Estimating volcanic ash emissions using retrieved satellite ash columns and inverse ash transport modelling using VolcanicAshInversion v1.2.1, within the operational eEMEP volcanic plume forecasting system (version rv4_17)

André R. Brodtkorb[1,2], Anna Benedictow[2], Heiko Klein[2], Arve Kylling[3], Agnes Nyiri[2], Alvaro Valdebenito[2], Espen Sollum[3], and Nina Kristiansen[4]

[1]OsloMet - Oslo Metropolitan University, Oslo, Norway
[2]The Norwegian Meteorological Institute, Oslo, Norway
[3]NILU - Norwegian Institute for Air Research, Kjeller, Norway
[4]Met Office, Exeter, UK

Correspondence: André R. Brodtkorb (andre.brodtkorb@oslomet.no)

Abstract. Accurate modelling of ash clouds from volcanic eruptions requires knowledge about the eruption source parameters including eruption onset, duration, mass eruption rates, particle size distribution, and vertical emission profiles. However, most of these parameters are unknown and must be estimated somehow. Some are estimated based on observed correlations and known volcano parameters. However, a more accurate estimate is often needed to bring the model into closer agreement to observations.

This paper describes the inversion procedure implemented at the Norwegian Meteorological Institute for estimating ash emission rates from retrieved satellite ash column amounts and a priori knowledge. The overall procedure consists of five stages: (1) generate a priori emission estimates; (2) run forward simulations with a set of unit emission profiles; (3) collocate/match observations with emission simulations; (4) build system of linear equations; and (5) solve overdetermined system. We go through the mathematical foundations for the inversion procedure, performance for synthetic cases, and performance for real-world cases. The novelties of this paper includes a memory efficient formulation of the inversion problem, a detailed description and illustrations of the mathematical formulations, evaluation of the inversion method using synthetic known truth data as well as real data, and inclusion of observations of ash cloud-top height. The source code used in this work is freely available under an open source license, and is possible to use for other similar applications.

1 Introduction

Volcanic ash is considered a significant hazard for aviation, since ash particles at flight altitudes can be ingested into the jet engine, leading to debris build-up, with potential damage and engine failure risk (Casadevall, 1994; Clarkson et al., 2016). For that reason, it is important to forecast the location and movement of volcanic ash in the atmosphere to help aid safe flight operations. Atmospheric Transport and Dispersion models (ATDMs) are widely used for this task (e.g., Steensen et al.

(2017b); Beckett et al. (2020); Osores et al. (2020); Chai et al. (2017); Gouhier et al. (2020)). The accuracy of the model forecasts depends critically on details of the eruption source parameters (ESPs) (e.g. eruption onset, duration, mass eruption rates (MER), vertical emission profiles, particle-size, fine ash fraction, etc.) which the models require as input. While some volcanoes are well monitored and observations of e.g., eruption onset and duration can be directly observed (Lowenstern et al., 2022), ~~some~~ other volcanoes (typically in remote areas) are not routinely monitored. Moreover, MER is a quantity that often cannot be obtained directly with current measurement techniques, but needs to be estimated indirectly using other observations. Plume height observations from radar has often been used in combination with an empirical relationship from Mastin et al. (2009) to estimate the MER. However, this approach has large uncertainties (for example it is biased towards large eruptions) and does not provide an estimate of the vertical distribution of the ash in the eruption column. A more robust approach to estimate ESPs is using data assimilation and inverse modelling techniques which combine satellite data and transport modelling with the aim to estimate ESPs that bring the model simulations in closer agreements with observations (Stohl et al., 2011; Heng et al., 2016; Chai et al., 2017; Zidikheri and Lucas, 2020; Pelley et al., 2021).

Since the 2010 eruption of the Eyjafjallajökull volcano in Iceland, significant innovations and improvements in the capability of detecting and quantifying volcanic clouds from space have been made (Mastin et al., 2022), and in particular the ability to estimate the ash cloud altitude (see e.g., Prata and Lynch (2019) and Prata et al. (2022)). Some satellite instruments are furthermore able to provide both volcanic ash mass loading and ash cloud height information, such as the SLSTR (Sea and Land Surface Temperature Radiometer) instrument on board the Sentinel-3 satellite. Still, there are issues and uncertainties in the detection of volcanic ash, particularly around the effect of water vapour, cloud identification and the presence of ice (Kylling, 2016). These uncertainties are important to characterise and account for in any inverse modelling method, or other data assimilation approaches (see e.g., Pardini et al. (2020); Osores et al. (2020); Tichý et al. (2020)).

In this paper, we present a Python software package that is a further development of an inversion method originally developed for volcanic ash emission estimation (Stohl et al., 2011). The main idea of the method is based upon the variational principle, i.e., to use a set of forward simulations with a set of unit emission profiles, and then attempt to find the linear combination of these that best matches the observed ash locations and total column loadings. The methodology used here has previously been used for sulphur dioxide emission, see e.g., Seibert (2000); Seibert et al. (2011) and Eckhardt et al. (2008). It was for the first time used for volcanic ash emission rate determination for the Eyjafjallajökull 2010 eruption (Stohl et al., 2011). Steensen et al. (2017a) presented an uncertainty assessment of the method. This work extends the approach to also incorporate ash cloud-top height as an observation in the inversion. The ash cloud-top height can be estimated using the SLSTR instrument~~or~~ , using photogrammetric parallax between different satellites (Zakšek et al., 2013), or other geometric approaches (Horváth et al., 2021) . Chai et al. (2017) also included satellite retrievals of ash cloud-top heights as an additional constraint in their volcanic ash inverse system by explicitly enforcing no-ash above the observed ash cloud-top height. However, they found that in their case study such extra constraints were not helpful for the inverse modelling. In this paper we use a similar approach by using ~~non-zero observations~~ observations of non-zero ash mass from the ground up to the detected plume top, and ~~zero observations~~ observations of zero ash mass above the observed ash cloud-top.

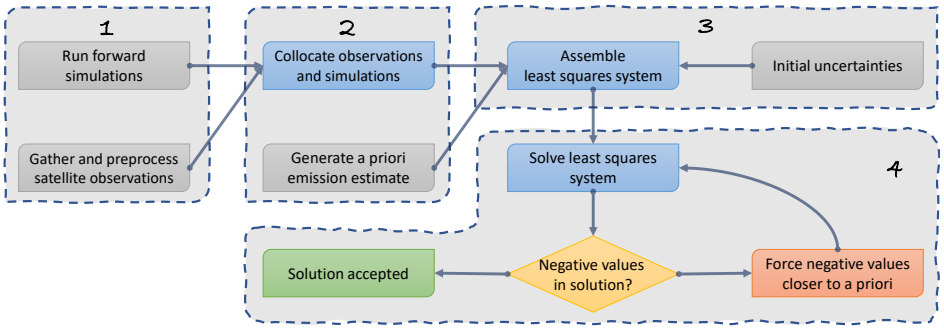


Figure 1. General solution framework for creating an accepted a posteriori emission estimate. The input to the algorithm are marked in gray (forward simulations, satellite observations, an a priori estimate, and initial uncertainties for both observations and a priori estimate). The algorithm runs sequentially through stages one to four marked as dashed boxes, and computes and iteratively refines the estimate in stage 4. The most computationally demanding stages are stage 1 (forward simulations), 2 (collocation of observations and simulations), and 3 (assembly of source-receptor matrix). Iteratively solving the least squares system is highly efficient.

The manuscript is organised as follows: In Section 2 the theory behind the inversion procedure is outlined. The atmospheric dispersion model is described in section 2.1. The synthetic benchmark cases and real-world cases are described in sections 3 and 4, respectively. The paper is summarized in section 6.

2 Mathematical formulation and solution framework

We base our inversion procedure on the approach taken by Seibert (2000), and adopted to ash emission by several authors (see e.g., Stohl et al. (2011) and the references therein). Figure 1 shows the general solution framework which we use. We first compute a set of forward runs, representing possible emissions, collocate these with ~~our~~ satellite observations, and finally assemble into a source-receptor matrix. We then formulate the solution as a least squares problem incorporating these collocated simulations/observations as well as an a priori estimate (initial guess) of ~~our~~ the emission. Together with the uncertainties of the observations and ~~our~~ an a priori estimate, we solve the least squares problem to find the emission that recreates ~~our satellite images~~ the observed satellite observations best (in a least squares sense). As there is no guarantee of non-negative values in this process, we may end up with non-physical emissions. Negative emissions are forced closer to ~~our~~ the (non-negative) a priori estimate by decreasing the uncertainty of these values, and solving the least squares problem again. This process is repeated until we have sufficiently few negative emissions. Other inversion methods (e.g., Pelley et al. (2021)) use an algorithm with a direct non-negative emissions constraint.

2.1 Atmospheric Dispersion Model

For this work, we use the eEMEP (Steensen et al., 2017b) model version rv4_17 ~~(EMEP MSC-W, 2018)~~, which is the *emergency* version of of the basic EMEP MSC-W chemical transport model (Simpson et al., 2012, 2018; EMEP MSC-W, 2018)

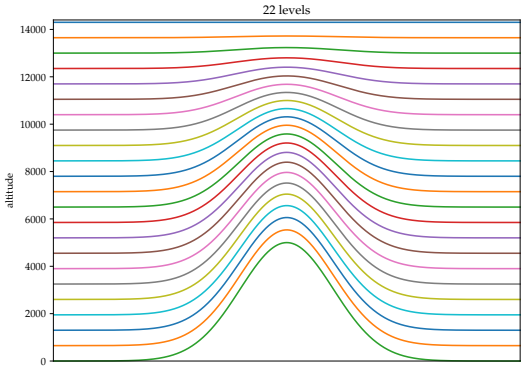


Figure 2. Vertical hybrid sigma levels for the inversion runs. Each level is designed to corresponds to roughly 650 meters of altitude given a ground pressure of 1013.25 hPa. The bottom layer shows a synthetic topography, and how this alters the altitude of the different layers. Please note that these layers are only represented as hybrid sigma coordinates, and never represent actual meters above sea level. The actual vertical definitions file is available in Brodtkorb et al. (2020)

. The model is based on an Eulerian advection model ~~based on~~ , uses the fourth order positive definite advection scheme of Bott (1989). ~~The eEMEP model is being~~ , an ash particle diameter distribution with nine bins from 4 to 25 μm, and has gravitational settling in all model layers. It is used operationally to generate volcanic ash forecasts at the Norwegian Meteorological Institute, ~~which~~ and these forecasts are published on the Avinor Internet Pilot Planning Center webpages for all pilots to use.

The current operational setup uses 48 vertical hybrid sigma layers from the ground up to 9.26 hPa (around 26 km above sea level) with ECMWF meteorology. The layer thickness is smallest close to the ground, and increases with altitude. For our use, this may not be the most efficient approach[1], and we instead use 22 vertical levels which are close to 650 meters thick each, shown in Figure 2. This is a trade-off between the number of levels to emit into (which corresponds to the system matrix size, see Section 2.2), and the computational time required to run the model. We restrict the vertical extent of our model to around 14 km due to limitations in the meteorological input fields available for the 2010 Eyjafjallajökull case.

An inversion run with 22 levels of emission every three hours for four days results in over 700 different unit emission scenarios that need to be simulated to generate the source-receptor matrix M. A regular run with the simulation model takes around 20 minutes to complete using 32 CPU-cores, which means that this represents over 40 weeks of CPU time. This is prohibitively expensive, and we therefore use a special version of the eEMEP model that can run up-to 19 tracers simultaneously. These tracers are independent, and reduces the number of simulator runs from 700 to 36. The major savings come from only having to

[1]The levels closest to ground are typically very thin, which is good from an accuracy point of view for concentrations close to ground. However, when we want to estimate the ash emissions it requires a huge computational effort to handle these layers without them having a significant effect on the result. Most of the volcanic ash we observe in our case studies is emitted high into the atmosphere.

read and process the meteorology 36 times instead of over 700 (which typically is the bottleneck for this kind of application)[2]. The numerical advection and writing results to file are not optimized by this approach. For a simpler setup, we also do not use the up-most 3 levels, so that the 19 altitudes at each emission time can be simulated with a single run. We end up with 33 runs that each take around 20 minutes to complete, and using the Nebula supercomputer[3] we are able to get all these inversion runs completed in less than one hour.

Figure 3 shows the result of emission simulations for different emission altitudes. The three simulations emit ash in level 1, 9, and 19 at midnight on April 14th. The simulation then progresses, and the plot shows how the ash is distributed in the vertical dimension as time goes on. Notice that large parts of the ash cloud leaves the simulation domain[4] just after midday on April 16th.

2.2 Source-receptor matrix

The inversion procedure is based upon creating a so-called source-receptor matrix, M, from a set of forward simulations with unit emissions (see Figure 3). We define the set of unit emissions as follows:

$$\begin{aligned}\{S\} &= \{S(\alpha_1,\beta_1),\ldots,S(\alpha_1,\beta_l),S(\alpha_2,\beta_1),\ldots S(\alpha_k,\beta_l)\}\\ &= \{S_1,\ldots,S_n\}.\end{aligned} \tag{1}$$

Here, α_k denotes *emission time* k and β_l denotes *emission level* l. An individual simulation S_j then contains time-dependent three-dimensional simulation results of a unit of ash emitted into the atmosphere at the given emission time and emission level.

We equivalently have the set of observations arising from individual pixels of processed satellite images,

$$\begin{aligned}\{O\} &= \{O(y_1,x_1,t_1),O(y_2,x_2,t_2)\ldots O(y_m,x_m,t_m)\}\\ &= \{O_1,\ldots,O_m\},\end{aligned} \tag{2}$$

in which (x_i,y_i) denotes the spatial coordinates, and (t_i) the observation time. The observations are sorted according to increasing time, t_i. We can assemble the matrix M so that element (i,j) of the matrix has simulation results from unit emission j at observation coordinate i,

$$M_{i,j} := S_j(y_i,x_i,t_i), \tag{3}$$

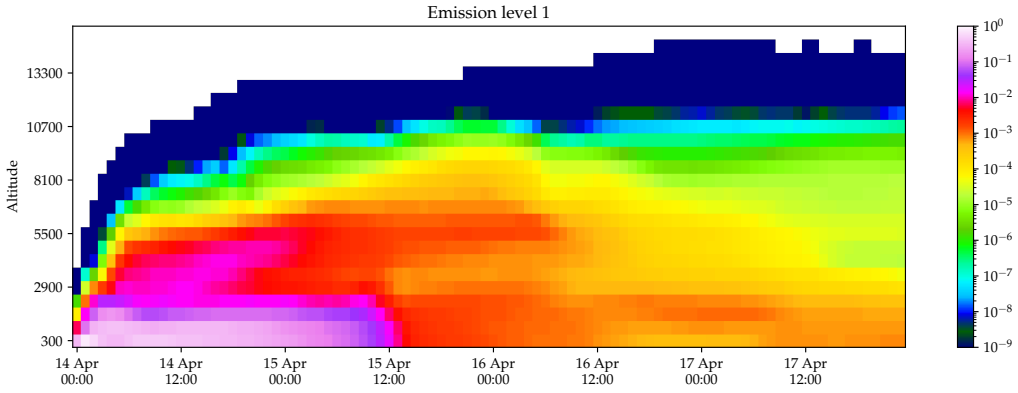
and similarly for the vector of observations so that element i corresponds to observation coordinate i,

$$y_{0_i} := O_i. \tag{4}$$

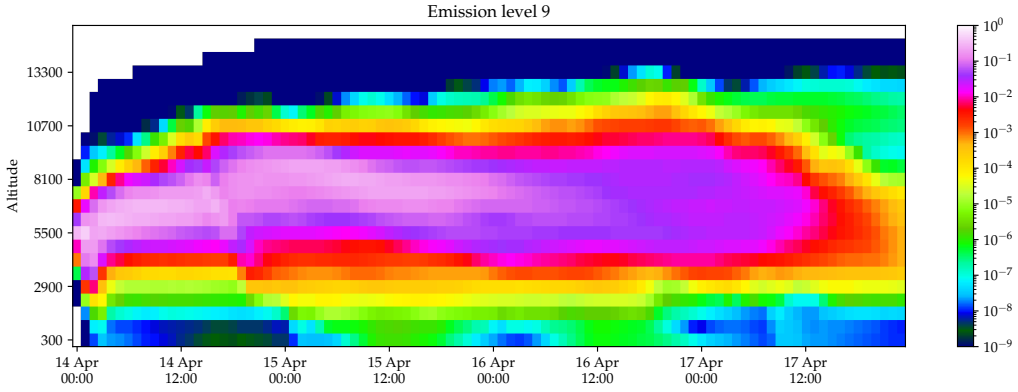
[2]By using tracers, the eEMEP program can advect up-to 19 independent tracers simultaneously in one simulation run, thus only having to read the meteorology once instead of 19 times.

[3]Nebula is a research and development cluster located at Linköping University, and is part of the MetCoOp supercomputing infrastructure shared by SMHI Sweden, Met Norway, and FMI Finland. The computer consists of 4352 cores partitioned over 136 nodes connected by an Intel OmniPath 100 Gbit network. Each compute node has 2 Intel Xeon Gond 6130 and 96 (thin node) or 384 (fat node) GiB memory, and is identical to the operational infrastructure.

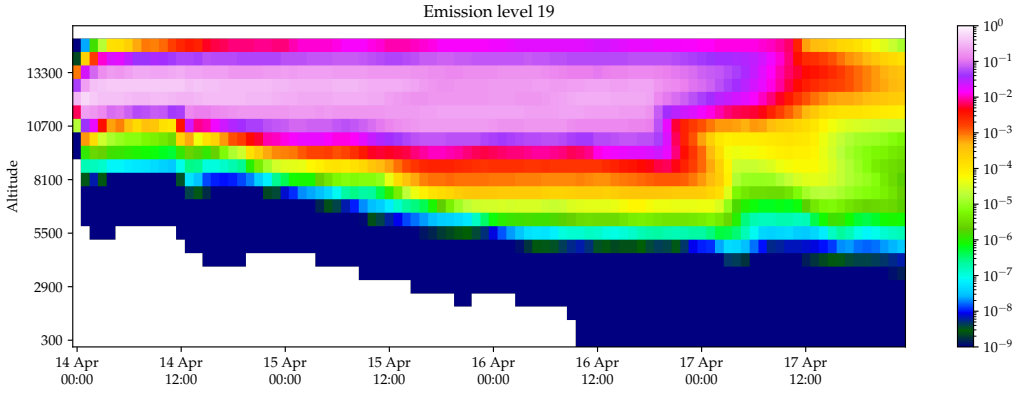
[4]The full simulation domain is shown in Figure 11, and covers the area from 30° north, 30° west to 76° north, 45° east.



(a)

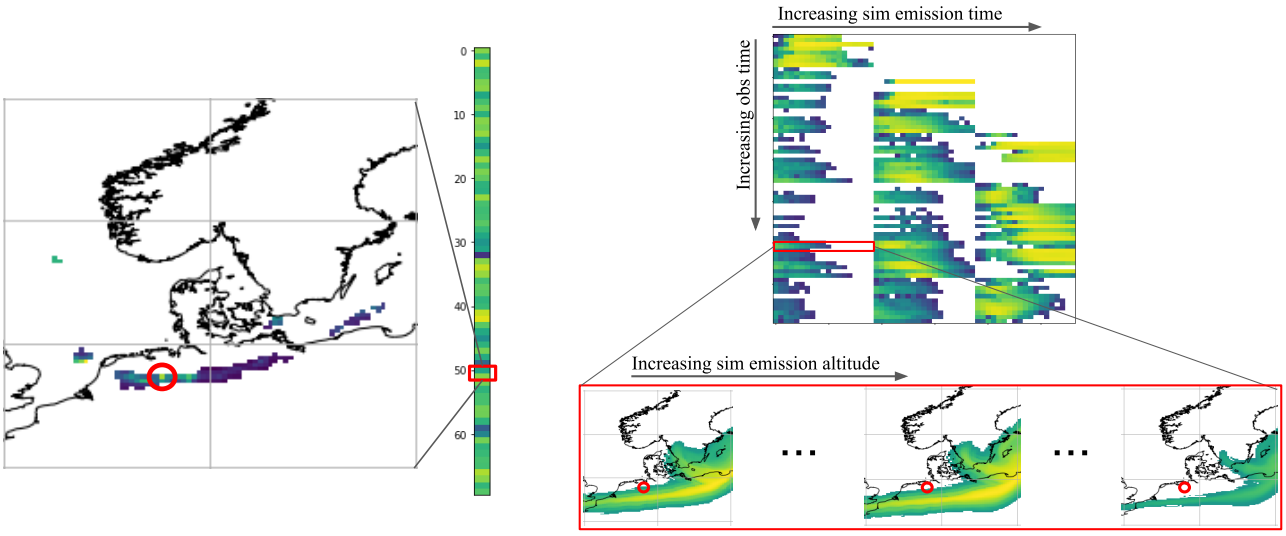


(b)



(c)

Figure 3. Unit emission profiles used in the inversion procedure. Here we emit a unit (1 teragram) of ash at different emission levels, and plot the amount of ash in different layers over time. The figure has been created by spatially integrating the ash mass in each layer, and thus shows how the ash emitted at a specific altitude disperses as time passes. The ash is released at the hybrid terrain-following vertical model levels shown in Figure 2. Notice that for emissions close to ground, the ash travels only for a short time.



(a) (b)

Figure 4. Linear system of equations. (a) shows the vector of observations originating from the SEVIRI instrument, in which a single location is highlighted with its corresponding location in the observation vector. Note that the image shows only part of the domain, and many of the detected ash pixels are outside the highlighted area. (b) shows the source-receptor matrix M. Each row in the matrix ~~corresponds to one~~ contains all simulated ash mass loadings at the observation coordinate of ~~ash, O_j~~ the observation in the same row in the observation vector (see also Equation (3)). Equivalently, ~~and~~ each column in the matrix corresponds ~~to~~ one individual emission simulation, S_j. ~~This~~ The matrix ~~shows~~ shown here has 60 observations (columns), three emission time points, and 19 emission altitudes (3×19 rows). White means no ash in the simulation, and the colored elements correspond to the concentration of ash in the individual simulations. See also Figure 5 which shows the full matrix.

This means that each row in the matrix corresponds to one observation, and each column corresponds to a single simulated unit emission.

Figure 4 shows a small source-receptor matrix, and part of the inputs used to create it. For each observation at (y_i, x_i, t_i), we find the simulated ash content at the same time and spatial coordinate for all of the different emission simulations.

With the source-receptor matrix assembled, the aim of the inversion procedure is to find the vector x so that

$$Mx = y_0. \tag{5}$$

Here, x is the linear combination of unit emissions that best reproduce the observations. In an ideal world, we could have hope of finding the vector of unit emissions, x, that reproduces our observations. However, in the presence of observation and simulation uncertainties, we instead have to resort to finding the vector that best represents our observations.

The size of the matrices and vectors involved in the computation is determined by the number of observations and unit emissions. The number of unit emissions, n, can typically be a few hundred to a few thousand, and the number of observations,

m can be hundreds of thousands or even millions. For example, for the beginning of the 2010 Eyjafjallajökull eruption (to be discussed later in Section 4), we have an emission simulation starting every three hours from 00:00 on April 14th UTC to 00:00 on April 18th, and emission heights every 650 meters up to about 12 km. This corresponds to 33 distinct times and 19 elevations, totaling to 627 unique unit emissions with one simulation each. The corresponding satellite images for this period has a total of 92403 observations, leading to an overdetermined system (i.e. 92403 constraints and 627 degrees of freedom).

2.3 Linear least squares with Tikhonov regularization

We cannot hope to solve the system in Equation 5 exactly as it is typically overdetermined and both the simulations and observations have errors. We therefore seek a solution vector x using linear least squares with Tikhonov regularization. The ~~Tikhonov regularization attempts~~ aim is to find a smooth solution that ~~allows both the observations and simulations to have~~ accounts for both observational and modeling errors.

Using (ground/visual) observations of the eruption, we create an a priori estimate of the emission, x_a, and incorporate this a priori knowledge into our least squares solver to give preference to solutions close to the a priori. We start by replacing our inverted emissions, x, with $\tilde{x} = x - x_a$,

$$\begin{aligned} Mx &= y_0, \\ M(x - x_a) &= y_0 - Mx_a, \\ M\tilde{x} &= \tilde{y}, \end{aligned}$$

to penalize solutions that lie far from our a priori. We still cannot hope to find an exact solution to this problem, but we can find the optimal solution, in a least squares sense, by minimizing

$$J_1 = ||M\tilde{x} - \tilde{y}||.$$

However, the observations are known to have measurement error, and this uncertainty can be included by assigning a weight to each observation

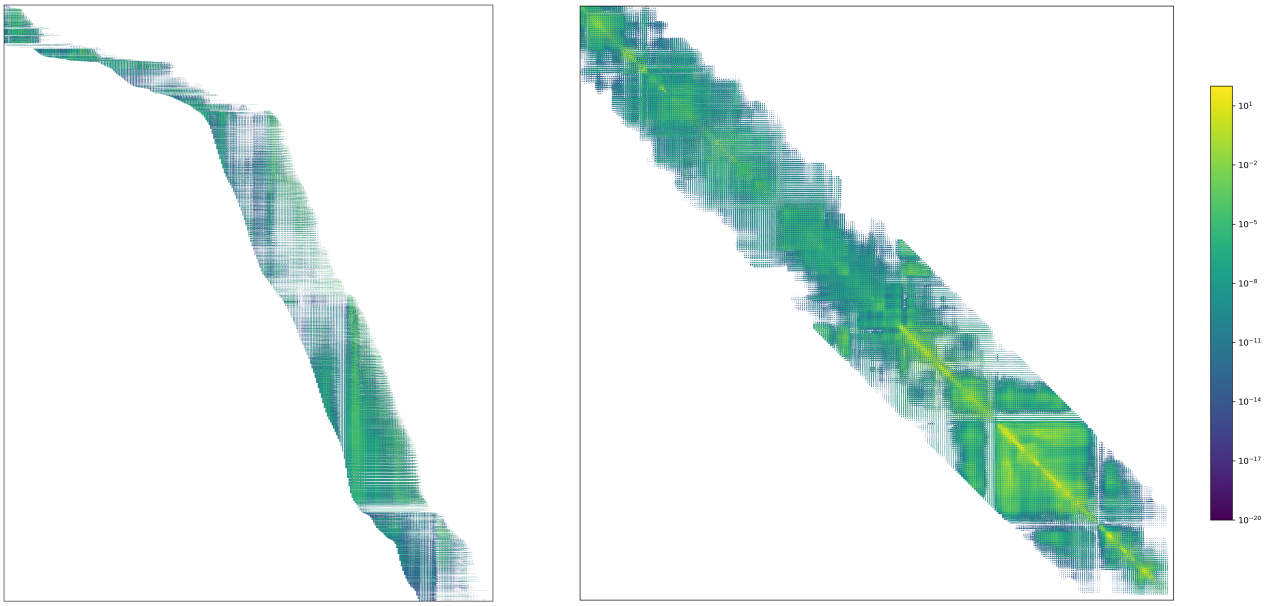
$$J_1 = ||\sigma_o^{-1}(M\tilde{x} - \tilde{y})|| \tag{6}$$

in which σ_o is a diagonal matrix with the standard error of observations to control how close we want our computed solution to match observation y_{0i}.

In the formulation above, we can also control how close we want our solution to lie to the observations. To control how close our solution should lie to the a priori knowledge, we add a second minimization term,

$$J_2 = ||\sigma_x^{-1}(x - x_a)|| = ||\sigma_x^{-1}\tilde{x}||, \tag{7}$$

where σ_x is a matrix with the estimated standard error of the a priori estimates on the diagonal. In our experiments, we have used $\sigma_x = \frac{1}{2}x_a$.



(a) (b)

Figure 5. Source receptor matrix M (left) and Least squares matrix G (right). The source-receptor matrix for the Eyjafjallajökull case without zero ash observations for the period April 14th - May 24th has 736494 rows (observations) and 6061 columns (ash emissions to estimate). It has approximately 521 million nonzero entries. The least squares matrix is a square matrix with 6061 rows and columns. The band width of the matrix is approximately 2000, which is determined by the maximum allowable time between an emission and observation of ash.

Unfortunately, solving this minimization problem often results in ~~solutions with sharp gradients, i.e., discontinuous~~ implausible solutions with non-physical discontinuities in the vertical dimension. To avoid such nonphysical solutions we can add a smoothness minimization term,

$$J_3 = \epsilon ||D\tilde{x}||, \tag{8}$$

in which D is a ~~diagonal~~ tridiagonal matrix that calculates the second derivative of $\tilde{x}$, and ϵ determines how smooth we want the solution to be. We have used $\epsilon = 1.0e^{-3}$, and this parameter is typically set by experimentation.

We can solve these combined minimization problems as a Tikhonov regularization problem (with Tikhonov matrix $Q = \sigma_x^{-2} + \epsilon D^T D$),

$$[M^T \sigma_o^{-2} M + \sigma_x^{-2} + \epsilon D^T D]\tilde{x} = M^T \sigma_o^{-2} \tilde{y},$$

in which the optimal solution x can be computed as

$$\begin{aligned} x &= x_a + [M^T \sigma_o^{-2} M + \sigma_x^{-2} + \epsilon D^T D]^{-1} M^T \sigma_o^{-2} \tilde{y} \\ &= x_a + G^{-1} M^T \sigma_o^{-2} \tilde{y} \\ &= x_a + G^{-1} B \end{aligned} \tag{9}$$

Figure 5 shows both the source-receptor matrix M and the least squares matrix G which then is used to compute the a posteriori emission estimates x. M is generally a large sparse matrix with m rows (one per observation O_i), and n columns (one per emission, S_j), whilst G is a relatively small banded matrix with n rows and columns. The band width depends on the maximum time from emission to observation, and is 1824 for our typical inversion parameters (19 emissions eight times per day for six days).

It is important that the units of the matrices and vectors are compatible when formulating the minimization problem. Our source-receptor matrix, M, and observation vector, ~~y, are concentrations per area scaled to~~ y_0, are column loadings expressed in kg/m^2. The a priori emission estimate, x_a is mass scaled to teragrams (10^{12} g), and our individual simulations S_j have a source term from the volcano that emits one teragram of ash at the given time point α and altitude β. This means that our computed solution x is given in teragrams of emitted ash.

2.4 Efficient computation of ~~LSQR~~ least squares matrix

One novelty in this paper is the efficient assembly and calculation of the least squares matrix, enabling the inversion to run on a regular laptop computer. The linear system in Equation 5 is solved using the Tinkhonov regularization, formulated as

$$x = x_a + G^{-1} B, \tag{10}$$

by inverting the least squares matrix G. G is a matrix of dimension $n \times n$, in which n is the number of emissions we want to estimate. This matrix is computed from the larger M matrix of size $m \times n$, in which m is the number of observations. For the full Eyjafjallajökull eruption case, that means that up-to 56536603 observations[5] are used to calculate the 6061 unknown emissions.

The 56.5 million observations are matched against the simulations and stored in the matrix M so that each row has up-to 912 nonzero entries (19 emissions eight times per day for six days), but an average of around 170 in our runs. The matrix then has approximately 9.7 billion non-zeroes, which corresponds to 72 gigabytes of data. This is prohibitively expensive for the inversion procedure, and limits the number of observations that are possible to assimilate. However, by carefully examining and restructuring the products involved, we can reformulate the procedure to compute G directly without storing M. This reduces the memory requirement from the aforementioned 72 gigabytes to less than 300 megabytes (a factor 260 improvement), enabling efficient inversion even on laptop computers with no special hardware requirements.

[5] The full Eyjafjallajökull eruption case has 876906 observations of ash and 55.7 million observations of no ash. In addition, there are uncertain observations e.g., due to cloud coverage, that are not used.

The matrix G in Equation 9 is computed using

$$G = M^T \sigma_o^{-2} M + \sigma_x^{-2} + \epsilon D^T D. \tag{11}$$

Here, M is the enormeous sparse matrix with 72 gigabytes of nonzero entries, whilst G is a relatively modest dense matrix. The classical way to compute this expression is to use the inner product,

$$\begin{aligned}\left(M^T \sigma_o^{-2} M\right)_{i,j} &= \text{diag}(\sigma_o^{-2}) \circ (M_{:,i} \cdot M_{:,j}) \\ &= \sum_{k=1}^{m} \sigma_{ok}^{-2} M_{k,i} \cdot M_{k,j} \\ &= \sigma_{o1}^{-2} M_{1,i} \cdot M_{1,j} + \cdots + \sigma_{om}^{-2} M_{m,i} \cdot M_{m,j}, \end{aligned} \tag{12}$$

in which $\text{diag}(\sigma_o^{-2})$ is the vector composed of the diagonal entries of σ_o^{-2}. Recall that m is here the number of observations.

By examining this inner product, we see that we can compute the matrix G as a sum of outer products for each row of M instead of the inner product formulation for each column above:

$$G = \sum_{k=1}^{m} \sigma_{ok}^{-2} M_{k,:} \otimes M_{k,:} + \sigma_x^{-2} + \epsilon D^T D. \tag{13}$$

By using the outer product, we can avoid having to store the large matrix M in memory, and instead store only the relatively small $n \times n$ matrix. For the full Eyjafjallajökull eruption case that means that we store a dense matrix of 6061×6061 elements which takes less than 300 megabytes of memory. In essence, this means that we assemble the matrix G using one observation at a time.

Unfortunately, this also means that we need to compute

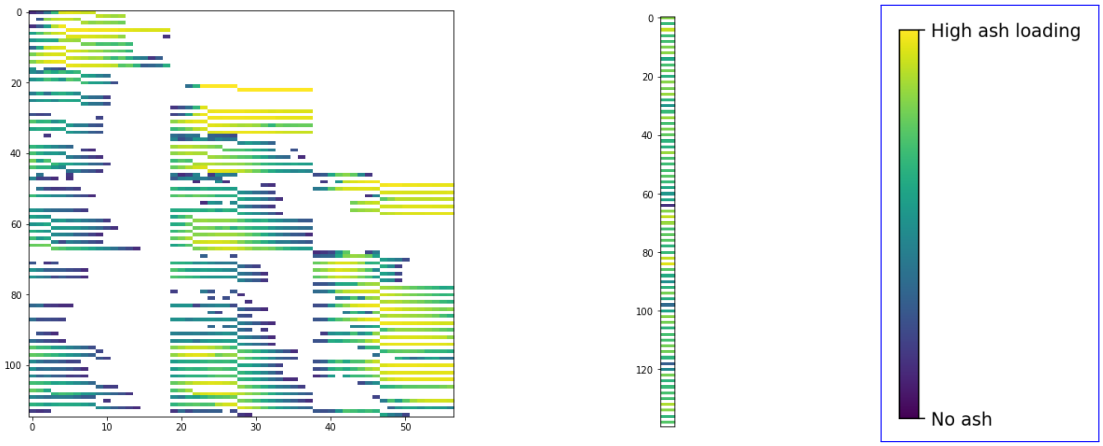
$$B = M^T \sigma_o^{-2} \tilde{y} = M^T \sigma_o^{-2} (y_0 - M x_a) \tag{14}$$

simultaneously. Whilst this does not seem to have a significant impact, it means that we can no longer change our a priori estimate, x_a, without also re-assembling the right hand side. However, it is trivial to use a matrix X_a, in which each column represents one a priori estimate, and solve for all these simultaneously.

Equations 12 and 13 represent two extremes in how to compute the matrix G, in which the former first assembles M and computes G using the (sparse) matrix product, and the latter assembles G directly one observation at a time using the outer product. We can also use a hybrid approach, in which we assemble a set of observations into G simultaneously. If we create a row partition of M, and call each partition $Q_{<1>}, \ldots, Q_{<p>}$, we can compute the same product as follows:

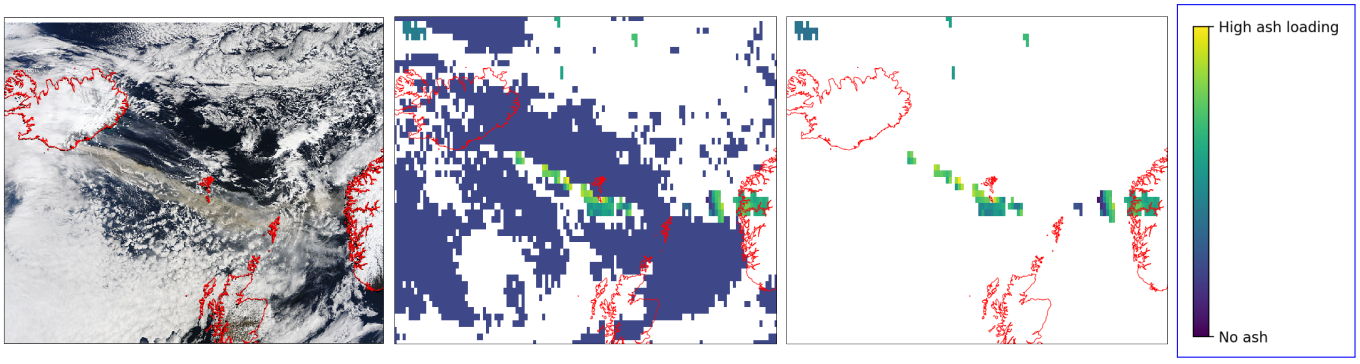
$$G = \sum_{k=1}^{p} Q_{<k>}^T \sigma_o^{-2} Q_{<k>} + \sigma_x^{-2} + \epsilon D^T D. \tag{15}$$

This is analogous of computing Equation 12 for a subset of the rows of M at a time, or a subset of observations at a time. The benefit of this hybrid approach is that the computational speed can increase dramatically, as we can use efficient dense linear algebra routines from Python libraries, thereby avoiding the prohibitive memory requirement of Equation 12 as well as the expensive ~~memory expansion~~ outer product formulation required for Equation 13. The total performance benefit of the hybrid approach is up-to 15 times in terms of wall clock time spent assembling and computing the ~~LSQR~~ least squares matrix.



(a) (b)

Figure 6. Linear system of equations with altitude information. (a) shows the source-receptor matrix M with altitude information, and (b) the corresponding vector of observations. Compare with Figure 4b and notice that each row is now split into two new rows. Even numbered rows now correspond to observations of ash (below the detected ash plume height), and odd numbered rows correspond to no ash. The rest of the solution procedure remains the same.



(a) (b) (c)

Figure 7. Satellite image and corresponding detection of ash. (a) is from NASA Terra/MODIS 2010/105 04/15/2010 11:35 UTC, and (b) is from the SEVIRI satellite instrument after detection of ash concentrations at 11:00. White pixels are unobserved or uncertain parts of the domain, and blue pixels are observations of zero ash. (c) shows only certain non-zero detection of ash.

2.5 Iterative inversion procedure

Because there are large uncertainties in both the meteorology and the satellite observations (see e.g., (Harvey et al., 2022)), we may end up with negative emission estimates at certain points, as there is nothing in our minimization problem that prohibits negative solutions. Negative values in the a posteriori are forced to lie closer to the a priori estimate by reducing the uncertainties, σ_x for these values and recomputing the solution. The iterative procedure repeats until the amount of negative ash emissions is reduced to a fraction (e.g., 1%) of the total a posteriori emission estimate.

2.6 Including ash cloud-top information

A novelty in this paper is the use of observed altitude from e.g., the SLSTR instrument. The dual-view capabilities of the SLSTR instrument may be used to detect the top of the ash plume, and thereby restrict the inversion procedure to give more correct altitudes for the a posteriori emissions. Mathematically, we formulate this by splitting each observation into two observations: one non-zero observation from the ground up to the detected plume height, and one zero observation from the plume height to the top of the model. In essence, we simply split each row in the source-receptor matrix, M into two as shown in Figure 6. This then doubles the number of rows in the matrix, whilst keeping the number of non-zeroes constant. The rest of the algorithm is unchanged.

The computational cost of this extra information is negligible for the whole algorithm as our final linear least squares system, G has the same dimension both with and without altitude information.

3 Verification using synthetic data

To check if the inversion procedure presented here works as intended is a non-trivial exercise, as there are large uncertainties in both the simulation model and observations being used. We have therefore checked our inversion procedure against a known truth, generated by the simulation model itself. We first generate an a priori emission estimate using radar observed plume heights and the mass eruption rate estimate in Mastin et al. (2009), and subsequently use this a priori estimate to generate ~~the synthetic truth consisting of~~ synthetic satellite images. The synthetic ~~truth is created by simply~~ satellite images are created by scaling the unit emission simulations with the a priori value, and then vertically integrate them to yield grams per square meters. ~~These synthetic satellite images observe the synthetic truth~~ We observe these satellite images at random locations in space and time to generate the truth. We then expect that our inversion procedure should generate an a posteriori which lies close to the a priori used to generate the truth. It should be noted that we do not expect a perfect inversion, as we try to estimate the vertical and time distribution of the ash emission whilst we only observe the vertically integrated ash concentration at certain points. We have used the Eyjafjallajökull 2010 eruption as a basis to generate a realistic scenario from April 14th to 18th. We also vary the a priori estimate and the ash cloud-top altitude to see the effect of these parameters on the solution. It should also be noted that synthetic benchmarks removes any uncertainty in both the transport model and the numerical weather forecast data.

Table 1. Summary of experiments using synthetic data. The observations column show what kind of synthetic observations are being used. The observations are generated using an eruption using the default a priori emission ("Default") or using the same amount of emission as the default a priori but with half the plume height ("Half altitude"). The a priori estimate used is either the default, or a constant estimate based on an 8.4 km plume height.

Case name	Observations	A priori	Altitude used?	Expected result / truth
A	Default	Default	Yes	Same as default a priori
B	Half altitude	Default	No	Same total emission as default a priori, but half altitude
C	Half altitude	Default	Yes	Same total emission as default a priori, but half altitude
D	Default	Constant 8.4 km plume	No	Same as default a priori
E	Default	Constant 8.4 km plume	Yes	Same as default a priori

We have devised several different experiments to test the inversion, which are summarized in the Table 1. The expected outcomes of the experiments are:

A: This case uses an a priori estimate that is identical to the truth, and the inversion procedure should be able to produce something close to the truth. This default a priori is created using an estimate of the eruption altitude and the relationship given by Mastin et al. (2009). Because we are using a Tikhonov ~~relaxation~~ regularization we will expect a small deviation between the default a priori and the computed a posteriori. If the inversion is unable to capture this state, we expect that something is wrong - either with the implementation or with the data.

B: This case uses the default a priori, but the truth has an emission altitude of half the a priori. As the inversion does not use altitude information in the inversion, we may not be able to resolve this case.

C: This case is identical to case B, but we use the altitude information in the inversion. Case B and C therefore showcase the added value of restricting the inversion with altitude information in the case we have a too large altitude estimate for the a priori.

D: This case shows how the inversion handles a constant a priori estimate without the use of altitudes in the inversion procedure. This is to simulate a "worst case scenario" in which we do not have a good description of the eruption, and therefore only have a very crude a priori estimate.

E: This case is identical to case D, but now using altitude information in the inversion.

All of these cases use synthetically generated satellite ~~images~~ observations (shown in Figure 8) in the inversion. These synthetic satellite ~~images~~ observations are generated by first multiplying the true a posteriori with the simulated emissions, followed by vertically integrating to achieve a column load for the whole domain. We finally sample at random locations to simulate true satellite images. The synthetic dataset consists of 235 observation timepoints with a total of 462654 ash observations, and covers the domain shown in Figure 8.

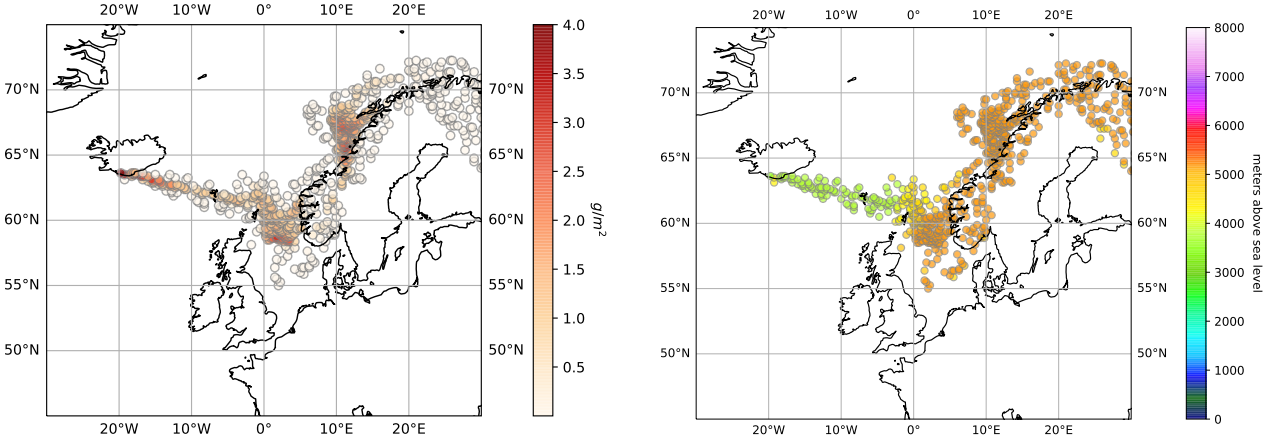


Figure 8. Synthetic satellite observations used in validation of inversion procedure with mass loading in g/m^2 (top) and ash cloud-top altitude in m (bottom). The synthetic observations are based upon a vertical integration of simulation results at specific coordinates. The figure shows a synthetic satellite observation from 10:00 on April 15 using the standard a priori estimate shown in Figure 9 (compare also with Figure 7 which shows actual satellite data). The coordinates are randomly generated, and consists of 774 observation points uniformly distributed in space.

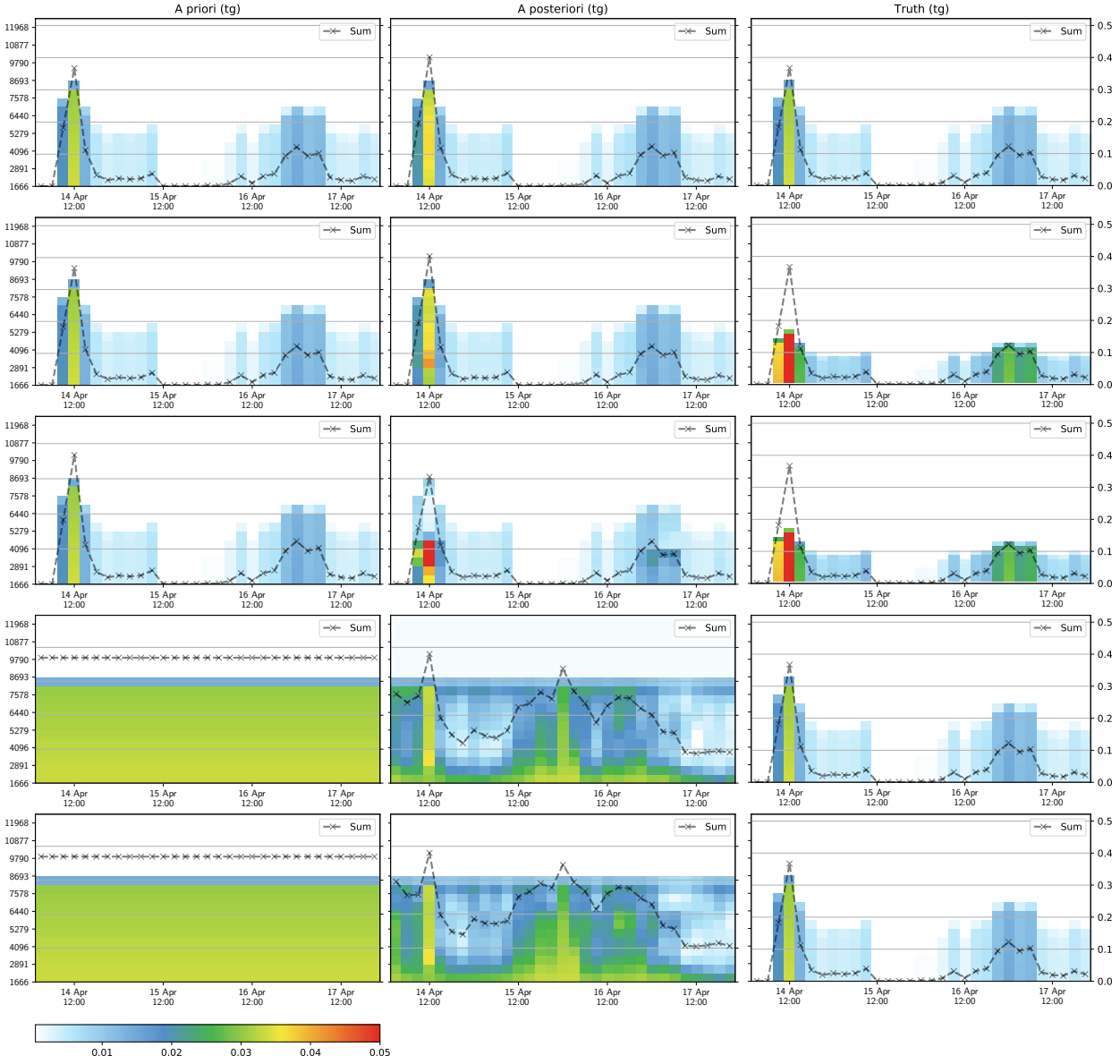


Figure 9. Verification of inversion code. The color bar is in teragrams of emitted fine ash, and the altitude is in meters above sea level (volcano altitude is 1666 meters above sea level). The rows (from top) represent the five cases A–E outlined in Table 1. The left column shows the a priori used in the inversion, the middle column shows the inverted a posteriori estimate, and the right column shows the truth. The left axis shows the emission altitude above sea level, and the right axis shows the total emission in teragrams. The first row uses an a priori equal the truth (case A), rows two and three use a truth in which the ash is emitted at half the altitude of the a priori (cases B and C), and rows four and five use a constant a priori estimate (cases D and E). Rows one (case A), three (case C) and five (case E) use ash cloud-top observations to restrict the inversion.

Figure 9 shows the result of running the verification experiments. For case A (the top row) the inversion algorithm is quite capable of recovering the truth. The inverted results show a slight increase in the total emission, in particular during the peak of the eruption on the 14th of April. However, the inverted results are overall in good agreement with the expected results.

For case B shown in row two, the effect of overestimating the altitude of the eruption in the a priori estimate is evident, illustrating that the inversion does not capture the true emission very well. However, the result is not very surprising as the inversion algorithm is only using observations of vertically integrated quantities. Without a restriction on the altitude in the inversion itself, there needs to be a significant difference in the prevailing winds at different altitudes to penalize emissions at wrong altitudes.

Case C in the third row shows the exact same experiment as in row two, but this time we use observations of the ash cloud-top to restrict the inversion. Compared with case B, we clearly see the benefit of using altitude observations. The inversion is capable of capturing the emission altitude very well during the peak of the eruption on April 14th, and also shows a relatively good reduction in higher altitude emissions during the eruption from April ~~16th–17th~~16th to 17th.

The D and E cases are shown in the final two rows which show how the inversion algorithm behaves when we do not have a good a priori estimate of the eruption. The results show that both with and without altitude observations, the algorithm is unable to find a satisfactory result. The two experiments are able to capture the peak of the eruption on April 14th quite well, but wrongfully estimates a larger eruption from April ~~15–16~~15th to 16th. This can be explained by our synthetic satellite ~~images~~ observations shown in Figure 8, which do not include observations of no ash. Thus, the algorithm does not have sufficient observations to restrict the inversion, and in particular the eruption free period from April ~~15–16~~15th to 16th.

4 Validation using the Eyjafjallajökull 2010 eruption

The previous section outlined some of the sensitivities of the inversion procedure on synthetic datasets, in which we have generated satellite ~~images~~ observations from a synthetic known truth. In this section, we apply the methodology to a real-world eruption, the Eyjafjallajökull 2010 eruption. The Eyjafjallajökull eruption is a well studied case which makes it possible to compare the quality of our results with other approaches.

We have run two scenarios with our inversion code. The first uses the same setup as Steensen et al. (2017a), and covers the onset of the eruption from April 14th to April 18th with an a priori estimated from eruption altitude and the relationship given by Mastin et al. (2009). The second mimics the setup of Stohl et al. (2011), and covers the whole eruption from April 14th to May 24th with a more complex a priori estimate. The full dataset consists of 959 observation timepoints and a total of 876906 non-zero observations (detected ash), and the eruption onset dataset is simply a subset of the 96 first observation timepoints with a total of 43241 non-zero observations. The extent of the dataset is shown in Figure 11.

4.1 Eyjafjallajökull eruption onset

In this case, we use the same setup as in Steensen et al. (2017a), and have, as far as possible, used the same parameters. We have used the same SEVIRI satellite data and simulated unit emission simulations with the eEMEP advection model.

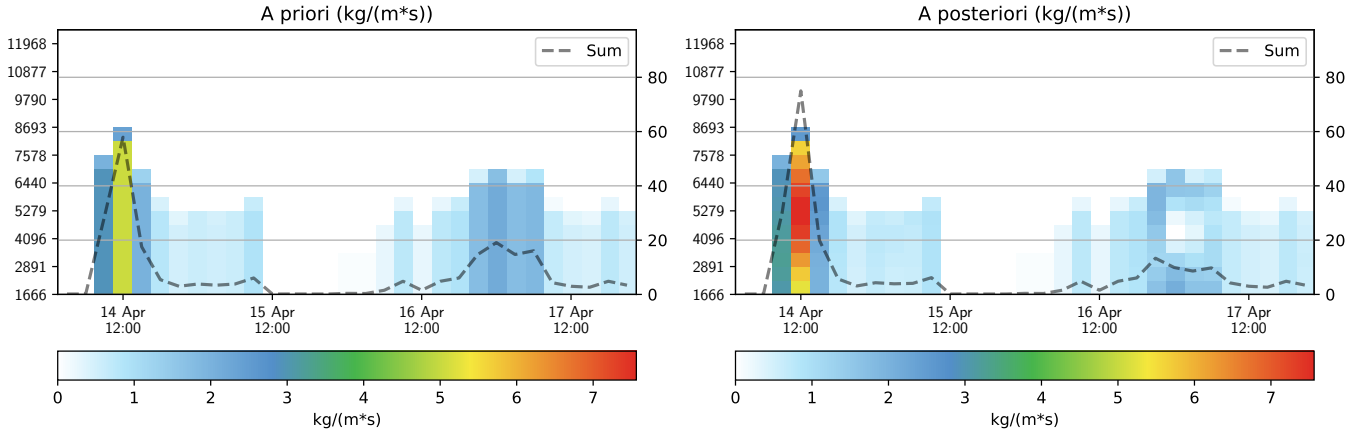


Figure 10. Comparison of the a priori estimate with the inverted results (in $kg/(m \cdot s)$) for the Eyjafjallajökull eruption onset. The altitude is in meters above sea level (volcano altitude is 1666 meters above sea level). The dashed line shows the total erupted mass in teragrams.

Figure 10 shows the result of the inversion algorithm. Overall our results match that of Steensen et al. (2017a), but there are some differences. We find that the main emissions for 14 April are fairly evenly spread from three to seven km and with a maximum around five km, while in Steensen et al. (2017a) the emissions peaks around 8 km and there is no emission below 3 km; 2) early on 17 April we find emissions below 6 km while none are reported by Steensen et al. (2017a). It is noted that results presented here agree better with those presented by (Stohl et al., 2011, their Fig. 2c). We have identified several possible reasons for the differences with Steensen et al. (2017a): this is a new implementation and we have used different parameters in the simulation runs with eEMEP[6]; there are different parameters used in the inversion run; there is a slightly different vertical discretization of the eEMEP model; and the advection model (eEMEP) has gone through significant upgrades and changes. As the details required to reproduce the runs of Steensen et al. (2017a) are not fully available, it is not possible to reproduce their results.

4.2 Eyjafjallajökull full eruption

We use a similar setup as in (Stohl et al., 2011) in the inversion of the full eruption period of Eyjafjallajökull, lasting from April 14 to around May 24th 2010. Our inversion results (Figure 12) show a clear reduction of the eruption strength around April 16, as well as the two high altitude eruptions on May 14th and 16th, and a significant increase of the eruption on May 12th – 13th. Our results generally match well with that of (Stohl et al., 2011), but there are some differences. In particular, the eruption onset on April 14th appears to be stronger in our results, and our results maintain more ash in the aposteriori below 3 km. There are several factors that contribute to these differences: different advection models are used (eEMEP and FLEXPART); different meteorological data sets (ECMWF and a combination of ECMWF and GFS); differences in analysis of satellite observations; and the a priori estimate is similar but not identical.

[6] ~~e.g., use of gravitational settling.~~

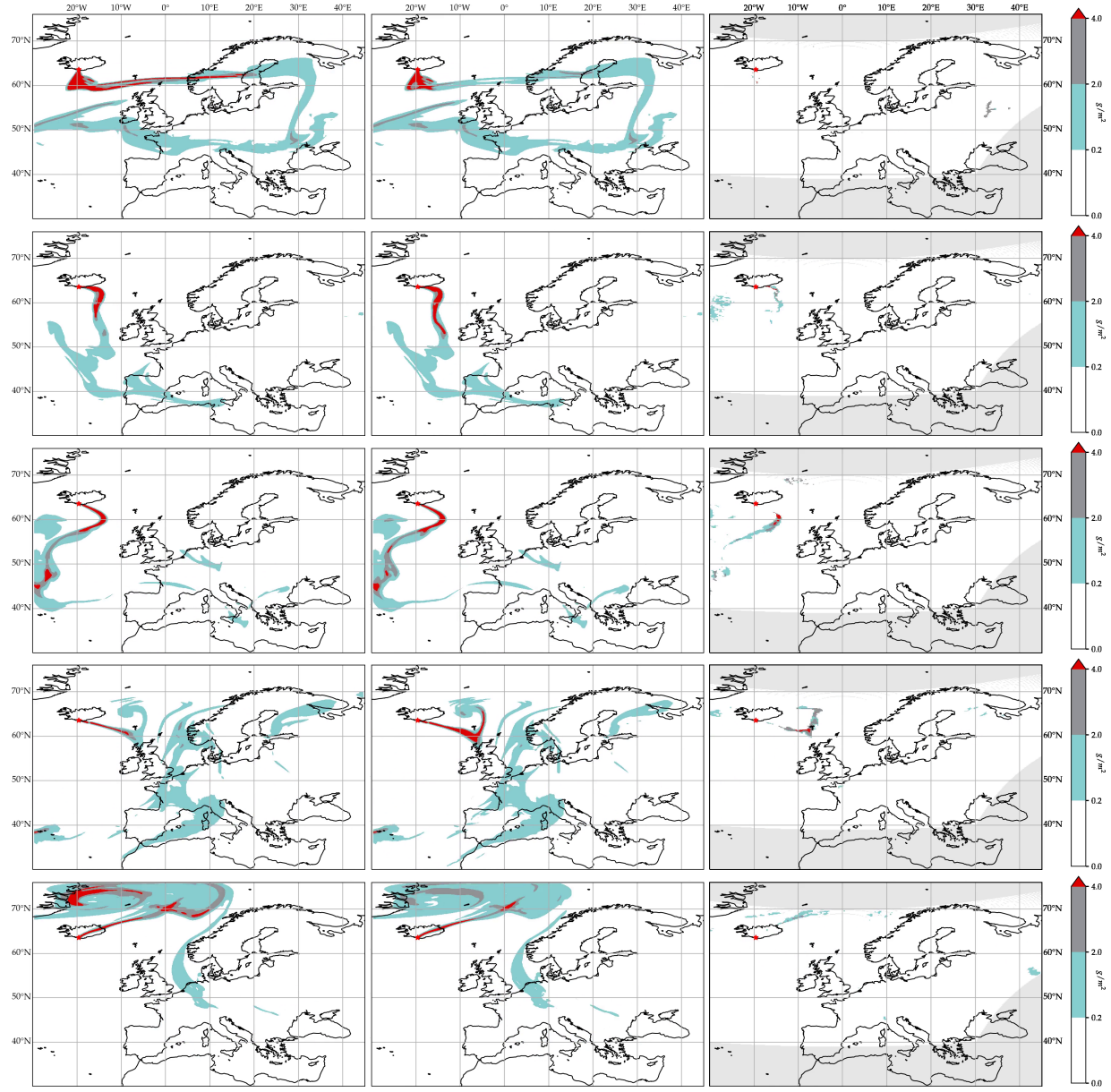


Figure 11. Modelled and observed ash clouds over Europe during the Eyjafjallajökull 2010 eruption at different times. From top to bottom, the figures represent 2010-04-17T19:00Z, 2010-05-06T10:00Z, 2010-05-08T19:00Z, 2010-05-13T03:00Z, and 2010-05-18T16:00Z. The leftmost column shows the modelled a priori estimate, the next column shows the modelled a posteriori estimate after inversion, and the rightmost column shows the satellite image. The colors in the figures correspond to the color codes used operationally on the Internet Pilot Planning Centre (IPPC) operated by Avinor Air Navigation Services AS (https://ippc.no/). The color codes groups into no ash ($< 0.2\ g/m^2$), low ash ($0.2 - 2.0\ g/m^2$), medium ash ($2.0 - 4.0\ g/m^2$), and high ash ($> 4.0\ g/m^2$) total column loadings. For the satellite images, not-analysed areas are in light gray.

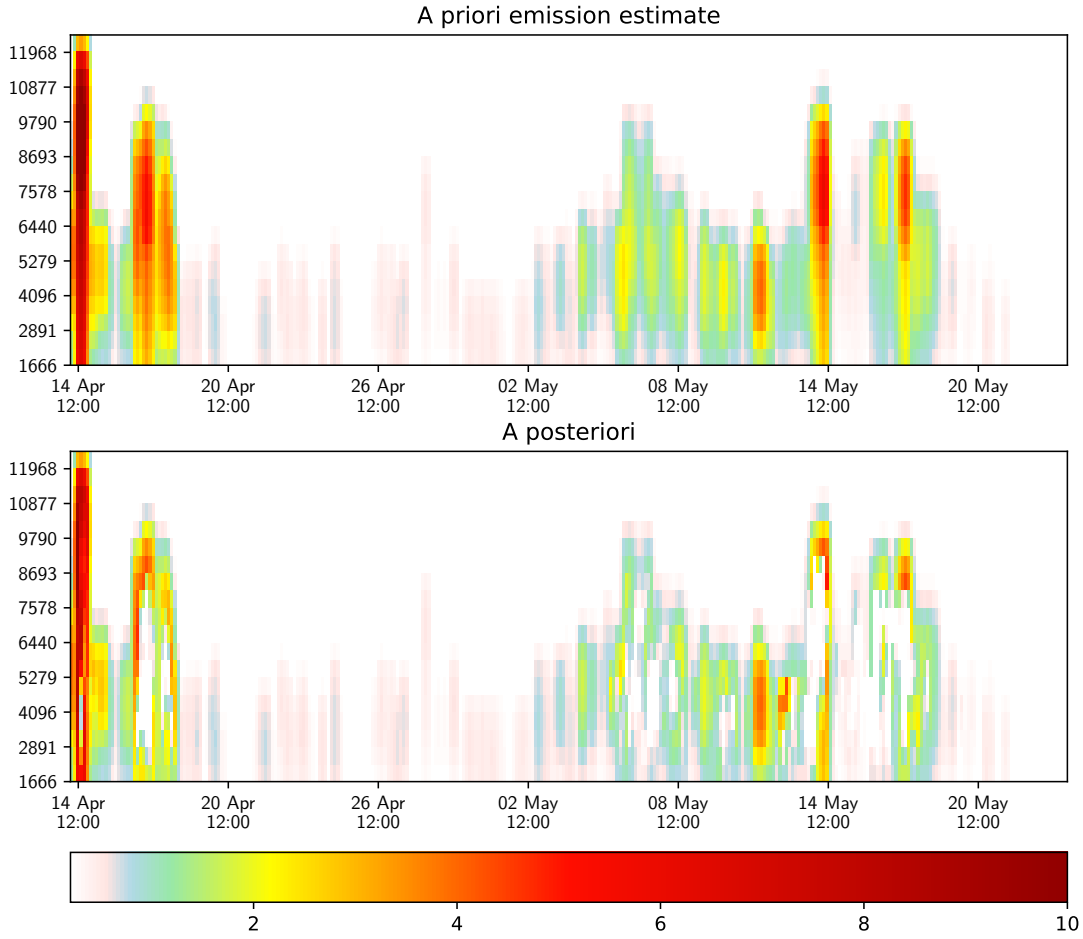


Figure 12. Eyjafjallajökull full eruption a priori vs inverted emission estimate (in $kg/(m \cdot s)$). The altitude is in meters above sea ~~lelve~~level, and the volcano summit is at 1666 meters above sea level.

Figure 11 compares our a priori and a posteriori results with the retrieved ash from satellite observations for five different time points during the eruption. The figure shows quantized results in the same color scheme as the official Norwegian volcanic ash reports used by airline pilots[6], with no ash ($< 0.2\ g/m^2$), low ash ($0.2 - 2.0\ g/m^2$), medium ash ($2.0 - 4.0\ g/m^2$), and high ash ($> 4.0\ g/m^2$) total column loadings. The first column shows the a priori estimate, the central column shows the a posteriori result, and the right column shows the satellite images.

The satellite images illustrate that the observations of ash are relatively sparse and that they suffer from false positives (including detected ash far from the simulated ash cloud, see also Figure 7). They also show that there are differences between the position of the modeled and observed ash clouds, as shown in the forth row close to the Faroe Islands. This is possibly due to errors in meteorology driving the advection model. Both false positives and positional errors between observed and

[6]The Norwegian volcanic ash reports are available on https://ippc.no/.

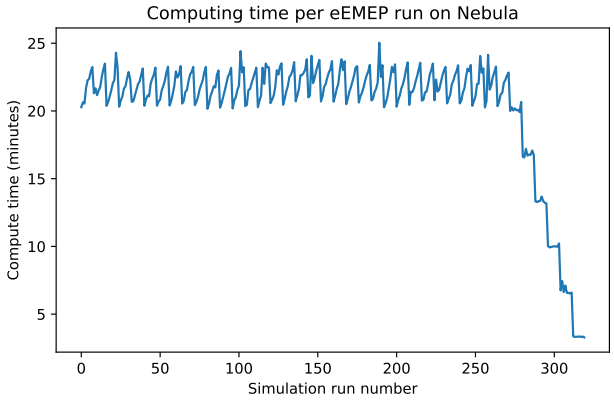


Figure 13. Computing time per forward simulation on the Nebula cluster. Notice that the last simulations take significantly less time as they approach the end simulation time (May 24th 2010). The regular sawtooth pattern for the other runs is similarly explained by each new simulation starting at 00:00, even though the emission may start at e.g., 21:00. This is due to a technical limitation in the eEMEP model.

modeled ash clouds make the inversion scenario significantly more challenging than the idealized data used for the verification in Section 3.

If we compare the a priori estimate with the a posteriori results, we see that the inversion algorithm reduces the areas with high ash column loadings during parts of the eruption (rows one, three and five), whilst increasing it at other time points (rows two and four). Overall, this matches well with the observations of ash in the satellite images.

5 Performance assessment

The inversion framework consists of the four main stages (Figure 1), and we have optimized each of these. The most novel improvements and optimizations have been in the assembly of the least squares system in stage three, but the other stages have also been improved.

In stage one, we simulate the time development of unit emissions for a given altitude and time. This requires over 6000 unique simulation results, which is reduced to 319 eEMEP runs by using tracers (see also Section 2.1). Still, each run can take up-to 40 minutes to complete, thus requiring over 200 computing hours on the Nebula supercomputer infrastructure. We have optimized the eEMEP runs by limiting the simulation time to a maximum of 6 days (thus assuming ash in the atmosphere older than 6 days will not be important for the inversion), which reduces the requirement to 110 computing hours on the cluster, as shown in Figure 13. By running five simulations in parallel on the cluster, the wall clock time spent for the forward simulations is reduced to ten hours. By fully utilizing the cluster, it would be technically possible to reduce the wall clock time to approximately one hour.

Stage two is also a computationally demanding stage, and consists of collocating all observations with the simulations created in stage one as well as generating the a priori estimate. The generation of the a priori estimate is not a computationally demanding challenge, but the collocation of observations is quite demanding as it requires us to match each observation to the

Table 2. Summary of wall clock time for the inversion procedure for the full Eyjafjallajökull eruption.

Stage	Old approach	New approach
One: forward simulations	19 hours	10 hours
Two: Collocation of observations and simulations	5 hours	5 hours
Three: Least squares assembly	19 hours	1 hour
Four: Iterative inversion	1 minute	1 minute
Whole inversion	43 hours	16 hours

unit emissions simulation that might have caused it. This means that for each of the 56.5 million observations we have to find the corresponding temporospatial location in 912 unique simulations (see also Subsection 2.4). We can perform this on the PPI cluster[7], and the process takes just shy of five hours wall clock time. This process is parallelizable (e.g., each of the 962 hourly satellite images can easily be processed independently), and it would be technically possible to significantly reduce the wall clock time by utilizing the full cluster.

Finally, stage three is the last of the computationally demanding stages, in which we assemble the collocated simulations/observations into the least squares system. This process takes over 19 hours to complete on the PPI cluster when first assembling the huge M matrix, and subsequently computing the matrix product ~~$M^t\sigma_o^{-2}M$~~$M^T\sigma_o^{-2}M$. The reformulation to compute G matrix directly reduces this process to just over one hour. This stage may in fact also be parallelized further by following an approach similar to that shown in Equation 15, i.e., partitioning the set of observations, assembling

$$G_q = Q_{<q>}^T \sigma_o^{-2} Q_{<q>} + \sigma_x^{-2} + \epsilon D^T D,$$

in parallel, and finally summing to compute the final matrix

$$G = \sum_{q=1}^{p} G_q.$$

The performance increases in this new approach are summarized in Table 2. In total, the inversion procedure runs approximately 2.7 times faster with the improvements discussed in this paper. Furthermore, an added benefit of the new inversion procedure is that it is extremely memory efficient, allowing it to run efficiently on a simple laptop[8].

6 Summary and discussion

We have presented an inversion algorithm and open source implementation for volcanic ash emission estimates based on satellite imagery. The main novelties of this paper includes a memory-efficient algorithm for assembly and calculation of the least squares system, inclusion of ash cloud-top altitude information in the inversion, and verification using synthetic data sets.

[7]PPI - the Post Processing Infrastructure, is a linux-based compute cluster at the Meteorological Institute of Norway that runs GridEngine. The nodes have 16/32 core AMD CPUs, 130 GiB memory, and are connected with 100 Gbps infiniband interconnect.

[8]The forward simulations using the eEMEP model must run on a supercomputing infrastructure with access to the relevant meteorology.

Our verification shows the potential benefit of using observations of the ash cloud-top to restrict the inversion, and we have also run our code on the 2010 Eyjafjallajökull eruption. Our results are mostly in agreement with previously presented results, but with some differences believed to be mainly due to differences in model setup and meteorology used.

Our approach to assembling and calculating the least squares system has reduced the computational time from over 19 hours to just one for this part of the inversion. For the simulation of the unit emissions, the computational time has gone from 19 hours to ten hours. These two stages constitute the most time demanding parts of the inversion, and reduce the total time to solution by a factor 2.7.

Our framework is also well tailored to an operational setting with an eruption going on for several weeks. It is technically possible to use a least squares matrix G_1 representing e.g., the first three days of the eruption, and simply add another matrix G_2 that may represent a subsequent four days (and equivalently for the B vector), instead of assembling the full seven day period from scratch. With satellite imagery coming in continuously, this may save large amounts of computational time to create up-to-date estimates.

Our results show that using ash cloud-top altitude information gives a clear benefit, which contrasts the findings of Chai et al. (2017), who report a negative impact of using this information in the inversion. Our explanation to this is that we have tested with synthetic data in which there is no meteorological error. In the presence of meteorological error, we agree that there may be circumstances where adding more constraints to the inversion may negatively impact the results. Nevertheless, our results show that using ash cloud-top observations can improve the inversion results under favourable conditions.

This work has several opportunities for further improvement that we see natural to pursue. First of all, the use of coordinates in the collocation of satellite observations and simulations makes the inversion procedure sensitive to slight differences in modeled versus true meteorology. It would be fruitful to explore other forms of collocation strategies, e.g., using polar coordinates centered on the erupting volcano, or streamline distance from it, to possibly reduce this sensitivity. The use of ensemble meteorology coupled with ensemble inversion will possibly also address the sensitivity to differences between modeled and observed meteorology (see e.g., Webster and Thomson (2022); Crawford et al. (2022)).

The current algorithm is somewhat robust when it comes to false positives (falsely detected ash) and false negatives (falsely detected no ash) in the satellite images. However, it will be valuable to quantify this sensitivity using synthetic satellite images, i.e., determining when false negatives and positives will affect the inverted results. On a similar note, full uncertainty quantification of the inversion will be important to determine when and where the inversion is most valid in the presence of different kinds of errors, including false positives, false negatives, meteorological errors, and inversion parameter settings.

Another aspect to explore is the use of iterative least squares to solve the inversion problem without ~~non-physical~~ nonphysical (negative) emissions. An alternate approach may be to use non-negative least squares or other optimization strategies to solve the system.

6.1 Code and data availability

All source code The source code and data used in this work is released under an open source license, and available on Github (Brodtkorb, 2020b). The datasets used are also available under open licenses (Brodtkorb et al., 2020; Brodtkorb, 2020a; Kylling and . licenses, and archived versions are available on Zenodo:

Software package: André R. Brodtkorb. (2022). VolcanicAshInversion: v1.2.1 (v1.2.1). Zenodo. https://doi.org/10.5281/zenodo.8073110.

Satellite observations: André Rigland Brodtkorb. (2020). Eyjafjallajökull satellite observations (1.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3855526.

Forward eEMEP simulations: André Rigland Brodtkorb, Alvaro Valdebenito, & eEMEP contrubutors. (2020). Three-hourly gridded volcanic ash emissions for the Eyjafjallajökull 2010 eruption [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3818196.

6.2 Acknowledgments

This work has partly been performed as part of project number NIT.09.16.05 funded by the Norwegian Space Agency. Simulations have been run on the research and development Nebula supercomputer funded by the MetCoOp HPC infrastructure.

Author contribution / CRediT statement

A. R. Brodtkorb: Methodology Development, Software, Validation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization; A. Benedictow: Resources; A. Kylling: Conceptualization, Writing - Review & Editing, Funding acquisition; H. Klein: Conceptualization, Writing - Review & Editing, Funding acquisition; A. Nyiri: Data Curation; A. Valdebenito: Methodology; N. Kristiansen: Methodology Development, Writing - Review & Editing;

References

Beckett, F. M., Witham, C. S., Leadbetter, S. J., Crocker, R., Webster, H. N., Hort, M. C., Jones, A. R., Devenish, B. J., and Thomson, D. J.: Atmospheric Dispersion Modelling at the London VAAC: A Review of Developments since the 2010 Eyjafjallajökull Volcano Ash Cloud, Atmosphere, 11, 352, https://doi.org/10.3390/atmos11040352, 2020.

Bott, A.: A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes, Monthly Weather Review, 117, 1006–1016, 1989.

Brodtkorb, A. R.: Eyjafjallajökull satellite observations, https://doi.org/10.5281/zenodo.3855526, 2020a.

Brodtkorb, A. R.: VolcanicAshInversion, doi: 10.5281/zenodo.3818001, 2020b.

Brodtkorb, A. R., Valdebenito, A., and eEMEP contributors: Eyjafjallajökull unit emission simulations, 10.5281/zenodo.3818196, 2020.

Casadevall, T.: Volcanic Ash and Aviation Safety: Proceedings of the First International Symposium on Volcanic Ash and Aviation Safety, 2047, 1994.

Chai, T., Crawford, A., Stunder, B., Pavolonis, M. J., Draxler, R., and Stein, A.: Improving volcanic ash predictions with the HYSPLIT dispersion model by assimilating MODIS satellite retrievals, Atmospheric Chemistry and Physics, 17, 2865–2879, https://doi.org/10.5194/acp-17-2865-2017, 2017.

Clarkson, R. J., Majewicz, E. J., and Mack, P.: A re-evaluation of the 2010 quantitative understanding of the effects volcanic ash has on gas turbine engines, Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 230, 2274–2291, https://doi.org/10.1177/0954410015623372, 2016.

Crawford, A., Chai, T., Wang, B., Ring, A., Stunder, B., Loughner, C. P., Pavolonis, M., and Sieglaff, J.: Evaluation and bias correction of probabilistic volcanic ash forecasts, Atmospheric Chemistry and Physics, 22, 13 967–13 996, https://doi.org/10.5194/acp-22-13967-2022, 2022.

Eckhardt, S., Prata, A., Seibert, P., Stebel, K., and Stohl, A.: Estimation of the vertical profile of sulfur dioxide injection into the atmosphere by a volcanic eruption using satellite column measurements and inverse transport modeling, Atmospheric Chemistry and Physics, 8, 3881–3897, 2008.

EMEP MSC-W: metno/emep-ctm: OpenSource rv4.17 (201802), https://doi.org/10.5281/zenodo.3355023, 2018.

Gouhier, M., Deslandes, M., Guéhenneux, Y., Hereil, P., Cacault, P., and Josse, B.: Operational response to volcanic ash risks using HOTVOLC satellite-based system and MOCAGE-accident model at the Toulouse VAAC, Atmosphere, 11, 864, 2020.

Harvey, N. J., Dacre, H. F., Saint, C., Prata, A. T., Webster, H. N., and Grainger, R. G.: Quantifying the impact of meteorological uncertainty on emission estimates and the risk to aviation using source inversion for the Raikoke 2019 eruption, Atmospheric Chemistry and Physics, 22, 8529–8545, 2022.

Heng, Y., Hoffmann, L., Griessbach, S., Rößler, T., and Stein, O.: Inverse transport modeling of volcanic sulfur dioxide emissions using large-scale simulations, Geoscientific Model Development, 9, 1627–1645, https://doi.org/10.5194/gmd-9-1627-2016, 2016.

Horváth, Á., Carr, J. L., Girina, O. A., Wu, D. L., Bril, A. A., Mazurov, A. A., Melnikov, D. V., Hoshyaripour, G. A., and Buehler, S. A.: Geometric estimation of volcanic eruption column height from GOES-R near-limb imagery–Part 1: Methodology, Atmospheric Chemistry and Physics, 21, 12 189–12 206, 2021.

Kylling, A.: Ash and ice clouds during the Mt Kelud February 2014 eruption as interpreted from IASI and AVHRR/3 observations, Atmospheric Measurement Techniques, 9, 2103–2117, 2016.

Kylling, A. and Sollum, E.: Hourly non-gridded volcanic ash properties retrieved from SEVIRI measurements for the Eyjafjallajökull 2010 eruption, https://doi.org/10.5281/zenodo.3830363, 2020.

Lowenstern, J. B., Wallace, K., Barsotti, S., Sandri, L., Stovall, W., Bernard, B., Privitera, E., Komorowski, J.-C., Fournier, N., Balagizi, C., et al.: Guidelines for volcano-observatory operations during crises: recommendations from the 2019 volcano observatory best practices meeting, Journal of Applied Volcanology, 11, 1–24, 2022.

Mastin, L., Guffanti, M., Servranckx, R., Webley, P., Barsotti, S., Dean, K., Durant, A., Ewert, J., Neri, A., Rose, W., Schneider, D., Siebert, L., Stunder, B., Swanson, G., Tupper, A., Volentik, A., and Waythomas, C.: A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions, Journal of Volcanology and Geothermal Research, 186, 10–21, 2009.

Mastin, L., Pavolonis, M., Engwell, S., Clarkson, R., Witham, C., Brock, G., Lisk, I., Guffanti, M., Tupper, A., Schneider, D., et al.: Progress in protecting air travel from volcanic ash clouds, Bulletin of Volcanology, 84, 1–9, 2022.

Osores, S., Ruiz, J., Folch, A., and Collini, E.: Volcanic ash forecast using ensemble-based data assimilation: an ensemble transform Kalman filter coupled with the FALL3D-7.2 model (ETKF–FALL3D version 1.0), Geoscientific Model Development, 13, 1–22, https://doi.org/10.5194/gmd-13-1-2020, 2020.

Pardini, F., Corradini, S., Costa, A., Esposti Ongaro, T., Merucci, L., Neri, A., Stelitano, D., and de' Michieli Vitturi, M.: Ensemble-Based Data Assimilation of Volcanic Ash Clouds from Satellite Observations: Application to the 24 December 2018 Mt. Etna Explosive Eruption, Atmosphere, 11, 359, https://doi.org/10.3390/atmos11040359, 2020.

Pelley, R. E., Thomson, D. J., Webster, H. N., Cooke, M. C., Manning, A. J., Witham, C. S., and Hort, M. C.: A Near-Real-Time Method for Estimating Volcanic Ash Emissions Using Satellite Retrievals, Atmosphere, 12, https://doi.org/10.3390/atmos12121573, 2021.

Prata, A. T., Grainger, R. G., Taylor, I. A., Povey, A. C., Proud, S. R., and Poulsen, C. A.: Uncertainty-bounded estimates of ash cloud properties using the ORAC algorithm: application to the 2019 Raikoke eruption, Atmospheric Measurement Techniques, 15, 5985–6010, https://doi.org/10.5194/amt-15-5985-2022, 2022.

Prata, F. and Lynch, M.: Passive Earth Observations of Volcanic Clouds in the Atmosphere, Atmosphere, 10, 199, https://doi.org/10.3390/atmos10040199, 2019.

Seibert, P.: Inverse Modelling of Sulfur Emissions in Europe Based on Trajectories, pp. 147–154, American Geophysical Union (AGU), https://doi.org/10.1029/GM114p0147, 2000.

Seibert, P., Kristiansen, N., Richter, A., Eckhardt, S., Prata, A., and Stohl, A.: Uncertainties in the inverse modelling of sulphur dioxide eruption profiles, Geomatics, Natural Hazards and Risk, 2, 201–216, 2011.

Simpson, D., Benedictow, A., Berge, H., Bergström, R., Emberson, L. D., Fagerli, H., Flechard, C. R., Hayman, G. D., Gauss, M., Jonson, J. E., Jenkin, M. E., Nyíri, A., Richter, C., Semeena, V. S., Tsyro, S., Tuovinen, J.-P., Valdebenito, A., and Wind, P.: The EMEP MSC-W chemical transport model – technical description, Atmospheric Chemistry and Physics, 12, 7825–7865, https://doi.org/10.5194/acp-12-7825-2012, 2012.

Simpson, D., Wind, P., Bergström, R., Gauss, M., Tsyro, S., and Valdebenito, Á.: Updates to the EMEP/MSC-W model, 2017–2018, in: Transboundary particulate matter, photo-oxidants, acidifying and eutrophying components. EMEP Status Report 1/2018, pp. 109–116, The Norwegian Meteorological Institute, Oslo, Norway, https://www.emep.int/, 2018.

Steensen, B., Kylling, A., Kristiansen, N., and Schulz, M.: Uncertainty assessment and applicability of an inversion method for volcanic ash forecasting, Atmospheric Chemistry and Physics, 17, 9205–9222, 2017a.

Steensen, B. M., Schulz, M., Wind, P., Valdebenito, Á. M., and Fagerli, H.: The operational eEMEP model version 10.4 for volcanic SO 2 and ash forecasting, Geoscientific Model Development, 10, 1927–1943, 2017b.

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

Tichý, O., Ulrych, L., Šmídl, V., Evangeliou, N., and Stohl, A.: On the tuning of atmospheric inverse methods: comparisons with the European Tracer Experiment (ETEX) and Chernobyl datasets using the atmospheric transport model FLEXPART, Geoscientific Model Development, 13, 5917–5934, https://doi.org/10.5194/gmd-13-5917-2020, 2020.

Webster, H. N. and Thomson, D. J.: Using ensemble meteorological datasets to treat meteorological uncertainties in a Bayesian volcanic ash inverse modelling system: A case study, Grímsvötn 2011, Journal of Geophysical Research: Atmospheres, https://doi.org/https://doi.org/10.1029/2022JD036469, 2022.

Zakšek, K., Hort, M., Zaletelj, J., and Langmann, B.: Monitoring volcanic ash cloud top height through simultaneous retrieval of optical data from polar orbiting and geostationary satellites, Atmospheric Chemistry and Physics, 13, 2589–2606, 2013.

Zidikheri, M. J. and Lucas, C.: Using Satellite Data to Determine Empirical Relationships between Volcanic Ash Source Parameters, Atmosphere, 11, 342, https://doi.org/10.3390/atmos11040342, 2020.