Tephra deposit inversion by coupling TEPHRA2 with the Metropolis-Hastings algorithm and its implications

Qingyuan Yang\textsuperscript{1,2,*}, E Bruce Pitman\textsuperscript{3,4}, Marcus Bursik\textsuperscript{5} and Susanna F Jenkins\textsuperscript{1,2}

\textsuperscript{1}Earth Observatory of Singapore, Singapore
\textsuperscript{2}Asian School of the Environment, Nanyang Technological University, Singapore
\textsuperscript{3}Department of Materials Design and Innovation, University at Buffalo, Buffalo, NY, USA
\textsuperscript{4}School of Engineering and Applied Sciences, University at Buffalo, Buffalo, NY, USA
\textsuperscript{5}Department of Geology, University at Buffalo, Buffalo, NY, USA

\textsuperscript{*qingyuan.yang@ntu.edu.sg}

June 24, 2020

Abstract

In this work we couple the Metropolis-Hastings algorithm with the volcanic ash transport model TEPHRA2, and present the coupled algorithm as a new method to estimate the Eruption Source Parameters of volcanic eruptions based on mass per unit area or thickness measurements of tephra fall deposits. Basic elements in the algorithm and how to implement it are introduced. Experiments are done with synthetic datasets. These experiments are designed to demonstrate that the algorithm works, and to show how inputs affect its performance. Results are presented as sample posterior distribution estimates for variables of interest. Advantages of the algorithm are that it has the ability to i) incorporate prior knowledge; ii) quantify the uncertainty; and iii) capture correlations between variables of interest in the estimated Eruption Source Parameters. A limitation is that some of the inputs need to be specified subjectively. How and why such inputs affect the performance of the algorithm and how to specify them properly are explained and listed. Correlation between variables of interest are well-explained by the physics of tephra
transport. We point out that in tephra deposit inversion, caution is needed in attempting to estimate Eruption Source Parameters, and wind direction and speed at each elevation level, as this increases the number of variables to be estimated. The algorithm is applied to a mass per unit area dataset of the tephra deposit from the 2011 Kirishima-Shinmoedake eruption. Simulation results from TEPHRA2 using posterior means from the algorithm are consistent with field observations, suggesting that this approach reliably reconstructs Eruption Source Parameters and wind conditions from deposits.

**Keywords**

TEPHRA2, Metropolis-Hasings algorithm, Markov Chain Monte Carlo methods, Tephra inversion, Bayesian inversion

1 Introduction

Quantifying Eruption Source Parameters (ESPs), such as eruption plume height, eruption duration and variability, and mass eruption rate or total eruption mass, is critical to studies of volcanic eruptions and their products (Newhall and Self, 1982; Pieri and Baloga, 1986; Armienti et al., 1988; Scarpati et al., 1993; Mastin et al., 2009; Stohl et al., 2011; Pouget et al., 2013; Madankan et al., 2014; Bear-Crozier et al., 2020). Knowing the values of ESPs helps reconstruct pre-historic and unobserved eruptions, and provides information for the characterization of potential future hazards (e.g., Carey and Sparks, 1986; Bursik et al., 1992b,a; Bonadonna et al., 1998; Sparks and Young, 2002; Hildreth, 2004; Bonadonna and Houghton, 2005; Neri et al., 2008; Jenkins et al., 2008; Bonasia et al., 2010; Jenkins et al., 2012; Bonadonna and Costa, 2012; Bonadonna et al., 2015; Bevilacqua et al., 2015; Engwell et al., 2015; Yang and Bursik, 2016; Bevilacqua et al., 2018; Yang et al., 2019; Biass et al., 2019). ESPs are commonly estimated by coupling field observations with expertise on the process being analyzed. Such expertise could be in the form of a quantitative or descriptive physical model, an empirical or semi-empirical relationship, or their combination. Estimating ESPs can be treated as an inverse problem (e.g., Tarantola, 2005; Kaipio and Somersalo, 2006), and requires the use of different statistical and engineering techniques (Klawonn et al., 2012, 2014; Biass et al., 2016; Poret et al., 2017; Koyaguchi et al., 2017; White et al., 2017). Previous workers have presented different methods to implement inversion to obtain ESPs from the characteristics of tephra deposits, such as...
deposit thickness and grain size. The simplex search algorithm, grid-search method, matrix inversion with Tikhonov regularization, and a regularized form of the Levenburg-Marquardt algorithm have been proposed (Connor and Connor 2006; Klawonn et al. 2012; Johnston et al. 2012; White et al. 2017; Moiseenko and Malik 2019). The efficiency and the ability to characterize uncertainty with various simplifications (such as those used to avoid attempting to solve ill-posed problems) are the main concerns in proposing these algorithms as alternatives to classical inversion.

The challenges in estimating ESPs derive from their (1) high-dimensionality (i.e., too many variables to be estimated) and (2) limited field observations (e.g. Green et al. 2016). Because of the ill-posedness of this inversion, it is important to quantify the uncertainty in the process of estimating ESPs. In addition, it has been shown that ESPs influence model prediction through interaction with other ESPs (Scollo et al. 2008). In tephra inversion, such interactions or coupling could potentially lead to correlated results, i.e., estimated ESPs are correlated with one another, which has never been investigated in a systematic and statistically formal manner. Markov Chain Monte Carlo (MCMC) methods have the ability to quantify inherent uncertainty and address the presence of correlation between variables of interest in the estimate.

In this work, we present and introduce an algorithm that couples the Metropolis-Hastings (M-H) algorithm (Hastings 1970), one of the most widely-used MCMC methods, with volcanic ash transport and deposition model TEPHRA2 (Bonadonna et al. 2005) for the estimation of ESPs of explosive volcanic eruptions. Main advantages in using MCMC methods to estimate ESPs under a Bayesian framework are that (1) the estimation can be denoted as a posterior probability distribution, which enables uncertainty quantification; (2) prior knowledge on ESPs and field observations can be combined in a statistically formal way to jointly determine the result; (3) correlations among ESPs and between ESPs and wind conditions can be captured by the algorithm in the estimated result. Because of these advantages, the algorithm presented herein has the potential to better differentiate and characterize sources of uncertainty, and detect insensitive variables of interest in tephra inversion.

In this work, we introduce and demonstrate the algorithm in the following way. We introduce the physical model TEPHRA2 (Bonadonna et al. 2005) first. We briefly explain Bayes’ rule, which is what the Metropolis-Hastings (M-H) algorithm (Hastings 1970) solves numerically. An intuitive interpretation of the M-H algorithm is given. Then we describe in detail the construction and implementation of the MCMC algorithm. We apply the algorithm to simulated (synthetic) datasets, i.e., datasets generated from TEPHRA2 with known ESPs and wind conditions, to validate the algorithm.
In validating the algorithm, the manner in which input parameters affect its performance is highlighted and explained. We adjust the number of input observations (dataset size), sample site locations and number of variables to be estimated, to showcase that the algorithm works under different scenarios.

In the discussion section, the main advantages and limitation of the algorithm are pointed out. Correlation in the posterior distribution between variables of interest in our experiments is explained by the physics of tephra transport. Whether a simplified wind profile should be adopted in tephra inversion is discussed. We apply the algorithm to a dataset consisting of observed mass per unit area data of the tephra deposit from the well-studied 2011 Kirishima-Shinmoedake eruption to estimate its ESPs. The results are in general consistent with observations and estimates from previous studies.

The algorithm is coded in python scripts, and is published in vhub (Yang et al., 2020). To make the work accessible to a broad audience, we minimize the use of mathematical and statistical terms in introducing Bayes’ rule and the M-H algorithm. We hope that the algorithm can benefit researchers with interest in estimating ESPs of volcanic eruptions regardless of their backgrounds, and the text can serve as a tutorial to potential users.

2 Volcanic ash transport model TEPHRA2

TEPHRA2 is a widely-used volcanic ash transport and deposition model (Bonadonna and Houghton, 2005). It has been coupled with different statistical and engineering techniques for forward and inverse modeling of tephra fall deposits and volcanic hazard analysis (Bonadonna et al., 2005; Connor and Connor, 2006; Volentik et al., 2010; Fontijn et al., 2011; Biass et al., 2012; Magill et al., 2015; Biass et al., 2016, 2017; Takarada, 2017; Wild et al., 2019; Connor et al., 2019). TEPHRA2 assumes that tephra particles with different grain sizes are released from a line source instantaneously, and their transport is subject to wind advection, horizontal turbulent diffusion, and falling at terminal velocities.

Inputs of TEPHRA2 include parameters to characterize the eruptive column and wind conditions. TEPHRA2 gives semi-analytical solution to the advection-diffusion equation, and its output is the tephra mass per unit area deposited and grain size distribution at user-specified locations. TEPHRA2 assigns the total erupted mass $M^0$ to grain size bins (in $\phi$ unit) based on the specified grain size distribution. The total mass for each grain size is distributed along the eruptive column (discretized to points). When tephra particles with grain size $\phi_j$ released at the height of $H_i$ (their total mass: $M_{i,j}$) settle and deposit on the ground, the corresponding spatial distribution of mass per unit area
The mass per unit area \((m_{i,j}(x,y))\) of the deposit is proportional to a 2D Gaussian function, and can be written as:

\[
m_{i,j}(x,y) = M_{i,j} f_{i,j}(x,y),
\]

where \((x,y)\) is the spatial coordinates, and \(f_{i,j}(x,y)\) is the 2D Gaussian function with its mean and variance depending on the grain size \(j\), released elevation \(H_i\), wind speed and direction, and parameters that characterize turbulent diffusion (e.g., turbulent diffusion coefficient). The total mass per unit area at \((x,y)\) is the sum of Eq. 1 for all particle sizes released from the eruptive column. As the eruptive column, a line source, is discretized into many point sources in TEPHRA2, the total mass per unit area can be written as:

\[
m(x,y) = \sum_{i=0}^{H_{max}} \sum_{j=\phi_{min}}^{\phi_{max}} M_{i,j} f_{i,j}(x,y).
\]

To run TEPHRA2, ESPs, wind conditions, and locations of interest need to be specified. TEPHRA2 discretizes the atmosphere into multiple horizontal layers. The number of horizontal layers and their elevations as well as the corresponding wind speeds and directions need to be determined or specified as wind conditions by users in order to run TEPHRA2. See more information on the use and implementation of TEPHRA2 in Courtland et al. (2012).

3

Not applicable

4 Inversion technique

4.1 Bayes’ rule

Simply put, Bayes’ rule states that our prior knowledge about certain quantities of interest can be updated based on new observations. Assuming that the quantities of interest (i.e., ESPs in this study) is a vector \(\mathbf{x}\), the prior knowledge on its value can be denoted as a prior probability distribution \(P(\mathbf{x})\). With a series of observations \(\mathbf{\theta}\) (i.e., mass per unit area of tephra deposit on the ground in this study), our understanding on \(\mathbf{x}\) could be updated, which is denoted by the posterior probability distribution
(\(P(x|\theta)\)). Bayes’ rule is written as:

\[
P(x|\theta) = \frac{P(\theta|x) \cdot P(x)}{P(\theta)},
\]

(3)

where \(P(\theta|x)\) is the likelihood function. It denotes the probability of observing \(\theta\) given \(x\). \(P(\theta)\) is the evidence \(P(\theta) = \int P(\theta|x)P(x)dx\), and is the total probability of the observations. Here we refer to probability density as probability for convenience.

We can “insert” our knowledge on the process being analyzed into Bayes’ rule with the help of the likelihood function. In our case, this knowledge is the model TEPHRA2. We use the simplest case with only one variable of interest unknown, say column height, and one mass per unit area observation \(\theta_\ast\) to explain the likelihood function. We could apply one value of column height \(h_\ast\) to TEPHRA2, and collect the corresponding output \(d(x = h_\ast)\) with \(d(\cdot)\) denoting the TEPHRA2 output (assume one location of interest in this example). Neglecting model uncertainty, knowing \(x = h_\ast\) is equivalent of knowing \(d(x = h_\ast)\). Further assuming that the likelihood function follows a Gaussian distribution, its mean value could be \(d(x = h_\ast)\), and its variance needs to be determined based on our understanding of the data (e.g., the variance should scale with measurement uncertainty; see [Kawabata et al., 2013; Green et al., 2016] for more information on selecting the likelihood function).

If the true column height that generates \(\theta_\ast\) is 10 km, we expect to see the likelihood function having a greater value if \(h_\ast\) is closer to 10 km—the probability of observing \(\theta_\ast\) is greater when \(h_\ast\) is closer to 10 km. In general, the likelihood function should have a greater value, if the model output is similar to the observation. This is also why a Gaussian distribution centered at the model output can be used as one form of the likelihood function. The scale of the likelihood function, which is standard deviation in this example, reflects the scale of measurement uncertainty in the present context. If multiple observations are made, by assuming that each observation is made independently, the likelihood function is simply the product of likelihood function for each observation.

It should be noted that the likelihood function could have different formats. Even for a Gaussian distribution-like likelihood function, it does not have to be centered at the model output \(d(\cdot)\). For example, if it is known that the studied tephra deposit has been eroded for about 5 units of mass per unit area everywhere, the likelihood function should be centered at \(d(\cdot) + 5\).

To obtain the posterior probability distribution, we need \(P(x), P(\theta),\) and the likelihood function \(P(\theta|x)\). The prior distribution \(P(x)\), the likelihood function \(P(\theta|x)\), and \(P(\theta)\) (by definition) need
to be defined beforehand based on prior knowledge about the ESPs and measurement uncertainty. The major difficulty in analytically deriving the posterior distribution comes from the difficulty in calculating the value of $P(\theta)$ even though it is a constant.

### 4.2 The M-H algorithm

MCMC methods, a class of methods that draw samples from a target distribution, can be used to sample from the posterior distribution based on $P(\theta|x)P(x)$, the numerator of the left-hand side of Eq. 3. In this way, the difficulty in calculating $P(\theta)$ is avoided. In volcanology, MCMC methods have been widely adopted for various purposes, such as estimating parameters and initial conditions of a physical model (which is similar to the goal of the present work), determining ages of volcanic events, and hazard forecasting (e.g., Green et al. 2016; Anderson et al. 2019; Covey et al. 2019; Lev et al. 2019; Jenkins et al. 2019; Wang et al. 2020; Liang and Dunham 2020). Green et al. (2016) used one MCMC method to estimate volumes of tephra fall deposits based on sparse and incomplete observations, and their work used a semi-empirical model to characterize tephra thickness distribution.

The procedure adopted in this work, the M-H algorithm, is one of the most widely used algorithms for the implementation of MCMC methods (Hastings 1970). A brief introduction to the algorithm and intuition on how it works are given below. More information about the algorithm can be found easily in textbooks and published articles (e.g., Chib and Greenberg 1995; Andrieu et al. 2003; Kaipio and Somersalo 2006).

The M-H algorithm draws a series of vector points (i.e., ESPs and wind conditions in our case) following certain rules. With sufficiently many draws, it is guaranteed that the distribution of points converges/approximates the target probability distribution, namely the posterior distribution in our case, regardless of the value of the starting point. How the algorithm works in its most simplified form can be described as follow.

With a random starting point $x_0$ and corresponding observations $\theta$, the algorithm proposes a new point $x^*_1$ using a proposal function that is known and easy to sample. In this work, we use one of its most common forms, a Gaussian probability density function centered at the previous point (i.e., $x_0$ for the first draw). The variance of the proposal function needs to be defined subjectively which will affect the efficiency of the M-H algorithm, and will be illustrated in later experiments. Proposing a new point $x^*_1$ is thus equivalent of drawing one sample from a Gaussian probability distribution.

By calculating and comparing $P(\theta|x_0)P(x_0)$ with $P(\theta|x^*_1)P(x^*_1)$, the algorithm decides whether
to accept or reject $x_1^*$. If it is rejected, $x_1 = x_0$. Otherwise, $x_1 = x_1^*$. The two procedures, namely drawing a new point and rejecting or accepting it, iterate, and after sufficient iterations, a chain of vectors $x_0, x_1, \ldots, x_n$ is obtained. By (1) excluding points with relatively small index (e.g., $x_0, \ldots, x_{499}$), and (2) taking points with a fixed interval along the chain (e.g., only taking $x_{500}, x_{600}, \ldots, x_{9900}$, $x_{10000}$), the target posterior distribution is obtained through sampling. The first measure is to make sure that the results are not affected by the value of the starting point, and the second is to avoid auto-correlation in the chain (see Chib and Greenberg 1995; Andrieu et al. 2003; Kaipio and Somersalo, 2006 for more details).

Whether to accept or reject a proposed point follows the rules below:

- If $P(\theta|x_1^*)P(x_1^*) > P(\theta|x_0)P(x_0)$, then the posterior probability is greater at $x_1^*$, and the proposed point will be accepted. That is, if the proposal has a higher posterior probability it is automatically accepted. Following this rule ensures that there will be more samples with greater posterior probability in the chain.

- If $P(\theta|x_1^*)P(x_1^*) < P(\theta|x_0)P(x_0)$, then the algorithm accepts $x_1^*$ with probability $P(\theta|x_1^*)P(x_1^*)/P(\theta|x_0)P(x_0)$. This allows the algorithm to occasionally sample points with low or relatively low posterior probability, in order to explore the entire possible domain of $x$. Following this rule implies that:

  - If $P(\theta|x_1^*)P(x_1^*)$ is a lot smaller than $P(\theta|x_0)P(x_0)$, the posterior probability at $x_1^*$ is small, and should be accepted with low probability. This ensures that there will be fewer points with low posterior probability in the chain.

  - If $P(\theta|x_1^*)P(x_1^*)$ is only slightly smaller than $P(\theta|x_0)P(x_0)$, the probability of accepting $x_1^*$ is relatively greater. In such a case, the algorithm encourages (with high probability of acceptance) keeping $x_1^*$ and further exploring points around (sharing similar values with) $x_1^*$.

The second rule is critical to the M-H algorithm. Instead of searching for the value that maximizes the posterior probability (maximum a posteriori estimation), the algorithm functions through sampling from the target distribution.
5 Setup and running the algorithm

5.1 Form of the likelihood function

In the present version of the algorithm, it is assumed that the likelihood function for each observation follows a Gaussian distribution with variable \( \log(\frac{\text{observation}}{\text{model output}}) \). The mean of the distribution centers at 0 such that the likelihood function peaks when the observation is identical to the model output. This setup is consistent with previous works (e.g., Connor and Connor, 2006; Kawabata et al., 2013; White et al., 2017). It states that measurement uncertainty scales with magnitude of the observation. The variance or scale of the likelihood function, which reflects measurement uncertainty, needs to be specified by users. Its effect on the results of the algorithm will be examined in the following experiments. It should be noted that the form of the likelihood function could be changed to different forms easily.

5.2 Two ways to specify the wind profile

The algorithm allows for two ways to specify and estimate the wind profile. In cases with limited observations, it is impossible to estimate the wind speed and direction at each elevation. In such cases, a simplified form of the wind profile can be adopted. It assumes that (1) the wind direction is constant, and does not change with elevation, and (2) wind speed increases from zero to a certain value with elevation, and then decreases to zero with elevation linearly. Four variables define such a simplified wind profile. They are the wind direction, maximum wind speed, elevation that corresponds to the maximum wind speed, and elevation that the wind speed reaches zero. In the second way, all values are possible for the wind direction and speed at each elevation (non-simplified wind profile), and users could estimate the wind direction and speed at all or certain specified elevations.

5.3 Running the algorithm

To run the algorithm, users first need to prepare input files. Examples of input files can be found in the Appendix. They have five columns including the “variable name” column. The “initial value” column specifies the starting value of each variable. If the value of one certain variable is known (not to be estimated), its initial value will be fixed during the implementation of the algorithm, and the “prior” column needs to be marked as “Fixed” correspondingly. For such variables, users could leave the last three columns blank. To estimate a non-simplified wind profile, users must specify the wind
direction and speed at several elevation levels, depending on the form of the assumed profile.

For variables to be estimated, their initial values do not affect the results as long as the specified number of draws is sufficiently large, and users only need to specify values that are physically reasonable and would not lead to error when running *TEPHRA2*. Forms of prior distributions for these variables need to be defined in the “prior” column. The current version of the algorithm supports “Gaussian” and “Uniform” distributions. If the former is specified, columns “parameter_a” and “parameter_b” should be filled with the mean and standard deviation of the prior distribution, respectively. Otherwise, the two columns correspond to the minimum and maximum of the uniform prior distribution. The last column “draw_scale” specifies the standard deviation of the proposal functions for each variable to be estimated.

After the input files are prepared, users need to run all python scripts in order to execute the algorithm. Other than setting up proper file paths to read input files and observed data, and store the results, users only need to specify two values, namely the scale of the likelihood function and the number of draws (length of the chain), to run the algorithm. How to determine the number of draws is universally challenging when working with MCMC methods. It is preferred to have a large number of draws so that measures can be done to reduce autocorrelation in the sampled chain (e.g., leave out samples at the beginning of the chain and taking samples along the chain with a fixed interval as the final sample distribution; see Chib and Greenberg 1995, Andrieu et al. 2003, Kaipio and Somersalo 2006 for more details), and the results are less likely to be divergent. Here we recommend the number of draws to be ranging from five thousand to one million. This is based on our experiments with the algorithm. Users could check for convergence by running two or more separate runs with identical inputs except for the starting values (the initial value column). If sample distributions from the runs are similar to each other, results from them converge. Otherwise, users need to increase the number of draws, and run the algorithm and check again following the same procedure. See Andrieu et al. (2003) and references within for more information on how to implement the M-H algorithm properly.

The primary output from the algorithm is the sampled chain. If 10000 draws are specified, the sampled chain will be a 10000-by-23 (19 variables for the eruptive column plus 4 variables for the wind profile) matrix if a simplified wind profile is adopted. Otherwise, with the non-simplified wind profile scenario, the result will be presented as three separate matrices for the eruptive column (10000-by-19 matrix) and wind speed and direction at each elevation (two 10000-by-40 matrices if the specified elevations are 1000, 2000,...,40000 m), respectively. Variables with known and fixed values will remain
constant in the corresponding columns.

The algorithm also produces the log-transformed values of the prior probability and likelihood function of the proposed values at each draw, log-transformed posterior probability for each point on the chain, and the number of acceptances from the run as outputs. They could potentially help users to debug and adjust parameters to run the algorithm.

The running time of the algorithm depends on the specified number of draws and the running time of TEPHRA2. On an iMac with a 3 GHz Intel Core i5 processor, implementing the algorithm with ten thousand draws takes 394 seconds (~6.5 minutes).

6 Algorithm demonstration

In this section, we generate simulation data using TEPHRA2. The simulated data are treated as field measurements for testing and validation of the algorithm. The specified ESPs and wind conditions to generate these “field measurements” are listed in Table 1. Twelve experiments are set up. We set column height and total eruption mass as our variables of interest, and their true values are 15000 m and $1.88 \times 10^{11}$ kg ($25.96 = \log(1.88 \times 10^{11})$ log-transformed kg), respectively. Values of other ESPs and wind conditions used to run TEPHRA2 are listed in Table 1 and are fixed in all following experiments. The simplified wind profile is adopted throughout the 12 experiments. Focusing on just two variables makes it easier for examining and interpreting the results. The 12 experiments are also designed to highlight how the algorithm is affected by its inputs. The first experiment, Experiment # 0, is used as reference for comparison with results from the rest. For the rest experiments, we just modify one input of the algorithm or the observation dataset, and keep the others the same as they are in Experiment # 0. In this way, the impact of each factor on the performance of the algorithm can be isolated and highlighted. Changes made in each experiment are listed in Table 2.

For the 12 experiments, the M-H algorithm is set to draw 10000 points (10001 points in the chain including the starting point). After each experiment is finished, we post-process the results by taking the first 1000 points out of the chain, and collecting samples from the rest of the chain by a 15-points interval (only taking 1015th, 1030th, ..., 10000th points in the chain) and discarding all other points. These measures are adopted to prevent the results from being affected by the initial starting point and autocorrelation as mentioned earlier. For two variables of interest, a chain with 10000 draws is sufficient enough for the samples to converge to the posterior distribution. Summary of the sampled
posterior distributions is given in Table 3.

In addition to the 12 experiments, we also test the performance of the algorithm in cases with non-simplified wind profile. An example of such an experiment is given. In all of our experiments, we find that the resultant sample posterior distributions are symmetric, and resemble Gaussian distributions. This allows us to use posterior means and standard deviations to characterize the results, and the posterior means also correspond to values with the maximum posterior probability.

6.1 Reference experiment

In Experiment # 0, ten observations are sampled at localities far from the source vent downwind (~14 – 30 km from the vent; sample sites shown in Fig. 1b). Their values range from 50-383 kg/m². We set the priors as Gaussian distributions for the column height and log-transformed total eruption mass. The means and standard deviations are 16000 m and 2000 m for column height, and 25.96 log-transformed kg (1.88 × 10¹¹ kg) for the eruption mass. It is noted that 16000 m is slightly greater than the specified column height used to generate the dataset, and the prior mean for the log-scaled eruption mass is identical to the true value (Table 2).

The results are shown and summarized in Table 3 and Fig. 2. The posterior means for the column height and log-scaled eruption mass are 15338 m and 25.928 log-transformed kg, respectively, and the corresponding posterior standard deviations (of the samples) are 1067 m and 0.066 log-transformed kg. Both posterior means are consistent to the true values, and the posterior standard deviations are much smaller than those of the priors (2000 m and 2 log-transformed kg, respectively). The result suggests that the algorithm works. With ten observations and the presented algorithm, the reduced standard deviations and consistency between the posterior means and the true values suggest that our knowledge about the column height and total eruption mass has correctly improved.

6.2 Effects from specifications required by the M-H algorithm

6.2.1 Scale of the likelihood function

In Experiment # 1, we increase the scale of the likelihood function from 0.05 to 0.2 (Table 2). A greater scale of the likelihood function implies that a greater measurement uncertainty is assumed (neglecting model uncertainty). The resulting posterior from Experiment # 1 is not significantly different from the prior for column height. Its posterior mean and standard deviation are 15990 m.
and 1775 m, respectively (prior mean and standard deviation: 16000 m and 2000 m). The larger measurement uncertainty means the measurements do not substantially change, and the posterior is not significantly different from the prior.

In other words, to the algorithm, new data are collected, but cannot be trusted due to the greater measurement uncertainty. This is also manifested in the acceptance rate of Experiment # 1, which is 84.0%. What the M-H algorithm does here is basically accepting or rejecting points based on the prior distribution, and the role of the likelihood function is largely ignored due to the greater scale of the likelihood function (i.e., greater measurement uncertainty). The posterior distribution of the log-transformed total eruption mass still centers close to the true value with lowered standard deviation compared to its prior, suggesting that estimating total eruption mass is less sensitive to measurement uncertainty.

6.2.2 Scale of the proposal function

In Experiment # 2, scales of the proposal function are increased from 500 m and 0.05 log-transformed kg to 2000 m and 0.2 log-transformed kg for column height and log-transformed total eruption mass, respectively (Table 3). Results from Experiments # 0 and 2 resemble each other (Table 3). With greater scales of the proposal functions, the algorithm is more likely to propose a new point that is far from the current point. Theoretically, this difference would not affect the posterior distribution as long as sufficient amount of draws are done. However, scales of the proposal functions would affect the efficiency of the algorithm and sometimes its performance when the number of draws is not large enough. The algorithm is being too “adventurous” with proposal functions characterized by greater scales: they tend to explore/propose values which are greatly different from the current point. Such values are less likely to be accepted especially when the current values are characterized by high posterior probability (acceptance rate: 13.6% for Experiment # 2). The greater probability of rejection slows down the process of exploring different values of the ESPs, and reduces the efficiency of the algorithm.

6.2.3 Prior distributions

Experiments # 3 and 4 are designed to test how prior distributions affect the posterior distribution. The prior means of the column height are set to be 14000 m and 12000 m, respectively, with standard deviation of 500 m for both experiments. The prior means for the log-scale eruption mass are 23.66
log-transformed kg, and the standard deviations are 0.5 log-transformed kg in the two experiments. The prior means are different from the true values of column height (15000 m) and total eruption mass (25.96 log-transformed kg). In the two experiments, we are incorrect (incorrect prior means) yet confident (small standard deviations for the priors) in our prior knowledge. As a result, the ten observations are not powerful enough to correct the prior. Posterior means of the two experiments are 14283 and 12666 m for the column height (std: 442 and 459 m), and 25.997 and 26.138 log-transformed kg for the log-scaled total eruption mass (std: 0.048 and 0.059 log-transformed kg), respectively. These two examples highlight the fact that the posterior distribution could be affected by the prior, and that the prior needs to be determined with caution.

6.2.4 Number of input observations

Experiments # 5- 9 share specifications with Experiments # 0-4, respectively, except that 20 more samples are included in the input. The 30 measurements (see Fig. [1] for sample localities) range from 32-383 kg/m². The results are compared pairwise, and summarized in this section.

Comparison between Experiments # 0 and 5 shows that they have similar posterior means that are consistent with true values of column height and log-transformed total eruption mass. The corresponding standard deviations are much smaller in Experiment # 5 (685 m and 0.039 log-transformed kg for column height and log-scaled eruption mass). This shows that more observations reduce the uncertainty in the posterior distribution. The same argument can be made for Experiment # 7 as it has the same specifications as Experiment # 5 does except for greater scales of the proposal functions (which is discussed in Experiment #2).

Even with more observations, the posterior distribution (mean: 15786 m; std: 1633 log-transformed kg) of column height in Experiment # 6 is not greatly updated from the prior. This is again due to the greater likelihood scale specified in Experiment # 6. Experiments # 8 and 9 with incorrect and confident priors have their posterior means (14483 and 13340 m for column height and 25.986 and 26.070 log-transformed ky for log-transformed total eruption mass) closer to the true values compared to results from Experiments # 3 and 4, and the posterior distributions in Experiments # 8 and 9 are also characterized by lower standard deviations (Table [2]). Even though, the posterior means are still not consistent with the specified values. It can be seen that more observations “drag” the means of the posterior distributions towards the true values. Results from experiments # 3, 4, 8, and 9 reflect the “wrestling” between prior knowledge and observations.
6.2.5 Sample site locations

In Experiments # 10 and 11, ten measurements at proximal (and closely-spaced) and medial localities are used as input observations, respectively (see sites in Fig. 1a). Results from these two experiments can be compared with that of Experiment # 0 to examine if sample site localities would affect the results. In this work, we only consider cases with sample sites located in the downwind area with respect to the source vent. The two experiments are set up to demonstrate that the algorithm could be used to study how different factors (sample site locations in the two experiments) affect the uncertainty in estimating ESPs in tephra inversion.

The posterior distributions from Experiment # 10 are not greatly updated. The posterior mean and standard deviation of the column height (16085 m and 1863 log-transformed kg, respectively) are close to the prior (Table 3). The posterior distribution of the log-transformed eruption mass is 26.106 log-transformed kg, and is characterized by a relatively greater standard deviation (0.280 log-transformed kg) compared to results from other experiments. The proximal observations are too close to each other, and as a result, they provide limited useful information to the algorithm to update the priors.

In Experiment # 11, the sample sites are spatially medial to the source vent, i.e., neither especially close nor far relative to the entire span of distances (Fig. 1a). The sites are distributed in a more scattered pattern than are those in Experiment # 10. These observations provide more useful information than do the proximal observations, which can be used to update the priors. As shown in Table 3, the posterior means of the column height and log-scaled eruption mass are consistent with the true values, and the posterior distributions are also characterized by smaller standard deviations compared to those in Experiment # 10 and the specified priors.

6.3 Experiment with non-simplified wind profile

Results from one experiment with non-simplified wind profile are reported in this section. The ESPs and wind conditions to generate the synthetic data are shown in Table 1. Change in wind speed and direction with elevation is taken into account in this experiment. The wind speed is specified to increase from 0 on the ground to a maximum of 70 m/s at 16 km, and then decrease to 10 m/s at 24 km. The wind direction is to the north on the ground, and gradually changes to eastwards with elevation.

We set our ESPs of interest to be column height and log-transformed total eruption mass. For
the wind profile, we choose to estimate wind directions and speeds at elevation levels 5, 14, and 18 km a.s.l. We do not choose to estimate wind direction and speed at each elevation (i.e., estimate the complete wind profile) because that would significantly increase the number of variables to be estimated. In such circumstances, the problem tends to be ill-posed, and whether it is solvable or not might become uncertain. Therefore, an experiment that estimates the wind direction and speed at each elevation cannot be used to justify the method works. Neither can it be used to justify that the method does not work. Here, the current experiment is designed to show that the method is able to estimate wind speed and direction at different elevations when the problem is known to be solvable (relatively fewer variables to be estimated). The problem and difficulty in estimating a non-simplified wind profile is discussed in the following section.

Elevation levels 5, 14, and 18 km are chosen because wind directions at these elevations are 25°, 65° and 85°, respectively. We “collect” tephra mass per unit area at 495 randomly selected sample sites (Fig. [1]). The dataset size is greater than commonly-seen thickness or mass per unit area datasets for tephra fall deposits. Estimating wind speed and direction at a few elevations and having 495 observations are not realistic in studies on tephra fall deposits. However, these are adopted such that we know that the problem is solvable, and can be used to validate the algorithm. In more realistic scenarios, whether the algorithm is able to correctly estimate the ESPs and wind conditions depends on model uncertainty, measurement uncertainty, number of observations, and potentially the unknown unknowns. Testing the algorithm with the current setup excludes such additional factors, and is considered as a more strict way of testing the algorithm.

Other ESPs, and wind directions and speeds at other elevations, are kept fixed throughout the implementation of the algorithm. Twenty thousand runs are done for three times with different starting points. In each run, the first 2000 runs are discarded to avoid auto-correlation, and samples with an interval of 20 points along the sample chain are subsetted as the final results. The resultant sample posterior distributions from the three runs are similar, suggesting that the results converge in each run. The posterior mean and standard deviation for each variable of interest are listed in Table [4]. They are highly consistent with specified values used to generate the dataset, and the posterior standard deviations are a lot smaller than those of the priors. The results suggest that the algorithm functions when a non-simplified wind profile is adopted. Greater uncertainty is obtained for wind direction and speed at higher elevations (i.e., 14 and 18 km). This is due to the fact that the wind speed at higher elevations is generally greater, and the estimated uncertainty scales with it.
7 Results

Not applicable

8 Discussion

In this work, we introduce an algorithm coupling the ash dispersal model *TEPHRA2* with the Metropolis-Hastings implementation of MCMC. We validated it with synthetic data generated by *TEPHRA2*. By varying the inputs to the algorithm and observation datasets one at a time, we examine and explain how they affect the performance and efficiency of the algorithm. Thirteen experiments are done. The first twelve of them focus on simple scenarios (i.e., with simplified wind profile) with two variables of interest (i.e., total eruption mass and column height) given ten to thirty observations. In these experiments, the algorithm is shown to work well, and has the ability to quantify the uncertainty in the estimate. In the thirteenth experiment, we set out to estimate two ESPs, and wind directions and speeds at three elevation levels, with sufficient observations. In this experiment, the wind direction is set to vary with elevation. The results again suggest that the algorithm could work well under such more complex scenarios. As the goal of this work is to present this algorithm and introduce and explain how it works, we avoid using real thickness or mass per unit area datasets of tephra fall deposits in testing the algorithm. This would avoid additional sources of uncertainty from affecting the results, and is considered as a stricter measure for testing and validating a method.

Our discussion here focuses on advantages and limitations of the algorithm, interpreting the posterior correlation between column height and total eruption mass in our experiments, and whether we should attempt to estimate wind direction and speed at each elevation in tephra inversion when working with *TEPHRA2*. The algorithm is then applied to the mass per unit area dataset of the 2011 Kirishima-Shinmoedake tephra deposit to infer the corresponding ESPs.

8.1 Advantages

The main advantages of the algorithm are that it makes use of prior knowledge on a deposit and eruption, and quantifies the uncertainty in the estimate of ESPs in a statistically formal manner. In studies on tephra fall deposits, previous knowledge plays a critical role in determining the ESPs and reconstruction of volcanic eruptions. Such knowledge has uncertainty within it. How to properly incorporate such uncertainty in the estimated results is challenging without a probabilistic Bayesian
framework. With the algorithm, prior knowledge about the studied deposit and eruption, and their associated uncertainties, is denoted as the prior probability distribution, and incorporated in the estimate.

Practically speaking, prior knowledge is being used consistently throughout the implementation of the algorithm. That is, the prior probability helps determine whether to accept or reject a proposed point in each draw of the algorithm. However, for a non-probabilistic inversion method, such as gradient methods, prior knowledge might be used only once in the inversion process—it helps determine the starting point.

In tephra deposit inversion, uncertainty in the estimated ESPs comes from the interplay of multiple sources, which include the uncertainty in the prior, measurement uncertainty, and potential model uncertainty. How they interact with each other and affect the final uncertainty has never been studied systematically for tephra inversion. The algorithm might be able to help address this concern in future work.

8.2 Limitations

Whether the priors, number of draws, and standard deviations of the proposal functions are specified properly or not affects the performance of the algorithm. This problem arises as long as the M-H algorithm is adopted, and there is no definitively correct way to determine some of such inputs to the algorithm. What we attempt to do, in this work, is to explain and highlight why and how all the inputs affect the performance of the algorithm through experiments, and present commonly adopted measures and references (Chib and Greenberg, 1995; Andrieu et al., 2003; Kaipio and Somersalo, 2006) on how to use the algorithm properly and how to check if the results converge. We hope that introducing the algorithm in this way helps users in understanding and implementing.

8.3 Correlation in the posterior distribution

From examining the sampled posterior distributions, we find that correlation exists between column height and total eruption mass. Here we focus on the correlations in Experiments # 0-11, as their setups are simpler. Correlations in these experiments are shown in Table 3 and the bivariate posterior distributions from selected experiments are shown in Fig. 3.

We find that both negative and positive correlations exist between column height and total eruption mass in the sampled posterior distributions, and whether the correlation is positive or not depends
on sample site localities. In Experiments # 10 and 11, the correlation is positive (0.986 and 0.831), and for the rest, the correlation (-0.474- -0.856) is negative. This can be explained by the physics of tephra transport. Considering all other ESPs fixed except for column height and total eruption mass, if observations are made at distal localities (Experiments # 0-9), the combination of a greater column height, which allows tephra to be dispersed farther downwind and leads to more tephra deposition at distal sites, and a smaller total eruption mass (eruption mass is proportional to tephra thickness/mass per unit area everywhere) leads to results similar to those for a lower column height and a greater total eruption mass (within the area where footprints of tephra deposits overlap). Therefore, the correlation is negative in Experiments # 0-9. However, a lower column height leads to a thicker tephra deposit at proximal sites because of less interaction with wind, and less time for turbulent diffusion to disperse tephra to distal localities. Total eruption mass is always proportional to tephra thickness and mass per unit area, i.e., scenarios with (1) greater column height and greater total eruption mass and (2) lower column height and lower total eruption mass lead to similar observed thickness or mass per unit area if the observations are all made at proximal sites. These relationships are well-known (e.g., Suzuki et al. 1983, Bonadonna et al. 2005), but how their interaction with sample site locations leads to and affects the correlation between the estimated column height and total eruption mass in tephra inversion has not been reported even in the simplest scenarios, i.e., our experiments with synthetic data and all other ESPs known.

It is worth noting that the correlation in Experiment # 10 is surprisingly high (0.986). This is due to the fact that the sample sites are too close to each other, and the “observed” tephra mass per unit area values at these sites also have similar values. To the algorithm, the ten observations in Experiment # 10 are almost the same (both their locations and their observed values). The high correlation in fact reveals that non-unique solutions exist in Experiment # 10. The presence of correlation in the posterior distribution suggests that the interaction of variables plays a role in tephra inversion. This is consistent with results from sensitivity analysis on TEPHRA2 in the work of Scollo et al. (2008).

This finding suggests that the algorithm has the potential to be used to discover intrinsic relationships (interactions) between variables of interest and wind conditions in TEPHRA2 and other dispersion models, and could thus improve our understanding on different sources of uncertainty in tephra inversion.
8.4 Whether to estimate the wind direction and speed at each elevation

The algorithm can be used with either simplified or non-simplified wind profiles. It is always desired and preferred to have a more exact and detailed understanding on wind conditions in tephra inversion. At the same time, we also should not ignore its corresponding practicality. It is common to estimate six or more ESPs (e.g., column height, total eruption mass, column mass distribution, mean and standard deviation of grain size distribution, and diffusion coefficient) in tephra inversion. If wind direction and speed are to be estimated at ten elevations, this adds up to at least $26 = 6 + 10 \times 2$ variables of interest. Considering just two values for each variable, this means $2^{26}$ (which is greater than 60 million) possible combinations for the 26 variables. The number of field measurements for tephra fall deposits rarely exceeds 300. In such circumstances, together with other sources of uncertainties, estimating the ESPs and the complete non-simplified wind profile would be challenging from a technical and non-volcanological perspective regardless of which inversion method is adopted. This concern motivates us to propose the two options to specify and estimate the wind profile. Estimating a simplified wind profile conforms to the Ocham’s razor.

However, as shown in previous works, estimating wind direction and speed at each elevation together with the ESPs can be done (e.g., [White et al., 2017]). Prior knowledge on local weather conditions and expertise on the studied tephra deposits (e.g., constructed isomass and isopach maps) provide additional constraints on the ESPs and wind conditions, which help make the problem solvable. A question naturally arises, that is, under what circumstances (e.g., spatial distribution and number of sample sites, observed values of tephra thickness, and prior knowledge on the wind condition), it is possible to estimate the ESPs and the non-simplified wind profile. This question will be crucial for the efficiency of tephra inversion and interpretability of its results. It involves factors such as measurement uncertainty (i.e., whether a few outlier observations are due to the complex wind profile or measurement error) and how much we trust the prior knowledge. This question cannot be answered without a probabilistic framework, and we think that the presented algorithm might be key to this question. The question of whether it would be practical and meaningful to estimate the ESPs and a non-simplified wind profile is comparable to the problem of detecting multiple lobes in tephra thickness distribution [Kawabata et al., 2013, 2015], which is solved by coupling a semi-empirical model of tephra thickness distribution with statistical methods [Kawabata et al., 2013, 2015]. In our case, however, the difference is that the forward model TEPHRA2 is characterized by a lot more variables of interest.
8.5 Application to the Kirishima-Shinmoedake dataset

In this section, we apply the M-H algorithm to a dataset containing mass of tephra per unit area for the 2011 Kirishima eruption. The Kirishima-Shinmoedake event took place from 26 to 29 January, 2011, with an eruption of the Shinmoedake volcano. The column height ranged from 6.2-8.6 km above the crater, based on different models and Doppler radar measurements (Shimbori and Fukui, 2012; Maeno et al., 2014). The total eruption mass was estimated to be $1.8 - 3.1 \times 10^{10}$ kg by Nakada et al. (2013). The mass of tephra erupted from the afternoon of 26 January to the early morning of 27 January, which corresponds to the current dataset, is about $1.4 - 2.5 \times 10^{10}$ kg (Maeno et al., 2013). The wind was blowing to the southeast, and the wind profile is reported in Hashimoto et al. (2012).

The tephra deposit data we are using are reported in Miyabuchi et al. (2013), and downloaded from White et al. (2017). The dataset was collected from the main tephra fallout deposit (emplaced from the evening of 26 January to the morning of 27 January). Tephra thickness and grain size distributions were measured at 55 locations downwind from the vent. In addition, tephra thickness was measured at another 63 locations. The thickness measurements were converted to mass per unit area. See Miyabuchi et al. (2013) for sample processing procedures.

We set the ESPs to be estimated to be column height, eruption mass, $\alpha/\beta$ ratio (which characterizes the mass distribution of tephra within the eruptive column), median and standard deviation of the grain size distribution, diffusion coefficient, fall time threshold, and densities of lithic (stony) and pumice fragments. The eddy constant is fixed as 0.04. For the wind condition, a simplified wind profile is adopted, to avoid overcomplication of the problem, as discussed before. We assume that the wind speed increases linearly from 0 to 11 km a.s.l., and then decreases with elevation to 24 km a.s.l. This setup is based on the wind speed profile reported in Hashimoto et al. (2012). Wind direction and maximum wind speed are the two variables to be estimated for the wind profile. This amounts to 11 variables of interest and 118 mass per unit area observations for the problem. The priors of these variables are inferred based on Shimbori and Fukui (2012); Nakada et al. (2013); Miyabuchi et al. (2013); Maeno et al. (2014); White et al. (2017), and are shown in Table 5.

We draw fifty thousand points; this process is repeated three times. Sample distributions from the three runs are almost identical to each other, suggesting that the results converge. The first 5000 samples are discarded to exclude the impact from the initial starting point. For the rest of the chain, we collect samples based on a 15-point interval to avoid autocorrelation.

The results are summarized in Table 5. Posterior means of column height and total eruption
mass are in general consistent with previous studies, and are updated from the priors with posterior standard deviations being much smaller. The simulated mass per unit area data from TEPHRA2 with posterior means as ESPs and wind conditions are plotted against field observations in Fig. 4 which suggests that TEPHRA2 could generally reproduce field observations based on estimated results from the algorithm. Posterior means of column height and total eruption mass are 7.3 km and $9.14 \times 10^9$ kg, respectively. The former is within the range of the observed column height, and the latter is slightly smaller than estimates from previous work. Posterior distributions of the other ESPs generally lie within commonly-seen ranges (Table 5), and are also altered from the corresponding priors. We note that the posterior mean of the median of the grain size distribution is finer than data reported by Miyabuchi et al. (2013). We think that this could be explained by the fact that data reported by Miyabuchi et al. (2013) represent the grain size distribution at certain sample sites, whereas our estimate focuses on the total grain size distribution.

This dataset has been used in White et al. (2017) for testing another tephra inversion method (i.e., Levenburg-Marquardt algorithm with Tikhonov and subspace regularization). We do not intend to compare our results with theirs for several reasons: (1) the goal of their method is to efficiently implement tephra inversion with the ability to quantify the uncertainty in the estimate, and additional simplifications are assumed in the method; (2) different priors are assumed, and their work estimates 10 ESPs and wind directions and speeds at 11 elevation levels. We think that the two algorithms have their own advantages, and which algorithm to use depends on users’ need.

9 Conclusions

In this work, we couple the well-known M-H algorithm with the tephra transport and deposition model TEPHRA2. The coupled algorithm can be used to infer ESPs of explosive volcanic eruptions and ambient wind conditions based on thickness or mass per unit area measurements of tephra fall deposits under a Bayesian framework. It allows users to include their prior knowledge on the eruption or deposit with field observations in a statistically formal way. The result of the algorithm is presented as sample posterior distributions for the variables of interest.

We introduce the model TEPHRA2 and basic elements of Bayes’ rule, and present intuitive interpretations on the M-H algorithm. How to implement the coupled algorithm and formats of the input files are also introduced.
The coupled algorithm is validated with 13 experiments. For the first twelve of them, we focus on two variables of interest. In these experiments, we vary values of the input and size of the synthetic observation dataset one at a time to show that the algorithm works, and also to show how inputs affect the performance of the algorithm. In the 13th experiment, we set eight variables of interest to be estimated, including not only two ESPs, but also wind directions and speeds at three elevation levels. This experiment is set up to show that the algorithm works with a complex wind profile. We do not use real datasets for validating the algorithm, because inversion results based on real datasets could be affected by additional sources of uncertainty (e.g., measurement uncertainty).

The advantages of the algorithm are that it has the ability to incorporate prior knowledge into the estimate in a statistically formal way, and to quantify the uncertainty in the estimate, and it captures correlation between variables of interest in the estimate. Because of these advantages, we think that the algorithm has the potential to improve our understanding on how different sources of uncertainty interact and affect the results in tephra inversion. The main limitation of the algorithm is that there are subjective choices in implementation, which affect its performance. How and why they affect the algorithm are introduced, and commonly adopted measures and references ([Chib and Greenberg 1995; Andrieu et al. 2003; Kaipio and Somersalo 2006]) on how to specify the inputs properly are given.

In the 12 experiments with simplified wind profiles, it is found that correlation exists between the two variables of interest, namely column height and total eruption mass, in the sample posterior distributions. Whether the correlation is positive or not depends on sample site locations. These could be well explained by the physics of tephra transport: a greater column height has a positive and negative relationship with tephra mass per unit area at distal and proximal sample sites, respectively, and total eruption mass is always proportional to tephra mass per unit area regardless of sample site location.

The algorithm supports specifying and estimating the wind profile in two ways. The first way takes advantage of a simplified wind profile based on four variables of interest, and assumes that the wind direction does not change with elevation. The second one allows users to estimate wind speed and direction at each elevation. We argue that users need to be cautious in choosing how to specify and estimate the wind profile, because the second way could introduce more variables to be estimated, and potentially make the problem extremely ill-posed. How to choose the appropriate way to specify and estimate the wind profile relies on factors such as prior knowledge of weather conditions and sample
site distributions. We think that by experimenting on appropriate synthetic data, this question can be addressed. We apply the algorithm to the 2011 Kirishima-Shinmoedake tephra dataset, and the results are in general consistent with observations and estimates from previous work. We hope that the present work benefits future studies that attempt to implement tephra inversion and quantify the associated uncertainty.

List of abbreviations

ESPs: Eruption Source Parameters
MCMC methods: Markov Chain Monte Carlo methods
M-H algorithm: Metropolis-Hastings algorithm

Declarations

Availability of data and materials

The data used in this work are mostly generated from synthesized experiments. How to generate these data is specified in the text. The dataset of the tephra deposit from the 2011 Kirishima-Shinmoedake eruption used in this work is from previous publications. Where to find the dataset and corresponding references are given in the text.

Competing interests

There is no competing interest involved in this work.

Funding

This work was partially supported by National Science Foundation Hazard SEES grant number 1521855 to G. Valentine, M. Bursik, E.B. Pitman, and A.K. Patra. This work comprises Earth Observatory of Singapore contribution no. 306. This research is partly supported by the National Research Foundation Singapore and the Singapore Ministry of Education under the Research Centres of Excellence initiative (Project Name: Evaluating Unrest and Potential Hazards at Changbaishan Volcano, China; Project Number: NRF2018NRF-NSFC003ES-010) to S.F. Jenkins and Q. Yang.
Authors’ contributions

The idea of this manuscript, i.e., coupling the Metropolis-Hastings algorithm with TEPHRA2 was proposed by E.B. Pitman. Q. Yang coded the algorithm in python scripts. Q. Yang wrote the manuscript with inputs from E.B. Pitman, M. Bursik, and S.F. Jenkins.

Acknowledgements

Not applicable

Tables

Table 1: Initial and wind conditions used to generate “field observations” for validation of the algorithm. Red- and green-striped cells correspond to ESPs for the twelve experiments with simplified wind profile and the experiment with the non-simplified wind profile, respectively. Yellow- and blue-striped cells correspond to wind profile specifications for the twelve experiments with simplified wind profile and the experiment with non-simplified wind profile, respectively.

| Variable name | Specified value (simplified wind profile) | Specified value (non-simplified wind profile) | Wind profile specification |
|---------------|------------------------------------------|-----------------------------------------------|----------------------------|
|               | Variable name                             | Elevation (m) Wind speed (m/s) Wind speed (°) Elevation (m) Wind speed (m/s) Wind speed (°) |
| Column height (m) | 25000                                    | 20000                                       | 1000 0 5 15000 10 90 |
| Total eruption mass (kg) | 1.88*10^11                              | 5*10^11                                    | 2000 0 5 20000 10 90 |
| α             | 5                                        | 1.5                                        | 3000 10 15 21000 10 90 |
| β             | 2                                        | 1                                          | 4000 10 15 25000 10 90 |
| Max grain size (φ) | -6                                      | -6                                         | 5000 20 25 25000 10 90 |
| Min grain size (φ) | 6                                        | 6                                          | 6400 20 25 26000 10 90 |
| Median grain size (φ) | 3                                        | 0                                          | 7000 10 15 26000 10 90 |
| Standard deviation of grain size | 1.5                                      | 2                                          | 8000 20 25 26000 10 90 |
| Vent coordinates (m) | (6, 6)                                   | (6, 6)                                     | 9000 40 45 27000 40 90 |
| Vent elevation a.s.l (m) | 1000                                     | 1000                                       | 10000 40 45 |
| Eddy constant | 0.24                                      | 0.04                                       | 13000 10 50 |
| Diffusion coefficient (m²/s) | 5000                                     | 3000                                       | 12000 10 50 |
| Fall time threshold | 1419                                     | 1419                                       | 13000 10 60 65 |
| Pumice density (kg/m³) | 1100                                      | 1000                                       | 15000 70 75 20000 10 65 |
| Column steps | 100                                      | 100                                        | 16000 70 75 |
| Particle steps | 2                                        | 2                                          | 18000 50 65 |
| Plume model | 2                                        | 2                                          | 18000 50 85 10000 |

Acknowledgements

Not applicable

Tables

Table 1: Initial and wind conditions used to generate “field observations” for validation of the algorithm. Red- and green-striped cells correspond to ESPs for the twelve experiments with simplified wind profile and the experiment with the non-simplified wind profile, respectively. Yellow- and blue-striped cells correspond to wind profile specifications for the twelve experiments with simplified wind profile and the experiment with non-simplified wind profile, respectively.
Table 2: Specifications and input observation data used to run Experiments # 0-11. Differences in the specification or input observation data in each experiment compared to Experiment # 0 (as reference experiment; marked as green cells) are highlighted in yellow.

| Run # | # of inputs | Column height prior (type, mean, std) | Eruption mass prior (type, mean, std; log-scale) | Scale of proposal function for column height | Scale of proposal function for log(eruption mass) | Likelihood scale | Note |
|-------|-------------|--------------------------------------|-----------------------------------------------|---------------------------------------------|-----------------------------------------------|-----------------|------|
| 0     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.05                                          |                 | Reference run (used for comparison with other simulations). |
| 1     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.2                                           |                 | Input observations are at ten sites close to the source vent. |
| 2     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 2000                                          | 0.2                                         | 0.05                                          |                 |                  |
| 3     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.05                                          |                 | Input observations are at ten sites medial to the source vent. |
| 4     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 2000                                          | 0.2                                         | 0.05                                          |                 |                  |
| 5     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.05                                          |                 |                  |
| 6     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.05                                          |                 |                  |
| 7     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 2000                                          | 0.2                                         | 0.05                                          |                 |                  |
| 8     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.05                                          |                 |                  |
| 9     | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.05                                          |                 |                  |
| 10    | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.05                                          |                 |                  |
| 11    | 10          | Gaussian(16000, 2000) Gaussian(log(1.88*10^{11}), 2) | 500                                           | 0.05                                        | 0.05                                          |                 |                  |
Table 3: Summary of results from Experiments # 0- 11.

| Run # | Posterior mean (column height; m) | Posterior std (column height) | Posterior mean and std (eruption mass; log-scale kg) | Posterior std (eruption mass; log-scale) | Correlation | Acceptance rate (%) |
|-------|---------------------------------|-------------------------------|------------------------------------------------------|------------------------------------------|------------|-------------------|
| True value | 15000 | - | log(1.88*10^{11})=25.960 | - | - | - |
| 0 | 15338 | 1067 | 25.952 | 0.066 | -0.808 | 53.4 |
| 1 | 15990 | 1775 | 25.928 | 0.174 | -0.474 | 84.0 |
| 2 | 15497 | 1106 | 25.945 | 0.066 | -0.810 | 13.6 |
| 3 | 14283 | 442 | 25.997 | 0.048 | -0.650 | 44.1 |
| 4 | 12666 | 459 | 26.138 | 0.059 | -0.785 | 42.0 |
| 5 | 15228 | 685 | 25.951 | 0.039 | -0.841 | 36.1 |
| 6 | 15786 | 1633 | 25.948 | 0.103 | -0.528 | 77.0 |
| 7 | 15285 | 707 | 25.950 | 0.041 | -0.832 | 58.9 |
| 8 | 14483 | 359 | 25.986 | 0.032 | -0.731 | 31.1 |
| 9 | 13340 | 388 | 26.070 | 0.039 | -0.856 | 28.8 |
| 10 | 16085 | 1863 | 26.106 | 0.280 | 0.986 | 42.9 |
| 11 | 15284 | 843 | 25.982 | 0.068 | 0.831 | 51.5 |
Table 4: True values, specified prior types and parameters, and posterior means and standard deviations for the experiment with the non-simplified wind profile.

| Variable name                  | True value | Prior type | Mean or minimum of the prior | Std or maximum of the prior | Posterior mean | Posterior std |
|--------------------------------|------------|------------|-----------------------------|-----------------------------|----------------|---------------|
| Column height (m)              | 20000      | Gaussian   | 19000                       | 3000                        | 20149          | 319           |
| Log(mass in kg)                | 26.94      | Gaussian   | 26.71                       | 0.5                         | 26.94          | 0.03          |
| Wind direction at 5 km (°)     | 25         | Uniform    | 0                           | 50                          | 24.82          | 1.85          |
| Wind direction at 14 km (°)    | 65         | Uniform    | 30                          | 80                          | 66.00          | 3.87          |
| Wind direction at 18 km (°)    | 85         | Uniform    | 60                          | 130                         | 79.26          | 4.42          |
| Wind speed at 5 km (m/s)       | 20         | Uniform    | 0                           | 50                          | 20.28          | 1.82          |
| Wind speed at 14 km (m/s)      | 60         | Gaussian   | 55                          | 10                          | 56.97          | 6.11          |
| Wind speed at 18 km (m/s)      | 50         | Uniform    | 0                           | 100                         | 48.3           | 8.34          |
Table 5: Priors and posterior means and standard deviations from applying the algorithm to the 2011 Kirishima-Shinmoedake eruption tephra mass per unit area dataset. Priors of column height, total eruption mass, and median and standard deviation of grain size distribution are referenced and inferred from (Shimbori and Fukui 2012; Nakada et al. 2013; Miyabuchi et al. 2013; Maeno et al. 2014; White et al. 2017). Priors of α/β ratio, diffusion coefficient, fall time threshold, and pumice and lithic densities are specified as commonly adopted ranges or maximum ranges possible.

| Variable name                      | Prior type | Prior mean/prior min | Prior std/prior max | Posterior mean | Posterior std |
|------------------------------------|------------|----------------------|---------------------|----------------|---------------|
| Column height (above vent; m)      | Gaussian   | 8000.00              | 1500.00             | 7372.00        | 538.00        |
| Log(mass in kg)                    | Gaussian   | 24.73 (54.96*10^9 kg) | 2.30                | 22.93 (9.14*10^9 kg) | 0.05          |
| Alpha/Beta                         | Uniform    | 0.01                 | 4.00                | 0.401          | 0.076         |
| Median grain size (φ)              | Gaussian   | -0.25                | 1.00                | 2.820          | 0.210         |
| Std of grain size (φ)              | Uniform    | 1.00                 | 4.00                | 2.040          | 0.240         |
| Diffusion coefficient (m^2/s)      | Uniform    | 1000.00              | 5000.00             | 2256.00        | 650.00        |
| Fall time threshold (s)            | Uniform    | 0.00                 | 6000.00             | 411.00         | 71.06         |
| Lithic density (kg/m^3)            | Gaussian   | 2500.00              | 300.00              | 2485.89        | 225.34        |
| Pumice density (kg/m^3)            | Gaussian   | 1000.00              | 250.00              | 985.47         | 279.78        |
| Maximum wind speed (m/s)           | Gaussian   | 80.00                | 15.00               | 65.62          | 5.76          |
| Wind direction (°)                 | Gaussian   | 135.00               | 20.00               | 124.60         | 0.33          |
Figures

Figure 1: a: mass per unit area distribution used for the validation of Experiments # 0-11 and sample site locations. White (larger), turquoise, yellow, and pink dots are the sample site locations used for Experiments # 0-4, 5-9, 10, and 11, respectively; b: mass per unit area distribution used for the experiment with the non-simplified wind profile. Small white points correspond to sample site locations. Mass per unit area distributions in a and b are in different resolutions. This difference is only for easier visualization (reducing the number of grid points to be plotted in b), and would not affect any arguments or conclusions from this work.
Figure 2: Selected posterior distributions of column height and log-scaled eruption mass. Results from Experiments # 0, 1, 2, and 4 are shown in a (column height) and c (log-scaled eruption mass), and results from Experiments # 5, 6, 7, and 9 are shown in b (column height) and d (log-scaled eruption mass). The blue dashed lines mark the true values of column height and log-transformed eruption mass used to generate the observation data. The red solid lines correspond to prior distributions assumed for all experiments except for Experiments # 4 and 9, and their priors are denoted as red dashed lines.
Figure 3: Selected sampled posterior distributions of column height and log-transformed total eruption mass in 2D. Dashed lines mark true values of column height and log-transformed total eruption mass. a: posterior distributions from Experiments # 0 (red; reference experiment) and 4 (blue; experiment with incorrect priors); b: posterior distributions from Experiments # 5 (red; experiment with 30 observations) and 9 (blue; experiment with 30 observations and incorrect priors). c and d display posterior distributions from Experiments # 10 and 11 (experiments with different sample site locations), respectively.
Figure 4: Simulated data from TEPHRA2 using posterior means as ESPs and wind conditions plotted against observation data of the tephra deposit from the 2011 Kirishima-Shinmoedake eruption under log-scale.

References

Anderson, K. R., Johanson, I. A., Patrick, M. R., Gu, M., Segall, P., Poland, M. P., Montgomery-Brown, E. K., and Miklius, A. (2019). Magma reservoir failure and the onset of caldera collapse at kilauea volcano in 2018. Science, 366(6470).

Andrieu, C., De Freitas, N., Doucet, A., and Jordan, M. I. (2003). An introduction to mcmc for
machine learning. *Machine learning*, 50(1-2):5–43.

Armienti, P., Macedonio, G., and Pareschi, M. (1988). A numerical model for simulation of tephra transport and deposition: Applications to may 18, 1980, mount st. helens eruption. *Journal of Geophysical Research: Solid Earth*, 93(B6):6463–6476.

Bear-Crozier, A., Pouget, S., Bursik, M., Jansons, E., Denman, J., Tupper, A., and Rustowicz, R. (2020). Automated detection and measurement of volcanic cloud growth: towards a robust estimate of mass flux, mass loading and eruption duration. *Natural Hazards*, 101(1):1–38.

Bevilacqua, A., Bursik, M., Patra, A., Bruce Pitman, E., Yang, Q., Sangani, R., and Kobs-Nawotniak, S. (2018). Late quaternary eruption record and probability of future volcanic eruptions in the long valley volcanic region (ca, usa). *Journal of Geophysical Research: Solid Earth*.

Bevilacqua, A., Isaia, R., Neri, A., Vitale, S., Aspinall, W. P., Bisson, M., Flandoli, F., Baxter, P. J., Bertagnini, A., Ongaro, T. E., et al. (2015). Quantifying volcanic hazard at campi flegrei caldera (italy) with uncertainty assessment: 1. vent opening maps. *Journal of Geophysical Research: Solid Earth*, 120(4):2309–2329.

Biass, S., Bonadonna, C., Connor, L., and Connor, C. (2016). Tephrapro: a matlab package for probabilistic hazard assessments of tephra fallout. *Journal of Applied Volcanology*, 5(1):10.

Biass, S., Bonadonna, C., and Houghton, B. F. (2019). A step-by-step evaluation of empirical methods to quantify eruption source parameters from tephra-fall deposits. *Journal of Applied Volcanology*, 8(1):1.

Biass, S., Frischknecht, C., and Bonadonna, C. (2012). A fast gis-based risk assessment for tephra fallout: the example of cotopaxi volcano, ecuador-part ii: vulnerability and risk assessment. *Natural hazards*, 64(1):615–639.

Biass, S., Todde, A., Cioni, R., Pistolesi, M., Geshi, N., and Bonadonna, C. (2017). Potential impacts of tephra fallout from a large-scale explosive eruption at sakurajima volcano, japan. *Bulletin of Volcanology*, 79(10):73.

Bonadonna, C., Biass, S., and Costa, A. (2015). Physical characterization of explosive volcanic eruptions based on tephra deposits: propagation of uncertainties and sensitivity analysis. *Journal of Volcanology and Geothermal Research*, 296:80–100.
Bonadonna, C., Connor, C. B., Houghton, B., Connor, L., Byrne, M., Laing, A., and Hincks, T. (2005). Probabilistic modeling of tephra dispersal: Hazard assessment of a multiphase rhyolitic eruption at tarawera, new zealand. *Journal of Geophysical Research: Solid Earth*, 110(B3).

Bonadonna, C. and Costa, A. (2012). Estimating the volume of tephra deposits: a new simple strategy. *Geology*, 40(5):415–418.

Bonadonna, C., Ernst, G., and Sparks, R. (1998). Thickness variations and volume estimates of tephra fall deposits: the importance of particle reynolds number. *Journal of Volcanology and Geothermal Research*, 81(3-4):173–187.

Bonadonna, C. and Houghton, B. (2005). Total grain-size distribution and volume of tephra-fall deposits. *Bulletin of Volcanology*, 67(5):441–456.

Bonasia, R., Macedonio, G., Costa, A., Mele, D., and Sulpizio, R. (2010). Numerical inversion and analysis of tephra fallout deposits from the 472 ad sub-plinian eruption at vesuvius (italy) through a new best-fit procedure. *Journal of Volcanology and Geothermal Research*, 189(3-4):238–246.

Bursik, M., Carey, S., and Sparks, R. (1992a). A gravity current model for the may 18, 1980 mount st. helens plume. *Geophysical research letters*, 19(16):1663–1666.

Bursik, M., Sparks, R., Gilbert, J., and Carey, S. (1992b). Sedimentation of tephra by volcanic plumes: I. theory and its comparison with a study of the fogo a plinian deposit, sao miguel (azores). *Bulletin of Volcanology*, 54(4):329–344.

Carey, S. and Sparks, R. (1986). Quantitative models of the fallout and dispersal of tephra from volcanic eruption columns. *Bulletin of Volcanology*, 48(2-3):109–125.

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335.

Connor, C. B., Connor, L. J., Bonadonna, C., Luhr, J., Savov, I., and Navarro-Ochoa, C. (2019). Modelling tephra thickness and particle size distribution of the 1913 eruption of volcán de colima, mexico. In *Volcán de Colima*, pages 81–110. Springer.

Connor, L. J. and Connor, C. B. (2006). Inversion is the key to dispersion: understanding eruption dynamics by inverting tephra fallout.
Courtland, L., Connor, C., Connor, L., and Bonadonna, C. (2012). Introducing geoscience students to numerical modeling of volcanic hazards: The example of tephra2 on vhub. org. *Numeracy*, 5(2).

Covey, J., Horwell, C. J., Rachmawati, L., Ogawa, R., Martin-del Pozzo, A. L., Armienta, M. A., Nugroho, F., and Dominelli, L. (2019). Factors motivating the use of respiratory protection against volcanic ashfall: A comparative analysis of communities in Japan, Indonesia and Mexico. *International Journal of Disaster Risk Reduction*, 35:101066.

Engwell, S., Aspinall, W., and Sparks, R. (2015). An objective method for the production of isopach maps and implications for the estimation of tephra deposit volumes and their uncertainties. *Bulletin of Volcanology*, 77(7):61.

Fontijn, K., Ernst, G. G., Bonadonna, C., Elburg, M. A., Mbede, E., and Jacobs, P. (2011). The 4-ka rungwe pumice (South-Western Tanzania): a wind-still plinian eruption. *Bulletin of Volcanology*, 73(9):1353–1368.

Green, R. M., Bebbington, M. S., Jones, G., Cronin, S. J., and Turner, M. B. (2016). Estimation of tephra volumes from sparse and incompletely observed deposit thicknesses. *Bulletin of Volcanology*, 78(4):25.

Hashimoto, A., Shimbori, T., and Fukui, K. (2012). Tephra fall simulation for the eruptions at Mt. Shinmoe-Dake during 26–27 January 2011 with JManhm. *Sola*, 8:37–40.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications.

Hildreth, W. (2004). Volcanological perspectives on Long Valley, Mammoth Mountain, and Mono Craters: several contiguous but discrete systems. *Journal of Volcanology and Geothermal Research*, 136(3-4):169–198.

Jenkins, S., Magill, C., McAneney, J., and Blong, R. (2012). Regional ash fall hazard i: a probabilistic assessment methodology. *Bulletin of Volcanology*, 74(7):1699–1712.

Jenkins, S., Magill, C., McAneney, J., and Hurst, T. (2008). Multistage volcanic events: tephra hazard simulations for the Okataina volcanic center, New Zealand. *Journal of Geophysical Research: Earth Surface*, 113(F4).
Jenkins, S. F., Goldstein, H., Bebbington, M., Sparks, R., and Koyaguchi, T. (2019). Forecasting explosion repose intervals with a non-parametric bayesian survival model: Application to sakurajima volcano, japan. *Journal of Volcanology and Geothermal Research*, 381:44–56.

Johnston, E., Phillips, J., Bonadonna, C., and Watson, I. (2012). Reconstructing the tephra dispersal pattern from the bronze age eruption of santorini using an advection–diffusion model. *Bulletin of Volcanology*, 74(6):1485–1507.

Kaipio, J. and Somersalo, E. (2006). *Statistical and computational inverse problems*, volume 160. Springer Science & Business Media.

Kawabata, E., Bebbington, M. S., Cronin, S. J., and Wang, T. (2013). Modeling thickness variability in tephra deposition. *Bulletin of Volcanology*, 75(8):738.

Kawabata, E., Cronin, S., Bebbington, M., Moufti, M., El-Masry, N., and Wang, T. (2015). Identifying multiple eruption phases from a compound tephra blanket: an example of the ad1256 al-madinah eruption, saudi arabia. *Bulletin of Volcanology*, 77(1):6.

Klawonn, M., Frazer, L., Wolfe, C., Houghton, B., and Rosenberg, M. (2014). Constraining particle size-dependent plume sedimentation from the 17 june 1996 eruption of ruapehu volcano, new zealand, using geophysical inversions. *Journal of Geophysical Research: Solid Earth*, 119(3):1749–1763.

Klawonn, M., Wolfe, C. J., Frazer, L. N., and Houghton, B. F. (2012). Novel inversion approach to constrain plume sedimentation from tephra deposit data: Application to the 17 june 1996 eruption of ruapehu volcano, new zealand. *Journal of Geophysical Research: Solid Earth*, 117(B5).

Koyaguchi, T., Anderson, K., and Kozono, T. (2017). Bayesian estimation of analytical conduit-flow model parameters from magma discharge rate observed during explosive eruptions. In *AGU Fall Meeting Abstracts*.

Lev, E., Conroy, C. J., Birnbaum, J., and Zhan, Y. (2019). Combining probabilistic methods and deterministic models to estimate uncertainties in lava flow parameters inferred from observations. *AGUFM, 2019:V34C–02*.

Liang, C. and Dunham, E. M. (2020). Lava lake sloshing modes during the 2018 kilauea volcano eruption probe magma reservoir storativity. *Earth and Planetary Science Letters*, 535:116110.
Madankan, R., Pouget, S., Singla, P., Bursik, M., Dehn, J., Jones, M., Patra, A., Pavolonis, M., Pitman, E. B., Singh, T., et al. (2014). Computation of probabilistic hazard maps and source parameter estimation for volcanic ash transport and dispersion. *Journal of Computational Physics*, 271:39–59.

Maeno, F., Nagai, M., Nakada, S., Burden, R. E., Engwell, S., Suzuki, Y., and Kaneko, T. (2014). Constraining tephra dispersion and deposition from three subplinian explosions in 2011 at shinmoedake volcano, kyushu, japan. *Bulletin of Volcanology*, 76(6):823.

Maeno, F., Nakada, S., Nagai, M., and Kozono, T. (2013). Ballistic ejecta and eruption condition of the vulcanian explosion of shinmoedake volcano, kyushu, japan on 1 february, 2011. *Earth, Planets and Space*, 65(6):12.

Magill, C., Mannen, K., Connor, L., Bonadonna, C., and Connor, C. (2015). Simulating a multi-phase tephra fall event: inversion modelling for the 1707 hoei eruption of mount fuji, japan. *Bulletin of Volcanology*, 77(9):81.

Mastin, L. G., Guffanti, M., Servranckx, R., Webley, P., Barsotti, S., Dean, K., Durant, A., Ewert, J. W., Neri, A., Rose, W. I., et al. (2009). 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(1-2):10–21.

Miyabuchi, Y., Hanada, D., Niimi, H., and Kobayashi, T. (2013). Stratigraphy, grain-size and component characteristics of the 2011 shinmoedake eruption deposits, kirishima volcano, japan. *Journal of volcanology and geothermal research*, 258:31–46.

Moiseenko, K. B. and Malik, N. A. (2019). Linear inverse problem for inferring eruption source parameters from sparse ash deposit data as viewed from an atmospheric dispersion modeling perspective. *Bulletin of Volcanology*, 81(3):19.

Nakada, S., Nagai, M., Kaneko, T., Suzuki, Y., and Maeno, F. (2013). The outline of the 2011 eruption at shinmoedake (kirishima), japan. *Earth, planets and space*, 65(6):1.

Neri, A., Aspinall, W. P., Cioni, R., Bertagnini, A., Baxter, P. J., Zuccaro, G., Andronico, D., Barsotti, S., Cole, P. D., Ongaro, T. E., et al. (2008). Developing an event tree for probabilistic hazard and risk assessment at vesuvius. *Journal of volcanology and geothermal research*, 178(3):397–415.
Newhall, C. G. and Self, S. (1982). The volcanic explosivity index (VEI) an estimate of explosive magnitude for historical volcanism. *Journal of Geophysical Research: Oceans*, 87(C2):1231–1238.

Pieri, D. C. and Baloga, S. M. (1986). Eruption rate, area, and length relationships for some Hawaiian lava flows. *Journal of Volcanology and Geothermal Research*, 30(1-2):29–45.

Poret, M., Corradini, S., Merucci, L., Costa, A., Andronico, D., Montopoli, M., Vulpiani, G., Scolo, S., and Freret-Lorgeril, V. (2017). Tephra dispersal and fallout reconstructed integrating field, ground-based and satellite-based data: Application to the 23rd November 2013 Etna paroxism. In *AGU Fall Meeting Abstracts*.

Pouget, S., Bursik, M., Webley, P., Dehn, J., and Pavolonis, M. (2013). Estimation of eruption source parameters from umbrella cloud or downwind plume growth rate. *Journal of Volcanology and Geothermal Research*, 258:100–112.

Scarpati, C., Cole, P., and Perrotta, A. (1993). The Neapolitan yellow tuff—a large volume multiphase eruption from Campi Flegrei, southern Italy. *Bulletin of Volcanology*, 55(5):343–356.

Scolo, S., Tarantola, S., Bonadonna, C., Coltelli, M., and Saltelli, A. (2008). Sensitivity analysis and uncertainty estimation for tephra dispersal models. *Journal of Geophysical Research: Solid Earth*, 113(B6).

Shimbori, T. and Fukui, K. (2012). Time variation of the eruption cloud echo height from Shinmoe-dake volcano in 2011 observed by Tanegashima and Fukuoka weather radars. Part II. *Rep. Coord. Comm. Predict. Volcan. Erup.*, 109:173–178.

Sparks, R. and Young, S. (2002). The eruption of Soufriere Hills volcano, Montserrat (1995–1999): overview of scientific results. *Geological Society, London, Memoirs*, 21(1):45–69.

Stohl, A., Prata, A., Eckhardt, S., Clarisse, L., Durant, A., Henne, S., Kristiansen, N. I., Minikin, A., Schumann, U., Seibert, P., et al. (2011). 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(9):4333–4351.

Suzuki, T. et al. (1983). A theoretical model for dispersion of tephra. *Arc volcanism: physics and tectonics*, 95:113.
Takarada, S. (2017). The volcanic hazards assessment support system for the online hazard assessment and risk mitigation of quaternary volcanoes in the world. *Frontiers in Earth Science*, 5:102.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*, volume 89. siam.

Volentik, A. C., Bonadonna, C., Connor, C. B., Connor, L. J., and Rosi, M. (2010). Modeling tephra dispersal in absence of wind: Insights from the climactic phase of the 2450 bp plinian eruption of pululagua volcano (ecuador). *Journal of Volcanology and Geothermal Research*, 193(1-2):117–136.

Wang, T., Schofield, M., Bebbington, M., and Kiyosugi, K. (2020). Bayesian modelling of marked point processes with incomplete records: volcanic eruptions. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 69(1):109–130.

White, J., Connor, C., Connor, L., and Hasenaka, T. (2017). Efficient inversion and uncertainty quantification of a tephra fallout model. *Journal of Geophysical Research: Solid Earth*, 122(1):281–294.

Wild, A., Wilson, T., Bebbington, M., Cole, J., and Craig, H. (2019). Probabilistic volcanic impact assessment and cost-benefit analysis on network infrastructure for secondary evacuation of farm livestock: A case study from the dairy industry, taranaki, new zealand. *Journal of Volcanology and Geothermal Research*, 387:106670.

Yang, Q. and Bursik, M. (2016). A new interpolation method to model thickness, isopachs, extent, and volume of tephra fall deposits. *Bulletin of Volcanology*, 78(10):68.

Yang, Q., Bursik, M., and Pitman, E. B. (2019). A new method to identify the source vent location of tephra fall deposits: development, testing, and application to key quaternary eruptions of western north america. *Bulletin of Volcanology*, 81(9):51.

Yang, Q., Pitman, E. B., Bursik, M. I., and Jenkins, S. (2020). Tephra inversion with tephra2 and the metropolis-hastings algorithm.