Estimation of Volcanic Ash Emissions Using Trajectory-Based 4D-Var Data Assimilation

S. LU, H. X. LIN, A. W. HEEMINK, AND G. FU

Delft University of Technology, Delft, Netherlands

A. J. SEGERS

TNO, Department of Climate, Air and Sustainability, Utrecht, Netherlands

(Manuscript received 18 May 2015, in final form 15 October 2015)

ABSTRACT

Volcanic ash forecasting is a crucial tool in hazard assessment and operational volcano monitoring. Emission parameters such as plume height, total emission mass, and vertical distribution of the emission plume rate are essential and important in the implementation of volcanic ash models. Therefore, estimation of emission parameters using available observations through data assimilation could help to increase the accuracy of forecasts and provide reliable advisory information. This paper focuses on the use of satellite total-ash-column data in 4D-Var based assimilations. Experiments show that it is very difficult to estimate the vertical distribution of effective volcanic ash injection rates from satellite-observed ash columns using a standard 4D-Var assimilation approach. This paper addresses the ill-posed nature of the assimilation problem from the perspective of a spurious relationship. To reduce the influence of a spurious relationship created by a radiate observation operator, an adjoint-free trajectory-based 4D-Var assimilation method is proposed, which is more accurate to estimate the vertical profile of volcanic ash from volcanic eruptions. The method seeks the optimal vertical distribution of emission rates of a reformulated cost function that computes the total difference between simulated and observed ash columns. A 3D simplified aerosol transport model and synthetic satellite observations are used to compare the results of both the standard method and the new method.

1. Introduction

Volcanic eruptions lead to the release of greenhouse gases, ashes, and chemicals that influences climate (Robock 2000). Volcanic ash impacts respiratory health of human beings, and can cause acute and chronic diseases (Horwell and Baxter 2006). People in a wide area covered by extensive ash fallout will be affected, and there are long-term health risks after ashfall. Despite the relatively rapid settling process of large particles (Rose et al. 2001), fine particles with small radius can float in the atmosphere for a long time, be transported, and spread out in a continental domain. Floating ash, containing silicate that can melt in the engines and cause devastating risk to jet planes, could affect commercial and military air operations, and disturb local economy

(Casadevall 1994). Therefore, rerouting and cancellation is necessary according to hazard assessment.

Volcanic ash forecasting is a crucial tool in hazard assessment and operational volcano monitoring. Analysts use volcanic ash transport models to determine the future location of ash clouds. In addition to validated atmospheric transport models that are capable of modeling the physical behavior of volcanic ash transport and dispersion process, accurate estimation of emission source terms is needed in order to obtain reliable prediction results. However, the source parameters, such as total mass of eruption products and the altitude at which they are effectively released into the atmosphere, are hardly known. Thus, the source term, which usually appears in forms of vertical distribution of emission rates in an aerosol transport model, must be accurately estimated.

It is a great challenge to accurately determine source parameters such as vertical distribution of volcanic emission plume. Volcanoes exhibit a wide range of eruption types, thus, it is difficult to make assumptions

Corresponding author address: S. Lu, Delft University of Technology, Mekelweg 4, 2628 CD Delft, Netherlands.
E-mail: s.lu-1@tudelft.nl

of emission parameters estimation for predictive purpose. Up to now, indirect methods are used to estimate volcanic ash emissions (Mastin et al. 2009). For instance, the empirical relationship between plume height above vent and total volume is built up to determine total emission rates from plume altitude observations. This plume height could be obtained from aircraft measurements (Mankin et al. 1992) or ground-based radar or lidar observations (Wang et al. 2008), which are often not available. Then, in practice, explicit assumptions on the vertical distribution have to be made such as a uniform-distributed (Dacre et al. 2011), umbrella-shaped (Stuefer et al. 2012; Webley et al. 2012), Poisson-distributed, or exponentially distributed plume (Searcy et al. 1998). Even if the vertical distribution is more or less correct, the relationships between plume height and total emission is loose, thus the vertical assignments of the total emission rates to each layer is poorly quantified.

Therefore, systematic and automatic estimation approaches are needed to correct emission parameters and improve forecast accuracy. In this study the use of data assimilation (DA) is considered. Data assimilation techniques have been successfully applied to meteorology and oceanography, to produce improved model states forecast or model parameters, by combining model and observations. There are two main categories of modern DA approaches: filtering (Evensen 1994) and variational methods (Le Dimet and Talagrand 1986). The former is a class of sequential methods, which seeks a balance between prior model states and observations by minimizing the updated error covariance matrix. The latter search an optimal set of control variables by minimizing a cost function measuring the difference between the model outputs and observations, which is preferable when estimating atmospheric emissions (Strunk et al. 2010). In this work we adopt a four-dimensional variational data assimilation (4D-Var) method to estimate vertical emission profiles as the forcing input terms of the model.

There are centers already employing 4D-Var to improve the forecast accuracy of air pollution, chemical transport, and other similar issues. However, the existing operational three-dimensional variational data assimilation (3D-Var) or 4D-Var systems in these centers are usually developed to estimate the initial conditions (initial model fields) using in situ data, and the number of applications to volcanic ash problem are quite limited. Valdebenito et al. (2010) assimilated the particle mass (PM) concentrations near the surface with observations provided by Dutch stations, and corrected the PM concentrations in other places according to the correlation between the observed state variables and the unobserved ones. The emission rates were adjusted through experience (not using systematic approach) by comparison of aerosol optical depth (AOD) values with

the simulated model forecasts and the retrievals from satellite instruments. Because of the insufficiency of in situ observations, and the small correlation between the state variables at the surface layer and the ones at higher layers, the estimation of volcanic ash concentrations at higher layers remains unreliable. Accurate estimation of volcanic ash cloud at higher altitudes is critical for aviation hazard information, which needs a more sufficient observation sets and an systematic way to assimilation them.

As a powerful tool of volcanic activity monitoring and model validation with a wide coverage and continuous access, satellite data have been considered to be assimilated in the volcanic cloud transport system by researchers. Flemming and Inness (2013) used satellite observations of sulfur dioxide to estimate total emission rate and injection height using a 4D “trajectory matching” method. However, the vertical distribution of emission rate remains uncertain given these two parameters: required assumptions such as Poisson distribution or umbrella-shaped distribution. The reason for this is that existing satellite data usually provide observation of total columns of volcanic ash only, without any vertical resolution. Furthermore, previous studies show that sulfur dioxide as a proxy for volcanic ash is inadvisable (Smith et al. 2014).

In this paper, we will demonstrate that standard 4D-Var is not able to estimate vertical distribution of effective volcanic ash injection rates from satellite-observed ash columns accurately. This problem will be illustrated both theoretically and experimentally.

A trajectory-based 4D-Var assimilation method is proposed to deal with the problem, which is more accurate to estimate emission rates of volcanic ash. The method reformulates the cost function in a regression type that computes the total difference between observed ash columns and a linear combination of simulated trajectories coupled with a priori emission knowledge (“background” term). To construct the cost function, it decomposes the system into subsystems that represent source–receptor relationships to each source term (ash injection at each layer), thus it does not lose meteorological information as with regular 4D-Var cost function. Twin experiments have been used to test the performance of the reformulated 4D-Var method. Although the computational effort required for the modified 4D-Var method increases fast with the number of parameters, the modified approach is adjoint free, so it does not require the implementation of the adjoint of the tangent linear model. This is a very attractive feature of the new approach.

The paper is organized as follows. In section 2 we give an overview of the standard 4D-Var methods and the ineffectiveness issue of using 4D-Var to estimate emission rates from satellite data. In section 3 the

trajectory-based 4D-Var is proposed to solve the inefficiency problem in our case (total column observations to estimate volcanic ash emissions). Then twin experiments are performed in section 4 for comparison and further analysis of both methods. Finally, in section 5 we summarize the conclusions of this study and discuss future work.

2. Standard 4D-Var assimilation methods

a. Preliminary knowledge

Variational methods determine an optimal combination of the prior information and observation data over an assimilation period by minimizing a cost function to estimate the control variables of a model. Control variables can include both state variables and system parameters, and in our case (total column observations to estimate volcanic ash emissions) they are the input of the system. A typical cost function would be the sum of the squared deviations of the analysis values from the observations weighted by the accuracy of the observations, plus the sum of the squared deviations of the background fields and the analyzed fields weighted by the accuracy of the background parameters.

In our case, suppose the volcanic ash model can be given in forms of

$$\mathbf{x}_k = M_k(\mathbf{x}_{k-1}, \mathbf{u}_k + \mathbf{w}_k), \quad (1)$$

$$\mathbf{y}_k = H_k(\mathbf{x}_k) + \mathbf{v}_k. \quad (2)$$

The subscript k represents time t_k . The term $\mathbf{x}_k \in R^n$ is the state vector; in volcanic ash transport model, this contains ash concentrations defined in a 3D array of grid cells. Here $\mathbf{u}_k \in R^p$ is the parameter vector including model parameters, inputs, and initials. Here we define $\mathbf{u}_k$ to contain the emission rate in each vertical layer. There are p ($p \ll n$) effective emission layers, which means the emission heights are no higher than the top of the p th layer above summit. Vector $\mathbf{y}_k \in R^q$ contains the observations. The noise terms $\mathbf{w}_k$ and $\mathbf{v}_k$ represent the model and observation uncertainty, respectively, both in terms of a Gaussian distribution with covariance matrices $\mathbf{B}_k$ and $\mathbf{R}_k$, respectively, where the model uncertainty lies in the parameter in this case. The term $M_k \in R^{n \times n}$ is the eventually transport model. Observation operator $H_k \in R^{q \times n}$ projects the state space into observation space; the inverse observation operator H_k^T that will be used in 4D-Var projects the observation space into the state space.

Once the observations are obtained, we could use 4D-Var data assimilation to seek optimal initial states and parameters that minimize the following cost function over an assimilation window $[t_0, t_N]$:

$$\begin{aligned} J(\mathbf{u}_k) = & \frac{1}{2} \sum_{k=0}^N (\mathbf{u}_k - \mathbf{u}_k^b)^T \mathbf{B}_k^{-1} (\mathbf{u}_k - \mathbf{u}_k^b) \\ & + \frac{1}{2} \sum_{k=0}^N (\tilde{\mathbf{y}}_k - \mathbf{y}_k)^T \mathbf{R}_k^{-1} (\tilde{\mathbf{y}}_k - \mathbf{y}_k) \\ = & J^b + J^o, \end{aligned} \quad (3)$$

where $\tilde{\mathbf{y}}_k = H_k(\mathbf{x}_k)$. The model described by Eqs. (1)–(2) is regarded as constraints to the minimization problem in Eq. (3).

The solution of the minimization problem could be obtained using gradient-based methods (conjugate gradient, quasi Newton, etc.), which are iterative methods by updating the parameters based on the gradients of the cost function with respect to the parameters. The adjoint model is used to compute the gradient (Talagrand and Courtier 1987; Courtier and Talagrand 1987). The incremental 4D-Var method is developed to accelerate the solution process of nonlinear problems (Courtier et al. 1994). However, no matter how advanced the approach is, the key point lies in how accurate the gradient represents the perturbations of parameters. When the gradient cannot represent the perturbations, it will never converge to the correct solution, as will be illustrated for the volcanic ash application in the next section.

b. Ineffectiveness of a standard 4D-Var methods

The standard 4D-Var has been successfully applied in air quality applications (Elbern et al. 2000), including source estimation (Elbern et al. 2007; Meirink et al. 2008). However, estimation of vertical distribution of ash emission rates from satellite data is problematic. Infrared satellite instruments from which AOD is retrieved only provide information on total column ash loadings. The instruments provide effective particle radius and AOD (τ) of volcanic ash clouds (Prata 1989). The column mass loading (Ω) [or $\mathbf{y}_k$ in Eq. (2)] is related to τ by $\Omega = 4\rho r_e \tau / (3Q_e)$, where ρ is the aerosol mass density at ambient relative humidity, r_e is the column-averaged effective radius, and Q_e is the column-averaged extinction efficiency with 40%–60% (Wen and Rose 1994; Fu et al. 2015). To transfer the ash concentrations $\mathbf{x}_k$ to ash mass loadings $\mathbf{y}_k$ [in Eqs. (1)–(2)] the observation operator is in the following form:

$$H_k(\mathbf{x}_k) = \sum_{i=1}^{N_z} \omega^i \times \mathbf{x}_k^i \cdot \Delta \mathbf{h}^i, \quad (4)$$

where N_z is the number of the vertical layers; $\cdot$ denotes element-wise multiplication; $\mathbf{x}_k^i$ and $\Delta \mathbf{h}^i$ represent the states and thicknesses of the grid cells in the i th layer, respectively; and ω^i is the weighing factor of the i th layer corresponding to mass extinction and the sensitivity of

FIG. 1. Spurious relationship between perturbed input parameter and updated parameters.

the specific satellite instrument with respect to height (Pierce 2013).

Since in this way the satellite observation computes the vertical integral of state variables along a column and has no vertical resolution, we cannot distinguish which of the vertical emissions could explain a mismatch between the model and the observation. This results in invalid gradients for the parameter update, and ill-conditioning issue in the optimization process.

To explain how the invalid gradients are formed, first we introduce the concept of “spurious relationship” created by observation. Suppose there is an observation variable formed by summation of state variables as $y = x_1 + x_2$. A change in y will be explained by changes in both x_1 and x_2 even it is introduced by only one of the states, because the transpose H^T assigns the observation value y to both variables. Therefore, there is a spurious correlation between x_1 and x_2 created through shared observation y .

Additionally, if x_1 and x_2 are sensitive to different parameters u_1 and u_2 through a model, respectively, then the spurious relationship may cause ineffective estimations of the parameters. If x_1 and x_2 are sensitive to the same parameter, then the spurious relationship may help to accelerate the convergence process. In this paper, in the former case we call it an “ineffective spurious relationship,” and in the latter case an “effective spurious relationship.” Figure 1 gives us an example of the former case, and illustrates how satellite observation y build up spurious relationships between states $[x_1, \dots, x_i, \dots, x_p]$, which are sensitive to different input parameters $[u_1, \dots, u_i, \dots, u_p]$. All parameters are updated no matter which parameter causes a perturbation in the observation.

Now we show how a spurious relationship occurs in our case. In atmospheric transport models, particles are transported with the wind to somewhere else other than the source position. Thus, the wind transports information of the emissions from the location of release to the place where the particles reach. Observations of concentration at certain position (in situ observation) could be used to trace back information of emissions through wind field. This is how the sensitivity of a state vector (concentrations) with respect to emissions is formed.

However, observations in the real world are usually not of a single state variable but a combination of several state variables or other variables, such as satellite data, which measures total column mass loadings of volcanic ash. It is problematic to use observations like this to trace back along wind and obtain the emission information. Wind fields bear the characteristic that horizontal winds are much stronger than vertical winds, thus states in one layer are sensitive to the input parameter in that layer. Subsequently, states along a column are sensitive to different input parameters locating at the corresponding layers, respectively. Therefore, the satellite observation operator creates a strong ineffective spurious relationship between originally unrelated or slightly related state variables, and between emission sources. Therefore, a perturbation in one emission layer may result in big values of the element in gradient corresponding to other layers.

3. Trajectory-based 4D-Var

a. Methodology

We can break up the ineffective spurious relationship artificially by perturbing parameters one by one, such that the states sensitive to the perturbed parameter are automatically activated and separated from other states. Therefore, we propose a modified 4D-Var method, which reformulates the least squares problem by splitting the spurious relationships. In this method, we seek a linear combination of states (or trajectories) generated by system simulations with perturbed input parameters, which best matches the observation data.

The trajectories consists of a reference simulation with background input parameters $M_k(\mathbf{x}_{k-1}, \mathbf{u}^b)$ and a set of simulations with perturbed parameters $M_k(\mathbf{x}_{k-1}, \mathbf{u}^b + \Delta \mathbf{u}^i)$ for $i = 1, \dots, p$, where $\Delta \mathbf{u}^i$ is a perturbation of the emission rate in the i th layer above summit, which is the i th effective emission layer (EEL) and denoted as i th EEL in this paper. Then we define the background measurements: $\mathbf{y}_k^0 = H_k[M_k(\mathbf{x}_{k-1}, \mathbf{u}^b)]$ and the snapshot measurements: $\Delta \mathbf{y}_k^i = H_k[M_k(\mathbf{x}_{k-1}, \mathbf{u}^b + \Delta \mathbf{u}^i)] - \mathbf{y}_k^0$. Suppose the parameter space can be spanned by the perturbed parameter sets, which means that the “truth” can be computed as a linear combination of perturbed parameters weighted by unknown parameters β^i : $\mathbf{u} = \mathbf{u}^b + \sum_{i=1}^p \beta^i \Delta \mathbf{u}^i$. Then the observations in Eq. (2) can be approximated by

$$\begin{aligned} \mathbf{y}_k &\approx H_k[M_k(\mathbf{x}_{k-1}, \mathbf{u}^b)] + \sum_{i=1}^p \beta^i \mathbf{H}_k \mathbf{M}_k(\mathbf{x}_{k-1}, \mathbf{u}^b) \Delta \mathbf{u}^i + \mathbf{v}_k \\ &\approx \mathbf{y}_k^0 + \sum_{i=1}^p \beta^i \{H_k[M_k(\mathbf{x}_{k-1}, \mathbf{u}^b + \Delta \mathbf{u}^i)] - \mathbf{y}_k^0\} + \mathbf{v}_k \\ &= \mathbf{y}_k^0 + \sum_{i=1}^p \beta^i \Delta \mathbf{y}_k^i + \mathbf{v}_k, \end{aligned} \quad (5)$$

where $\mathbf{H}_k$ and $\mathbf{M}_k$ denote the tangent linear model of H_k and M_k , respectively. Therefore, the coefficients β can be computed by minimizing the reformulated cost function given by

$$\begin{aligned} J(\beta) &= \frac{1}{2} \sum_{k=1}^{N_t} \left(\sum_{i=1}^p \beta^i \Delta \mathbf{y}_k^i + \mathbf{y}_k^0 - \mathbf{y}_k \right)^T [\mathbf{R}_k]^{-1} \left(\sum_{i=1}^p \beta^i \Delta \mathbf{y}_k^i + \mathbf{y}_k^0 - \mathbf{y}_k \right) \\ &\quad + \frac{1}{2} \sum_{k=1}^{N_t} (\mathbf{u} - \mathbf{u}^b)^T [\mathbf{B}_k]^{-1} (\mathbf{u} - \mathbf{u}^b) + \mu \|\nabla \mathbf{u}\|^2 \\ &= J^o + J^b + J^r, \end{aligned} \quad (6)$$

where J^o and J^b are defined similar to those in cost function Eq. (3), and J^r is a regularization term in case the large uncertainties in the model or observations lead to an unstable solution or a negative solution.

When the number of parameters is small (less than 100) we can transform the optimization problem into an algebraic problem and solve it directly. As it can be observed that the cost function in Eq. (6) is convex, according to the first-order optimality condition, to find the minimum is to find the solution to the linear problem:

$$\frac{dJ}{d\beta} = \frac{dJ^o}{d\beta} + \frac{dJ^b}{d\beta} + \frac{dJ^r}{d\beta} = 0. \quad (7)$$

Simple derivation and moving all the terms independent of β to the right-hand side yields the following:

$$(\mathbf{A}^o + \mathbf{A}^b + \mathbf{A}^r)\beta = \mathbf{b}^o + \mathbf{b}^b + \mathbf{b}^r, \quad (8)$$

where

$$\begin{aligned} \mathbf{A}^o &= \begin{bmatrix} \Delta \mathbf{y}_1^1 & \dots & \Delta \mathbf{y}_1^p \\ \vdots & \ddots & \vdots \\ \Delta \mathbf{y}_{N_t}^1 & \dots & \Delta \mathbf{y}_{N_t}^p \end{bmatrix}^T \begin{pmatrix} [\mathbf{R}_1]^{-1} & & 0 \\ & \ddots & \\ 0 & & [\mathbf{R}_{N_t}]^{-1} \end{pmatrix} \begin{bmatrix} \Delta \mathbf{y}_1^1 & \dots & \Delta \mathbf{y}_1^p \\ \vdots & \ddots & \vdots \\ \Delta \mathbf{y}_{N_t}^1 & \dots & \Delta \mathbf{y}_{N_t}^p \end{bmatrix}, \\ \mathbf{b}^o &= \begin{bmatrix} \Delta \mathbf{y}_1^1 & \dots & \Delta \mathbf{y}_1^p \\ \vdots & \ddots & \vdots \\ \Delta \mathbf{y}_{N_t}^1 & \dots & \Delta \mathbf{y}_{N_t}^p \end{bmatrix}^T \begin{pmatrix} [\mathbf{R}_1]^{-1} & & 0 \\ & \ddots & \\ 0 & & [\mathbf{R}_{N_t}]^{-1} \end{pmatrix} \begin{bmatrix} \mathbf{y}_1 - \mathbf{y}_1^0 \\ \vdots \\ \mathbf{y}_{N_t} - \mathbf{y}_{N_t}^0 \end{bmatrix}, \\ \mathbf{A}^b &= [\Delta \mathbf{u}^1 \quad \dots \quad \Delta \mathbf{u}^p]^T \left(\sum_{k=1}^{N_t} \mathbf{B}_k \right)^{-1} [\Delta \mathbf{u}^1 \quad \dots \quad \Delta \mathbf{u}^p], \end{aligned}$$

$\mathbf{b}^b = 0$, $\mathbf{A}^r = \mu \mathbf{L}$, $\mathbf{b}^r = 0$, and regularization operator $\mathbf{L}$ is defined in section 3b.

The Trj4DVar algorithm is thus described as the following sequence of operations:

- 1) Specify the background emission rates $\mathbf{u}^b$ and run the model to obtain the model state $\mathbf{x}_k = M_k(\mathbf{x}_{k-1}, \mathbf{u}^b)$ and background measurement $\mathbf{y}_k^0 = H_k(\mathbf{x}_k)$ ($k = 1, \dots, N_t$).
- 2) Perturb the emission rates $\mathbf{u}^i = \mathbf{u}^b + \Delta \mathbf{u}^i$ and perform the forward model runs $\mathbf{x}_k^i = M_k(\mathbf{x}_{k-1}^i, \mathbf{u}^i)$, then compute the snapshot measurements $\Delta \mathbf{y}_k^i = H_k(\mathbf{x}_k^i) - \mathbf{y}_k^0$ ($k = 1, \dots, N_t, i = 1, \dots, p$).
- 3) Solve the linear equations [Eq. (8)] to obtain the coefficients $\beta = [\beta^1, \dots, \beta^p]^T$.
- 4) Update emission rates $\mathbf{u} = \mathbf{u}^b + \sum_{i=1}^p \beta^i \Delta \mathbf{u}^i$.

Note that Trj4DVar is adjoint free, since it needs only forward model runs to generate the trajectories and the effort to solve the transformed algebraic problem without adjoint model runs to compute the gradients. This saves the effort in the formulation and implementation of the tangent linear model and the adjoint model for the transport model, which is usually required for a 4D-Var application.

Since the wind fields vary at each layer, the position and shape of the ash plume bear information on its emission altitude, and vice versa. The snapshots $\Delta \mathbf{y}_k^i$ contain meteorological information with vertical difference, which benefits in emission profile estimation with this trajectory-based 4D-Var (Trj4DVar) method. Trj4DVar would not work with a vertically constant meteorological condition, which fortunately, is not the

case in the real world. However, wind velocity fields might differ very little at some layers close to each other, which resulted in unstable solution and this problem can be solved by regularization as mentioned above.

Notice that although the formulation is similar, our method is different from ensemble-based 4D-Var (En4DVar) (Tian et al. 2008; Liu et al. 2009; Tian et al. 2011) in three ways. First, this method aims at solving the ill-conditioning problem in our case, whereas En4DVar focuses on obtaining a flow-dependent error covariance matrix. Second, the generation of trajectories in this method is specially designed according to a priori knowledge of the model property and the characteristics of the observation, but En4DVar adopts a Monte Carlo method, which technically can be applied to any case. Third, this method is not scalable to the cases with a huge amount of parameters and it is specially designed to efficiently estimate vertical emission profile with a relative small number of parameters, while En4DVar naturally reduces the rank of the problem and can be scalable to high-dimensional cases. However, in our case, the utilization of En4DVar with a set of randomly generated ensembles would lead to inaccurate estimations as with the standard 4D-Var (see Figs. 4 and 10).

b. Regularization

There are three types of ill posedness of an inverse problem (the optimization problem in our case): existence, uniqueness, and stability of the solution (Doicu et al. 2010). In our case we have the third type, the stability of the solution. Note that due to the severe ill conditioning of the Hessian $\text{Hess} = \{[\sum_{k=1}^{N_t} (\Delta \mathbf{y}_k^i)^T (\Delta \mathbf{y}_k^j)]_{ij}\}$ in the term J^o of the cost function in Eq. (6) and the noise in the observations $\mathbf{y}_k$, the solution of minimizing J^o is not a meaningful approximation of the truth. To be able to obtain an accurate approximation of the truth, a regularization term is added in the cost function, whose solution is less sensitive to perturbations in $\mathbf{y}_k$ (Reichel and Ye 2008).

Actually, the “background” term J^b plays a role as regularization term (Cacuci et al. 2013), but the ill posedness of the problem is still unresolved although somewhat improved since we cannot arbitrarily enlarge the proportion it takes in the cost function in Eq. (6). First in our case the background emissions have a large uncertainty, a larger proportion of J^b leads to a solution closer to the very inaccurate background emissions; and second the weighting factors of J^o and J^b (i.e., $[\mathbf{R}_k]^{-1}$ and $[\mathbf{B}_k]^{-1}$) are generated according to their statistic properties through Bayesian theory in 4D-Var approaches, hence, increasing the weighting

for J^b will make it statistically unrealistic. Therefore, an extra term J^r for regularization is introduced in the cost function in Eq. (6).

In the Trj4DVar approach, the popular Tikhonov regularization is used, where a regularization term is in the following form:

$$J^r(\beta) = \mu \|\mathbf{L}\beta\|^2, \quad (9)$$

where the matrix $\mathbf{L} \in R^{k \times p}$, is referred to as the regularization operator and the scalar $\mu \leq 0$ is referred to as the regularization parameter.

Common choices of regularization operators are the identity matrix and a scaled finite-difference approximation of a derivative $\|\nabla \beta\|^2$ (equals $\|\nabla \mathbf{u}\|^2$ in our case). They both have a smoothing effect on the solution and make the solution more stable. However, when μ is large, using identity matrix results in a solution close to zero, while using the derivative can preserve the total emission rate. Therefore, in this paper, a scaled finite-difference approximation of a derivative in the following forms is used:

$$\mathbf{L}_1 := \frac{1}{2} \begin{bmatrix} 1 & -1 & & & 0 \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ 0 & & & 1 & -1 \end{bmatrix} \in R^{(p-1) \times p} \quad (10)$$

or

$$\mathbf{L}_2 := \frac{1}{4} \begin{bmatrix} -1 & 2 & -1 & & 0 \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \\ 0 & & & -1 & 2 & -1 \end{bmatrix} \in R^{(p-2) \times p}. \quad (11)$$

c. Computational performance and acceleration scheme

Most of the computational cost of the Trj4DVar method lies in the generation of the trajectories. Therefore, the total cost is linearly proportional to the number of forward model simulations denoted by $n_{\text{sim}}^{\text{ref}}$. The number of ensembles $n_{\text{sim}}^{\text{ref}}$ is equal to the number of input parameters p according to section 3a. For standard 4D-Var method the computation of gradients takes most of the time, which can also be quantified in terms of forward model simulations $n_{\text{sim}}^{\text{trad}}$. Since one gradient computation requires one forward run and one adjoint run, this can be counted approximately as two forward simulations. Thus, $n_{\text{sim}}^{\text{trad}}$ is equal to the twice the number of iterations n_{itr} . Therefore, when p and $2n_{\text{itr}}$ are almost the same, the computational costs

TABLE 1. Input parameters for the 14–19 Apr 2010 period of activity at Eyjafjallajökull [taken from Webley et al. (2012)].

Start time (UTC)	Duration (h)	End time (UTC)	Height (km MSL)	Eruption rate ($\text{m}^3 \text{s}^{-1}$)	Eruption rate (kg s^{-1})
0900 UTC 14 Apr 2010	10	1900 UTC 14 Apr 2010	9	219.553	5.71×10^5
1900 UTC 14 Apr 2010	9	0400 UTC 15 Apr 2010	5.5	14.884	3.87×10^4
0400 UTC 15 Apr 2010	39	1900 UTC 16 Apr 2010	6	199.503	6.44×10^4
1900 UTC 16 Apr 2010	35	0600 UTC 18 Apr 2010	8.25	140.325	3.65×10^5
0600 UTC 18 Apr 2010	17	2300 UTC 18 Apr 2010	5	8.335	2.17×10^4
2300 UTC 18 Apr 2010	1	0000 UTC 19 Apr 2010	4	1.898	4.93×10^3

of the two methods are comparable. In our experiment, p is smaller than $2n_{\text{itr}}$ and thus the reformulated method is more efficient in this case.

However, $n_{\text{sim}}^{\text{ref}}$ increases with p , thus refining grid cells vertically could make the reformulated 4D-Var less efficient than the standard 4D-Var. Moreover, to deal with the nonlinearity of a more realistic model, the reformulated 4D-Var scheme should be repeated several times with new ensembles generated from an updated background state $\mathbf{x}_k^{\text{updated}}$ and parameters $\mathbf{u}^{b, \text{updated}}$ in each iteration. This problem can be reduced by a simple acceleration scheme through generating multiple ensembles in one forward simulation. This can be done in the following way:

- 1) In one simulation, we perturb multiple input parameters which are physically unrelated or slightly related to each other. For instance we perturb the input variables located in every m layers (i.e., with a distance of m layers) as follows $\mathbf{u} = \mathbf{u}^b + \sum_{i=1}^{(p-1)/m} \Delta \mathbf{u}^{m \times i + 1}$, and compute $\Delta \mathbf{x}_k = M_k(\mathbf{x}_{k-1}, \mathbf{u}) - M_k(\mathbf{x}_{k-1}, \mathbf{u}^b)$.
- 2) After the forward simulation, we generate the ensembles corresponding to each perturbed parameter by computing the column integral of the layers where the states are sensitive to the specific parameter according to Eq. (4) as follows:

$$\Delta \mathbf{y}_k^l = \sum_{j=\max\{1, l-(m+1)/2\}}^{\min\{p, l+(m+1)/2\}} \omega^j \times \Delta \mathbf{x}_k^j * \Delta \mathbf{h}^j, \\ l = m \times i + 1, \quad i = 1, \dots, (p-1)/m. \quad (12)$$

- 3) Repeat steps 1) and 2) to compute ensembles $\Delta \mathbf{y}_k^{m \times i + j}, i = 1, \dots, (p-j)/m$, with perturbed parameters $\mathbf{u} = \mathbf{u}^b + \sum_{i=1}^{(p-j)/m} \Delta \mathbf{u}^{m \times i + j}, j = 2, \dots, m$.

In this way, we can efficiently obtain the entire approximated ensembles with m model simulations, which is independent of p .

4. Experiments and discussion

a. Experimental setup

A number of twin experiments have been performed to compare the performance of the standard and the

trajectory-based 4D-Var methods, on a PC with an Intel Xeon E3-1240 V2 processor with 8 cores, a total memory of 32 GB (approximately 5% was taken during experiments), and a clock speed of 3.4 GHz. The experiments are conducted with observation data generated by implementation of the model with truth emission rates. Then we estimate the emission rate from the background and the observation through both the standard 4D-Var and Trj4Dvar.

A 3D aerosol transport model is used to simulate the eruption of the Eyjafjallajökull volcanic activity during 14–19 April 2010. For simplicity, the transport model includes only the advection process. Wind fields are obtained from 3-hourly meteorological data from the European Centre for Medium-Range Weather Forecasts (ECMWF), which is interpolated to hourly resolution.

The emission information of the first few days of explosive eruption is taken from Webley et al. (2012) and shown in Table 1. The eruption is described in terms of parameters such as total emission rate and plume height, which are assumed to be constant during an emission episode of several hours. Poisson distribution is used to generate the truth emission rates in vertical layers according to emission information in Table 1. To give the background' emission rates, we first use a total emission rate assumed to be underestimated from the truth, then generate the plume height from the empirical relationship and use the Poisson distribution to obtain the vertical profile.

To compare the influence of assimilation window length, assimilation windows of 1, 3, or 6 h are used, which are not longer than one emission episode. In practice, assimilation window of constant length should be used within one episode. To deal with the change of episodes, the length should be adaptively shortened within one episode. Since in our case new released particles will not spread farther than 50 grid cells horizontally within 6 h according to local wind velocity, we use a small domain (Iceland area in the rectangular; Fig. 2) for the assimilation in order to keep a low computational expense. In practice, an extended domain (European area; Fig. 2) should be used for forecast of one day or longer after assimilation.

FIG. 2. European domain for forecast and Iceland domain in the rectangular for assimilation.

Because this study focuses on obtaining a theoretical base for future implementations, we do not use real data here but synthetic observations ($\mu\text{g m}^{-2}$) generated by assuming $\omega = 1$ in Eq. (4):

$$y^i = \sum_{j=1}^{N_z} x^{i,j} \times \Delta h^{i,j}, \quad (13)$$

where y^i is the i th observation variable, $x^{i,j}$ is the ash density at the grid cell horizontally corresponding to y^i and at the j th vertical layer, and $\Delta h^{i,j}$ is the thickness of the grid cell where $x^{i,j}$ locates. In reality a satellite instrument will not observe this quantity directly but the optical properties and the observation operator for the model state should include a radiative transfer model with assumptions on, for example, the ash density ρ .

In the experiments, both deterministic model and stochastic models are used. The former is used to compare the difference of the standard and trajectory-based 4D-Var method. The latter is used to show the influences of uncertainties or other aspects to the performance of Trj4DVar. For noise-free model experiments, observations are obtained by direct realization of the model with the truth emission rates. Since the emission parameters $\mathbf{u}$ and the observation variables $\mathbf{y}_k$ are assumed to be independent, the background covariance matrices $\mathbf{B}$ and observation covariance matrices $\mathbf{R}_k$ are diagonal [in Eqs. (1)–(2)], in forms of $\mathbf{B} = \sigma_b^2 \mathbf{I}$ and $\mathbf{R}_k = \sigma_o^2 \mathbf{I}$, respectively. For experiments using models with uncertainties, 50% noise is added to observation data according to realistic satellite instrument retrieval

FIG. 3. Estimation result of emission rates using standard 4D-Var with complete observations and 1-h assimilation window.

uncertainties in section 2b, and to investigate the influence of noise for solutions 10% noise is used in comparison experiments. For the construction of $\mathbf{B}$ in the cost function in Eq. (3) of the stochastic model, the uncertainties in the background parameters are always assumed to be 50%.

In this paper, we consider observations with a very high temporal resolution of 15 min such as potentially available from the geostationary SEVERI instrument (Millington et al. 2012; Schmets et al. 2002). A spatial resolution of $0.25^\circ \times 0.25^\circ$ is assumed, which is equivalent to the spatial resolution of the meteorological data used in this study.

b. Results and discussion of standard 4D-Var method

First, a perfect model and complete observation that measures the state variables in all the grid cells are used to testify whether the model is physically well conditioned. The results with a 1-h assimilation window are shown in Fig. 3, where the “analysis” perfectly matches the truth, and the results with 3- and 6-h assimilation windows are similar. This means that emission rates can be well estimated using standard 4D-Var according to model dynamic. Therefore, once the estimation result with another observation operator is undesirable, it is caused by the observation but not the model.

Next, experiments with satellite ash column observations are performed conducted by both the noise-free model and stochastic model, and 50% noise is added in the observation data of the stochastic case. The noise-free model is used here to exclude other reasons (such as uncertainty in observations) for the ineffectiveness problem. However, in

FIG. 4. Estimation result of emission rates using standard 4D-Var with satellite observations and assimilation window of (a),(b) 1 and (c),(d) 6 h, conducted by both the (left) noise-free model and (right) stochastic model.

the noise-free model with a zero observational error covariance matrix, the inverse of the matrix will be infinity and the background term vanishes compared to the “observation” term. Thus, the stochastic model is used to include the influence of the background term.

Results in Fig. 4 shows that standard 4D-Var is incapable of identifying the injection height but can correct the emissions as a whole. The ineffective estimation problem is not caused by observation uncertainty but by the observation itself (cf. Figs. 4a and 3), and it will not be improved by the background term (cf. the left columns and the right columns in Fig. 4). Enlarging the assimilation window and increasing the amount of data do not have obvious improvements on the situation. Therefore, after data assimilation, although the predicted ash columns are closer to the observation data, shapes and concentrations of ash cloud at each layer do not fit the truth, especially at the layer with the largest emission rate. This would mean that, in an application with real data, even if we obtain an analysis that matches

satellite image, there is a high probability that we may still give a bad and misleading forecast, as is illustrated in Figs. 5 and 6. We can see that after 4D-Var assimilation, the forecast better matches the observation generated from the truth (Figs. 5a–c), but little improvement has been made of the ash could at the seventh EEL (Figs. 6a–c).

To illustrate how the ineffective spurious relationship and invalid gradient lead to the problem, a single-input-perturbation experiment is conducted. In the experiment, perturbation is given in a single input parameter and gradients are computed and normalized with both complete observations and column observations. The former reflects model behavior and the comparison with the perturbation vector illustrates whether the model is physically well conditioned to identify the perturbed parameter, which will be called the normalized “model gradient” in this paper. The latter compared with the normalized model gradient illustrates whether the observation creates an ineffective spurious relationship or leads to an invalid gradient, which can be called a

FIG. 5. Forecast of ash columns generated with emission rates of (a) the truth, (b) the background, (c) estimated through the standard 4D-Var method, and (d) estimated through the Trj4DVar method at 1900 UTC 14 Apr 2010.

normalized “observation gradient” accordingly. The seventh parameter is perturbed, which bears the biggest difference between the background and the truth input, and the both normalized gradients are shown in Fig. 7. It is observed that the model gradient can well represent the perturbation of the input but the observation gradient evens out the perturbation, and moreover, the value of the seventh element of the observation gradient is smaller than some other elements. This is consistent with the estimation results, which means the estimation result will never reach the truth regardless of the number of iterations.

c. Results and discussion of trajectory-based 4D-Var method

The Trj4DVar method is first applied to the noise-free model with the same background emissions as in

the previous experiments, and the comparison with the standard 4D-Var method shows that this method can help solve the ill-conditioned problem caused by the radiation observation (cf. Figs. 8 and 4d). However, when it is applied to a stochastic model, an unstable solution with big oscillations occurs especially with a small assimilation window (Fig. 9a), which is caused by the similarity of meteorological fields in each layer. Therefore, the regularization term is used according to a priori knowledge of the model that the input should be smoothed.

To test the influence of noise, assimilation window length, and regularization term on the performance of Trj4DVar, we apply Trj4D-Var with 50% model uncertainties, 10% and 50% observation uncertainties, and assimilation windows of 1, 3, and 6 h, respectively, and the results are shown in Fig. 9 where “est_reg” is the estimate with regularization.

FIG. 6. Forecast of ash clouds in the seventh EEL with emission rates of (a) the truth, (b) the background, (c) estimated through the standard 4D-Var method, and (d) estimated through the Trj4Dvar method at 1900 UTC 14 Apr 2010.

It can be noticed that the basic vertical profile of inputs is preserved in Fig. 9, especially the effective injection layer is well estimated, but oscillations in the solution occur when noise is added to observations. We observe large oscillation in 1-h cases especially with 50% observation noise in Fig. 9d. The situation is improved with a smaller observation uncertainty or a larger assimilation window. This can be explained from the perspective of statistics that a larger amount of data can reduce the disturbance of noise and provide more stable solution.

Moreover, by comparing the results with 10% noise and 50% noise, we can observe obviously in Figs. 9e and 9f big bumps around the fifth layer where the maximal input of background lies. This is because with a larger observation uncertainty the

background impacts more in the solution. Enlarging the assimilation window helps improve the situation, which can be observed by comparing the three figures with 50% uncertainty. However, enlarging the assimilation window will increase computational time exponentially with both increased DA time and enlarged DA domain, thus a balanced decision should be made between a stable solution and computational efficiency.

Regularization can smooth the solution and reduce oscillation (Fig. 9a), which improves the accuracy of parameter estimation. However, sometimes it works in a negative way such that it smooths the solution too much and thus decreases the accuracy (Fig. 9f). Therefore, a compromise should be made between the oscillation problem and the smoothing problem.

FIG. 7. Normalized gradient generated by the perturbation in a single input parameter at the seventh EEL with a 1-h assimilation window.

The impact of regularization decreases with the length of assimilation window, which can be observed by comparing the 1-, 3-, and 6-h cases. It would not be a problem, considering that as the estimation becomes more accurate with a bigger assimilation window, regularization becomes less important to further improve the accuracy.

Moreover, with a more accurate estimation of the vertical profile of emission rates, a better forecast of state fields will be obtained. Although the forecasts of ash mass loadings using both Std4DVar and Trj4DVar methods are largely improved (Fig. 5), Trj4DVar gives a more reliable ash field in layers (Fig. 6) as advisory information.

Additionally, free forecasts are conducted using the state fields at the end of assimilation time (1900 UTC 14 April 2010, in this experiment) as the initial condition and emissions in Table 1 for further forecast (till 1000 UTC 15 April 2010, in this experiment). Results show that an initial condition computed with a more accurate emission yields a more faithful forecast (Fig. 10). It can also be observed that ash clouds between Figs. 10c and 10d are more distinguishable in terms of the shapes and values than those between Figs. 5c and 5d. This is because, although they bear a great resemblance from the top view at one moment, the ash clouds are located in different altitudes, and will eventually transport in different patterns due to the meteorological fields. It confirms the conclusion mentioned in section 4b that an inaccurate estimate of vertical distribution is likely to result in a misleading forecast in the long run.

FIG. 8. Estimation result using the Trj4DVar method with a 1-h assimilation window and a noise-free model.

5. Conclusions and future work

This paper first discussed the ineffectiveness of gradient-based standard 4D-Var in estimating the vertical distribution of the emission rate with satellite ash column observations. It was explained how ineffective spurious relationship was built up and affected the update of the parameters. Twin experiments of a real study case of Eyjafjallajökull with a 3D simplified volcanic transport model were performed to further illustrate the problem. It was demonstrated that standard 4D-Var was unable to identify the injection layer. It corrected the total emission rate as a whole and not the vertical profile. This made the estimated ash fields better match the satellite data column-wisely, but it was insufficient to provide reliable information for each layer.

To solve the ineffectiveness problem, a trajectory-based 4D-Var method was proposed to reduce the impact of ineffective spurious relationship by dividing the states sensitive to different input parameters into separate ensembles. By taking advantage of the fact that the time evolution of ash clouds in satellite data bear vertical wind information, Trj4DVar showed good performance in the estimation of the emissions with a perfect model. However, the ill-conditioned physical condition with similar wind fields in each layer can lead to a set of similar snapshot observations and resulted in an unstable solution. Therefore, regularization terms were introduced in the cost function to include additional information. Furthermore, Trj4DVar is adjoint free but might require more computational effort than the standard 4D-Var method. An acceleration technique was

FIG. 9. Trj4DVar results of twin experiments with (a)–(c) 10% and (d)–(f) 50% observation noise and assimilation window of (a),(d) 1; (b),(e) 3; and (c),(f) 6 h.

discussed to solve the problem of large computational costs.

A stochastic model was also used to test the performance of Trj4DVar method. Although basically preserving the vertical profile of emission rates, oscillations in solution occurred with noise, and large noise led to large oscillations. Regularization helped to smooth the solution and improve the estimation accuracy. With a larger assimilation window, more observations were available, which diminished the influence of noise and

resulted in a more stable and accurate solution. Moreover, in the case of larger assimilation window, the regularization had less impact on the solution, and thus the more accurate solution could be better preserved. However, enlarging the assimilation window would cause an exponential increase in computational expense, and decisions should be made between a more accurate estimation and less computational effort.

The Trj4DVar method showed a good performance in estimating the vertical distribution of emission rate of

FIG. 10. Free forecast of ash columns after DA at 1000 UTC 15 Apr 2010: generated with emission rates for the first 10 h of (a) the truth, (b) the background, (c) estimated through the standard 4D-Var method, and (d) estimated through the Trj4Dvar method.

volcanic ash with artificial satellite total-column observations and a simplified transport model and one has yet to test it on operational models and real-life volcanic ash emissions. Future work will be focused on application to real data with a full model including other physical ash transport processes besides advection and diffusion, such as dry deposition and wet deposition.

Acknowledgments. We sincerely thank the three anonymous reviewers for their constructive comments that helped to improve this paper.

REFERENCES

- Cacuci, D. G., I. M. Navon, and M. Ionescu-Bujor, 2013: *Computational Methods for Data Evaluation and Assimilation*. Chapman and Hall/CRC, 373 pp.
- Casadevall, T. J., 1994: Alaska Volcano Observatory—The 1989–1990 eruption of Redoubt Volcano, Alaska: Impacts on aircraft operations. *J. Volcanol. Geotherm. Res.*, **62**, 301–316, doi:10.1016/0377-0273(94)90038-8.
- Courtier, P., and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Quart. J. Roy. Meteor. Soc.*, **113**, 1329–1347, doi:10.1002/qj.49711347813.
- , J. N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387, doi:10.1002/qj.49712051912.
- Dacre, H. F., and Coauthors, 2011: Evaluating the structure and magnitude of the ash plume during the initial phase of the 2010 Eyjafjallajökull eruption using lidar observations and NAME simulations. *J. Geophys. Res.*, **116**, D00U03, doi:10.1029/2011JD015608.
- Doicu, A., T. Trautmann, and F. Schreier, 2010: *Numerical Regularization for Atmospheric Inverse Problems*. Springer Praxis, 426 pp.
- Elbern, H., H. Schmidt, O. Talagrand, and A. Ebel, 2000: 4D-variational data assimilation with an adjoint air quality model for emission analysis. *Environ. Model. Software*, **15**, 539–548, doi:10.1016/S1364-8152(00)00049-9.
- , A. Strunk, H. Schmidt, and O. Talagrand, 2007: Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmos. Chem. Phys.*, **7**, 3749–3769, doi:10.5194/acp-7-3749-2007.

- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, doi:[10.1029/94JC00572](https://doi.org/10.1029/94JC00572).
- Flemming, J., and A. Inness, 2013: Volcanic sulfur dioxide plume forecasts based on UV satellite retrievals for the 2011 Grímsvötn and the 2010 Eyjafjallajökull eruption. *J. Geophys. Res. Atmos.*, **118**, 10 172–10 189, doi:[10.1002/jgrd.50753](https://doi.org/10.1002/jgrd.50753).
- Fu, G., H. X. Lin, A. W. Heemink, A. J. Segers, S. Lu, and T. Palsson, 2015: Assimilating aircraft-based measurements to improve forecast accuracy of volcanic ash transport. *Atmos. Environ.*, **115**, 170–184, doi:[10.1016/j.atmosenv.2015.05.061](https://doi.org/10.1016/j.atmosenv.2015.05.061).
- Horwell, C., and P. Baxter, 2006: The respiratory health hazards of volcanic ash: A review for volcanic risk mitigation. *Bull. Volcanol.*, **69**, 1–24, doi:[10.1007/s00445-006-0052-y](https://doi.org/10.1007/s00445-006-0052-y).
- Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–110, doi:[10.1111/j.1600-0870.1986.tb00459.x](https://doi.org/10.1111/j.1600-0870.1986.tb00459.x).
- Liu, C., Q. Xiao, and B. Wang, 2009: An ensemble-based four-dimensional variational data assimilation scheme. Part II: Observing system simulation experiments with Advanced Research WRF (ARW). *Mon. Wea. Rev.*, **137**, 1687–1704, doi:[10.1175/2008MWR2699.1](https://doi.org/10.1175/2008MWR2699.1).
- Mankin, W. G., M. T. Coffey, and A. Goldman, 1992: Airborne observations of SO₂, HC₁, and O₃ in the stratospheric plume of the Pinatubo Volcano in July 1991. *Geophys. Res. Lett.*, **19**, 179–182, doi:[10.1029/91GL02942](https://doi.org/10.1029/91GL02942).
- Mastin, L. G., and Coauthors, 2009: A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions. *J. Volcanol. Geotherm. Res.*, **186**, 10–21, doi:[10.1016/j.jvolgeores.2009.01.008](https://doi.org/10.1016/j.jvolgeores.2009.01.008).
- Meirink, J. F., and Coauthors., 2008: Four-dimensional variational data assimilation for inverse modeling of atmospheric methane emissions: Analysis of SCIAMACHY observations. *J. Geophys. Res.*, **113**, D17301, doi:[10.1029/2007JD009740](https://doi.org/10.1029/2007JD009740).
- Millington, S. C., R. W. Saunders, P. N. Francis, and H. N. Webster, 2012: Simulated volcanic ash imagery: A method to compare NAME ash concentration forecasts with SEVIRI imagery for the Eyjafjallajökull eruption in 2010. *J. Geophys. Res.*, **117**, D00U17, doi:[10.1029/2011JD016770](https://doi.org/10.1029/2011JD016770).
- Pierce, R. B., G. Carmichael, P. Sadie, G. Grell, M. Stuefer, M. Pavolonis, and K. Chance, 2013: Development of data assimilation capabilities for SEVIRI volcanic ash retrievals. *NOAA Satellite Science Week Virtual Meeting*, NOAA. [Available online at <http://www.goes-r.gov/downloads/ScienceWeek/2013/presentations/03-20/01-R-Pierce.pdf>.]
- Prata, A. J., 1989: Infrared radiative transfer calculations for volcanic ash clouds. *Geophys. Res. Lett.*, **16**, 1293–1296, doi:[10.1029/GL016i011p01293](https://doi.org/10.1029/GL016i011p01293).
- Reichel, L., and Q. Ye, 2008: Simple square smoothing regularization operators. *Electron. Trans. Numer. Anal.*, **33**, 63–83, doi:[10.1.1.418.5500](https://doi.org/10.1.1.418.5500).
- Robock, A., 2000: Volcanic eruptions and climate. *Rev. Geophys.*, **38**, 191–219, doi:[10.1029/1998RG000054](https://doi.org/10.1029/1998RG000054).
- Rose, W. I., G. J. S. Bluth, D. J. Schneider, G. G. J. Ernst, C. M. Riley, L. J. Henderson, and R. G. McGimsey, 2001: Observations of volcanic clouds in their first few days of atmospheric residence: The 1992 eruptions of Crater Peak, Mount Spurr Volcano, Alaska. *J. Geol.*, **109**, 677–694, doi:[10.1086/323189](https://doi.org/10.1086/323189).
- Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *Bull. Amer. Meteor. Soc.*, **83**, 977–992, doi:[10.1175/1520-0477\(2002\)083<0977:AITMSG>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2).
- Searcy, C., K. Dean, and W. Stringer, 1998: PUFF: A high-resolution volcanic ash tracking model. *J. Volcanol. Geotherm. Res.*, **80**, 1–16, doi:[10.1016/S0377-0273\(97\)00037-1](https://doi.org/10.1016/S0377-0273(97)00037-1).
- Smith, A., T. Sears, G. Thomas, E. Carboni, and D. Grainger, 2014: Can we use sulphur dioxide as a proxy for volcanic ash in aviation hazard avoidance? Research Highlights 2014, National Centre for Earth Observation, National Environment Research Council, 8–8. [Available online at http://www.nceo.ac.uk/assets/documents/NCEO_highlights_2014.pdf.]
- Strunk, A., A. Ebel, H. Elbern, E. Friese, N. Goris, and L. Nieradzik, 2010: Four-dimensional variational assimilation of atmospheric chemical data application to regional modelling of air quality. *Large-Scale Scientific Computing*, I. Lirkov, S. Margenov, and J. Waśniewski, Eds., Lecture Notes in Computer Science, Vol. 5910, Springer, 214–222.
- Stuefer, M., S. R. Freitas, G. Grell, P. Webley, S. Peckham, and S. A. McKeen, 2012: Inclusion of ash and SO₂ emissions from volcanic eruptions in WRF-CHEM: Development and some applications. *Geosci. Model Dev. Discuss.*, **5**, 2571–2597, doi:[10.5194/gmdd-5-2571-2012](https://doi.org/10.5194/gmdd-5-2571-2012).
- Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328, doi:[10.1002/qj.49711347812](https://doi.org/10.1002/qj.49711347812).
- Tian, X., Z. Xie, and A. Dai, 2008: An ensemble-based explicit four-dimensional variational assimilation method. *J. Geophys. Res.*, **113**, D21124, doi:[10.1029/2008JD010358](https://doi.org/10.1029/2008JD010358).
- , —, and Q. Sun, 2011: A POD-based ensemble four-dimensional variational assimilation method. *Tellus*, **63A**, 805–816, doi:[10.1111/j.1600-0870.2011.00529.x](https://doi.org/10.1111/j.1600-0870.2011.00529.x).
- Valdebenito, A. M., S. Tsyro, M. Kahnert, and H. Heiberg, 2010: The EMEP data assimilation system: Technical description and first results. Tech. Rep. 4/2010, Norwegian Meteorological Institute, 30 pp.
- Wang, X., and Coauthors, 2008: Volcanic dust characterization by EARLINET during Etna's eruptions in 2001–2002. *Atmos. Environ.*, **42**, 893–905, doi:[10.1016/j.atmosenv.2007.10.020](https://doi.org/10.1016/j.atmosenv.2007.10.020).
- Webley, P. W., T. Steensen, M. Stuefer, G. Grell, S. Freitas, and M. Pavolonis, 2012: Analyzing the Eyjafjallajökull 2010 eruption using satellite remote sensing, lidar and WRF-Chem dispersion and tracking model. *J. Geophys. Res.*, **117**, D00U26, doi:[10.1029/2011JD016817](https://doi.org/10.1029/2011JD016817).
- Wen, S., and W. I. Rose, 1994: Retrieval of sizes and total masses of particles in volcanic clouds using AVHRR bands 4 and 5. *J. Geophys. Res.*, **99**, 5421–5431, doi:[10.1029/93JD03340](https://doi.org/10.1029/93JD03340).