Atmos. Chem. Phys. Discuss., 11, C1777–C1793, 2011
www.atmos-chem-phys-discuss.net/11/C1777/2011/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Determination of time- and height-resolved volcanic ash emissions for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull eruption" *by* A. Stohl et al.

A. Stohl et al.

ast@nilu.no

Received and published: 13 April 2011

We thank reviewer 2 for the many constructive comments on our paper. We think they give us a good chance to clarify technical aspects of the inversion and other issues. We will also make many small changes in the paper but wish to keep the paper concise, especially since details on the method have been given previously and should not be repeated. Still, the review certainly has helped improving the paper! Below, we repeat the reviewer's comments in italics and add our answers in normal font.

My main concern is that it remains unclear how the inversion algorithm determines the a posteriori emission profiles from satellite observations from two different space

borne instruments and the FLEXPART model results. I know that the method was described in previous studies but here the algorithm acts as a "black box" which produces a presumably better estimate of the volcanic emissions than they could be derived by using the "standard" procedures that rely mainly on the observed plume height. What in particular confuses me is that the a posteriori emissions are lower than the a priori emissions AND the satellite observations. Fig. 7 shows that the columnar ash totals are higher in both satellite pictures than in the model results based on the a posteriori emissions. In the text it is written that the inversion algorithm leads to lower a posteriori emissions compared to the a priori estimate. If the satellite pictures show higher columnar ash totals than calculated by the model, how can they influence the emissions to go down? Figure A1 to A5 show that the a priori columnar values fit better to the SEVIRI observations than the a posteriori values.

The method, including all mathematical details, is described at length in two previous papers, so we don't think it acts as a "black box". It just doesn't make sense to repeat everything in this paper. We specify all changes that have been made to the inversion. However, these changes are relatively minor and so we kept this short. Readers who are interested in the algorithm are recommended to read Eckhardt et al. (2008). However, we will add a few lines on the principles of the inversion in the revised version of the paper.

Regarding the examples shown in Fig. 7 and A1 to A5, the a priori emissions are lower than the a posteriori emissions most of the time (see Fig. 2). We describe one such case (12-13 May) in the paper, but this holds also for other days in May. Thus, the a priori assumption for the cases shown constrains the inversion solution to lower, not higher, values. We, thus, expect the model solution for the cases shown in the Figures to be biased low compared to the satellite observations. The opposite is the case in April and for the full period of the eruption (April and May combined). Indeed, if we consider all data, the a posteriori model results are biased high by 7% (not low) compared to all the satellite observations.

Furthermore, the a posteriori model results may “look” more strongly biased low in Fig. 7 and A1 to A5 than is actually the case. The satellite data (and the SEVIRI data in particular) are quite noisy and the reader’s eyes are easily drawn to the highest observed values, which FLEXPART does not fully reproduce. FLEXPART produces a smoother version of the plume where many of the zero values observed by SEVIRI are “filled in”, resulting overall in a simulation that is only weakly biased (due to the too low apriori emissions in May). Notice, for instance, in Fig. 7 that SEVIRI observes no ash immediately downwind of the volcano and only scattered ash detections are made in the westernmost and southernmost ash bands. The inversion optimizes the agreement of the model with all the observations (also the observations of zero ash), not only with the high ones. This certainly can be critical, as high ash loadings may be underestimated. However, it also seems that SEVIRI sometimes has too high ash loadings compared to IASI, a result of the generally noisy nature of the SEVIRI data. We assume that, overall, SEVIRI data contain no bias. If better satellite data and improved retrieval methods will provide a more consistent view of the ash plume in the future, this problem should become much smaller.

It also needs to be explained how you get improved vertical profiles by adding information about the columnar ash values. I can imagine that it might be related to the different dispersion in different altitudes due to wind shear but it is not said in the paper and maybe I am wrong and there is additional information from the satellites used.

There is no additional information from the satellites used. The height information indeed comes from different dispersion patterns due to vertical shear of the horizontal wind. This is the topic of the papers by Eckhardt et al. (2008) and Kristiansen et al. (2010). For the revised version of the paper, we have added: “In previous studies (Eckhardt et al., 2008; Kristiansen et al., 2010), we developed an inversion algorithm to calculate the vertical distribution of sulfur dioxide emission rates for instantaneous volcanic eruptions, using only total column observations of sulfur dioxide. The algorithm extracts emission height information from the horizontal dispersion patterns, which de-

pend on altitude because of the vertical shear of the horizontal wind. By matching the observed plume with many simulations initialized at different altitudes, a best-fitting vertical emission profile is obtained as a linear combination of all emission height scenarios."

Title: You might think once more about it: Now it implies that the paper presents a way to determine emissions that may be used by others in their models. This is certainly true bout only half of what is shown here because you use these emissions in your model, too, and you draw conclusions from this subsequent model run (e.g. how much of the European air space was affected by ash concentrations above certain limit values. You could e.g. replace the "for quantitative ash dispersion modeling" by "and their application in a quantitative ash dispersion model”.

This is a good suggestion and we will change the title to: Determination of time- and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: The 2010 Eyjafjallajökull eruption

Line 4: I do not see "dramatic" improvements. There are some larger differences in the a priori and a posteriori emission estimates, particularly in the initial phase of the eruption, but I did not see a comparison to independent observations that proves that there are "dramatically" better results with the a posteriori emissions.

The comparison to the independent aircraft measurements shows large improvements, both in terms of bias and correlation (see Fig. 11). However, we have changed the word ”dramatic” to ”large”.

Line 26/27: "However, these relationships are subjective ...": Why are they subjective? Assumptions may be subjective but the method presented here also relies on a number of assumptions. I do not see a general difference between empirical relationships to estimate mass fluxes from plume heights and this method that would allow to call this method "objective" while others are "subjective". To me, the term "subjective" is a negative qualifier that should not be used in this context.

Agreed. We have changed the word ”subjective” to ”loose”.

Section 2.1.1: It is said that the SEVIRI subregion is 30 W to 30 E and 40 N to 70N. In Figure 7 and A1 to A5 a different region is shown. This is confusing. Which region was taken for the SEVIRI evaluation?

For the inversion, we only used SEVIRI data in the domain 30 W to 30 E and 40 N to 70 N. In contrast, to produce Figure 7 and A1 to A5, we also used SEVIRI data from outside this domain. This will be explained in the revised version of the paper in the caption of Fig. 7.

End of page 3: "a very large table": What does that imply? In which sense is the table large and why is it important to emphasize this?

We agree this should not be emphasized and have removed the words ”very large”.

End of page 3: "that can be interpolated ...": to what or between what? The description of what has been done here is not clear enough.

We have replaced the line: "The result of these calculations is a very large table of simulated TOA brightness temperatures for the two SEVIRI channels that can be interpolated based on the water vapour corrected SEVIRI data."

With " The result of these calculations is a table of simulated TOA brightness temperatures for the two SEVIRI channels and for each of the combinations of cloud top and surface temperature. Each entry in the table appropriate for the scene cloud top and surface temperature, is indexed by optical depth and effective particle radius. Based on the water vapour corrected SEVIRI data, the table is interpolated in terms of brightness temperatures to find the closest entries corresponding to the retrieved optical depth and effective particle radius. This procedure uses both the 11 and 12 μm brightness temperatures."

Line 60: "errors of 40 - 60 contribute to this overall error? Is it considered that the refractive index of andesite may not be correct for this ash? Which particle shape was

considered? Which error in the density of the ash was assumed?

We based this statement on the prior work of Wen and Rose (1994), suitably referenced. That paper devotes 5 pages to a sensitivity analysis and it seems inappropriate to repeat their findings here. We feel this reference adequately describes the sensitivities to the most important parameters in the retrieval process and that our summary statement based on their work is adequate. The interested reader may explore this matter in more detail by consulting the paper by Wen and Rose (1994).

Line 63: I don’t like the term "meteorological cloud". I would assume you mean a water cloud or an ice cloud. Could you please avoid this term in the whole document?

We agree and have changed all occurrences of ”meteorological cloud” to ”water and ice clouds”.

Line 98: "The exact conversion factor ... was calibrated to match the SEVIRI retrievals ..." What is the interplay between the satellite products? Is one combined product generated? Can the IASI data used independently? (Obviously yes, as it is demonstrated later, but is it in that case calibrated to the SEVIRI data, too?)

The IASI data can be used independently but even in that case it needs to be calibrated with the SEVIRI data since no actual IASI retrieval was made. Thus, the IASI data are never completely independent of the SEVIRI retrievals.

Line 108: "We assumed that 10% of the erupted mass was fine ash ...": What is the basis for this assumption?

We know that the fraction is small and depends on the eruptive style. Rose et al. (2000) report that between 0.04 and 2.6% of the erupted mass in volcanic eruptions may be in the fine ash range. The fraction cannot only be variable but also depends on how exactly "fine" ash is defined (i.e., the respective size range), and also how the total erupted mass flux is derived. In our case, this is the mass of tephra erupted into the atmosphere. The difficulty in quantifying the mass flux of fine ash with volcanological methods is actually an important motivation for our study.

Line 112: We all know that the term "reasonable agreement" is difficult because it is not clear what is exactly meant. The situation does not improve by putting it in quotes.

Agreed. We have removed the quotes.

Model simulations: It is not clear why the runs to derive the a posteriori emissions need more disk space "keep the model output at a manageable size) than the run over the whole time period of 41 days. Don't you have to consider all times and layers for the run with the a posteriori emissions, too?

The difference is that for the sensitivity runs, the output of all 6232 runs must be stored separately, while the "final" run using the a posteriori emission data is a single model simulation, for which output volume is less critical.

Line 165: "... modeled particle size distributions with a maximum modal diameter below 7 micrometer are inconsistent with downwind in situ measurements of ash particle size distributions ...". I cannot follow that so strictly. Schumann et al. report 12 cases, in 5 of these cases the "diameter of maximum coarse mode volume-size spectrum" (I assume this what you refer to) is below 7 micrometer.

Yes, but the cases with smaller diameter of maximum coarse mode volume-size spectrum reported by Schumann et al. (2011) are all cases with low ash concentrations. The cases with the highest ash concentrations all had values between 9 and 11 micrometer. Furthermore, the size distribution is shifted to smaller particles with time, as larger particles are lost preferentially. A value of 7 micrometer measured on the aircraft would indicate that the emitted size distribution must have had a larger diameter of maximum coarse mode volume-size spectrum.

Inversion algorithm: As already said above, the inversion appears to be a "black box". I know that it has already been discussed in other papers but because it is central for the results of this paper, it needs to be explained and discussed more. It particular

the reader should be informed if there could be particular difficulties when applying the algorithm to ash particles instead of SO2. Why are the a posteriori emissions lower than before if the satellite images point to higher columnar ash values. Could it be a problem of the algorithm that the emissions in lower altitudes are reduced in the a posteriori case (e.g. is there not enough information from the satellite instruments in these altitudes because clouds prevent the observation of ash in lower altitudes or the sensitivity is reduced there?)?

The algorithm works the same for SO_2 and for ash. Only the model simulations are different, with the ash simulations accounting for different removal processes than the SO_2 simulations. This will be explained better in the revised version of the paper. As explained before, we cannot see inconsistencies. For the days before 10 May shown in Fig. 7, the a posteriori emissions are higher (not lower) than the a priori emissions, despite the fact that the total emissions for April and May are decreased by the inversion. The same holds for some of the other comparison cases that the reviewer refers to. As mentioned above, we now explain a little more the principles of the inversion (i.e., use of vertical wind shear for retrieving height information, etc.).

Line 184: "The model results for all scenarios were matched with about 2.3 million satellite observations ...": What does "matched" mean? Is 2.3 million a lot? Does it reduce the uncertainty to have many observations?

"Matched" means that we are selecting co-located values. We have added: "(i.e., ensuring spatio-temporal co-location)" to explain this. Uncertainties are reduced by using many observations but not in proportion to the number of observations, since the information content is partly redundant. However, especially given the noisy nature of the SEVIRI data, use of as many observations as possible can help to reduce the effect of noise in individual observations. A complete error characterization of individual observations is unfortunately beyond the scope of this paper and may not even be possible.

A posteriori emissions: Are the results plausible? (From the satellite pictures I would say no). Maybe you could explain a bit what could be expected from the satellite pictures and what the inversion than gave you.

We don't understand why the reviewer may think that the results are not plausible. We believe this results from the fact that the reviewer thinks that the emissions were always reduced by the inversion, whereas for the examples discussed (e.g., Figure 7), the emissions during the preceding days have been increased by the inversion.

Line 248 and Fig. 4: Since the differences cannot be seen, it is not necessary to show Fig.4.

Careful inspection of the figure does reveal very minor differences, but the purpose of the figure is demonstrating that the differences are small. Still, we will consider removing the figure.

Line 260: "The main reason for this is that the number of gridded IASI observations is about an order of magnitude smaller than the number of SEVIRI observations, thus providing a weaker constraint on the emissions, which therefore remain closer to the a priori values." Wouldn't it depend in the first instance on the columnar ash values and not on the number of observations how close the a posteriori emissions are to the a priori emissions?

It depends both on the columnar ash values and the number of observations, although the solution will converge towards a certain value (which then would also be independent from the a priori emission) for a very large number of observations. The number of IASI observations is insufficient to reach such convergence.

Line 266: "All inversions also lead to substantial emission increases for 12-13 May and to a general shift of the ash emissions to higher altitudes." Isn't it mainly a decrease of emissions in lower altitudes while those in higher altitudes are less influenced? At least for the ECMWF-based inversion with SEVIRI data, there is almost no increase of

the emissions in higher altitudes in Fig 5a.

As can be seen in Fig. 5d, also the SEVIRI-only inversion more than doubles the emission on 12-13 May, and this is mainly a result of emissions being increased at high altitudes. This cannot be seen in Figure 5a referred to by the reviewer, since this panel gives the emissions averaged over the entire eruption period. Again, one needs to look at different periods separately. Even though total a posteriori ash emissions are reduced, this does not mean that they are always reduced! In particular, this was not the case on 12-13 May.

Line 293: "Furthermore, in general, the a posteriori plume heights are in better agreement with the webcam observations than the a priori plume heights." Can that be underlined by some numbers?

We have tried doing so, but partly this remains a matter of subjective judgment, since the two webcams are in so poor agreement with each other. Furthermore, the fact that often only minimum heights can be given, makes a statistical comparison very difficult.

Line 320: " ...the comparison with the model is qualitative and further complicated by meteorological clouds, which produce similar backscatter signals as ash clouds." Above you said (line 310) "This signal responds to aerosols (including volcanic ash) as well as meteorological clouds which in many cases can be distinguished." Can they be distinguished or is it difficult? How is it done? Can't you use the depolarization signal to distinguish them?

We have used the depolarization signal, but even that does not always allow a clear separation between ash and clouds. Sometimes there remains doubt about the nature of the signal. We have selected cases for presentation where we are comfortable with our interpretation of the CALIPSO data, but there are many cases which are not very clear. In some cases, ash and water or ice clouds also appeared to be mixed.

Line 332: "While meteorological clouds often complicate the comparisons, qualitatively *we find that the a posteriori results are in better agreement with the CALIPSO data (Fig. A1-A5)." It is not easy to see that. What is your criterion to see it?*

The comparisons are qualitative and there is no objective measure of success. However, these are our reasons to believe that the a posteriori simulations improved the model-measurement comparison with respect to CALIPSO: Fig. A1: The vertical extension of the plume to 10 km near 47 N is captured by the a posteriori, but not by the a priori. Fig. A2: Again, the plume top near 50 N is correct in the a posteriori, but more than 1 km too low in the a priori. Also, some other CALIPSO features between 55 and 60 N are at least partially overlapped by the a posteriori model output but missed completely by the a priori results. Fig. A3: Again, the a posteriori plume near 38 N is higher than the a priori plume height, in agreement with a faint observed feature which extends even higher. Furthermore, white isolines (indicating high modeled concentrations) exist near observed ash maxima. For instance, the major observed maximum near 51 N is completely outside even of the black isoline in the a priori results but partly covered by a white isoline in the a posteriori results. Fig. A4: plume top near 50 N better captured by a posteriori simulation; simulated high a priori concentrations (white isolines) near 35 and 40 N (where no or little ash is observed) considerably reduced. Fig. A5: indeed little difference between both simulations, but both capture observed structures.

Line 333: "We also evaluated our results against quantitative vertical ash concentration profiles obtained from lidar measurements over Europe (Fig. A6-A8) and find that the modeled a posteriori ash concentrations are similar to the observed concentrations." Unfortunately the observations you compare your results with are not shown in the figures. Taking Ansmann's plots, a visual inspection indicates quite some differences, in particular at 13 UT on 16 April the a priori emissions seem to give better results and at the other times, the differences between the different emission cases are rather small. I think your finding is unbalanced towards the a posteriori emissions.

Line 355: "The model captured these ash layers and there is relatively good quantita-

tive agreement between the a posteriori model results and the measurements." Could you give numbers and compare both the a priori and the a posteriori runs to the observations.

We did not have access to the measurement data, so we could not use them directly and since these data are not critically important to our paper, we made no effort to get them (or invite station PIs to become co-authors) but rather put our model results into the appendix. The lidar profiles are all published and the readers can just compare the plots. We agree it is difficult from these few plots to demonstrate an improvement due to the inversion. However, importantly, measurements and model results show similar concentrations, confirming at least that the model is getting the right order of magnitude and that the altitudes are reasonably accurate.

Line 355: "The model captured these ash layers and there is relatively good quantitative agreement between the a posteriori model results and the measurements." Could you give numbers and compare both the a priori and the a posteriori runs to the observations.

We have shown GFS data in all plots where we could add GFS results without requiring a new panel or figure (e.g., in Fig. 8-10). We have not shown GFS data in figures where we would have had to produce separate figures (e.g., Fig. 7, 11, A1-A5), as this would have increased the length of the manuscript substantially and we wanted to keep this paper concise. ECMWF data are our primary data source, as they allow a more accurate simulation of dispersion (also thanks to higher resolution than for GFS data), so it is more important to show ECMWF than GFS results.

Line 375: "The data for the comparison was selected by screening the entire observation data set (including gas-phase measurements) for volcanic plumes. An unbiased but imperfect model would underestimate the observations in such a comparison, since slight displacements of the modeled plumes would lead to sampling lowerconcentration parts of the plume in the model as compared to the observations." Could you explain *this a bit? Couldn't also the aircraft fly in an area of lower ash concen- trations compared to a region e.g. 100 km away that is captured by the model? Why should the aircraft always sample the region with the highest concentrations?*

This is not a matter of how the aircraft sampled the ash plumes, but how ash plume sections of the flight have been selected for further study in Schumann et al. (2011). Very low concentrations of ash may not have been recognized as ash by Schumann et al. (2011) or may have been mixed with other aerosols or cloud particles and, thus, not been used for further study. The same holds true for sections of the flights with no ash at all, which were also not used. This causes a bias of the selected data sub-set towards high observed values, for which ash could be identified reliably.

Comparing all measurement data along the entire flight tracks with model results would have avoided this, but this was not possible because of cloud contamination of the data. The instruments are simply not specific to ash only but also observe other types of particles, causing high observed values which are no ash. Thus, we could only use data for confirmed ash plumes, thus causing the bias.

Line 393: "Patches of highly concentrated ash were present over Europe (10W-30E, 36N-60N) during both April and May, and it is important for aviation to avoid them." How do you know? I think this is nothing you could state here. You can say that according to your model results the limit values were exceeded in some (few) areas but nothing more.

Well, there have supposedly been aircraft observations of high ash concentrations, but these data are not published yet so we cannot refer to them. Furthermore, the SEVIRI data does show small areas of high ash concentrations. Partly, these high values reflect the noisy nature of the data but partly this may also indicate small-scale variability not captured by the model.

Our statement is also meant to avoid false security about the model predictions. It is quite obvious that, below the model resolution, there will be local concentrations

higher than predicted by the model. Whether such small pockets of concentrated ash are dangerous to aviation or not, is a question that the aviation industry will have to answer.

Line 397: "In this paper we have, for the first time, objectively determined the ash emissions of a volcanic eruption as a function of time and altitude." I do not agree that other emission estimates are less objective. Maybe they are less accurate, but your method also includes assumptions that others may call "subjective".

To our knowledge, there is not a single study that has determined volcanic ash emission as a function of time and altitude. Especially the height distribution has never been determined. However, maybe the reviewer's confusion comes from the word "objectively". We will therefore remove this word in the revised version of the paper.

Line 404: " ... ground-based and space-based lidar observations ..." The ground based observations need to be quantified and a figure could be shown in the paper itself if the original data could be plotted in the figures, too.

The lidar data are published data and everyone can look them up. We have shown our model profiles corresponding to the published lidar data only in the appendix, as we want to keep the paper itself as concise as possible. Showing a comparison with the lidar data in the main part of the paper would require adding explanations of the data and text describing the comparison. This would run against our intention of keeping the paper short. Furthermore, there are several submitted papers, which will present such model-measurement comparisons.

Figures: Which heights are given? Are they above sea level or above ground?

All heights are given as meters above sea level. We will add a clarifying statement to the caption of Figure 2.

Fig. 4 may be omitted. It does not contain new results.

We will consider this. However, the purpose of the figure is to convince the reader how small the differences are and to illustrate the sensitivity studies we have made.

Fig. 7 and A1-A5: Comment also the low clouds in 1-2 km altitude.

We have added in the caption of Figure 7 (and a similar note in appendix caption): Also the layer extending from 30-60°N near 2 km altitude are clouds.

Fig 8, Fig. 9: I am quite surprised about the very good timing of the ash cloud in the model and the observations. Is there a reason why this looks almost perfect?

Notice that most of the observations of high ash concentrations were made during ascents or descents. This means that both the modeled horizontal as well as vertical position of ash are simulated with quite high accuracy. The timing itself is less important, as the aircraft covers large distances in only a few minutes. We think the good agreement is a result of our carefully determined source term – to some extent this is valid even for our a priori source term, which should be better than what is often used in comparable model studies.

Fig. 11: You could another scatter plots with the GFS results.

We will consider this. However, for consistency we would then also have to add additional panels/plots for the lidar comparisons (Fig. 7, A1-A5) and our aim was to keep the paper short.

Line 7 "a posteriori model": better "model results with a posteriori ash emissions".

Agreed. We will change this as suggested.

Line 43: on Iceland

Agreed. We will add ”on Iceland”.

Section 2.1.1: something's wrong with the line numbering

This seems to have been a strange Latex problem in the submitted manuscript, but it is OK in the published ACPD version.

Line 74: "improvement on". Better: improvement compared to

Agreed. We will change this.

Line 203: "In total, 2.3 million observation cases were used for the inversion." This has been said before.

Yes, but only a few words are repeated and we think it improves clarity to repeat this information.

Line 341: Now "Schumann et al. (2011)".

Yes, this will be updated.

Include the year 2010 in the figures (or captions) where necessary.

Adding the year to the labels would require more space and, thus, we would have to remove many labels, or labels would have to be placed very close to each other. We therefore added the year to some of the figure captions.

Fig.1: A legend for the thin blue lines would be nice to have.

Since there are two thin blue lines, this cannot be done unambiguously in the legend. But using an additional color or line style would also not be a good solution since there are already quite a few different lines, so we leave this as being explained in the figure caption.

Fig.5: Use the same dates (every 7 days) as in Figure 2 and 3.

Agreed. We have changed the figure.

Fig. 6: include a legend for the x-axis

There appears to be too much text underneath the figure (the symbol legend and the label bar) to add yet more text. However, we have added the words ”as a function of date” to the figure caption to make this more clear.

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 5541, 2011.