. Atmos. Chem. Phys. 12, 11417–11434. doi:10.5194/acp-12-11417-2012

Carboni, E., Grainger, R.G., Mather, T.A., Pyle, D.M., Thomas, G., Siddans, R., Smith, A., Dudhia, A., Koukouli, M.L., Balis, D., 2016. The vertical distribution of volcanic SO2 plumes measured by IASI. Atmos. Chem. Phys. 16, 4343–4367. doi:doi:10.5194/acp-16-4343-2016

Elbern, H., Strunk, A., Schmidt, H., Talagrand, O., 2007. Emission rate and chemical state estimation by 4-dimensional variational inversion. Atmos. Chem. Phys. 7, 3749–3769. doi:10.5194/acpd-7-1725-2007

Flemming, J., Inness, A., 2013. Volcanic sulfur dioxide plume forecasts based on UV-satellite retrievals for the 2011 Grímsvötn and the 2010 Eyjafjallajökull eruption. J. Geophys. Res. Atmos. 118. doi:10.1002/jgrd.50753

Kristiansen, N.I., Stohl, A., Prata, A.J., Richter, A., Eckhardt, S., Seibert, P., Hoffmann, A., Ritter, C., Bitar, L., Duck, T.J., Stebel, K., 2010. Remote sensing and inverse transport modeling of the Kasatochi eruption sulfur dioxide cloud. J. Geophys. Res. 115, 1–18. doi:10.1029/2009JD013286

Lu, S., Lin, H.X., Heemink, A.W., Fu, G., Segers, A.J., 2016. Estimation of Volcanic Ash Emissions Using Trajectory-Based 4D-Var Data Assimilation. Mon. Weather Rev. 144, 575–589. doi:10.1175/MWR-D-15-0194.1

Meirink, J.F., Bergamaschi, P., Krol, M.C., 2008. Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion. Atmos. Chem. Phys. 8, 6341–6353. doi:10.5194/acpd-8-12023-2008

Müller, J.F., Stavrakou, T., 2005. Inversion of CO and NOx emissions using the adjoint of the IMAGES model. Atmos. Chem. Phys. 5, 1157–1186. doi:10.5194/acp-5-1157-2005

Rix, M., Valks, P., Hao, N., Loyola, D., Schlager, H., Huntrieser, H., Flemming, J., Koehler, U., Schumann, U., Inness, A., 2012. Volcanic SO2 , BrO and plume height estimations using GOME-2 satellite measurements during the eruption of Eyjafjallajökull in May 2010. J. Geophys. Res. Atmos. 117. doi:10.1029/2011JD016718

Seibert, P., Kristiansen, N.I., Richter, A., Eckhardt, S., Prata, A.J., Stohl, A., 2011. Uncertainties in the inverse modelling of sulphur dioxide eruption profiles. Geomatics, Nat. Hazards Risk 2, 201–216. doi:10.1080/19475705.2011.590533

Stohl, A., Prata, A.J., Eckhardt, S., Clarisse, L., Durant, A., Henne, S., Kristiansen, N.I., Minikin, A., Schumann, U., Seibert, P., Stebel, K., Thomas, H.E., Thorsteinsson, T., Tørseth, K., Weinzierl, B., 2011. Determination of time- and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull eruption. Atmos. Chem. Phys. 11, 4333–4351. doi:10.5194/acp-11-4333-2011

Vira, J., Sofiev, M., 2012. On variational data assimilation for estimating the model initial conditions and emission fluxes for short-term forecasting of SOx concentrations. Atmos. Environ. 46, 318–328. doi:10.1016/j.atmosenv.2011.09.066

Winiarek, V., Bocquet, M., Duhanyan, N., Roustan, Y., Saunier, O., Mathieu, A., 2014. Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations. Atmos. Environ. 82, 268–279. doi:10.1016/j.atmosenv.2013.10.017

Variational assimilation of IASI ~~SO2~~SO$_2$ plume height and total-column retrievals in the 2010 eruption of Eyjafjallajökull using the SILAM v5.3 chemistry transport model

Julius Vira[1], Elisa Carboni[2], Roy G. ~~Grainger[3]~~Grainger[2], Mikhail Sofiev[1]

[1] Finnish Meteorological Institute, Erik Palménin aukio 1, FI-00560 Helsinki, Finland
[2] COMET, Atmospheric, Oceanic and Planetary Physics, University of Oxford, Parks Road, Oxford, OX1 3PU, U.K.
~~[3] National Centre for Earth Observation, Atmospheric, Oceanic and Planetary Physics, University of Oxford, Parks Road, Oxford OX1 3PU, U.K.~~

Correspondence to: J. Vira, julius.vira@fmi.fi

Abstract

This study focuses on two new aspects on inverse modelling of volcanic emissions. First, we derive an observation operator for satellite retrievals of plume height, and second, we solve the inverse problem using an algorithm based ~~the~~ on ~~the~~ 4D-Var data assimilation method. The approach is first tested in a twin experiment with simulated observations and further evaluated~~demonstrated~~ by assimilating IASI ~~SO2~~SO$_2$ plume height and total column retrievals in a source term inversion for the 2010 eruption of Eyjafjallajökull. The inversion resulted in temporal and vertical reconstruction of the ~~SO2~~SO$_2$ emissions during the 1-2~~0~~1 May, 2010 with formal vertical and temporal resolutions of 500 m and 12 hours.

The plume height observation operator is based on simultaneous assimilation of the plume height and total column retrievals. The plume height is taken to represent the vertical centre of mass, which is transformed into the first moment of mass (centre of mass times total mass). This makes the observation operator linear and simple to implement. The necessary modifications to the observation error covariance matrix are derived.

Regularisation by truncated iteration is investigated as a simple and efficient regularisation method for the 4D-Var based inversion. In the twin experiments, the truncated iteration was found to perform similarly to the commonly used Tikhonov regularisation, which in turn is equivalent to a Gaussian a priori source term. However, the truncated iteration allows the level of regularisation to be determined a posteriori without repeating the inversion.

In the twin experiments, assimilating the plume height retrievals resulted in 5-20% improvement in root mean squared error but simultaneously introduced a 10-20% low bias on the total emission depending on assumed emission profile. The results are consistent with those obtained with real data. For Eyjafjallajökull, the comparison between results with and without assimilation of plume height retrievals shows that the estimated injection height was mostly constrained by the inversion even using only total column retrievals. However, comparison with the profile observations from the CALIOP instrument showed that assimilating the plume height retrievals improved the vertical distribution during episodes when the estimated injection height was not otherwise not sufficiently constrained.

32 Regularisation by truncated iteration is investigated as a simple and efficient regularisation method for the 4D-Var based
33 inversion. In an experiment with synthetic observations, the truncated iteration was found to perform similarly to the
34 commonly used Tikhonov regularisation. However, the truncated iteration allows the amount of regularisation to be varied a
35 posteriori, without repeating the inversion. For inverting the Eyjafjallajökull SO_2 emission at the temporal and vertical
36 resolution used in this study, the 4D-Var method required about 70% less computational effort than commonly used methods
37 based on performing a separate model simulation for each degree of freedom in the estimated source term.

38 Compared to the inversion using only total column retrievals, assimilating the plume height resulted in a vertical emission
39 profile more closely matching the ash plume heights observed by radar. The a posteriori source term for Eyjafjallajökull
40 consisted of gave an estimate of 0.29 Tg (with total column plume height retrievals) or 0.33 Tg (with total column retrievals
41 only) erupted SO_2 of which 95% was injected below 11 (12) km.

1   Introduction

43   Sulphur dioxide (SO_2) is one of the major gas-phase species released in volcanic eruptions. Large SO_2 releases
44 pose a hazard to aviation, decrease air quality, and as precursors to sulphate aerosols, have a potential impact on the Earth's
45 radiative balance (Bernard and Rose, 1990; Robock, 2000; Schmidt et al., 2015). Volcanic SO_2 plumes can be detected
46 by satellite instruments measuring in either UV or IR wavelengths - however, reliably forecasting the atmospheric transport
47 of volcanic plumes is hindered by the lack of in-situ measurements to characterise the emission fluxes of volcanic species
48 (Carn et al., 2009; Stohl et al., 2011; Zehner, 2012).

49   While methods based purely on satellite retrievals (Theys et al., 2013 and references therein) exist for inferring the total
50 SO_2 flux for a given eruption, a successful prediction of volcanic tracers generally requires information also on the
51 vertical profile of emissions. An important technique for assessing both vertical and temporal distribution of the emission
52 fluxes is provided by inverse dispersion modelling, first demonstrated for volcanic emissions by Eckhardt et al. (2008).

53   The previous studies on iInverse modelling of volcanic emissions has have been based on using total column retrievals of
54 SO_2 or volcanic ash together with a Lagrangian (Kristiansen et al., 2010; Stohl et al., 2011) or Eulerian (Boichu et al.,
55 2013; Boichu and Clarisse, 2014) dispersion models. In addition, Flemming and Inness (2013) devised a trajectory based
56 scheme to evaluate the vertical emission profile, which was used together with assimilation of SO_2 retrievals with the
57 IFS (Integrated Forecast System) weather prediction system. The previous studies have demonstrated that the vertical
58 distribution of emissions can be inferred from total column retrievals in presence of sufficient vertical wind shear. However,
59 in the case of the Eyjafjallajökull eruption in 2010, both Boichu et al. (2013) and Flemming and Inness (2013) pointed out a
60 lack of wind shear and a subsequent difficulty at estimating the vertical distribution of emissions.

61   Retrievals of SO_2 plume height have been performed with various satellite instruments (Carboni et al., 2012; Rix et
62 al., 2012). Nevertheless, only a few studies have incorporated these data into models. Wang et al. (2013) derived a three-
63 dimensional SO_2 distribution from retrievals by the Ozone Monitoring Instrument (OMI), and used the distribution to

2

64 initialize CTM simulations for the 2008 eruption of Kasatochi. Wilkins et al. (2015) used 1D-Var ash retrievals for
65 initialising dispersion simulations. However, neither of the studies used plume height retrievals in inverse modelling of
66 volcanic emissions.

67 The first objective of the present paper is to assess the usefulness of assimilating ~~SO₂~~SO_2 plume height retrievals from the
68 Infrared Atmospheric Sounding Interferometer (IASI) instrument in a source term inversion. Throughout this paper, the term
69 plume height will refer to the vertical centre of mass, which is consistent with the IASI retrievals of this study. Following
70 this definition of plume height, we introduce i~~I~~n Section 3.2 ~~we develop~~ an observation operator for the vertical centre of
71 mass.

72 Since the observation operator only depends on the centre of mass and column loading, the vertical profile is only partly
73 constrained. However, in contrast to the previous studies, this approach makes no further assumptions about the shape or
74 thickness of the ~~SO₂~~SO_2 layer. This could be advantageous, since volcanic ash or ~~SO₂~~SO_2 layers vary considerably in depth
75 (Dacre et al., 2014) and can be emitted in multiple, overlapping layers (Kristiansen et al., 2010). Although the variability of
76 the vertical profiles may introduce uncertainty into the retrieval of the plume height, by assimilating only the centre of mass,
77 we avoid forcing the model into a prescribed vertical profile whose uncertainty may be difficult to quantify. In
78 ~~addition~~contrast, our approach makes full use of the retrieval error estimates provided with the IASI data for both column
79 mass and plume height, including the estimated correlation between plume height and mass errors.

80 The second objective of this paper is to explore the connection between the source term inversion and the 4D-Var data
81 assimilation widely used in numerical weather prediction. Elbern et al. (2000) showed that the 4D-Var assimilation method
82 (Le Dimet and Talagrand, 1986) can be easily extended into estimating emission fluxes with a chemistry transport model.
83 Elbern et al. (2007) further evaluated the joint estimation of emission flux and airborne concentration as a strategy for
84 improving air quality forecasts. However, in this study, ~~the~~ the 4D-Var method is formulated to include only the emission
85 forcing, which ~~Under the assumption of a linear dispersion model and observations, the 4D-Var formulation~~ results in a least
86 squares problem similar to that solved by many existing inversion algorithms. ~~However, T~~the iterative solution employed in
87 4D-Var favours a different regularisation approach, which is ~~discussed~~ compared in Section ~~4.~~ 4 to a more classical
88 regularisation technique.

89 Finally, we test the variational inversion method and assimilation of plume height retrievals for estimating temporal and
90 vertical distribution of ~~SO₂~~SO_2 emission during the 2010 eruption of Eyjafjallajökull. Results of the inversion, presented in
91 Section 5, indicate that although the vertical distribution of emissions is mostly constrained by the total column retrievals
92 and the meteorological conditions, assimilation of plume height retrievals results in more vertically concentrated emission
93 profile. In particular, emissions above 8-10 km between 5 and 9 May are reduced substantially, which is consistent with
94 ~~radar-based estimates~~the observations of the eruption column height.

3

2   Model setup and observational data

2.1   Dispersion model

The transport and removal of SO_2 was evaluated using the dispersion model SILAM (System for integrated modelling of atmospheric composition; Sofiev et al., 2015, http://silam.fmi.fi) version 5.3. The model includes chemical removal of SO_2 as described by Sofiev (2000) with the OH climatology of Spivakovsky et al. (2000). The computations were driven by the ERA-Interim meteorological reanalysis (Dee et al., 2011) except for evaluating the simulated satellite retrievals described in Section 4, where operational ECMWF forecasts were used.

SILAM includes a variational data assimilation module, which was previously used for assimilation of air quality monitoring data of SO_2 by Vira and Sofiev (2012). The same 4D-Var implementation, including the adjoint codes, is used in this studyThis study uses the same assimilation system, but instead of estimating a refinement for a regional emission inventory, we seek to reconstruct the emissions for a single volcanic eruption as a function of time and injection height.

The model was configured for a domain covering 50°E to 30°W and 30°N to 80°N. Horizontal resolution of 0.5° was used for the inversion, while the a posteriori simulations were run with a higher 0.25° resolution. The vertical discretisation grid consists of 34 terrain-following z-levels with a 500 m resolution at the top of the domain increasing to 50 m near the surface.

2.2   Observations: tThe IASI dataset

IASI is an infrared Fourier transform interferometer that measures in the spectral range 645–2760 cm^{-1} with spectral sampling of 0.25 cm^{-1} (apodized spectral resolution of 0.5 cm-1) and has global coverage every 12h. The lev1b dataset from EUMETSAT/CEDA archive is used in this study.

The algorithm and the and the IASI SO_2 dataset (column amount and altitude)dataset are explained in more detail by Carboni et al. (2012). The same algorithm has been applied to other volcanic eruptions and successfully compared with other datasets (Carboni et al., 2016; Fromm et al., 2014; Koukouli et al., 2014; Schmidt et al., 2015; Spinetti et al., 2014).

The main points of the retrieval scheme are:

Retrievals are performed for the pixels that were identified by the SO_2 detection scheme (Walker et al 2011, 2012).

All the channels between 1100-1200 and 1300-1410 cm^{-1} are used in the iterative optimal estimation retrieval scheme to obtain SO_2 column amount and altitude of the plume (in pressure, under the assumption that the vertical concentration of SO_2 follows a Gaussian distribution) together with the surface temperature. The scheme determines the column amount and effective altitude (mean of a Gaussian profile) of the SO_2 plume with high precision (up to 0.3 DU error in SO_2 amount if the plume is near the tropopause), and it is well suited for plumes in lower troposphere.

The IASI SO_2 retrieval is not affected by underlying cloud. If the SO_2 is within or below an ash or cloud layer its signal will be masked and the retrieval will underestimate the SO_2 amount. In the case of ash this is discernible a posteriori by the value of the cost function. The altitude retrieved for the Eyjafjallajökull eruption plume (using the same

4

dataset as in this paper) in the presence of underlying cloud is consistent with the CALIPSO vertical backscatter profile (Carboni et al 2016, Figs. 1,2,3).

A comprehensive error budget for every pixel is included in the retrieval. This is derived from an error covariance matrix S_ε that is based on the SO_2-free climatology of the differences between the IASI and forward modelled spectra.

Note that the error covariance, S_ε, is defined to represent the effects of atmospheric variability not represented in the forward model, as well as instrument noise. This includes the effects of cloud and trace-gases which are not explicitly modelled. The matrix is constructed from differences between forward model calculations (for clear-sky) and actual IASI observations for wide range of conditions, when we are confident that negligible amounts of SO_2 are present. It follows that a rigorous error propagation, including the incorporation of forward model and forward model parameter error, is built into the system, providing quality control and error estimates on the retrieved state. The retrieval state error covariance matrix, used for the assimilation in this work, is directly provided as output of the retrieval pixel by pixel.

2.3   Other observations

Section 5 presents comparisons of the a posteriori simulation and the source term with the IASI plume height and total column observations. However, additional datasets required used for evaluating vertical structure of the inversion results. Due to the scarcity of vertically resolved SO_2 data, the comparison is based on aerosol observations. The vertical profiles of the emitted plumes are compared with the backscatter profiles by a satellite-borne lidar, and the SO_2 injection height is compared to plume top time series obtained with a C-band weather radar. The potentially different emission and transport of volcanic ash and SO_2 introduces some ambiguity to the comparisons; however, as found in Section 5, the different data sources together with the IASI retrievals nevertheless form a fairly coherent picture. This supports the conclusion of Thomas and Prata (2011), who found that ash and SO_2 were mostly collocated with each other during the Eyjafjallajökull eruption.

The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument (Winker et al., 2009) on board the CALIPSO satellite is near-nadir viewing, two-wavelength, polarisation-sensitive lidar. The comparisons in this study are shown for the 532 nm total backscatter. Hence, two main challenges are involved in using lidar data for evaluation of simulated SO_2 plumes. First, the comparison relies on the assumption that the SO_2 plume is collocated with an aerosol plume consisting either of primary particles (mainly volcanic ash) emitted in the eruption, or secondary particles (mainly sulphates) formed during the transport. Second, the volcanic plumes need to be distinguished from water or ice clouds. Although the vertical feature mask available with the CALIOP products provides a classification of aerosol and cloud types, as pointed out (Liu et al., 2009) and (Winker et al., 2012), thick volcanic ash plumes are frequently misclassified as ice clouds by the standard algorithm.

The comparisons shown in Section 5 and Appendix A consist of CALIOP overpasses intersecting the simulated Eyjafjallajökull plumes. Cases where the CALIOP track is parallel to the plume are omitted, because this makes the profiles extracted from the model very sensitive to horizontal displacement errors. Three of the CALIOP profiles have been

5

collected with the IASI retrievals under the criteria of less than 2 h time difference and less than 150 km horizontal displacement. The three collocated CALIOP tracks were previously analysed for SO_2 by (Carboni et al., (2016) along with two additional ones for May 14 and 16; these tracks only intersected the edge of the SO_2 plume and did not offer a useful comparison with the model.

The estimated SO_2 injection height is compared to the observations of plume top described by (Arason et al., (2011). The dataset includes two plume top time series, one estimated from a C-band weather radar located at the Keflavik airport 155 km from the volcano, and one estimated from imagery taken with a web camera located 34 km from the volcano. The 5-minute radar data and the hourly web camera data are averaged in time to facilitate the comparison with the estimated emission. The radar data include values which indicate presence of a plume below the lowest observed height, and in order to maintain consistency with the published 6-hourly time series (Arason et al., 2011; Petersen et al., 2012a), and to avoid a high bias in the averaged values, the altitude of 2.5 km above sea level is assigned to these points.

Both datasets represent the highest altitude with measurable signal from the volcanic plume, and thus, the observed plume height might differ from the midpoint of the emitted layer. The radar data are consequently compared with 80th and 95th percentiles (altitudes with 80 or 95 % mass emitted below) of the emission.

2.32.4 Inversion experiments

The inversion algorithm is evaluated with two sets of experiments based on the eruption of Eyjafjallajökull in 2010, The eruption has been described in detail by Gudmundsson et al. (2012). The experiments covered the time between 1 and 21 May, 2010, which as shown by Flemming and Inness (2013) included the most significant SO_2SO_2 releases.

The observation operator and the variational inversion technique were first evaluated in experiments with synthetic data (Section 4), where the simulated observations mimicking the IASI retrievals are extracted from a model simulation. The simulations are repeated for several assumed artificial source terms. The synthetic experiments evaluate the impact of assimilating plume height retrievals in addition to total columns, and additionally compare two options for regularising the inverse problem.

The IASI data were subsequently assimilated to invert for the SO_2 emissions forin the Eyjafjallajökull (2010) eruption. The inversion was performed both with and without assimilation of the plume height retrievals keeping the setups otherwise identical.

3   In all inversion experiments, the emission flux density (kg m^{-1} s^{-1}) was estimated for each model level in steps of 12 hours. The model setup used in the synthetic experiments was otherwise identical to that used with the IASI data, but a lower vertical resolution of 1 km was used to reduce the computational cost.

6

4 Assimilation and inversion methods

The forward problem for volcanic tracer transport is defined by the advection-diffusion equation: given the emission forcing f, solve

$$(1) \qquad \frac{\partial c}{\partial t} + \nabla \cdot (c\mathbf{V}) - \nabla \cdot (K\nabla c) = f(x,t) - s(c,x,t),$$

where c is the tracer concentration, $\mathbf{V}$ is the wind vector, K is the turbulent diffusivity tensor, and $s(c,x,t)$ denotes the chemical and other sinks, which in this study include the wet and dry deposition of SO_2 and its chemical conversion to SO_4.

3.1 Variational source term inversion

The inverse problem discussed in this paper is to determine the forcing f, given a set of observations depending on c. We assume that Eq. (1) has been discretised, and following the common notation in data assimilation literature, we denote the tracer concentrations, collectively for all time steps, with the state vector $\mathbf{x}$. The state vector is related to the unknown parameter vector $\mathbf{f}$ by the model operator $\mathcal{M}$, and to the observations $\mathbf{y}$ by the observation operator $\mathcal{H}$ as $\mathbf{y} = \mathcal{H}(\mathbf{x}_t) + \epsilon$, where $\mathbf{x}_t$ denotes the true state. The random vector ϵ includes the effect of observation errors as well as the possible representativeness or model errors associated with $\mathcal{H}$.

Finally, the vector $\mathbf{y}$ of observations is given by the possibly non-linear observation operator $\mathcal{H}$ as $\mathbf{y} = \mathcal{H}(\mathbf{x}) + \varepsilon$, where ε denotes the observation error.

If the errors follow a multivariate normal distribution with covariance matrix $\mathbf{R}$, then a solution to the inverse dispersion problem can be sought by maximising the likelihood function, which is equivalent to minimising the cost function

$$(2) \qquad J(\mathbf{f}) = \frac{1}{2}(\mathbf{y} - \mathcal{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}(\mathbf{x})),$$

where $\mathbf{x} = \mathcal{M}(\mathbf{f})$. Model errors are not explicitly included in the cost function.

The cost function assumes that the airborne concentrations, which comprise the state vector $\mathbf{x}$, are completely determined by the emission. Therefore, contrary to chemical data assimilation studies such as (Elbern et al., 2007), no term corresponding to the concentration in the beginning of assimilation is included. This is reasonable, since the inversion is performed in a single step, and the state and observation vectors in Eq. (2) cover the whole simulated period. The total SO_2 loading was low in the beginning of the assimilation due to the inactive phase of eruption and initial state was therefore unlikely to affect the inversion for the emission forcing.

7

218 Model errors are not explicitly included in the cost function, as the relation between concentrations x and the emission f
219 is taken as a strong constraint. Arranging the inversion into a sequence of shorter assimilation windows with a background
220 term for the initial state would relax this constraint at the boundaries of assimilation windows. However, this would still not
221 allow for model errors arising within the assimilation window, and problematically, the emitted mass would no longer be
222 conserved between the assimilation windows. Consequently, we use a single assimilation window and adopt the approach of
223 previous studies (Seibert et al., 2011; Stohl et al., 2011), where the model uncertainty is incorporated to the observation error
224 covariance matrix R. The form of R is explained in more detail in Sections 3.2 and 3.3.
225 Model errors are not explicitly included in the cost function. f

226 If the model and observation operators are linear, represented by matrices M and H, then (2) becomes a linear least-
227 squares problem. For volcanic eruptions with a known location, the emission vector f is zero almost everywhere, which
228 makes it feasible to evaluate the matrix HM and solve (2) algebraically. This is the basis for inversion methods of Boichu et
229 al. (2013), Eckhardt et al. (2008), and Lu et al. (2016).

230 As an alternative to the algebraic solution, the minimisation problem (2) can be solved with gradient-based, iterative
231 algorithms, which avoids evaluating the matrix HM. In this study, the cost function is minimized using the L-BFGS-B (the
232 limited memory Broyden-Fletcher-Goldfarb-Shanno algorithm with bound constraints) algorithm of Byrd et al. (1995) which
233 allows constraining the solution to non-negative values. Evaluating the gradient requires solving the adjoint problem for Eq.
234 (1).

235 The iteration is continued until a stopping criterion is satisfied, e.g. until the norm of the gradient is reduced by a
236 prescribed factor. However, in Section 4 we will discuss truncating the iteration before formal convergence in order to
237 control the regularization.

238

239 **4.23.2 Assimilation of plume height retrievals**

240 Given the tracer concentration $c(x, y, z)$ in three dimensions, the observation operator for column integrated mass m_{ij} is
241 given by

242 (33)
$$m_{ij} = \sum_{k=1}^{N} w_k c(x_i, y_j, z_k)$$

243 where x_i, y_j and z_k are the gridpoint coordinates and w_k denotes the thickness (in meters) of the kth model level. The layer
244 concentrations are often weighted with an averaging kernel (Eskes and Boersma, 2003) to account for the vertically varying
245 sensitivity of the satellite retrieval. In this work, weighting is not applied because the IASI retrievals treat the plume height
246 explicitly.

247 In the retrievals, the plume height is represented by its centre of mass.

8

248   (4) $\qquad Z_{CM.ij} = \dfrac{1}{m_{ij}} \sum_{k=1}^{N} z_k w_k c_{ijk}.$

249 It would be possible to develop an observation operator for z_{CM}, however, the operator would be nonlinear and only defined

250 for nonzero columns. These problems can be overcome by replacing the centre of mass with the first moment of mass mZ_{CM}.

251 Then, the observations consist of pairs $(m_{ij}, m_{ij}Z_{CM.ij})$ given by

252 (54) $\qquad \begin{pmatrix} m_{ij} \\ m_{ij}Z_{CM.ij} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{N} w_k c_{ijk} \\ \sum_{k=1}^{N} z_k w_k c_{ijk} \end{pmatrix},$

253 where z_k is the height of the kth model level and i and j refer to the horizontal coordinates. Transforming the observations of

254 Z_{CM} into mZ_{CM} changes the magnitudes of observation errors, and introduces a correlation between the observation

255 components m and mZ_{CM}. However, this effect can be evaluated and included into the observation operator.

256 The mean and standard deviation of m and Z_{CM} are denoted as μ_m, σ_m and μ_z, σ_z respectively. ~~Denote the means~~

257 ~~and standard deviations of m and Z_{CM} with μ_m, μ_z and σ_m and σ_z.~~ Assuming that the errors of m and Z_{CM} are normally

258 distributed, it can be shown that the variance of first moment equals

259 (6) $\qquad \begin{aligned} \mathrm{Var}[mZ_{CM}] &= \mu_m^2 \sigma_z^2 + \mu_z^2 \sigma_m^2 + \sigma_m^2 \sigma_z^2 \\ &+ 2\mu_m \mu_z \mathrm{Cov}[m, Z_{CM}] \\ &+ \mathrm{Cov}[m, Z_{CM}]^2. \end{aligned}$

260 ~~(5)~~

261 Under similar assumptions, the covariance of m and mZ_{CM} becomes

262 (7) $\qquad \mathrm{Cov}[m, mZ_{CM}] = \sigma_m^2 \mu_z + \mu_m \mathrm{Cov}[m, Z_{CM}].$

263 ~~(6)~~

264 Finally, the expectation of mZ_{CM} is shifted due to the correlation between retrievals of m and Z_{CM}:

265 (8) $\qquad \mathrm{E}[mZ_{CM}] = \mu_m \mu_z + \mathrm{Cov}[m, Z_{CM}]$

266 ~~(7)~~ .

267 The retrieval errors of different pixels are assumed to be uncorrelated. The observation error covariance matrix $\mathbf{R}$ is

268 therefore block-diagonal, and its entries can be evaluated using Eqs. (6) and (7) from the retrieval error estimates σ_m, σ_z

269 and $\mathrm{Cov}[m, Z_{CM}]$, which are all included in dataset used in this study. ~~the known covariances of m and Z_{CM} using Eqs. (5)~~

270 ~~and (6).~~ However, even if the standard deviations m Z_{CM} are known accurately, the means μ_m and μ_z need to be

9

Formatted: Indent: First line: 0 cm

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Formatted: MTDisplayEquation

Field Code Changed

Formatted: Indent: First line: 0,4 cm

Formatted: MTDisplayEquation

Field Code Changed

Formatted: MTDisplayEquation

Field Code Changed

Formatted: Indent: First line: 0 cm

Field Code Changed

271 substituted with the observed values of m and Z_{CM} . ~~R~~ The impact of this approximation is evaluated numerically in
272 Section 4.

273   Assimilation schemes commonly assume uncorrelated and unbiased observation errors. A non-diagonal $\mathbf{R}$ can be
274 introduced with a transformation of variables: define

$$\mathbf{L}^T\mathbf{L} = \mathbf{R}^{-1}$$
275 (9) $$\tilde{\mathbf{y}} = \mathbf{L}(\mathbf{y} - \mathbf{b})$$
$$\tilde{\mathbf{H}} = \mathbf{L}\mathbf{H}$$

276 (8)

277 where $\mathbf{L}^T\mathbf{L}$ is the Cholesky factorisation of the inverse observation error covariance matrix $\mathbf{R}^{-1}$ and $\mathbf{b} = (0, \mathrm{Cov}[m, Z_{CM}])$
278 corrects for the bias according to Eq. (8) ~~(7)~~. Then, substituting the transformations of Eq. (9) ~~(8)~~ into the cost function (2)
279 shows that assimilation of $\mathbf{y}$ with the original $\mathbf{R}$ is equivalent to assimilation of $\tilde{\mathbf{y}}$ using the transformed observation
280 operator $\tilde{\mathbf{H}}$ with unit matrix in place of $\mathbf{R}$.

281   The above formulas can be implemented as a preprocessing step for the observations. In summary, the procedure is then
282 as follows:

283   1. For each available IASI pixel i , evaluate the tuple $\mathbf{y}_i - \mathbf{b}_i = (m_i, m_i Z_{CM,i} - \mathrm{Cov}[m_i, Z_{CM,i}])$ and the corresponding
284   2x2 covariance matrix $\mathbf{R}_i$.

285   2. Factorise $\mathbf{R}_i^{-1} = \mathbf{L}_i^T\mathbf{L}_i$ and transform the observations according to Eq. (9) ~~(8)~~.

286   3. Store the transformed observations $\tilde{\mathbf{y}}_i$ with their pixel-specific vertical weighting functions given by rows of the
287   matrix $\tilde{\mathbf{H}} = \mathbf{L}_i\mathbf{H}$.

288   After the transformation, the observations are handled identically to regular column observations with a vertical weighting
289 function.

290 ### 4.3 3.3 Observation errors

291   The IASI retrievals used in this study include pixel-specific error estimates for total column and plume height retrievals.
292 The estimates are derived statistically (Carboni et al., 2012) from differences between the transmission spectra computed by
293 a forward model and those observed by IASI. Together with estimates for the correlation between plume height and total
294 column retrieval errors, this provides the necessary input for equations (6) and (7) ~~(5) and (6)~~.

295   The retrieval error estimates are only provided for pixels with positive ~~SO₂~~ SO_2 detection. For the non-~~SO₂~~ SO_2 pixels,
296 which are assimilated as zero values, a different estimate is used, based on the detection limits estimated by Walker et al.
297 (2012). The detection limit was translated into a standard deviation of a Gaussian random variable assuming, conservatively,
298 a probability of 0.95 for a correct detection.

299 However, performing the inversions with $\mathbf{R}$ defined only by retrieval errors resulted in poor a posteriori agreement with
300 the IASI data, which suggested that the retrieval errors are not sufficient to describe the discrepancy between the simulated
301 and observed values. As will be found shown in with the synthetic experiments, the impact of model uncertainty is
302 significant compared to the retrieval errors, and it needs to be taken into account. The problem of model errors affecting the
303 inversion is discussed by Boichu et al. (2013), who found the impact to depend strongly on treatment of zero-value
304 observations, and consequently chose to keep only every tenth zero-valued observation.

305 In this study, the model errors are included by modifying the observation error covariance matrix, which is set to
306 $\mathbf{R} = \mathbf{R}_{obs} + \mathbf{R}_{model}$, where $\mathbf{R}_{model}$ is constant, diagonal and determined experimentally. The model error standard deviation for
307 total column observations is set to 2 DU for both the experiments using synthetic data (Section 4) and for the inversion for
308 Eyjafjallajökull (Section 5), while the model error for the plume height retrievals was set to 2 km for the synthetic
309 experiments and 1 km for the Eyjafjallajökull inversion. Reducing the plume height standard deviation to 1 km in the
310 synthetic experiments resulted in large negative bias in the total emission, while increasing the standard deviation to 2 km
311 did not significantly change the total emission in the inversion for Eyjafjallajökull.
312 and corresponds to experimentally determined constant standard deviations of 2 DU for total column and 1 km for the
313 plume height.
314 The model errors for plume height and total column are assumed uncorrelated and independent of the observation errors.
315 However, their effect is propagated to the covariance matrix for first moment according to Eq. (6) (5). The actual model
316 errors evolve dynamically and are likely to be variable and correlated in space and between the plume height and total
317 column components; however, a more advanced treatment including these effects appears difficult in the current inversion
318 approach.

319 **4.4 3.4 Regularization**

320 The least squares problem (2) has a unique solution only if the matrix $\mathbf{H}\mathbf{M}$ is of full (numerical) rank. Furthermore, if
321 $\mathbf{H}\mathbf{M}$ is close to singular, the problem remains ill-posed, which results in a noisy solution. Consequently, some form of
322 regularisation has been employed in all previous works based on the least-squares approach.

323 A common option is the Tikhonov regularisation (Tikhonov, 1963; Engl et al., 2000; Tikhonov, 1963), which introduces a
324 penalty term into the cost function (2), which in the simplest form becomes

325 $$(10) \qquad J(\mathbf{f}) = \frac{1}{2}(\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) + \alpha^2 \sum_{k,n} w_k \mid f_{k,n} \mid^2$$

326 (9)

327 where the summation is over levels k and timesteps n. The weights w_k in Eq. (10)(9) are set equal to the thickness of each
328 model layer; this makes the penalty term consistent with its continuous counterpart $\int f(z,t)^2 \, dt \, dz$, which in turn ensures that
329 the regularisation term does not depend on the vertical discretisation.

11

330 The penalty term can be modified to include a non-zero a priori source term. However, this approach is not taken in the
331 present work. Instead, we aim to choose the level of regularisation optimally, so as to avoid excessive bias in the regularised
332 solution. The need for regularisation depends on the coverage of observations, accuracy of the forward model as well as on
333 the meteorological conditions controlling the dispersion. Thus, the regularisation parameter α^2 cannot be chosen a priori.

334 In this work, a criterion known as the L-curve (Hansen, 1992) is used for determining the amount of regularisation. In the
335 L-curve approach, the inversion is performed with various values of α^2, and the residual $\|y - Hx\|$ is plotted as a function of
336 the solution norm $\|f\|$. For ill-posed inverse problems, the curve is typically L-shaped. The residual initially reduces quickly
337 as the regularization is relaxed, however, for some value of α^2, the curve flattens and reducing the regularization further
338 only marginally improves the fit. This point, where L-curve reaches its maximum curvature, is taken to represent the optimal
339 regularisation. In the present study, the L-curve is evaluated without the frequently used logarithmic transformation.

340 The main advantage of the L-curve method is that it does not rely on a priori estimates for the observation error. This is
341 useful, since in practice the discrepancy between simulated observations and the data is also affected by model errors, which
342 are poorly known. The L-curve was, in effect, used in inverse modelling of volcanic SO₂SO₂ also by Boichu et al. (2013).

343 Changing the regularisation parameter requires the minimisation to be started over, which is costly in the variational
344 inversion scheme where each iteration requires a model integration. However, as noted by Fleming (1990) and Santos
345 (1996), the iteration itself forms a sequence of solutions with decreasing regularisation. Thus, instead of minimising the
346 regularised cost function (10)(9), we iterate to minimise the original cost function (2), and truncate the iteration according to
347 the L-curve criterion. This approach, known as regularisation by truncated iteration (Kaipio and Somersalo, 2006), or
348 iterative regularisation (Hansen, 2010), provides a computationally efficient method to regularise large-scale inverse
349 problems. In the following section, we show experimentally that the truncated iteration such an approach results in similar
350 solutions for the source term inversion as the more common Tikhonov regularisation.

351 **54 Experiments with synthetic data**

352 Regularisation by truncated iteration has been studied in detail especially for Krylov subspace based algorithms (Calvetti
353 et al., 2002; Fleming, 1990; Kilmer and O'Leary, 2001). The effect of truncated iteration on quasi-Newton minimisation
354 methods, such as the L-BFGS-B algorithm used in this work, has been studied less extensively. To evaluate the truncated
355 iteration in comparison to Tikhonov regularisation for inverse modelling of volcanic emissions, we performed an experiment
356 with synthetic observations extracted from forward model simulations. In addition to the comparison of regularisation
357 methods, the synthetic experiments enable us to evaluate robustness of the L-curve method and to assess the impact of
358 assimilation of plume height retrievals, and to quantify how model errors affect the source term estimate.
359 For the sake of computational convenience, the experiments in this section are not performed using the variational method
360 described in Section 3.1, but instead the forward sensitivity matrix **HM** is evaluated by running a separate model simulation

12

Field Code Changed

for each component of the emission vector f . The sensitivity matrix is subsequently used for evaluating the cost functions (Eq. (2) for truncated iteration, Eq. (10) for Tikhonov regularisation) and the respective gradients required by the L-BFGS-B minimisation code. Evaluating the sensitivity matrix also provided an opportunity to numerically confirm the equivalence of the matrix-based and variational inversion methods.

The experiments with synthetic data were set up for the same time (1 to 21 May, 2010) as the inversion for Eyjafjallajökull. The synthetic observations were evaluated by running forward simulations with a set of artificial source profiles (cases A to D) shown in the leftmost column of Figure 1. The synthetic observational data (total columns and first moments as explained in Section 3.2) correspond to the locations and times covered by the IASI overpasses during the simulated period.

The artificial source terms A and B are defined arbitrarily, while cases C and D are realisations of a stochastic process where the total flux (kg/s) is given by a lognormal, temporally correlated random variable and the eruption height follows the relation of Mastin et al. (2009). At each time, a piecewise constant vertical profile is assumed with a transition at 75% of height. The emission rate is distributed evenly between the two sections.

The simulations with artificial source terms were driven by the meteorological data valid for the simulated period. However, Ttwo sets of meteorological input were used: the synthetic observations were generated using the operational ECMWF forecast fields, and but to simulate the effect of model errors, the sensitivity matrix for used in the inversions was evaluated using the ERA-Interim as the meteorological driver. Although changing the meteorological driver does not cover all sources of model error, we expect the resulting perturbation to have statistical properties similar to the real model uncertainty.

The effect of retrieval errors was simulated by

For the sake of computational convenience, the experiments in this section were carried out by pre-evaluating the forward sensitivity matrix **HM** by running a separate model simulation for each component of the emission vector f . In order to simulate the effect of model errors, the matrix **HM** was evaluated with both the ERA-Interim and operational ECMWF forecast fields as meteorological drivers.

The sensitivity matrix for inversions was extracted from the run with ERA-Interim meteorological data. The set of synthetic observations of the SO_2 column density, on the other hand, was evaluated from the model run based on the operational meteorological fields and used as the data for the inversion experiments. The perturbing simulated the extracted (simulated) observations were perturbed with additive Gaussian noise. In order to perturb the simulated plume height retrievals, the unperturbed simulated first moments and total columns were first converted back to the centre of mass and total column for the pixels with column density higher than 0.2 DU in the forward run. Then, both the simulated centre of mass and the total column were perturbed and transformed back to the (perturbed) total columns and first moments. The total columns were perturbed with standard deviation equal to 0.1 DU + 10 % of the true values; the centres of mass were perturbed with a constant standard deviation of 1 km. A negative correlation coefficient of -0.9 was assumed between the perturbations to the total column and centre of mass.

13

The error covariance matrix used in the inversion was supplemented with 2 DU and 2 km "model error" as described in Section 3.3. For the inversions using simulated plume height retrievals, the observation error covariance matrices were transformed according to Eqs.(6) –(8) using the perturbed centre of mass and total column values for μ_z and μ_m μ_z.

with standard deviation equal to 1 DU + 30% of the true value. The observation error covariance matrix used in the inversion was supplemented with 2 DU "model error" as described in Section 3.3.

The residual and solution norms, which define the L-curves, are evaluated consistently to the penalized cost function (10) (9):

(11)
$$\|\mathbf{H}\,\mathbf{x} - \mathbf{y}\| = \sqrt{(\mathbf{H}\,\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{H}\,\mathbf{x} - \mathbf{y})}$$
$$\|\mathbf{f}\| = \sqrt{\sum_{k,n} w_k \mid f_{k,n} \mid^2}$$

(10)
$$\|\mathbf{H}\,\mathbf{x} - \mathbf{y}\| = \sqrt{(\mathbf{H}\,\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{H}\,\mathbf{x} - \mathbf{y})}$$
$$\|\mathbf{f}\| = \sqrt{\sum_{k,n} w_k \mid f_{k,n} \mid^2},$$

where $\mathbf{f}$ denotes the emission, $\mathbf{x} = \mathbf{M}\,\mathbf{f}$ and w_k is the thickness of the kth model layer. To evaluate the L-curve for Tikhonov-regularisation, the parameter α^2 was incremented in discrete steps given by $\alpha_i^2 = 10^7 \cdot 2^{-i}$ for $i = 0, 1, 2, \dots$. The L-BFGS-B minimization method with non-negativity constraint was used for both Tikhonov regularisation and the truncated iteration; in the case of Tikhonov regularisation, the iteration was continued until convergence for each α_i^2 either until convergence or for maximum of 50 iterations. A zero-valued solution was always used as the first guess in the iteration. With the truncated iteration, the weights w_k, required by Eqs. (10) and (11) (9) and (10), are not explicitly included in the cost function. Instead, the same effect is achieved by transforming the parameter vector as $f'_{k,n} = w_k^{1/2} f_{k,n}$.

The point where the L-curve flattens, which is taken as the final solution, was determined numerically. First, the points $(\|\mathbf{f}\|, \|\mathbf{H}\,\mathbf{x} - \mathbf{y}\|)$ are sorted according to increasing $\|\mathbf{f}\|$. Then, the points where the residual increases are removed, and finally, the optimal point is chosen using the "triangle" algorithm of Castellanos et al. (2002).

Figure 1Figure 1 presents the inversion results using Tikhonov regularisation with total column observations, truncated iteration with total column observations, and truncated iteration with total column and centre of massplume height observations. Regardless of the assumed source term or inversion method, the emission timing is well captured within the 12 h resolution. The overall vertical profiles are also recovered, however, spurious features are present especially in cases B and C.

For comparison, Figure 2 presents the solution corresponding to the case B in Figure 1Figure 1 but evaluated without model errors – that is, using the same sensitivity matrix $\mathbf{H}\mathbf{M}$ for both evaluating the observations and performing the inversion. In this case, regularisation was not needed, and the true solution was recovered almost perfectly despite the noisy

14

observations. Thus, the noise present in the estimated solutions in Figure 1~~Figure 1~~ is mainly due to model error, which affects the elements of matrix $\mathbf{M}$. All other results presented in this section are obtained in presence of model errors.

Numerical evaluation of the inversion results in terms of RMSE and relative bias is presented in Table 1. The scores are evaluated for both truncated iteration and Tikhonov regularisation, each with and without plume height observations. Furthermore, two numbers are given for each case: the optimal value, corresponding to the regularisation (for Tikhonov, the value of α^2, for truncated iteration, the iteration number) with lowest RMSE, and the L-curve value corresponding to the choice of regularisation as determined from the L-curve explained above. Clearly, the regularisation with optimal RMSE is not necessarily optimal with respect to bias.

For all cases, the optimally truncated iteration had lower RMSE than the optimally tuned Tikhonov regularisation. However, this advantage was not always realised when the truncation was determined from the L-curves, which are shown in Figs. 3 and 4. For the Tikhonov regularisation, the L-curve solution was generally closer to the optimal. The difference is caused by differing features of the L-curves for the two regularisation methods: for the Tikhonov regularisation, the L-curve forms a convex graph varying smoothly with α^2, while the curves formed by the L-BFGS-B iterates are neither smooth nor even monotonous. Although points where the residual increases are omitted from the search, points with a locally large curvature remain in the curve, and such points are responsible for the under-regularised L-curve solutions in cases A and D when only total column was assimilated.

~~In 34 cases out of 40, the RMSE of the solution determined from the L-curve was within 20% from the optimally regularised solution. Of the remaining six cases, two were over-regularised and four were under-regularised.~~

In Figs. 3 and 4, the root mean squared error (RMSE) of the solution is shown next to each L-curve as a function of the regularisation parameter. As expected, the RMSE initially drops as the regularisation is relaxed, reaches a minimum, and eventually increases as the solution becomes contaminated by noise. This behaviour was especially clear when only total column observations were assimilated. When also centres of mass were assimilated, the minima in RMSE become weaker, and the RMSE with maximum number of iterations was only slightly higher than optimal. Thus, assimilating the centres of mass had the unintended but potentially useful side effect of making the inversion less sensitive to under-regularisation.

Since the regularised cost function (10) favours solutions with a small squared norm, the inversion is expected to underestimate the true emission. If only total column observations are used, the underestimation remains small, being 5 – 10% for the L-curve solutions with truncated iteration, and up to 15 % for the corresponding Tikhonov regularised solutions. However, when the plume height observations were included, the negative biases increased to 15-25% even when using truncated iteration.

Magnitude of the negative bias turned out to be sensitive to the assumed model uncertainty as described by the covariance matrix $\mathbf{R}_{mdl}$. Reducing the standard deviation for plume height errors to 1 km resulted in negative biases between 25 and

15

35%. σ_z ±† As a further sensitivity test, we evaluated the effect of approximating the true values for total column and plume height with the respective observed values when transforming the observation error covariance matrix, as explained in Section 3.2. Using, unrealistically, the true values in the inversion, the relative biases were reduced to 16-21%. The RMSE was reduced by up to ~15%. It can be noted that none of the tested setups describe an observation error covariance matrix that would perfectly match the perturbations applied the simulated observations, since the model errors, simulated by using a different meteorological driver, are not well described by additive, white noise. Taking the cross-correlations and spatial variation of model errors into account might lead into different optimal $\mathbf{R}_{mdl}$.

The total emitted mass is underestimated by < 10 % for the solution from truncated and by up to about 15 % for the Tikhonov regularised solution. The underestimation is expected due to the form of cost function (9). However, the inversion results show that the negative bias is not necessarily large unless the problem is regularised too strongly.

For comparison, Figure 2 presents the solution corresponding to the case B in Figure 1 but evaluated without model errors – that is, using the same sensitivity matrix $\mathbf{H}\mathbf{M}$ for both evaluating the observations and performing the inversion. In this case, regularisation was not needed, and the true solution was recovered almost perfectly despite the noisy observations. Thus, the noise present in the estimated solutions in Figure 1 is mainly due to model error, which affects the elements of matrix $\mathbf{M}$.

The L curves corresponding to each case in Figure 1 are shown in Figs. 3 and 4. The root mean squared error (RMSE) of the solution is shown next to each L curve as a function of the regularisation parameter. When measured by the solution RMSE, an optimal regularisation indeed existed in each case. In case A, where the solution varies smoothly in time and space, the solution error is only moderately sensitive to the regularisation. The L curve formed by the L-BFGS-B iterates is shallow in this case, which caused the algorithm to choose an unnecessarily high number of iterations. However, the negative effect on the solution quality was small. For the Tikhonov regularisation, the regularisation parameter was determined almost optimally.

The choice of regularisation was more critical in the remaining cases. In cases B and C, the L curve has a clear plateau after initial decrease, and the chosen corner point is close to optimal for the both regularisation methods. In case D, the truncated iteration leads to a somewhat under-regularised solution similar to case A.

Outcome of the four experiments indicates that the need for regularisation varies strongly depending on the true source, whose characteristics also affect how accurately the algorithm determines the optimal regularisation. We used the stochastic source terms to evaluate this more quantitatively. Figure 5 presents the RMSE as a function of the iteration number for 40 realisations of the stochastic source term used in cases C and D. The optimal iteration numbers chosen from each L curve are marked with stars.

487 ~~The RMS errors shown in Figure 5 are normalised by the minimum error for each inversion, which shows that in most~~
488 ~~cases, the inversion was only moderately sensitive to the exact point of truncation. In 34 cases out of 40, the RMSE of the~~
489 ~~solution determined from the L-curve was within 20% from the optimally regularised solution. Of the remaining six cases,~~
490 ~~two were over-regularised and four were under-regularised.~~

491 While the experiments in this section were performed by pre-evaluating the matrix $\mathbf{HM}$, in 4D-Var, the multiplications
492 by $\mathbf{HM}$ and its transpose are replaced by forward and adjoint model evaluations. Although the approaches are formally
493 equivalent, this change results in a slightly different sequence of iterations from which the L-curve is evaluated. To
494 investigate this difference, we performed the inversion using the real IASI data using both approaches. The two solutions are
495 shown in Figure 5~~Figure 6~~. The total released mass differs by less than 1% between the solutions, and the emission patterns
496 are qualitatively similar. The differences for individual values, although larger, appear small compared to the inversion
497 errors.

498 In summary, the experiment with synthetic data showed that the truncated iteration resulted in solutions similar to those
499 obtained with the more common Tikhonov regularisation. This makes the truncated iteration, in combination with the L-
500 curve, an attractive option for regularising the variational source term inversion. On the other hand, ~~the overall need for~~
501 ~~regularisation depended strongly on the assumed source term.~~ ~~n~~No regularisation was needed in absence of model error~~,~~
502 which indicates that the need for regularisation is likely to also depend on quality of the forward model. This emphasizes the
503 need for a robust method to determine the appropriate regularisation according to the situation at hand.

504 ## ~~6~~5 Inversion results for Eyjafjallajökull

505 Optimising the source term following the regularisation strategy (truncated iteration) described in Section 3.4 results in
506 satellite-derived estimates on the temporal and vertical emission profiles, as well as on the total emitted amount. The
507 solutions presented here correspond to iterates chosen from the L-curve ~~as~~ using the algorithm described in Section 3.4. For
508 assimilation of column mass only, the 9th iterate was chosen; with column mass and plume height assimilation, the 13th
509 iterate was chosen. Similarly to the synthetic experiments, the initial iterate was a zero solution. The L-curves are shown in
510 the supplementary information.
511

512 ~~Figure 7~~Figure 6 shows the temporal and vertical distribution of the SO_2~~SO_2~~ emission obtained both with and without
513 assimilation of plume height. The plume height time series estimated from radar and camera observations (Petersen et al.,
514 2012b) are plotted on top of the emission distributions. Both the camera and radar observations represent the top of the
515 visible plume, and ~~e~~Even if the visible plume does not necessarily coincide with the SO_2~~SO_2~~ plume, the plume height
516 observations provide an indication of the eruption activity.

517 Figure 7 shows the vertical profile of emissions integrated over the whole period. The bulk of emissions are between 2
518 and 8 ~~kilometres~~km even if only column density is assimilated. Assimilating the plume height retrievals further decreases the

17

fraction of emissions above 8 km. When the plume height is assimilated, about 85% of total emission is estimated below 8 km and about 95% below 11 km. Without assimilation of plume heights, the 95% level raises to 12 km.

The total released mass of SO_2 is 0.33 Tg when plume height is not assimilated and 0.29 Tg when plume height is assimilated. Figure 9 depicts the emission flux as a function of time and shows that while the largest difference in emission rate is during the peaks of 6th May, the assimilation of plume heights tends to decrease the emission rate throughout the eruption.

The strongest emission occurred during 6th May. However, the vertical distribution of the peak depends on whether the plume height is assimilated. While the maximum occurs at 5-6 km, if plume height is not assimilated, secondary maxima appear at 11 km, reaching 13 km on 9th May. If plume height retrievals are assimilated, the emission above about 8 km is strongly suppressed. Similarly, on 18th May, the isolated emissions at 10 and 15 km are essentially largely removed when the plume height is assimilated.

A more quantitative view on the effect of assimilating the plume height retrievals is given by Figure 8, which compares the estimated centre of mass of the SO_2 emission with the retrieved plume heights. The plume heights are shown as averages within both 50 and 500 km radius from the volcano. The averages over wider area have better temporal coverage and they are likely to be less affected by unresolved temporal or spatial variations in the plume height. The retrievals with estimated error larger than 5 km are excluded from the averaging (although used in assimilation).

In addition, Figure 8 includes radar and camera observations of the plume top which are compared with the 80th and 95th percentiles of the emission. The 95th percentile, although formally more representative of the top of emissions, shows very large fluctuations compared to both observations and the 80th percentile, which suggests that the highest percentiles might not be a robust way to characterise the plume top in the inversion results.

Over the whole period, the inversion results show a larger variability of injection height in comparison to both IASI and the radar or camera time series. Between May 4 and 5 and later May 10 and 17, the average IASI retrievals and the emission centre of mass agree mostly within 1-2 km, as do the radar observations with the 80th percentile of emission. An exception is the evening of May 11 when the injection height appears overestimated, however, the total emission rate was low at that time. Assimilation of plume height retrievals had little impact on the injection height during these times.

Between May 6 and 10, the injection height is overestimated in comparison with both IASI and radar observations. Assimilating the plume height retrievals improves the comparison, but the injection height remains 2-5 km too high compared to the averaged IASI retrievals. A similar overestimation occurs on May 17 and 18. Assimilating the plume height again reduces the overestimation significantly on those days, however, both the centre of mass and the percentiles remain overestimated.

The total released mass of SO_2 is 0.33 Tg when the plume height is not assimilated and 0.29 Tg when the plume height is assimilated. Figure 8d, which depicts the emission flux as a function of time, shows that while the largest difference in

18

emission rate is during the peaks of 6th May, the assimilation of plume heights tends to decrease the emission rate throughout the eruption.

~~Figure 8 shows the vertical profile of emissions integrated over the whole period. The bulk of emissions are between 2 and 8 kilometres even if only column density is assimilated. Assimilating the plume height retrievals further decreases the fraction of emissions above 8 km. When the plume height is assimilated, about 85% of total emission is estimated below 8 km and about 95% below 11 km.~~
~~The total released mass of SO_2 is 0.33 Tg when plume height is not assimilated and 0.29 Tg when plume height is assimilated. Figure 9 depicts the emission flux as a function of time and shows that while the largest difference in emission rate is during the peaks of 6th May, the assimilation of plume heights tends to decrease the emission rate throughout the eruption.~~
~~The inversion results of Figure 7 can be compared with those in Figure 10, which are obtained by assimilating both total column and plume height but neglecting all off-diagonal observation error covariance matrix elements. The distribution of emissions differs strongly from both cases in Figure 7, and the vertical distribution remains as spread as with assimilation of total column only. The treatment of observation errors as described in Section 3.2 is therefore a necessary step for successful assimilation of the plume height retrievals.~~

The SO_2 column densities simulated a posteriori are shown for 5-10 May in Figure 9 ~~Figs. 11 and 12,~~ along with the corresponding IASI retrievals. The overall patterns are well reproduced, although the column density is underestimated for some parts of the plume, especially on 6th ~~7th~~ and 7th ~~8th~~ of May. Due to the smaller total emission, the column densities are slightly lower when plume height is assimilated~~, however, the difference is small~~. Comparisons of the total columns for all 21 days are presented in the supplementary material.

Figure 10 shows the simulated plume height (evaluated as centre of mass) for 7-9 May, which corresponds to the period of overestimated injection height shown in Figure 8. Compared to IASI, the inversion using only total columns tends to overestimate the plume height for all three days. As expected from Figure 8, when the plume height retrievals are assimilated, the overestimation is reduced, but not entirely removed.

A more detailed evaluation of the vertical profiles is enabled by comparison with the CALIOP lidar backscatter data. It should be noted that the most prominent features in the CALIOP data are regular clouds; in particular, this includes the near-constant layers located at 1-2 km altitude.

In Figure 11, the simulated SO_2 concentration is plotted as contours together with the CALIOP attenuated backscatter data collected on May 6 and 8, 2010. On both days, the track segment intersects the SO_2 plume near its source. On May 6, this part of the volcanic plume is obscured by a cloud, but a distinctive aerosol layer is visible south of 60° N. This layer is reproduced by the model, however, the observed vertical extent is much thinner than modelled, indicating that the vertical variation of the transport was not sufficient to resolve the emission vertically. The plume height for the thickest part of the

19

plume is nevertheless reproduced within ~2 km, and hence, assimilating the plume height retrievals had only little impact on the simulated plume.

On May 8, the highest simulated concentrations coincide with a strong backscatter signal at 3-4 km altitude close to the emission (near 62° N). The altitude is consistent with the averaged IASI plume height retrievals shown in Figure 8, whereas the simulated vertical extent between 2 and 7.5 km is again too wide. While a second layer between 8 and 12 km is present in the CALIOP data, the horizontal extent of this feature is far too wide to represent the volcanic plume. A third simulated SO_2 layer is present at 13 km only if plume height retrievals are not assimilated; this demonstrates the difference of injection heights seen in Figure 8.

The CALIOP track on May 8 also crosses an older SO_2 plume around 48° N, where the simulated vertical extent is compatible with the CALIOP data. However, a prominent layer extending between 50° and 55° N is present in the CALIOP data. The layer is classified partly as cloud and partly aerosol in the CALIOP vertical feature mask (not shown), but the layer does not coincide with the simulated SO_2 plume. However, Figures 9 and 10 indicate that the simulated plume was erroneously displaced towards west during the evening of May 7. Taking this into account, it is feasible that the observed backscatter would be caused by the volcanic plume. The 3-4 km altitude of the layer would agree with the IASI plume height retrievals (Figure 10) and support the below 5 km injection heights indicated by the IASI and radar data in Figure 8.

Figures 12 through 14 combine the simulated SO_2 profiles and the CALIOP data with collocated IASI total column and plume height retrievals. The simulated vertical distributions are mostly consistent with both the CALIOP and the IASI data. In Figure 12, the 3-4 km mean altitude of the peak reaching 20 DU according to the IASI data is reproduced by the model. The altitude of the plume extending towards south (between 48-50° N) is also reproduced given the higher retrieval uncertainty. The column densities up to 20 DU, however, are not reproduced: the highest simulated values are displaced towards west and remain below 10 DU.

Figures 13 and 14 show generally similar level of agreement in the vertical structures. In both figures, the northern part of the plume (55-60° N) is partly obscured by a cloud, which is reflected by the large retrieval error estimates. In both figures, assimilating only total column retrievals resulted in several isolated SO_2 layers between altitudes of 10-15 km. Presence of these layers is supported by neither IASI nor CALIOP data. Even if the corresponding SO_2 emissions did not coincide with ash emissions, some CALIOP signal could be expected due to the sulphate particles forming in the plume. Altogether, the comparisons in Figs. 12 through 14 and the comparison of the emission profiles (Figure 8) support the conclusion that the emissions above 8-10 km on 6-9 May were an artefact and probably related to insufficient wind shear.

Further comparisons with CALIOP data on 14 to 17 May are shown in Appendix A. The simulated vertical distributions generally coincide with layers observed by CALIOP; however, assimilation of plume height retrievals had little impact on the simulated plumes at those times.

The plume height, evaluated as centre of mass, for 6-9 May is shown in Figure 13. Compared to IASI, the simulation based only on assimilation of total columns tends to overestimate the plume height for all four days. When the plume height retrievals are assimilated, the overestimation is reduced consistently, although not entirely removed.

20

76 Discussion

No a priori assumptions regarding shape the emission profile were made in this study. If only total column retrievals are used in the inversion, the estimated source term includes isolated emissions reaching up to 15 km. With plume height assimilation, the vertical distribution becomes more concentrated and also more consistent with the plume observed with the radar, which suggests that the vertical distribution SO_2 and ash emissions was mostly similar.

The centre of mass retrievals only partly constrain the vertical distribution, and hence some emission remains between 8 and 12 km, and the overestimation of plume height is reduced but not removed in the a posteriori simulations. However, given the about 1 km uncertainty in the IASI plume height retrievals and the 1 km assumed model uncertainty (Section 3.3) included into the observation errors, the inversion results for plume height are consistent with the assumptions of the inversion. No a priori assumptions regarding shape the emission profile were made in this study. The comparison with the IASI retrievals, CALIOP data and weather radar observations of the plume shows that the resulting vertical distributions were frequently in good agreement with the observations even if only total column retrievals were used in the inversion. The most notable exception were the emissions between 6 and 10 May, when the injection height was strongly overestimated, and although assimilating the plume height retrievals improved the agreement, the discrepancy was not fully resolved. Since the plume height retrievals are introduced as a weak constraint, a complete match between the inversion results and the observation data is not expected. However, some of the discrepancies remain too large to be explained by retrieval errors even together with the assumed model 1 km uncertainty.

Generally, two factors could lead to an inaccurate reconstruction of the vertical profile from the total column observations. First, the horizontal transport patterns on different altitudes might be too similar for resolving the vertical structure. Second, the simulated horizontal patterns might be too inaccurate due to errors or low resolution of the transport model or its input data. Since the inversion does not allow for systematic model errors, including the plume height retrievals in the inversion is expected to improve the vertical profile mainly in the first case. The discrepancy remaining between the observed and modelled plume heights suggests that model errors were at least partly responsible for the overestimation of injection heights on 6-10 May.

The main effect of assimilating the plume height retrievals was the reduction of emissions above 10-12 km. Although these emissions are not large compared to the total emission, this outcome has some qualitative significance, since without assimilation of plume heights, some emissions would be assigned above the tropopause. In addition to the data presented in the previous section, pPrevious studies based on Lidar lidar observations data (Ansmann et al., 2010) or, aircraft measurements (Schumann et al., 2011) or inverse modelling (Stohl et al., 2011) do not suggest significant injection above the 10 km altitude. However, these studies were mainly focused on volcanic ash instead of SO_2SO_2, and as shown by Thomas and Prata (2011) , ash and SO_2 were not always transported together. On the other hand, In contrast, tthe SO_2SO_2 plume height estimates derived from the GOME-2 satellite instrument by Rix et al. (2012) do indicate heights above 10 km and up to 13 km on 5th of May. However, Neither our data nor inverse modelling reproduces this result, as the plume heights

21

retrieved from IASI data are below 6 km for that day, which agrees with the modelled plume heights (not shown) even when only total column retrievals are included in the inversion.

Among the previous emissions estimates for Eyjafjallajökull, Flemming and Inness (2013) estimated a 0.25 Tg total SO_2 release using GOME-2 satellite retrievals, and 0.14 Tg using the OMI retrievals. Our estimates of 0.29-0.33 Tg are higher, especially compared OMI, but this is consistent with the higher total SO_2 burden estimated (Carboni et al., 2012) from the IASI data used in this study. Using the GOME-2 data, Flemming and Inness (2013) furthermore estimated SO_2 injection heights (defined as centres of 2-3 km thick layers) to mostly between 4 and 6 km above sea level with a peak reaching 10 km on May 19th. This agrees reasonably well with our mean profile (Figure 7), although contrary to our results without plume height assimilation, Flemming and Inness (2013) did not obtain the injection heights above 6 km on May 6th and 7th.

Boichu et al. (2013) used the IASI retrievals of Clarisse et al. (2012) to invert for temporally resolved SO_2 emissions of Eyjafjallajökull between May 1th and 12th, 2010, and estimated a total emission of about 0.17 Tg. Our inversion yielded for the same time 0.21 (total column and plume height retrievals) or 0.23 (total column only) Tg of SO_2. The larger total emission in our study might be due to assumptions regarding plume height in the IASI retrievals. The retrievals used by Boichu et al. (2013) assumed constant 7 km plume height, while the retrieved plume heights in this study were frequently lower especially near the volcano, and this would result in a higher retrieved values for the total column. For the emission, Boichu et al. (2013) assumed a constant injection height of 6 km, which turns out to coincide with the maximum of the mean profile (Figure 7) obtained in this study.

Stohl et al. (2011) determined the temporal and vertical distribution of volcanic ash emissions for the Eyjafjallajökull eruption with an inversion constrained by SEVIRI ash retrievals and an a priori source derived from plume top observations. Although the ash and SO_2 emissions cannot be compared quantitatively, the mean vertical profile obtained using ECMWF meteorological data (Fig. 2 in Stohl et al. (2011) is not very different from the one in Figure 7. In both profiles, the emissions are restricted mainly below 8 km and have maxima at 6 km.

Including the plume height retrievals in the inversion resulted in a total emission 12% lower than with total column retrievals only. Similar differences were observed in the experiments with synthetic data discussed in Section 4, where the inversion results were biased low by 15-20% using both plume height and total column retrievals and by only 2-10% using total columns only.

In ideal conditions, assimilating the plume height information should not affect the simulated total columns. However, adding a vertical constraint to the inversion can never improve the agreement for total columns, and in presence of realistic model uncertainty, a negative effect can be expected. The systematic tendency towards smaller emission may be caused by the regularisation, which penalises the quadratic norm of the solution. The synthetic experiments indicated that introducing the plume height retrievals did not allow relaxing the regularisation, since the optimal level (as identified from the parameter α^2) was similar with and without the plume height observations.

22

686　　On the other hand, the synthetic experiments also indicated that the estimation error for the total emission was only
687　moderately sensitive to the differences of the assumed source terms. The estimate for total emission was also robust with
688　regard to the vertical resolution, as halving the vertical resolution of the reconstruction (compare Figs. 5 and 6) resulted in
689　only minimal change in the total emission. The estimated total emission could, nevertheless, be affected by biases in the
690　satellite retrievals, or by model errors not exposed by the change of meteorological driver.

691　　α^2

692　　The experiments with synthetic data furthermore (Section 4) showed that the need for regularisation, or in Bayesian terms,
693　the need for a priori information, strongly depends on the situation. In addition, the need for regularisation, was stronglyor in
694　Bayesian terms, the need for a priori information, was strongly affected by uncertainty of the forward model., and Tthe
695　efforts needed to handle zero-valued observations in this and other studies (Boichu et al., 2013; Seibert et al., 2011) support
696　this conclusion. The errors arising from the dispersion model are likely to be correlated in space, and therefore, introducing
697　the corresponding non-diagonal elements in the error covariance matrix $\mathbf{R}$ could improve the inversion results.
698　The model errors resulted in noisy temporal and vertical emission profiles in the synthetic experiments and probably also in
699　the real inversions. However, the estimates for total emission were fairly robust regardless of the assumed source term or
700　perturbations to the forward model. Also, halving the vertical resolution of the reconstruction (compare Figs. 6 and 7)
701　resulted in only minimal change in the total emission. Nevertheless, the estimates of the total emission could be affected by
702　biases in the satellite retrievals, or by model errors not exposed by the change of meteorological driver.
703　　While the regularisation used in this work is equivalent to a zero-valued a priori source, a more informative a priori source
704　could be accommodated with a change of variables. Other forms of regularisation proposed for the volcanic source term
705　inversion include second-order temporal smoothing (Boichu et al., 2013), which also could be handled by truncated iteration
706　as discussed by Calvetti et al. (2002).
707　　The variational inversion method is computationally efficient if high temporal or vertical resolution is desired for the
708　reconstruction. In the current configuration, the reconstructed solution had formally 1360 degrees of freedom. Each iteration
709　consisting of one forward and one adjoint integration, the 25 iterations would require model integrations equivalent to about
710　about 1000 simulated daysdays to be simulated. In comparison, evaluating the matrix $\mathbf{HM}$ directly would require 1360
711　model integrations, and if the sensitivity was evaluated in windows of e.g. 72 hours, almost 4000 simulated days would be
712　required. The matrix-based approach is, however, more easily parallelised, while the parallelisation of the variational method
713　relies on the dispersion model. In our configuration, one iteration took about 5 minutes wall clock time on a 20-core node of
714　a Cray XC30 supercomputer.
715　　A drawback of the 4D-Var inversion method is that the a posteriori error covariance matrix for the source term is difficult
716　to evaluate. However, Monte Carlo techniques could be used to sample the a posteriori uncertainty.

8   Conclusions

We have presented an observation operator for retrievals of the vertical centre of mass of a tracer plume. The operator is based on transforming the centre of mass into first moment of mass using the retrieval of total column. The approach was tested by performing a source term inversion using both artificial data and the SO_2 retrievals from the IASI instrument during the Eyjafjallajökull eruption in May 2010. The inverse problem was solved with the 4D-Var method embedded into the SILAM dispersion model, and the truncated iteration is proposed as an efficient regularisation method for the 4D-Var inversion. Using both real and synthetic data, the 4D-Var method was shown to produce a similar solution as the more common algebraic method, but at lower computational cost.

The inversion results for Eyjafjallajökull were compared to radar based ash plume observations and CALIOP lidar profiles. The comparisons show that assimilating the plume height retrievals reduced the overestimation of injection height during individual periods of 1-3 days. However, for most of the simulated 21 days, the injection height was constrained by meteorological conditions and assimilation of the plume height retrievals had only small impact.

Assimilating the plume height retrievals reduced the vertical spread of the SO₂ injection. When the plume height was assimilated, about 85% of the 0.29 Tg total emission was below 8 km and about 95% was below 11 km. Compared to previous modelling studies (Boichu et al., 2013; Flemming and Inness, 2013), the total emission is 15-20% larger taking into account the differences in temporal coverage of the studies.

Introducing the plume height retrievals in the inversion may have an adverse effect on the estimated total emission. In the experiment with artificial observations, the inversions with only total column data had a negative bias of 2-10% which increased to 15-20% when the plume height observations were included. In the inversion for Eyjafjallajökull, performing the inversion using only total column retrievals resulted in ~15% larger total emission, which is consistent with the experiments with simulated observations.

The injection profile obtained by assimilating the plume height retrievals is more consistent with the radar and camera based observations of the ash plume.

The inverse problem was solved with the 4D-Var method embedded into the SILAM dispersion model. Truncated iteration is proposed as an efficient regularisation method for the 4D-Var inversion. Using both real and synthetic data, the 4D-Var method was shown to produce a similar solution as the more common algebraic method, but at considerably lower computational cost.

Experiments with both synthetic and real data suggest that the inversion is sensitive to errors in the forward model, and to their assumed uncertainty. Methods more robust to model errors are a topic suitable for future research.

Code availability

The source code for SILAM v5.3, including the data assimilation component, is available on request from the authors (julius.vira@fmi.fi, mikhail.sofiev@fmi.fi).

25

756

757

Appendix A: Additional comparisons with CALIOP data
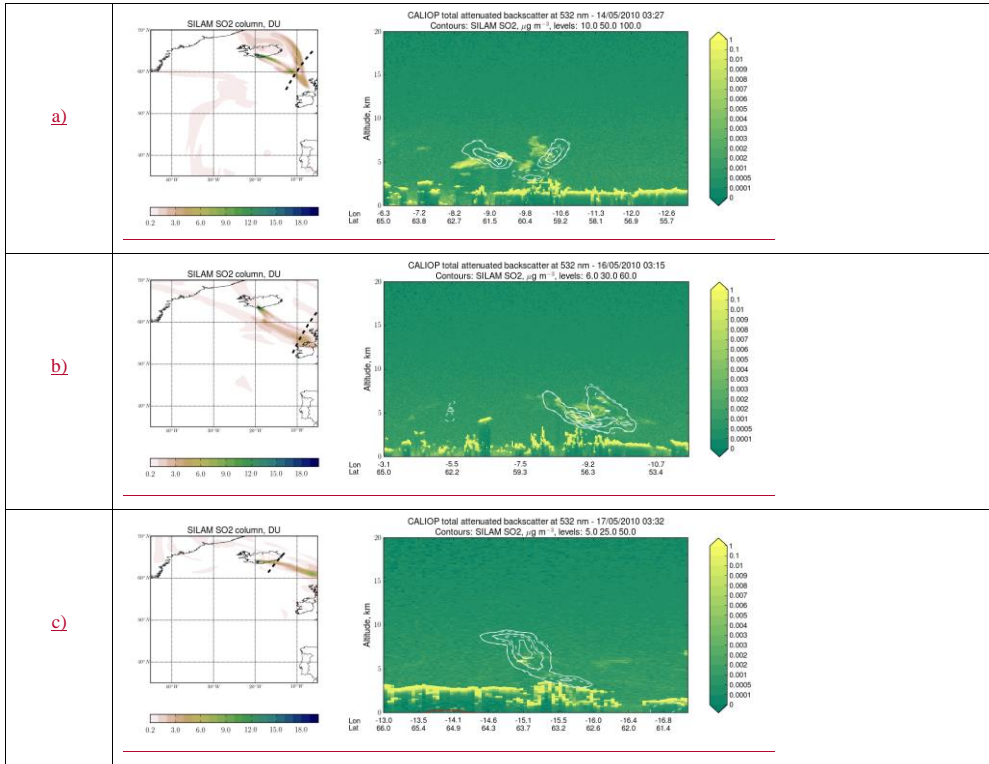
a)

b)

c)

759

Figure A1. Comparison of simulated SO$_2$ concentration (µg m^{-3}) compared to CALIOP total backscatter at 532 nm on 14 (panel a), 16 (b) and 17 (c) May, 2010. The inversion with only total column retrievals is shown in dashed contours. The contour levels 10, 50, and 100 in panel a, 6, 30 and 60 in panel b and 5, 25 and 50 in panel c.

763

Appendix B: moments of products of correlated Gaussian random variables

Let X and Y be scalar random variables with means and variances μ_X, μ_Y, σ_X^2 and σ_Y^2. Then, it follows from the definitions for variance and covariance that

(12) $$\mathrm{Var}[XY] = \sigma_X^2\sigma_Y^2 + \mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2 - 2\mu_X\mu_Y\mathrm{Cov}[X,Y] - \mathrm{Cov}[X,Y]^2 + \mathrm{Cov}[X^2,Y^2]$$

and

(13) $$\mathrm{Cov}[X,XY] = E[X^2]E[Y] + \mathrm{Cov}[X^2,Y] - E[X]E[XY] \ .$$

To expand $\mathrm{Cov}[X^2,Y^2]$ and $\mathrm{Cov}[X^2,Y]$ we assume that X and Y are normally distributed. We first define normalized auxiliary variables

(14) $$\tilde{X} = \frac{X - \mu_X}{\sigma_X}, \tilde{Y} = \frac{Y - \mu_Y}{\sigma_Y} \ .$$

Then, by expressing $\tilde{Y}$ as

(15) $$\tilde{Y} = c\tilde{X} + \sqrt{1 - c^2}\,\tilde{Z} \ ,$$

where $c = \mathrm{Cov}[\tilde{X},\tilde{Y}]$ and $\tilde{Z} \sim \mathcal{N}(0,1)$ independent of $\tilde{X}$, it is simple to verify that

(16) $$\mathrm{Cov}[\tilde{X}^2,\tilde{Y}^2] = 2c^2$$
$$\mathrm{Cov}[\tilde{X}^2,\tilde{Y}] = 0.$$

For the original random variables X and Y, we find by substituting (14) into the definition, expanding the terms, and using identities (16) that

(17) $$\mathrm{Cov}[X^2,Y^2] = 2\mathrm{Cov}[X,Y]^2 + 4\mu_X\mu_Y\mathrm{Cov}[X,Y]$$

and

(18) $$\mathrm{Cov}[X^2,Y] = 2\mu_X\mathrm{Cov}[X,Y] \ .$$

Formulas (5) and (6) now follow by combining Eqs. (17) and (18) with (12) and (13).

789

References

Ansmann, A., Tesche, M., Groß, S., Freudenthaler, V., Seifert, P., Hiebsch, A., Schmidt, J., Wandinger, U., Mattis, I., Müller, D., Wiegner, M., 2010. The 16 April 2010 major volcanic ash plume over central Europe: EARLINET lidar and AERONET photometer observations at Leipzig and Munich, Germany. Geophys. Res. Lett. 37, 1–5. doi:10.1029/2010GL043809

Arason, P., Petersen, G.N., Bjornsson, H., 2011. Observations of the altitude of the volcanic plume during the eruption of Eyjafjallajökull, April–May 2010. Earth Syst. Sci. Data 3, 9–17. doi:10.5194/essdd-4-1-2011

Bernard, A., Rose, W.I., 1990. The injection of sulfuric acid aerosols in the stratosphere by the El Chichón volcano and its related hazards to the international air traffic. Nat. Hazards 3, 59–67. doi:10.1007/BF00144974

Boichu, M., Clarisse, L., 2014. Improving volcanic sulfur dioxide cloud dispersal forecasts by progressive assimilation of satellite observations. Geophys. Res. Lett. 41, 2637–2643. doi:10.1002/2014GL059496.Abstract

Boichu, M., Menut, L., Khvorostyanov, D., Clarisse, L., Clerbaux, C., Turquety, S., Coheur, P.-F., 2013. Inverting for volcanic ~~SO2~~SO_2 flux at high temporal resolution using spaceborne plume imagery and chemistry-transport modelling: the 2010 Eyjafjallajökull eruption case study. Atmos. Chem. Phys. 13, 8569–8584. doi:10.5194/acp-13-8569-2013

Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput. 16, 1190–1208.

Calvetti, D., Lewis, B., Reichel, L., 2002. GMRES, L-curves, and discrete ill-posed problems. BIT Numer. Math. 42, 44–65.

Carboni, E., Grainger, R., Walker, J., Dudhia, a., Siddans, R., 2012. A new scheme for sulphur dioxide retrieval from IASI measurements: application to the Eyjafjallajökull eruption of April and May 2010. Atmos. Chem. Phys. 12, 11417–11434. doi:10.5194/acp-12-11417-2012

Carboni, E., Grainger, R.G., Mather, T.A., Pyle, D.M., Thomas, G., Siddans, R., Smith, A., Dudhia, A., Koukouli, M.L., Balis, D., 2016. The vertical distribution of volcanic ~~SO2~~SO_2 plumes measured by IASI. Atmos. Chem. Phys. 16, 4343–4367. doi:doi:10.5194/acp-16-4343-2016

Carn, S.A., Krueger, A.J., Krotkov, N.A., Yang, K., Evans, K., 2009. Tracking volcanic sulfur dioxide clouds for aviation hazard mitigation. Nat. Hazards 51, 325–343. doi:10.1007/s11069-008-9228-4

Castellanos, J.L., Gómez, S., Guerra, V., 2002. The triangle method for finding the corner of the L-curve ☆. Appl. Numer. Math. 43, 359–373.

Clarisse, L., Hurtmans, D., Clerbaux, C., Hadji-Lazaro, J., Ngadi, Y., Coheur, P.-F., 2012. Retrieval of sulphur dioxide from the infrared atmospheric sounding interferometer (IASI). Atmos. Meas. Tech. 5, 581–594.

Dacre, H.F., Grant, A.L.M., Harvey, N.J., Thomson, D.J., Webster, H.N., Marenco, F., 2014. Volcanic ash layer depth: Processes and mechanisms. Geophys. Res. Lett. 42. doi:10.1002/2014GL062454

Dee, D.P., Uppala, S.M., Simmons, a. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. a., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, a. C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, a. J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, a. P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q. J. R. Meteorol. Soc. 137, 553–597. doi:10.1002/qj.828

Eckhardt, S., Prata, A.J., Seibert, P., Stebel, K., Stohl, A., 2008. Estimation of the vertical profile of sulfur dioxide injection

828   into the atmosphere by a volcanic eruption using satellite column measurements and inverse transport modeling.
829   Atmos. Chem. Phys. 8, 3881–3897. doi:10.5194/acpd-8-3761-2008

830   Elbern, H., Schmidt, H., Talagrand, O., Ebel, a., 2000. 4D-variational data assimilation with an adjoint air quality model for
831   emission analysis. Environ. Model. Softw. 15, 539–548. doi:10.1016/S1364-8152(00)00049-9

832   Elbern, H., Strunk, A., Schmidt, H., Talagrand, O., 2007. Emission rate and chemical state estimation by 4-dimensional
833   variational inversion. Atmos. Chem. Phys. 7, 3749–3769. doi:10.5194/acpd-7-1725-2007

834   Engl, H.W., Hanke, M., Neubauer, A., 2000. Regularization of Inverse Problems, Mathematics and Its Applications.
835   Springer Netherlands.

836   Eskes, H.J., Boersma, K.F., 2003. Averaging kernels for DOAS total-column satellite retrievals. Atmos. Chem. Phys. 3,
837   1285–1291.

838   Fleming, H.E., 1990. Equivalence of regularization and truncated iteration in the solution of Ill-posed image reconstruction
839   problems. Linear Algebra Appl. 130, 133–150. doi:10.1016/0024-3795(90)90210-4

840   Flemming, J., Inness, A., 2013. Volcanic sulfur dioxide plume forecasts based on UV-satellite retrievals for the 2011
841   Grímsvötn and the 2010 Eyjafjallajökull eruption. J. Geophys. Res. Atmos. 118. doi:10.1002/jgrd.50753

842   Fromm, M., Kablick, G., Nedoluha, G., Carboni, E., Grainger, R., Campbell, J., Lewis, J., 2014. Correcting the record of
843   volcanic stratospheric aerosol impact: Nabro and Sarychev Peak. J. Geophys. Res. Atmos. 119, 10,343-10,364.
844   doi:10.1002/2014JD021507

845   Gudmundsson, M.T., Thordarson, T., Höskuldsson, A., Larsen, G., Björnsson, H., Prata, F.J., Oddsson, B., Magnússon, E.,
846   Högnadóttir, T., Petersen, G.N., Hayward, C.L., Stevenson, J. a, Jónsdóttir, I., 2012. Ash generation and distribution
847   from the April-May 2010 eruption of Eyjafjallajökull, Iceland. Sci. Rep. 2. doi:10.1038/srep00572

848   Hansen, P.C., 2010. Discrete Inverse Problems: Insight and Algorithms, Fundamentals of Algorithms. Society for Industrial
849   and Applied Mathematics.

850   Hansen, P.C., 1992. Analysis of Discrete Ill-Posed Problems by Means of the L-Curve. SIAM Rev. 34, 561–580.
851   doi:10.1137/1034115

852   Kaipio, J., Somersalo, E., 2006. Statistical and Computational Inverse Problems, Applied Mathematical Sciences. Springer
853   New York.

854   Kilmer, M.E., O'Leary, D.P., 2001. Choosing Regularization Parameters in Iterative Methods for Ill-Posed Problems. SIAM
855   J. Matrix Anal. Appl. 22, 1204–1221. doi:10.1137/S0895479899345960

856   Koukouli, M.E., Clarisse, L., Carboni, E., van Gent, J., Spinetti, C., Balis, D., Dimopuolos, S., Grainger, R., Theys, N.,
857   Tampellini, L., Zehner, C., 2014. Intercomparison of Metop-A SO_2 measurements during the 2010-2011 Icelandic
858   eruptions. Ann. Geophys. Fast Track. doi:10.4401/ag-6613

859   Kristiansen, N.I., Stohl, A., Prata, A.J., Richter, A., Eckhardt, S., Seibert, P., Hoffmann, A., Ritter, C., Bitar, L., Duck, T.J.,
860   Stebel, K., 2010. Remote sensing and inverse transport modeling of the Kasatochi eruption sulfur dioxide cloud. J.
861   Geophys. Res. 115, 1–18. doi:10.1029/2009JD013286

862   Le Dimet, F.-X., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations:
863   theoretical aspects. Tellus A 38A, 97–110. doi:10.1111/j.1600-0870.1986.tb00459.x

864   Liu, Z., Vaughan, M., Winker, D., Kittaka, C., Getzewich, B., Kuehn, R., Omar, A., Powell, K., Trepte, C., Hostetler, C.,
865   2009. The CALIPSO lidar cloud and aerosol discrimination: Version 2 algorithm and initial assessment of
866   performance. J. Atmos. Ocean. Technol. 26, 1198–1213. doi:10.1175/2009JTECHA1229.1

867   Lu, S., Lin, H.X., Heemink, A.W., Fu, G., Segers, A.J., 2016. Estimation of Volcanic Ash Emissions Using Trajectory-

868   Based 4D-Var Data Assimilation. Mon. Weather Rev. 144, 575–589. doi:10.1175/MWR-D-15-0194.1

869   Mastin, L.G., Guffanti, M., Servranckx, R., Webley, P., Barsotti, S., Dean, K., Durant, a., Ewert, J.W., Neri, a., Rose, W.I.,
870   2009. A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and
871   dispersion during eruptions. J. Volcanol. Geotherm. Res. 186, 10–21. doi:10.1016/j.jvolgeores.2009.01.008

872   Petersen, G.N., Bjornsson, H., Arason, P., 2012a. The impact of the atmosphere on the Eyjafjallajökull 2010 eruption plume.
873   J. Geophys. Res. Atmos. 117, 1–14. doi:10.1029/2011JD016762

874   Petersen, G.N., Bjornsson, H., Arason, P., von Löwis, S., 2012b. Two weather radar time series of the altitude of the
875   volcanic plume during the May 2011 eruption of Grímsvötn, Iceland. Earth Syst. Sci. Data 4, 121–127.
876   doi:10.5194/essdd-5-281-2012

877   Rix, M., Valks, P., Hao, N., Loyola, D., Schlager, H., Huntrieser, H., Flemming, J., Koehler, U., Schumann, U., Inness, A.,
878   2012. Volcanic ~~SO2~~SO_2 , BrO and plume height estimations using GOME-2 satellite measurements during the
879   eruption of Eyjafjallajökull in May 2010. J. Geophys. Res. Atmos. 117. doi:10.1029/2011JD016718

880   Robock, A., 2000. Volcanic eruptions and climate. Rev. Geophys. 38, 191–219.

881   Santos, R.J., 1996. Equivalence of regularization and truncated iteration for general ill-posed problems. Linear Algebra
882   Appl. 236, 25–33.

883   Schmidt, A., Leadbetter, S., Theys, N., Carboni, E., Witham, C.S., Stevenson, J.A., Birch, C.E., Thordarson, T., Turnock, S.,
884   Barsotti, S., Delaney, L., Feng, W., Grainger, R.G., Hort, M.C., Höskuldsson, Á., 2015. Satellite detection, long-range
885   transport and air quality impacts of sulfur dioxide from the 2014-2015 flood lava eruption at Bárðarbunga (Iceland). J.
886   Geophys. Res. Atmos. 120, 9739–9757. doi:10.1002/2015JD023638.Received

887   Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, a., Forster, C., Baumann, R., Sailer, T., Graf, K.,
888   Mannstein, H., Voigt, C., Rahm, S., Simmet, R., Scheibe, M., Lichtenstern, M., Stock, P., Rüba, H., Schäuble, D.,
889   Tafferner, a., Rautenhaus, M., Gerz, T., Ziereis, H., Krautstrunk, M., Mallaun, C., Gayet, J.-F., Lieke, K., Kandler, K.,
890   Ebert, M., Weinbruch, S., Stohl, a., Gasteiger, J., Groß, S., Freudenthaler, V., Wiegner, M., Ansmann, a., Tesche, M.,
891   Olafsson, H., Sturm, K., 2011. Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space
892   closure in April and May 2010. Atmos. Chem. Phys. 11, 2245–2279. doi:10.5194/acp-11-2245-2011

893   Seibert, P., Kristiansen, N.I., Richter, A., Eckhardt, S., Prata, A.J., Stohl, A., 2011. Uncertainties in the inverse modelling of
894   sulphur dioxide eruption profiles. Geomatics, Nat. Hazards Risk 2, 201–216. doi:10.1080/19475705.2011.590533

895   Sofiev, M., 2000. A model for the evaluation of long-term airborne pollution transport at regional and continental scales.
896   Atmos. Environ. 34, 2481–2493. doi:10.1016/S1352-2310(99)00415-X

897   Sofiev, M., Vira, J., Kouznetsov, R., Prank, M., Soares, J., Genikhovich, E., 2015. Construction of an Eulerian atmospheric
898   dispersion model based on the advection algorithm of M . Galperin : dynamic cores. Geosci. Model Dev. Discuss. 8,
899   2905–2947. doi:10.5194/gmdd-8-2905-2015

900   Spinetti, C., Salerno, G.G., Caltabiano, T., Carboni, E., Clarisse, L., Corradini, S., Grainger, R.G., Hedelt, P.A., Koukouli,
901   M.E., Merucci, L., Siddans, R., Tampellini, L., Theys, N., Valks, P., Zehner, C., 2014. Volcanic ~~SO2~~SO_2 by UV–TIR
902   satellite retrievals: validation by using ground-based network at Mt. Etna. Ann. Geophys. Fast Track. doi:10.4401/ag-
903   6641

904   Spivakovsky, C.M., Logan, J.A., Montzka, S.A., Balkanski, Y.J., Foreman-Fowler, M., Jones, D.B.A., Horowitz, L.W.,
905   Fusco, A.C., Brenninkmeijer, C.A.M., Prather, M.J., Wofsy, S.C., McElroy, M.B., 2000. Three-dimensional
906   climatological distribution of tropospheric OH: Update and evaluation. J. Geophys. Res. 105, 8931–8980.

907   Stohl, A., Prata, A.J., Eckhardt, S., Clarisse, L., Durant, A., Henne, S., Kristiansen, N.I., Minikin, A., Schumann, U., Seibert,
908   P., Stebel, K., Thomas, H.E., Thorsteinsson, T., Tørseth, K., Weinzierl, B., 2011. Determination of time- and height-
909   resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull

910   eruption. Atmos. Chem. Phys. 11, 4333–4351. doi:10.5194/acp-11-4333-2011

911   Theys, N., Campion, R., Clarisse, L., Brenot, H., van Gent, J., Dils, B., Corradini, S., Merucci, L., Coheur, P.-F., Van
912   Roozendael, M., Hurtmans, D., Clerbaux, C., Tait, S., Ferrucci, F., 2013. Volcanic ~~SO2~~SO$_2$ fluxes derived from
913   satellite data : a survey using OMI , GOME-2 , IASI and MODIS. Atmos. Chem. Phys. 13, 5945–5968.
914   doi:10.5194/acp-13-5945-2013

915   Thomas, H.E., Prata, A.J., 2011. Sulphur dioxide as a volcanic ash proxy during the April–May 2010 eruption of
916   Eyjafjallajökull Volcano, Iceland. Atmos. Chem. Phys. 11, 6871–6880. doi:10.5194/acp-11-6871-2011

917   Tikhonov, A., 1963. Solution of Incorrectly Formulated Problems and the Regularization Method. Sov. Math. Dokl. 5,
918   1035–1038.

919   Vira, J., Sofiev, M., 2012. On variational data assimilation for estimating the model initial conditions and emission fluxes for
920   short-term forecasting of SOx concentrations. Atmos. Environ. 46, 318–328. doi:10.1016/j.atmosenv.2011.09.066

921   Walker, J.C., Carboni, E., Dudhia, a, Grainger, R.G., 2012. Improved detection of sulphur dioxide in volcanic plumes using
922   satellite-based hyperspectral infrared measurements: Application to the Eyjafjallajökull 2010 eruption. J. Geophys.
923   Res. Atmos. 117, n/a-n/a. doi:10.1029/2011JD016810

924   Wang, J., Park, S., Zeng, J., Ge, C., Yang, K., Carn, S., Krotkov, N., Omar, a. H., 2013. Modeling of 2008 Kasatochi
925   volcanic sulfate direct radiative forcing: assimilation of OMI ~~SO2~~SO$_2$ plume height data and comparison with MODIS
926   and CALIOP observations. Atmos. Chem. Phys. 13, 1895–1912. doi:10.5194/acp-13-1895-2013

927   Wilkins, K.L., Watson, I.M., Kristiansen, N.I., Webster, H.N., Thomson, D.J., Dacre, H.F., Prata, A.J., 2015. Using data
928   insertion with the NAME model to simulate the 8 May 2010 Eyjafjallajökull volcanic ash cloud. J. Geophys. Res.
929   Atmos. 121. doi:10.1002/2015JD023895

930   Winker, D.M., Liu, Z., Omar, A., Tackett, J., Fairlie, D., 2012. CALIOP observations of the transport of ash from the
931   Eyjafjallajökull volcano in April 2010. J. Geophys. Res. Atmos. 117, 1–12. doi:10.1029/2011JD016499

932   Winker, D.M., Vaughan, M.A., Omar, A., Hu, Y., Powell, K.A., Liu, Z., Hunt, W.H., Young, S.A., 2009. Overview of the
933   CALIPSO mission and CALIOP data processing algorithms. J. Atmos. Ocean. Technol. 26, 2310–2323.
934   doi:10.1175/2009JTECHA1281.1

935   Zehner, C. (Ed.), 2012. Monitoring volcanic ash from space. ESA-EUMETSAT workshop on the 14 April to 23 May 2010
936   eruption at the Eyjafjöll volcano, South Iceland (ESA/ESRIN, 26-27 May 2010). ESA Publication STM-280.

937

938

939

940
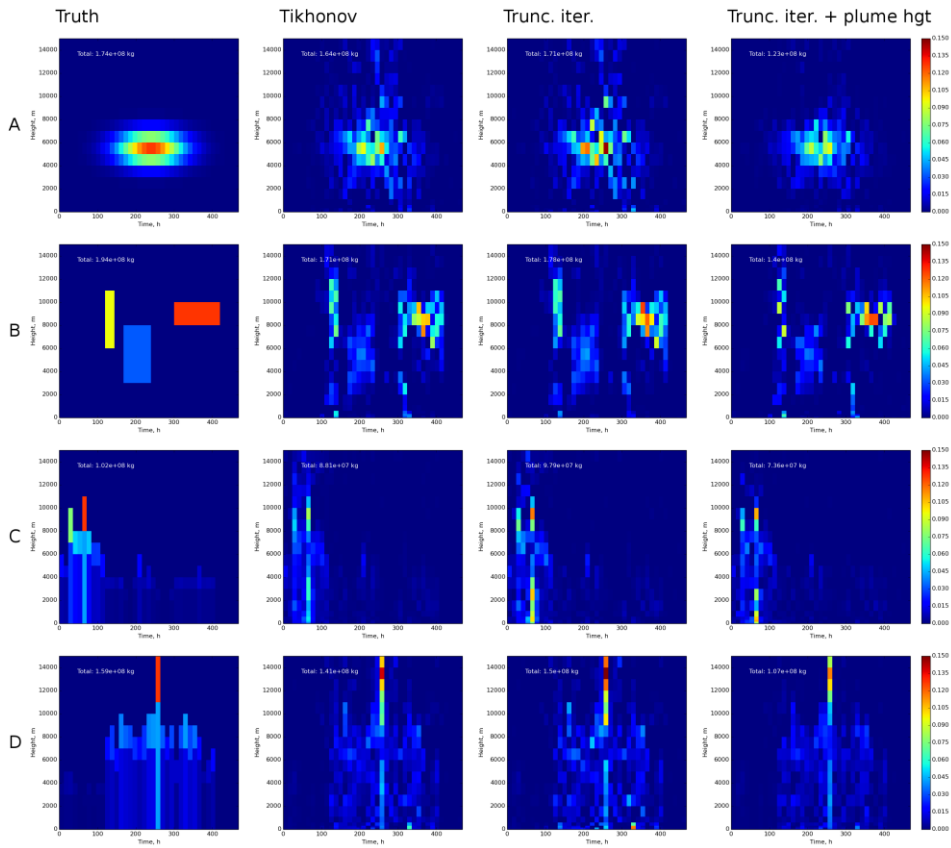
Figure 11. Estimated emission flux (kg m⁻¹ s⁻¹) in source term inversions with simulated data. True source terms for the four cases (A to D) are shown in the left column. ~~Solutions using truncated iteration are shown in the middle column, solutions using Tikhonov regularisation are shown in the right column.~~ The remaining columns show the inversion results using Tikhonov regularisation, using truncated iteration with total column data, and using truncated iteration with total column and plume height data.

946

947

948 **Figure 2~~2~~. Estimated emission flux with synthetic data: inversion results for the case B in ~~Figure 1~~Figure 1 assuming a perfect**
949 **forward model.**

950 Table 1. Bias and RMSE with respect to the true source term (case A…D) in experiments with synthetic data with assimilation of
951 total column (TC) and total column and ~~centre of mass~~plume height (TC+CM). Values are shown for both optimal regularisation
952 (regularisation parameter or iteration number with the lowest RMSE) and for the regularisation chosen from L-curve. Relative
953 bias is defined as the difference between estimated and true total emission divided by the true total emission.

Case		Tikhonov regularisation				Truncated iteration			
		RMSE		Relative bias		RMSE		Relative bias	
		Optimal	L-curve	Optimal	L-curve	Optimal	L-curve	Optimal	L-curve
A	TC	48.0	48.0	-5 %	-5 %	45.2	51.2	-3 %	-2 %
	TC+CM	39.8	39.8	-19 %	-19 %	36.5	36.7	-17 %	-17 %
B	TC	65.1	65.6	-8 %	-12 %	61.4	61.9	-8 %	-8 %
	TC+CM	59.3	60.2	-18 %	-23 %	56.9	58.4	-18 %	-17 %
C	TC	21.1	21.1	-13 %	-13 %	20.6	21.9	-8 %	-4 %
	TC+CM	18.5	18.6	-20 %	-24 %	17.8	18.1	-17 %	-17 %
D	TC	32.4	33.6	-15 %	-11 %	31.1	38.0	-8 %	-6 %
	TC+CM	29.3	29.5	-27 %	-24 %	27.3	28.0	-24 %	-21 %

954

A

B

TC and CM — TC and CM, Tikhonov reg.
TC only — TC only, Tikhonov reg.

957

Figure 33. L-curve (left) and RMS error (right) for inversions with simulated data for cases A and B in Figure 1Figure 1. The
iterate (for truncated iteration) or the regularisation parameter (for Tikhonov regularisation) chosen from the L-curve is marked
with a star.

Formatted: Not Highlight
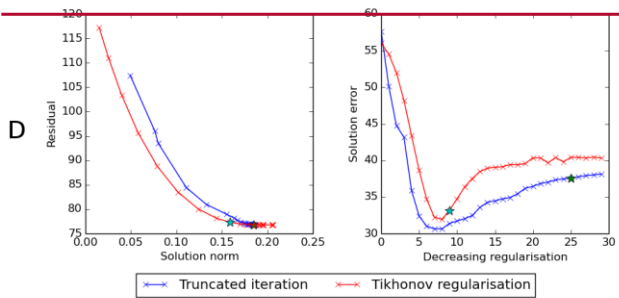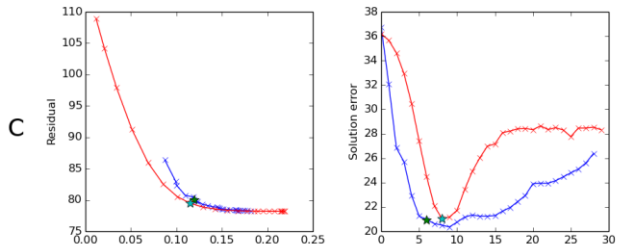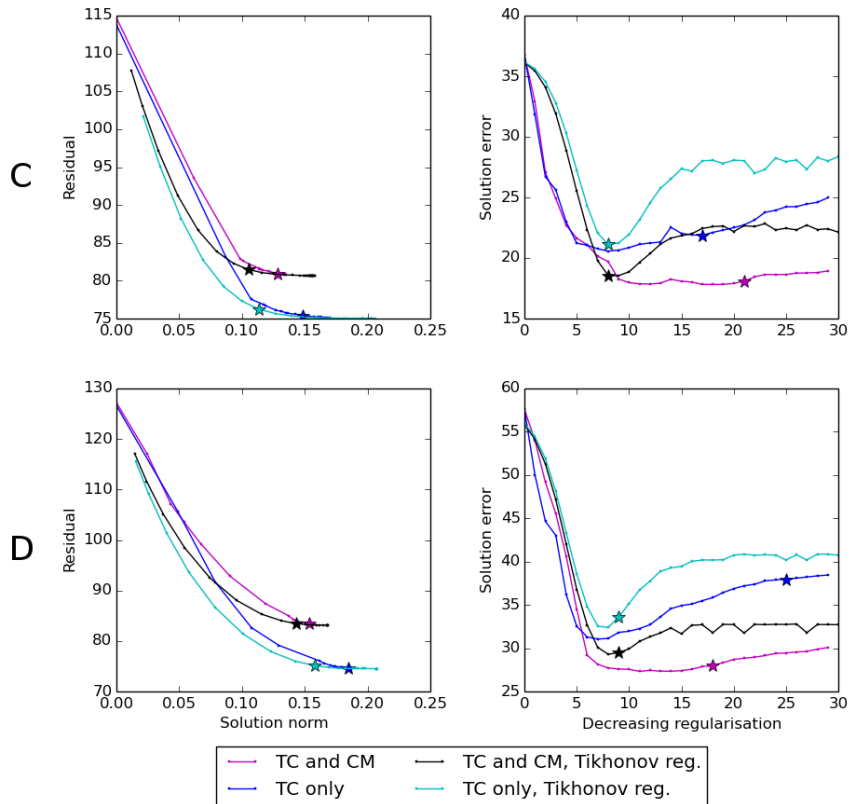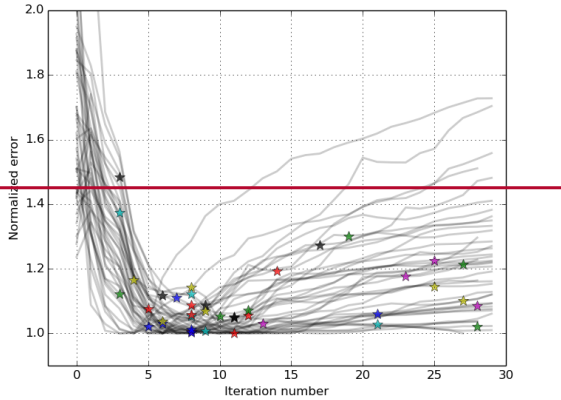
Figure 44. L-curve (left) and RMS error (right) for inversions with simulated data for cases C and D in Figure 1Figure 1. The iterate (for truncated iteration) or regularisation parameter (for Tikhonov regularisation) chosen from the L-curve is marked with a star.

967

968
969
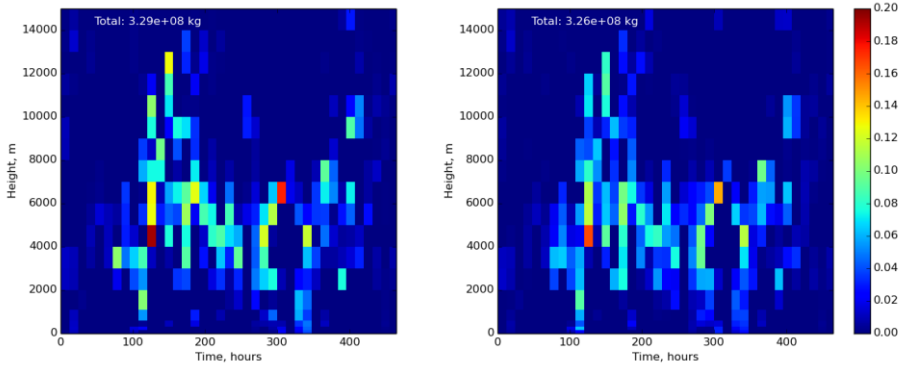
970 **Figure 5̶6. Inversion results with real observations: emission flux (kg m^{-1} s^{-1}) obtained using 4D-Var (left) and by evaluating the**
971 **sensitivity matrix (right). The inversions are based on total column observations.**
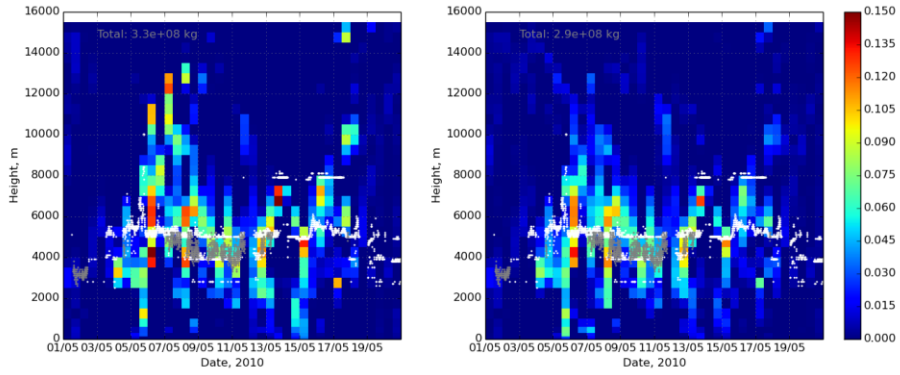
972
973
974   **Figure 67. Inversion results for Eyjafjallajökull. Left: emission flux (kg m^{-1} s^{-1}) with assimilation of column mass only. Right:**
975   **assimilation of column mass and plume height with full observation error covariance matrix. White dots denote plume height**
976   **observations by radar, grey dots denote plume height observations with a camera.**
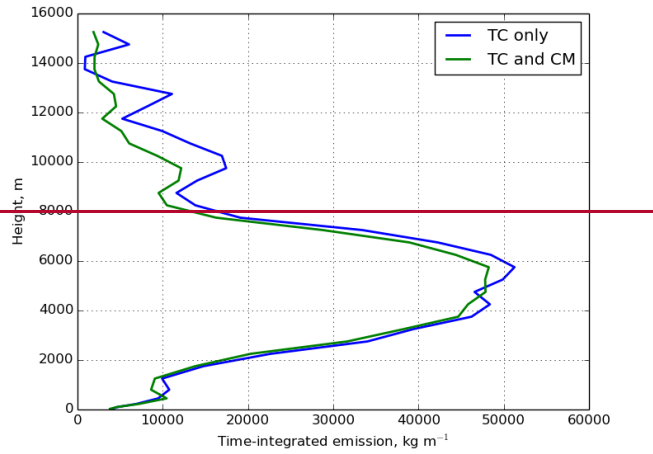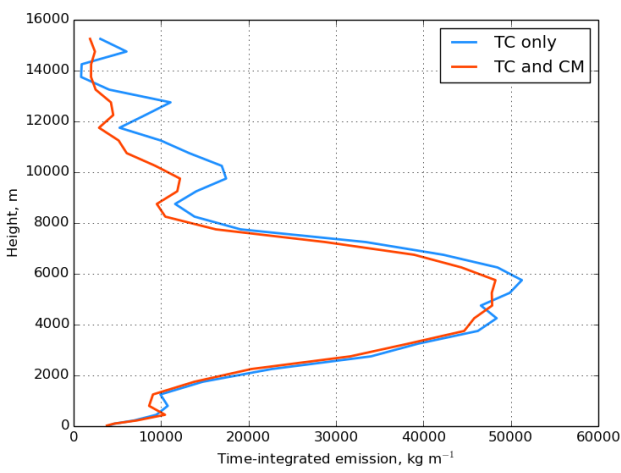
979 **Figure 7~~8~~. Time-integrated emission of ~~SO₂~~SO₂ (kg m⁻¹) during the simulated period as function of height (m) for the source term**
980 **inversions with (~~green~~red) and without (blue) plume height assimilation.**
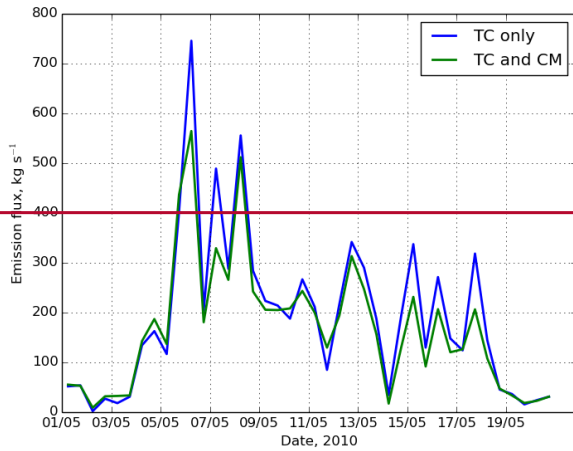
982

983　**Figure 9. Estimated SO₂ emission flux (kg s⁻¹) as function of time with (green) and without (blue) assimilation of plume height**
984　**retrievals.**
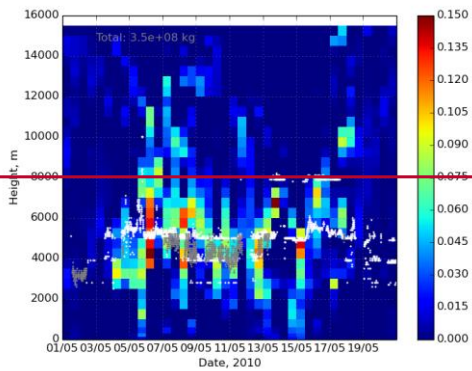
985

986　Figure 10. Inversion results for Eyjafjallajökull: emission flux (kg m⁻¹ s⁻¹) with assimilation of column density and plume

987　height but neglecting off diagonal elements in the observation error covariance matrix. White dots denote plume height

988　observations by radar, grey dots denote plume height observations with a camera.
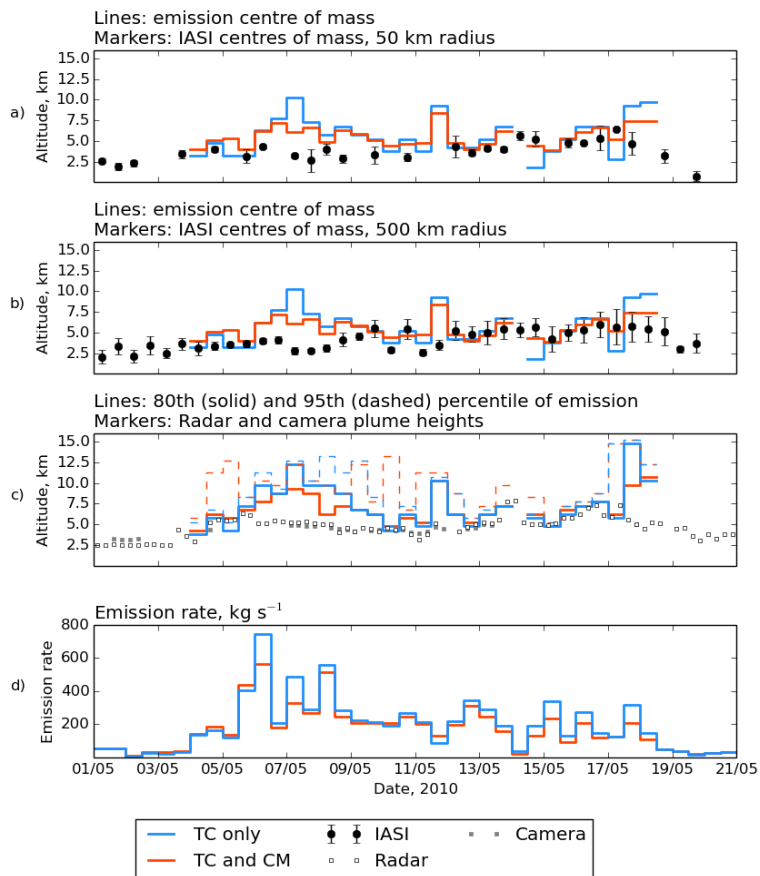
Figure 8. Inversion results for Eyjafjallajökull. Panels a and b: centre of mass of SO$_2$ injection and average IASI plume height within 50 and 500 km from the volcano; panel c: 95th and 80th percentiles of SO$_2$ injection and the plume top altitudes observed by radar and camera; panel d: estimated emission rate (kg s^{-1}). Inversions using only total column retrievals are plotted in blue; inversions using total column and plume height retrievals are plotted in red. Fully correlated errors are assumed for evaluating the error bars for IASI data. The data with retrieval error estimate larger than 5 km not included. The radar and camera observations are averaged to time steps of 6 hours. The centres of mass and percentiles of the inversion results are evaluated for the 12 hour steps emitting at least 1% of the total emission. All altitudes are above sea level.

2010-05-05 22:00   2010-05-06 22:00   2010-05-07 22:00

MODEL

2010-05-05 18:00 -->   2010-05-06 18:00 -->   2010-05-07 18:00 -->
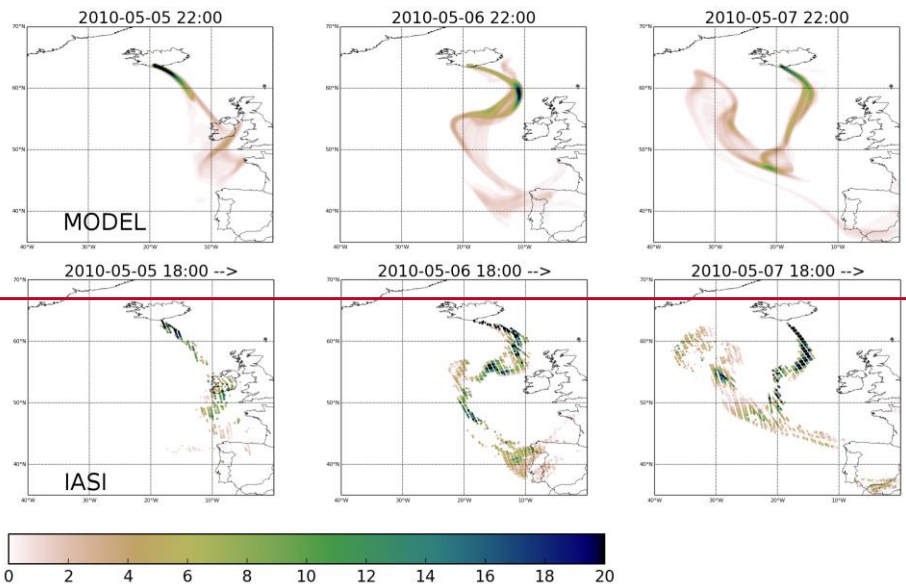
IASI

1001
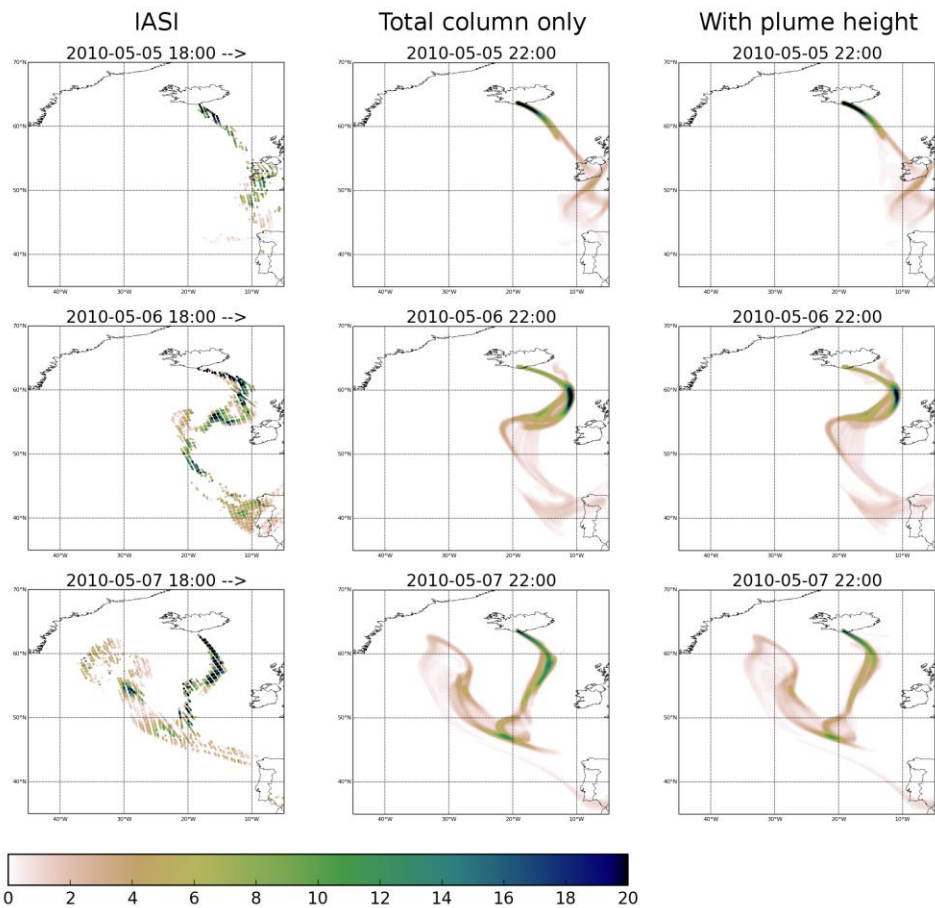
43

Figure 11. SO₂ column loading (DU) for the a posteriori simulation with assimilation of plume height (top) and for the IASI column retrievals (bottom row). Results for 5, 6 and 7th May, 2010 are shown in the columns from left to right. The evening overpasses are shown for IASI, the model fields are valid at 22 UTC.

Figure 9. SO₂ column loading (DU) for the IASI column retrievals (left column), for the a posteriori simulation with assimilation of total column only (middle) and with assimilation of total column and plume height retrievals. Results for 5, 6 and 7th May, 2010 are shown in the rows from top to bottom. The evening overpasses are shown for IASI, the model fields are valid at 22 UTC.

Formatted: Normal

1011

1012
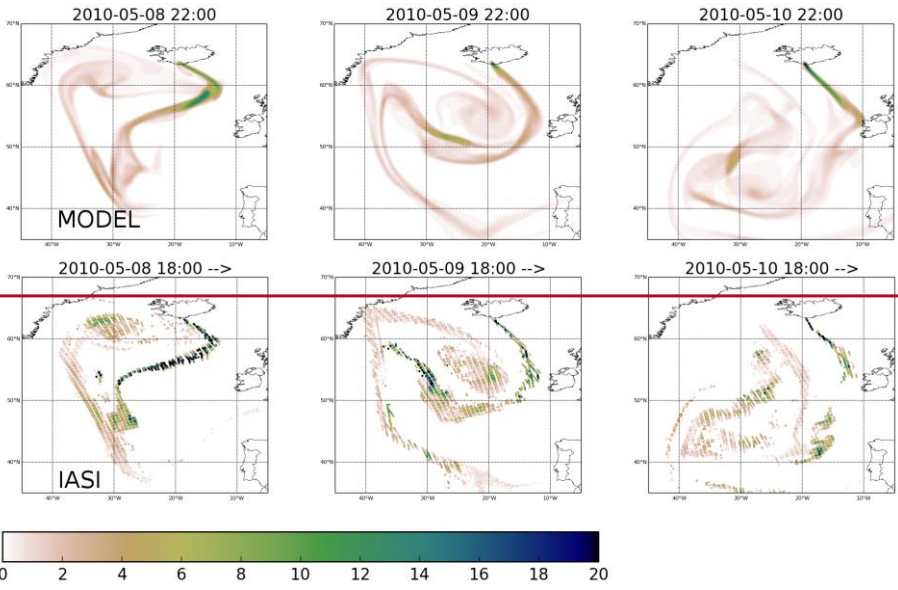
1013

1014   Figure 12. As Figure 11, but 8-10 May, 2010.
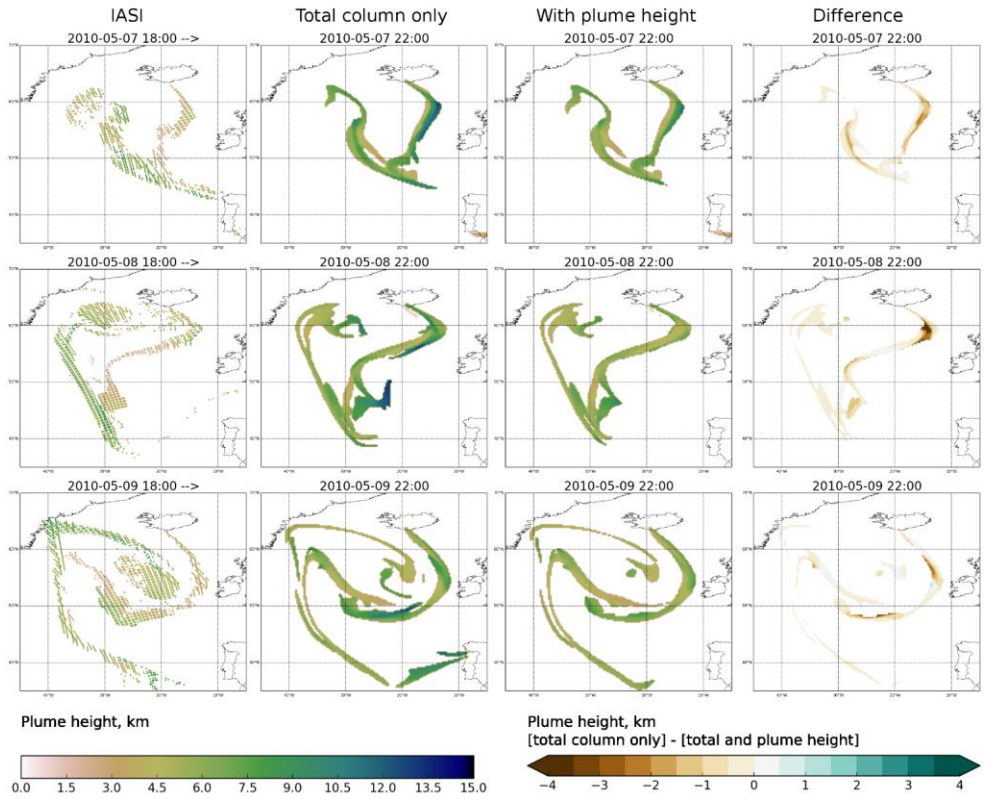
Figure 10. Retrieved SO$_2$ plume height (km, left column) and the simulated plume height (as centre of mass) without and with assimilation of plume height retrievals for 7-9 (top to bottom row) May, 2010. The difference (without plume height – with plume height) of the simulations is shown in the rightmost column.
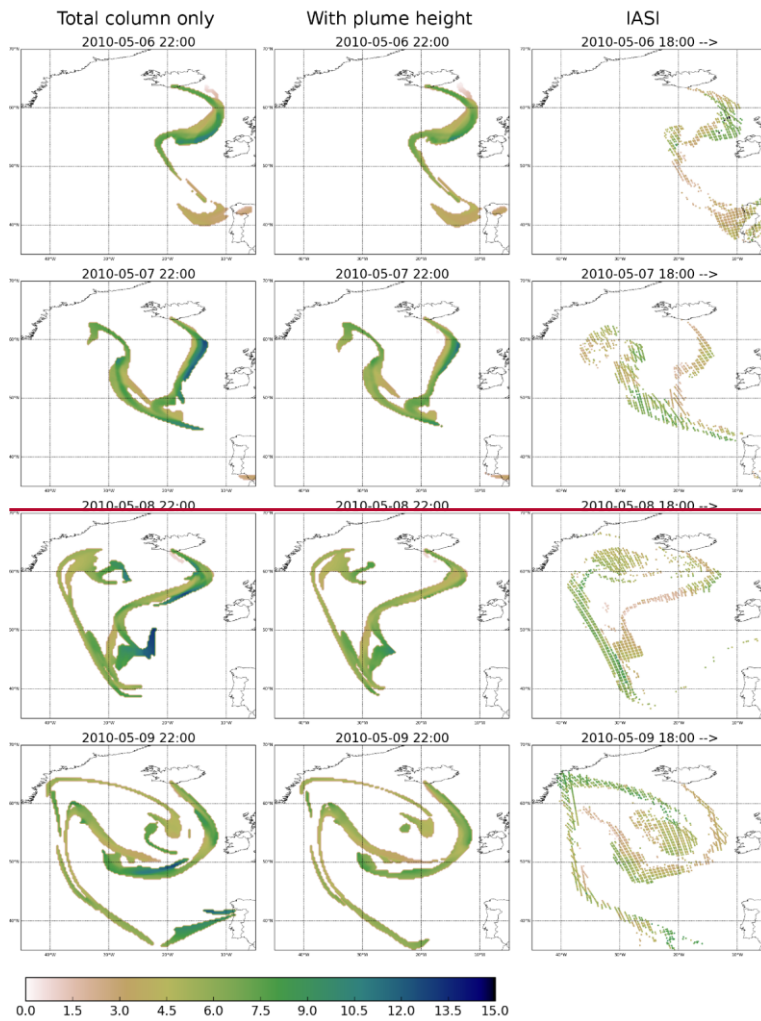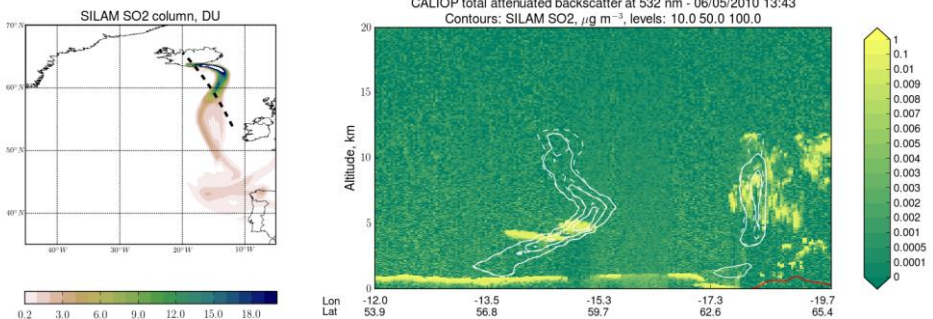
Figure 13. Simulated SO₂ plume height (centre of mass, km) without (left) and with (middle) assimilation of plume height retrievals for 6-9 (top to bottom row) May, 2010. The corresponding IASI retrievals are shown in the right column.

Formatted: Indent: First line: 0 cm

47

1026
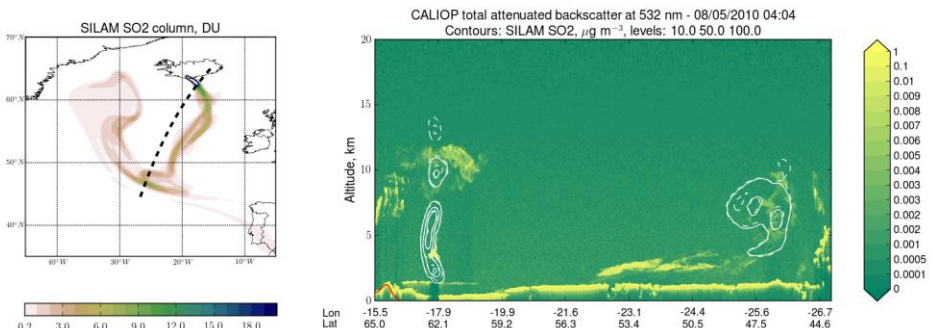
1028

Figure 11. Comparison of simulated SO$_2$ with CALIOP data for 14 UTC on 6 May (top) and 04 UTC on 8 May, 2010 (bottom). Left: the simulated SO$_2$ total column (DU, with assimilation of both total column and plume height) with the CALIPSO track plotted with dashed line. Right: CALIOP total attenuated backscatter at 532 nm with the simulated SO$_2$ concentration represented by contours. The solid contours correspond to assimilation of both total column and plume height, the dashed contours correspond to assimilation of total column only. The contour levels are 10, 50 and 100 µg m^{-3}.

1034
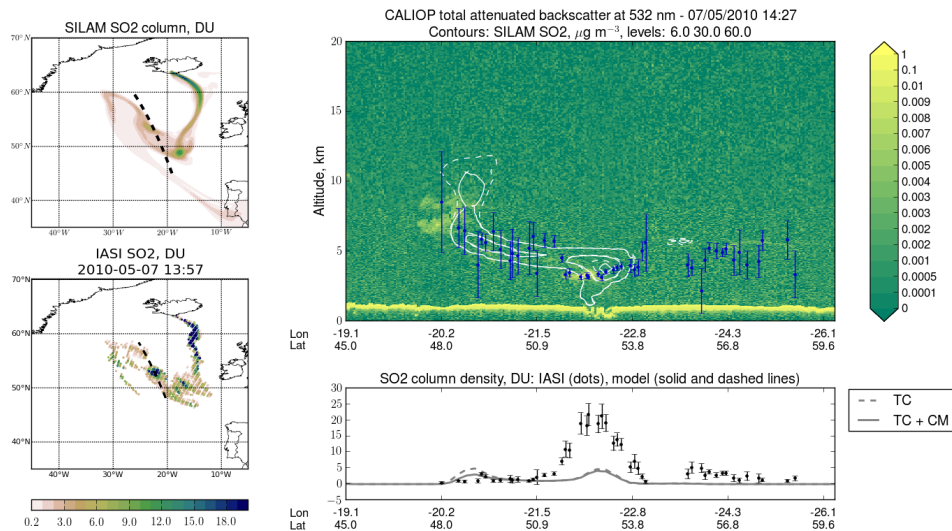
Figure 12. CALIOP total attenuated backscatter, simulated SO₂ concentration (contour levels indicated on the figure title) and collocated IASI plume height retrievals at ~14 UTC on May 7, 2010. The solid lines and contours correspond to inversion using total column and plume height retrievals, dashed lines and contours correspond to inversion using total column retrievals only. The modelled and retrieved column densities are shown in maps on the left and as a 1D plot along the CALIOP track on the bottom. The full CALIOP track segment is marked in the map of simulated SO₂ columns (top-left), the track segment where the collocated IASI data are extracted is shown in the map of retrieved SO₂ columns (bottom-left). The model SO₂ columns shown in the map are from the inversion using both total column and plume height retrievals.
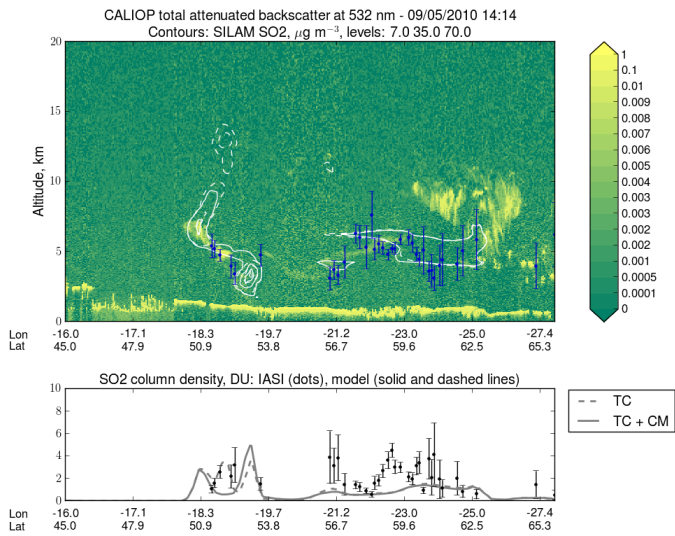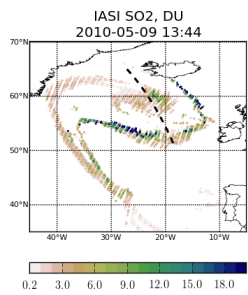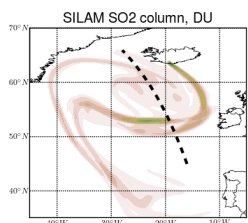
Formatted: Caption

49

CALIOP total attenuated backscatter at 532 nm - 09/05/2010 14:14
Contours: SILAM SO2, μg m^{-3}, levels: 7.0 35.0 70.0

SILAM SO2 column, DU

IASI SO2, DU
2010-05-09 13:44

SO2 column density, DU: IASI (dots), model (solid and dashed lines)

1044

1045   **Figure 13. As Figure 12 but for May 9, 2010.**
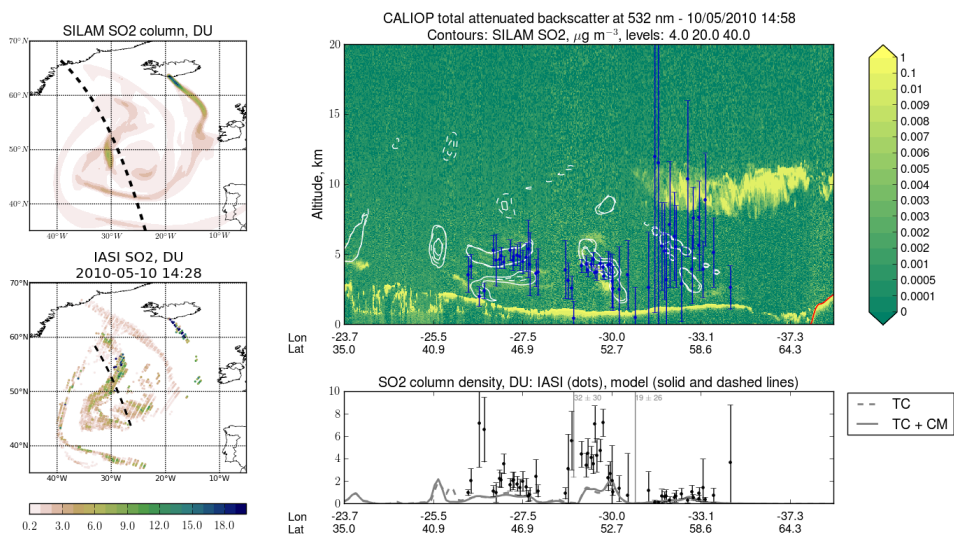
Formatted: Caption

1046

1047 **Figure 14. As Figure 12 but for May 10, 2010. In the 1D column density plot below the CALIOP curtain, two IASI data points with**
1048 **values 32±30 and 19±26 DU are outside the plot range.**