Sensitivity analysis and uncertainty estimation for tephra dispersal models

Scollo, Simona; Tarantola, Stefano; Bonadonna, Costanza; Coltelli, Mauro; Saltelli, Andrea

Sensitivity analysis and uncertainty estimation for tephra dispersal models

Simona Scollo,[1] Stefano Tarantola,[2] Costanza Bonadonna,[3] Mauro Coltelli,[1] and Andrea Saltelli[2]

Received 18 November 2006; revised 23 November 2007; accepted 16 January 2008; published 10 June 2008.

[1] Sensitivity analysis and uncertainty estimation are crucial to the validation and calibration of numerical models. In this paper we present the application of sensitivity analyses, parameter estimations and Monte-Carlo uncertainty analyses on TEPHRA, an advection-diffusion model for the description of particle dispersion and sedimentation from volcanic plumes. The model and the related sensitivity analysis are tested on two sub-plinian eruptions: the 22 July 1998 eruption of Etna volcano (Italy) and the 17 June 1996 eruption of Ruapehu volcano (New Zealand). Sensitivity analyses are key to (1) constrain crucial eruption parameters (e.g., total erupted mass) (2) reduce the number of variables by eliminating non-influential parameters (e.g., particle density) and (3) investigate the interactions among all input parameters (plume height, total grain-size distribution, diffusion coefficient, fall-time threshold and mass-distribution parameter). For the two test cases, we found that the total erupted mass significantly affects the model outputs and, therefore, it can be accurately estimated from field data of the fallout deposit, whereas the particle density can be fixed at its nominal value because it has negligible effects on the model predictions.

Citation: Scollo, S., S. Tarantola, C. Bonadonna, M. Coltelli, and A. Saltelli (2008), Sensitivity analysis and uncertainty estimation for tephra dispersal models, *J. Geophys. Res.*, *113*, B06202, doi:10.1029/2006JB004864.

1. Introduction

[2] Recent advances in particle dispersal models have significantly improved our capability to describe the plume dynamics and to compute the accumulation of tephra deposits [*Bursik et al.*, 1992a, 1992b; *Searcy et al.*, 1998; *Koyaguchi and Ohno*, 2001a, 2001b; *Bonadonna et al.*, 2005a; *Costa et al.*, 2006]. Model calibration and validation are usually carried out by comparing computed and observed data of well known eruptions [e.g., *Bonadonna et al.*, 2002; *Scollo et al.*, 2007]. Nevertheless, this process is often complicated by the lack of comprehensive data sets (including information on column height and field data collected in both proximal and distal area) and by the use of various assumptions (e.g., constant mass eruption rate during the eruption) required to parameterize volcanic processes.

[3] Another complication is given by the fact that some input parameters of tephra dispersal models are typically considered as "true values" even though they are affected by uncertainty. These uncertainties might significantly affect the model outputs. For example, the total erupted mass is not easy to determine because of the non-linearity of the function that describes the thickness or mass variation with the distance from the eruptive vent. Several techniques for extrapolation of tephra deposit thickness/mass with distance from the vent have been suggested [*Pyle*, 1989; *Fierstein and Nathenson*, 1992; *Froggatt*, 1982]. However, numerical simulations have shown that curve-fitting techniques can be misleading when applied to poor data sets and may generate large errors depending on the extrapolation method considered because sedimentation processes in proximal and distal areas are characterized by different fallout regimes [*Bonadonna and Houghton*, 2005]. As a result, the total erupted mass derived from even abundant field data might be affected by large uncertainties.

[4] In this paper, we present an iterative approach based on the sequential combination of sensitivity analysis [*Saltelli et al.*, 2008], parameter estimation procedure [*Beven and Binley*, 1992], and Monte Carlo-based uncertainty analysis, applied to the advection-diffusion model TEPHRA [*Bonadonna et al.*, 2005a]. For this model, the input data necessary to simulate the dispersal of volcanic clouds are: plume height, total mass, total grain-size distribution, density of lithics and pumices, diffusion coefficient, fall-time threshold and plume ratio (a factor describing the mass distribution in the plume). Values of these data are not often available during an eruption and increase the uncertainty in forecasting the plume dispersal. Hence we use sensitivity analysis to identify the input parameters that can be fixed in order to simplify the model and those that have a chance of being estimated. Subsequently, we estimate input parameters through an uncertainty estimation proce-

[1] Istituto Nazionale di Geofisica e Vulcanologia, Catania, Italy.
[2] Joint Research Centre of the European Commission, Ispra, Italy.
[3] Centre d'Etude des Risques Géologiques (CERG), Université de Genève, Geneva, Switzerland.

0148-0227/08/2006JB004864$09.00

dure between model predictions and field data. Finally, we use a Monte Carlo procedure to obtain the uncertainty distribution of the total mass per unit area deposited on the ground as predicted from the model. This approach has been tested on two sub-plinian eruptions: the 22 July 1998 eruption of Etna volcano (Italy), and the 17 June 1996 eruption of Ruapehu volcano (New Zealand), where comprehensive data sets are available (i.e., information on column height, wind profile, isomass maps covering both proximal and distal areas and based on field data collected within a few days after the eruption) [*Andronico et al.*, 1999; *Bonadonna et al.*, 2005b]. Both eruptions had a well characterized fallout pattern representative of typical explosive eruption styles of basaltic (Etna) and andesitic (Ruapehu) magmas. As examples we used the well studied basaltic eruption of Etna (22 July 1998), and the andesitic eruption of Ruapehu (17 June 1996) that has a strongly wind-skewed eruption cloud. These eruptions were characterized by two typical plume styles: strong plume (i.e., with a vertical convective phase) and weak plume (i.e., bent over by the wind). In particular, the data sets include a large number of locations sampled both for mass/area and grain-size (35 and 114 samples for Etna and Ruapehu respectively) and an observed height of the eruption column, strategic information often not available for small and medium size explosive eruptions.

2. The Proposed Approach

[5] The aim of sensitivity analysis is to quantify the uncertainty contribution of input data and parameters (all indicated here as input factors) to the overall uncertainty in the model prediction. Whereas, uncertainty analysis propagates the uncertainty of the input factors to the model outputs [*Helton*, 1993]. We review the components of our approach in the following sections.

2.1. Sensitivity Analysis

[6] Formal approaches of sensitivity analysis have been applied in various fields where models are employed, ranging from physics to economics [*Saltelli et al.*, 2000] and several methods are available, ranging from differential to Monte Carlo analysis, from response surface methodology to regression and correlation techniques [see *Helton* [1993] and *Saltelli et al.* [2000] for a review]. The simplest approach is when the input factors are varied "One At a Time", better known as the OAT method. The OAT method is easy to implement, computationally cheap, and has been frequently applied to tephra dispersal models to extrapolate the main features of sedimentation processes [*Macedonio et al.*, 1988; *Hurst and Turner*, 1999; *Bonadonna et al.*, 2002; *Pfeiffer et al.*, 2005; *Scollo et al.*, 2007]. However, the OAT approach suffers a number of limitations which can lead to inaccurate results especially when the numerical models are highly non-linear, such as models of tephra dispersal used for hazard studies.

[7] In this paper we focus on global sensitivity analysis techniques, which allow the simultaneous exploration of the space of the uncertain inputs over the whole domain of uncertainty. In particular, we will adopt the class of variance-based techniques, which are based on the decomposition of the prediction variance into components that quantify the importance of the single model input factors. Therefore let us consider a mathematical model f whose output variable Y is a nonlinear deterministic function of its k input factors:

$$Y = f(x_1, x_2, \ldots, x_k) = f(x) \tag{1}$$

where $x = [x_1, x_2, \ldots x_k]$ are the uncertain input factors of the model in the domain $\Omega \in [0;1]^k$ (e.g., diffusion coefficient, grain-size or particle density) and f is the model, i.e., the set of equations that links the input factors x to the model prediction Y (e.g., the total mass of tephra accumulated on the ground per unit area). We assume that the input factors are independent and we treat them as if they were random variables. We characterize the uncertainty of Y by its variance. The expected value $E(Y)$ and the variance $V(Y)$ are given by:

$$E(Y) = \int_{\Omega} f(x_1, x_2, \ldots x_k) dx \tag{2}$$

and

$$V(Y) = \int_{\Omega} [f(x_1, x_2, \ldots x_k) - E(Y)]^2 dx \tag{3}$$

while the expected value of Y conditional on x_i is given by:

$$E(Y|x_i) = \int_{\Omega} f(x_1, x_2, \ldots, x_k) \frac{dx}{dx_i} \tag{4}$$

[8] The variance of this conditional expectation is $V[E(Y|x_i)]$, and is called the "top marginal variance" (or the explained variance) [*Jansen et al.*, 1994]. Such conditional variance is a measure of importance of factor x_i [*Saltelli et al.*, 2000]. $V[E(Y|x_i)]$ is the expected reduction of the variance of Y in the case that x_i would become fully known, whereas the other input factors remain as uncertain as before. The top marginal variance is linked to the unconditional variance $V(Y)$ via the identity:

$$V(Y) = V[E(Y|x_i)] + E[V(Y|x_i)] \tag{5}$$

[9] The main effect (indicated by the first order sensitivity index S_i) is the top marginal variance of factor x_i divided by the total variance $V(Y)$:

$$S_i = \frac{V[E(Y|x_i)]}{V(Y)} \tag{6}$$

[10] Higher order sensitivity indices quantify the sensitivity of the model predictions to interactions among pairs, or groups, of input factors and are evaluated as:

$$S_{ij} = \left\{ V\left[E\left(Y|x_i;x_j\right)\right] - V[E(Y|x_i)] - V\left[E\left(Y|x_j\right)\right] \right\} / V(Y) \tag{7}$$

[11] The calculation of all sensitivity indices would require a relevant amount of computational work (in a

model with k input factors, there are 2^k-1 sensitivity indices) and, due to the large number of combinations for higher order sensitivity indices, the analysis is often stopped at the indices of the second order.

[12] In our analysis we also use a compact characterization of the factor importance, given by the total sensitivity index S_{Ti}, which embeds in one single number both the main effect and the higher order effects for a given input factor [*Homma and Saltelli*, 1996]. The S_{Ti} for the factor x_i is defined as the sum of all the sensitivity indices (of any order) that include x_i. We can estimate the total sensitivity index without needing to estimate each order index through the formula:

$$S_{Ti} = \frac{E[V(Y|x_{-i})]}{V(Y)} \tag{8}$$

where x_{-i} represents the totality of input factors excluding x_i.

[13] The procedure to calculate S_i and S_{Ti} is illustrated by *Saltelli* [2002]. Here, we just remind the reader that we need to generate a random sample of N points for the input factors and execute the simulation model N times:

$$Y_j = F\left(x_{j_1}, x_{j_2}, \ldots, x_{jk}\right)\ j = 1 \ldots N. \tag{9}$$

[14] Using the results of the N simulations (N is generally of the order of a few thousands), the computation of the multidimensional integrals (6) and (8) is carried out efficiently using SIMLAB, a software code designed for uncertainty and sensitivity analysis.

2.2. SIMLAB Software

[15] SIMLAB is a free software package for global uncertainty and sensitivity analysis developed at the Joint Research Centre of Ispra (Italy) and distributed for free at http://simlab.jrc.ec.europa.eu. This software allows the user to specify the distributions for each input factor of a model and generate a sample of elements of a given size N from a distribution set up previously. The sample generation can be made using a variety of methods, i.e., random sampling, quasi-random sampling, replicated Latin Hypercube, classic and extended FAST (Fourier Amplitude Sensitivity Test) and the Morris design. The matrix of values with N rows (each sample point) and k columns (the input factors) used in our study as input for the TEPHRA model has been generated using the Sobol' method [*Sobol'*, 1990]. Then, TEPHRA is run on the N sample points to provide the predictions. Finally, SIMLAB is supplied with the sample points and the corresponding predictions to execute the sensitivity analysis. The sensitivity analysis can be carried out by using different techniques, e.g., factor screening, standardized regression, and variance-based analysis. In our case the first order S_i sensitivity indices and the total sensitivity indices S_{Ti} are obtained using the algorithm proposed by *Saltelli* [2002].

2.3. Uncertainty Estimation Procedure

[16] The starting point of the uncertainty estimation procedure is based on the assumption that any combination of input factors can be considered as a possible simulator of the system [*Beven and Binley*, 1992]. For simplicity, observations are considered to be true in the frame of this work, even though both observations (e.g., mass/area collected at each location), and model predictions (e.g., simulations obtained from tephra dispersal models, on which the estimations of input factors are based) are subject to errors. As a result, the best fit values are determined as the closest values to the collected field data. Several simulations are performed on different combinations of these input factors (that are the same used in the previous sensitivity analysis). A measurement of how well the model simulations fit the measurements is done by the following best fit function:

$$f_b = \exp\left(-\sum_{l=1}^{Nobs}\left(Y(x_j) - O_l\right)^2 / std(Y - O_l)\right) \tag{10}$$

where x_j is the j-th element of the sample of the model input factors, O_l is the vector of observations, $Y(x_j)$ is the j-th model run and $std(Y - O_l)$ is the standard deviation of the model errors with respect to the observations.

[17] The best descriptor of the system is the input factor set corresponding to the highest value of equation (10).

[18] The final step is a Monte Carlo-based uncertainty analysis that selects probabilistically the input factor set corresponding to the highest values of the equation (10) and then uses the model results to evaluate the uncertainty. In fact, a new sampling of these input factors using Monte Carlo method will generate new model outputs, on which the mean value and the relative uncertainty (standard deviation/mean) will be evaluated. The method is computationally intensive, because a large number of runs are needed. However, the computational time has become feasible using parallel processing techniques implemented in the TEPHRA model [*Bonadonna et al.*, 2005a].

3. TEPHRA Model

[19] TEPHRA is a two dimensional advection-diffusion model that describes sedimentation processes of particles from volcanic plumes [*Bonadonna et al.*, 2005a]. Such a model is the result of the integration of several theories and approaches including a grain-size dependent diffusion [*Suzuki*, 1983], particle density variations [*Bonadonna and Phillips*, 2003], a stratified atmosphere [*Macedonio et al.*, 1988; *Connor et al.*, 2001; *Bonadonna et al.*, 2002], and terminal settling velocity that accounts for the variation of particle Reynolds number [*Bonadonna et al.*, 1998]. The model has been recently applied to assess the hazard of tephra dispersal at Tarawera volcano, New Zealand [*Bonadonna et al.*, 2005a]. One of the advantages of TEPHRA is the parallelization of the code that allows for a large number of fast simulations, critical in hazard assessments. A comprehensive evaluation of the tephra hazard at Tarawera volcano was hence obtained applying a probabilistic approach to both inputs (i.e., sampling of probability density functions accurately describing input parameters) and outputs (i.e., probability maps and hazard curves).

[20] The model solves the mass conservation equation:

$$\frac{\partial C_j}{\partial t} + w_x\frac{\partial C_j}{\partial x} + w_y\frac{\partial C_j}{\partial y} - V_{j,i}\frac{\partial C_j}{\partial z} = K_x\frac{\partial^2 C_j}{\partial x^2} + K_y\frac{\partial^2 C_j}{\partial y^2} + K_z\frac{\partial^2 C_j}{\partial z^2} \tag{11}$$

where x, y, z are the spatial coordinates (m), C_j is the particle concentration (kg m^{-3}), t is the time (s), w_x and w_y describe

the horizontal wind speed field (m s^{-1}), $V_{j,i}$ is the settling velocity of volcanic particles (m s^{-1}) and K_x, K_y, K_z are the three components of the diffusion coefficient (m^2 s^{-1}). Vertical diffusion coefficient K_z is assumed to be negligible and horizontal diffusion is considered constant and isotropic ($K = K_x = K_y$).

[21] Volcanic particles with dimension j are released instantaneously from a point source i situated along a hypothetical vertical line (the eruptive column) centered on the volcanic vent. Particles are characterized by a horizontal motion that depends on wind speed, atmospheric turbulence and plume spreading, and by a vertical motion, controlled by terminal settling velocities of volcanic particles $V_{j,i}$. Terminal settling velocity is function of Reynolds' number and is calculated, assuming particles of spherical shape, using the analytical expression of *Kunii and Levenspiel* [1969]. In addition, aggregation processes are neglected for simplicity. In fact, even though aggregation processes significantly affect sedimentation of fine particles [*Bonadonna et al.*, 2002; *Carey and Sigurdsson*, 1982; *Brazier et al.*, 1982], they can be neglected when describing sedimentation of tephra characterized by coarse grain-size distributions as shown by numerical simulations [e.g., *Bonadonna and Phillips*, 2003]. For both Etna and Ruapehu eruptions the sedimentation was not significantly affected by aggregation processes.

[22] The computation domain is divided into N_l layers on which the wind speed and the atmospheric diffusion coefficient are assumed constant. Inside each horizontal layer, the transport is described by a Gaussian distribution that expands due to the turbulent diffusion, and translates due to the wind for the time δt_j spent by volcanic particles inside a given layer. δt_j is function of the layer thickness δz and the terminal settling velocity $V_{j,i}$ of the particle and it is equal to $\delta z/V_{j,i}$. The center of the Gaussian distribution is shifted layer by layer following $\delta x_j = w_x\ \delta t_j$ and $\delta y_j = w_y\ \delta t_j$, on the ($x$, y). The width of the Gaussian distribution is controlled by the atmospheric diffusion.

[23] The total mass accumulated per unit area (kg m^{-2}) is the sum of the contribution of each point source i for each particle grain-size class j:

$$M(x,y) = \sum_{i}^{H} \sum_{\Phi_{\min}}^{\Phi_{\max}} \frac{M_{ij}^{0}}{2\pi\sigma_{ij}^{2}} \exp\left\{ -\frac{\left(x-\bar{x}_{ij}\right)^2+\left(y-\bar{y}_{ij}\right)^2}{2\sigma_{ij}^{2}} \right\} \quad (12)$$

where H is the total height of the eruption column, $\Phi_{\min}$ and $\Phi_{\max}$ ($\Phi = -\log_2(d)$ with d particle diameter in mm) are the minimum and maximum diameters of the total grain-size distribution, M_{ij}^0 (kg) is the total mass fraction of a particle size j that falls from the point source i at the height z_i, σ_{ij}^2 is the distribution variance. The variance σ_{ij}^2 is function of both turbulent atmospheric diffusion and plume gravity. For particles having small fall time t_{ij}, the diffusion can be described by a linear function with t (Fick's law):

$$\sigma_{ij}^2 = 2K\left(t_{ij} + t_i'\right) \quad (13)$$

where K (m^2 s^{-1}) is the *Diffusion Coefficient* and t_i'(s) is the horizontal diffusion time in the vertical plume which accounts for its change in width [*Woods et al.*, 1995; *Ernst et al.*, 1996]. The horizontal diffusion time is given by the integration of *Morton et al.* [1956] and *Bonadonna and Phillips* [2003] models and by observations of plume spreading [*Sparks and Wilson*, 1982]:

$$t_i' = \frac{0.0032 z_i^2}{K} \quad (14)$$

[24] For particle fall time of hours (for example, particles with diameter < 1 mm falling from a 30 km high plume), the variance σ_{ij}^2 is described by a power law [*Suzuki*, 1983]:

$$\sigma_{ij}^2 = \frac{4C}{5}\left(t_{ij} + t_i'\right)^{2.5} \quad (15)$$

where C is an empirical constant [*Suzuki*, 1983]. The combination of two diffusion laws can generate a break-in-slope in the thinning trend which, as a result, depends on the *Fall Time Threshold (FTT)*. Large values of the *FTT* mean that the diffusion of more particles can be described by the linear function, resulting in a thick and narrow deposit in the proximal area centered along the dispersal axis.

[25] The distribution of the mass inside the eruption column is described by two different plume models. The first considers a uniform mass distribution along the plume where the top and the bottom of the particle-source area are defined by the *Plume Ratio (PR)*; the second model describes the mass distribution in the plume as a lognormal distribution depending on a geometrical parameter A. In our study we use only the first model in which the *PR* factor is varied. *PR* is a dimensionless parameter that specifies the bottom of the particle-release region as a ratio of the total height. For example, if *PR* is 0.8 and the *Plume Height (H)* is 20 km, particles start being releases at 16 km.

[26] Input factors of TEPHRA that were considered in the following sensitivity analysis are:

[27] • *Plume Height (H)*: maximum height reached by the eruption column. It can be constrained with direct ground-based observations, satellite imagery and/or derived from tephra deposit data [e.g., *Carey and Sparks*, 1986; *Sparks*, 1986; *Wilson and Walker*, 1987; *Holasek and Self*, 1995].

[28] • *Total Mass (TM)*: the total erupted mass can be empirically determined from an empirical power law function of H [*Carey and Sigurdsson*, 1982; *Wilson and Walker*, 1987] or from field data using curve-fitting techniques [*Pyle*, 1989; *Fierstein and Nathenson*, 1992; *Bonadonna and Houghton*, 2005].

[29] • *Total Grain-Size distribution (TGS)*: the grain-size distribution of the whole deposit can be determined using various methods of integrations [see *Bonadonna and Houghton* [2005] for a review]. *TGS* is used in TEPHRA model as a Gaussian distribution characterized by a mode and a standard deviation.

[30] • *Density of Lithics and Pumices (DL, DP)*: the density of volcanic particles released from the column. The density of clasts varies widely from ~500 kg m^{-3} in highly vesicular clasts to ~2700 kg m^{-3} in dense ones. The density of lithics and pumices is typically measured in the laboratory for fragments down to about 2 mm. Density of smaller fragments can be described using the simple parameterization of *Bonadonna and Phillips* [2003], where

density of pumices is assumed to decrease with increasing size for particles having diameter <2 mm (−1 Φ) becoming equal to the lithic density when the particle size is below 0.0078 mm (>7 Φ).

[31] • *Diffusion Coefficient (K)*: this coefficient accounts for atmospheric diffusion and for the gravitational spreading of the volcanic cloud. Consequently, diffusion coefficients of advection-diffusion models are typically larger than standard atmospheric diffusion coefficients [e.g., *Armienti et al.*, 1988; *Bonadonna et al.*, 2002; *Pfeiffer et al.*, 2005].

[32] • *Fall Time Threshold (FTT)*: empirical threshold that defines the transition (expressed in particle-fall time) between the two different laws of particle diffusion used in the TEPHRA model [equations (13) and (15)].

[33] • *Plume Ratio (PR)*: factor describing the mass distribution in the plume.

[34] The model output is the mass per unit area (kg m^{-2}) deposited on grid points around the volcano. *Bonadonna et al.* [2005a] calibrated the TEPHRA model on the pyroclastic deposits produced by A.D. 1315 Kaharoa eruption and 17 June 1996 Ruapehu eruption (New Zealand) by varying one at time *K*, *FTT*, and *PR* factors (OAT method). They calculated a best fit function to evaluate the input factors that give the best agreement between the computed and the field data. Nevertheless, uncertainties of some input factors were not analyzed because they were considered as true values (i.e., the total erupted mass, the total grain-size distribution and the density of volcanic particles).

4. Case Studies

4.1. 1998 Etna Eruption

[35] The 22 July 1998 eruption was one of the largest explosive events at Etna volcano [*Andronico et al.*, 1999; *Aloisi et al.*, 2002]. The eruption column reached an altitude of 12 km (a.s.l.), as measured from rectified camera pictures. A thin tephra fallout deposit blanketed the south-east flank of the volcano. During the eruption, the wind blew almost constantly to the SE for altitudes up to 10 km (about 140 degree from the north). Above 10 km, the wind direction shifted to the NE (about 50 degree from the north). Wind speed ranged between 4 and 6 m s^{-1} at about 1 km (a.s.l) and between 9 and 11 m s^{-1} at about 16 km (a.s.l) [*Aloisi et al.*, 2002].

[36] The 35 samples used in our study were collected a few hours after the eruptive episode (Figure 1). *Andronico et al.* [1999] estimated the total erupted mass of the tephra deposit to be about 1.3×10^9 kg using the method of *Pyle* [1989] and classified the eruption as sub-plinian following the classification scheme of *Walker* [1973]. The total grain-size distribution estimated using the Voronoi method [*Bonadonna and Houghton*, 2005] is centered on 2.3 Φ with a standard deviation equal to 1.5 Φ [*Scollo*, 2006]. Input data used by tephra dispersal models are summarized in Table 1.

4.2. 1996 Ruapehu Eruption

[37] On 17 June 1996, Ruapehu volcano in New Zealand produced an andesitic sub-plinian eruption [*Hurst and Turner*, 1999; *Cronin et al.*, 2003]. This eruption developed two pulsating plume phases that were bent over by the strong winds (maximum wind speed of 40 m s^{-1}) and lasted for about 8 h each. The maximum column height of 8.5 km was constrained using satellite data [*Prata and Grant*, 2001]. The associated deposit on land was reconstructed by the analysis of 114 locations around the volcano [B. F. Houghton et al., unpublished data, 1996; Figure 1b]. The deposit extended to the northeast coast of New Zealand (i.e., 200 km from the vent) and showed a sinusoidal pattern [*Bonadonna et al.*, 2005b]. It was also characterized by a typical bent-over-plume sedimentation, with a sharp thickness decreasing in proximal area and a more gradual thinning in the distal area. The deposit was thoroughly sampled within a few days from the eruption, making this data set ideal for modeling calibration and sensitivity analysis. The total erupted mass of volcanic particles deposited on the ground was estimated about 4.6×10^9 kg by both exponential and power law method [*Bonadonna and Houghton*, 2005]. The total grain-size distribution, estimated by Voronoi method, is centered on −0.8 Φ with 2.4 standard deviation [*Bonadonna and Houghton*, 2005]. Input data used by tephra dispersal models are summarized in Table 1.

5. Results

5.1. Prior Parameter Distributions

[38] We have considered as uncertain the eight input factors of the TEPHRA model described in section 3. The distributions and range over which the input factors were sampled, were taken from the available literature based on the typical features of sub-plinian eruptions (Table 2). We have considered a logarithmic scale for the diffusion coefficient *(K)* and the total erupted mass *(TM)* due to the wide ranges of *K* values, between 0.0001 and 10000 m^2 s^{-1}, and of *TM* values, between 10^8 and 10^{13} kg.

[39] We considered a Gaussian distribution for the total grain-size distribution *TGS*:

$$p(\phi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\phi-\mu}{\sigma}\right)^2\right] \tag{16}$$

where μ is the mean and σ is the standard deviation. The other input factors, the fall time threshold *FTT*, the density of lithics and pumices *DL* and *DP*, the plume ratio *PR*, the plume height *H*, were considered as uniform distributions between the limits of the input factors.

[40] The sampling strategy is suited to the estimation of the sensitivity indices considered in the Sobol' method [*Sobol'*, 1990]. The Computational Cost *CC* of this method, in terms of model simulations, is proportional to the number of input factors *k*. The proportionality coefficient *N*, called base sample size, can be chosen in terms of the desired accuracy of the sensitivity estimates. *N* is selected as a power of two to guarantee uniform coverage of the sampling space when we use the quasi-random number generator proposed by Sobol' [*Sobol'*, 1967]. The algorithm generates uniformly distributed quasi-random sequences within the hypercube $\Omega = \{[0;1] \times [0;1]\ldots\}$ of unit volume. The sample is formed by a quasi-regular grid of points and has an extra uniformity condition known as Property A [*Sobol'*, 1967]. Geometrically, if the hypercube Ω is divided up by planes of equation $x_j = 1/2$ into 2^k equally sized cubes, then, given a sequence of 2^k points, each point belongs to

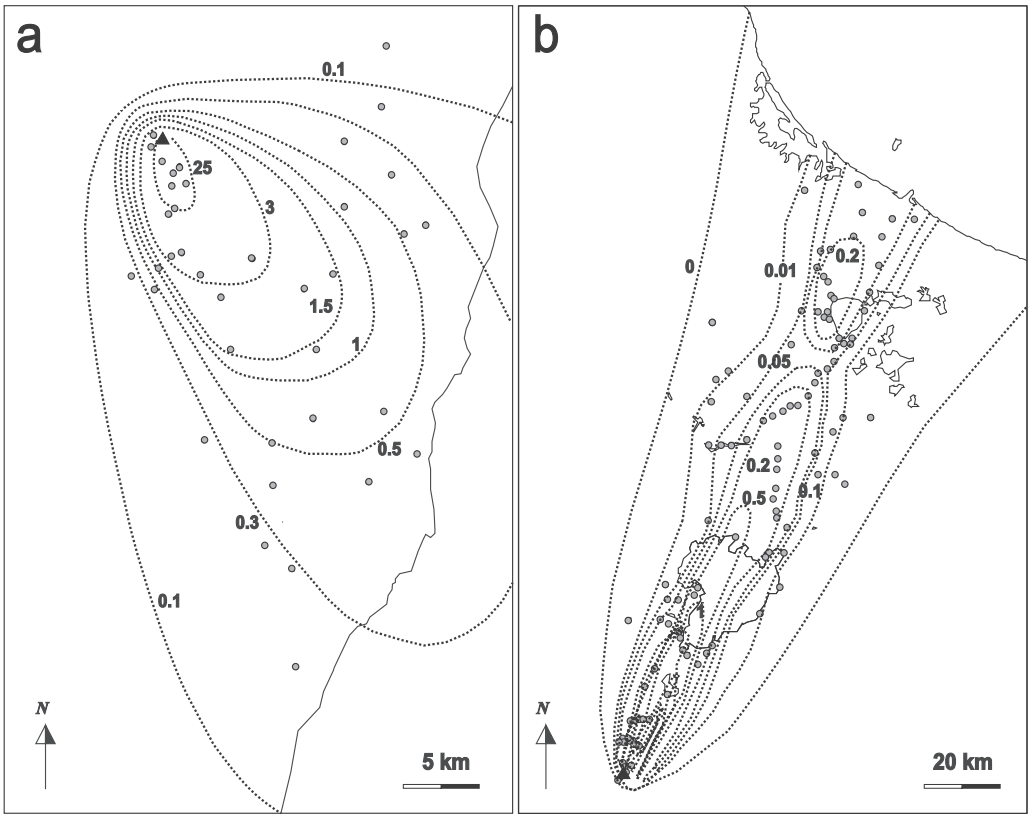

Figure 1. (a) Data points and isomass curves (kg m^{-2}) of the tephra deposit generated by the 22 July 1998 eruption of Etna volcano (Italy) [*Andronico et al.*, 1999]; (b) Data points and isomass curves (kg m^{-2}) of the tephra deposit generated by the 17 June 1996 eruption of Ruapehu volcano (New Zealand) [*Bonadonna et al.*, 2005b; B. F. Houghton et al., unpublished data, 1996]. The triangle indicates the position of the eruptive vent; the circles indicate the locations where tephra samples were collected.

one and only one cube. The prior distributions for the input factors are obtained by applying a suitable transformation to the uniformly distributed sample.

5.2. Sensitivity Analysis

[41] We applied a variance-based method of sensitivity analysis to the TEPHRA model. In particular, we apply an extension of the technique proposed by *Sobol'* [1990] and modified by *Saltelli* [2002].

[42] We have generated a sample in the input factor space by selecting $N = 64$. This value of N was selected as first guess with the hope that it guarantees statistically significant estimates of sensitivity indices; whether N is sufficiently large will be discovered later. The total computational cost CC of sensitivity analysis is $(2 \cdot k + 2) \cdot N$, where k is number of input factors (eight for both Etna and Ruapehu tests). We have run the TEPHRA model to obtain the corresponding 1152 outputs, i.e., the accumulated mass per unit area was computed at the locations where tephra samples were collected at Etna and Ruapehu tephra deposits (Figure 1): 35 locations for Etna and 114 locations for Ruapehu. At each location we calculated both the first order and the total sensitivity indices (S_i and S_{Ti}).

[43] For some input factors we estimated negative values of S_{Ti} in several locations. This is an indication that the initial value set for N should increase, as the total sensitivity indices cannot be negative by definition. So, we decided to increase N to 256 obtained by $64 \times 2 \times 2$ (remembering that N is a power of two necessary to guarantee uniform coverage of the sampling space). Accordingly, the number of simulations required to run the sensitivity analysis increases to 4608.

[44] The results of sensitivity analysis are shown in Figure 2 for Etna and Figure 3 for Ruapehu. For both Etna and Ruapehu we observe that the total erupted mass (*TM*) has the highest value of first order index at all measurement points. When an input factor has a high value of S_i and S_{Ti} it means that it is influent on the model output on its own but

Table 1. Input Data Used by Tephra Dispersal Models of the 22 July 1998 Eruption of Etna Volcano, Italy [*Andronico et al.*, 1999] and the 17 June 1996 Eruption of Ruapehu Volcano, New Zealand [*Bonadonna et al.*, 2005b], B. F. Houghton et al. (unpublished data, 1996)

Input Data	Etna (Italy) 22 July 1998	Ruapehu (New Zealand) 17 June 1996
Eruption style	sub-plinian	sub-plinian
Plume height	12 km (a.s.l)	8.5 km (a.s.l)
Total erupted mass	1.3×10^9 kg	4.6×10^9 kg
Total grain-size distribution (mean)	2.3 Φ	−0.8 Φ
Lithic density	2600 kg m^{-3}	2650 kg m^{-3}
Pumice density	900 kg m^{-3}	1100 kg m^{-3}

Table 2. Input Factors of the TEPHRA Model Considered in the Sensitivity Analysis

Input Factors		Prior Distribution	Distribution Ranges
Total grain-size distribution	*TGS*	normal	−7–9 Φ
Plume height	*H*	uniform	8000–25000 m
Diffusion coefficient	*K*	log-uniform	0.001–10000 m^2 s^{-1}
Fall time threshold	*FTT*	uniform	0–7200 s
Lithic density	*DL*	uniform	500–2300 kg m^{-3}
Pumice density	*DP*	uniform	2300–2500 kg m^{-3}
Plume ratio	*PR*	uniform	0.01–1
Total mass	*TM*	log-uniform	10^8–10^{13} kg

not through interactions with other parameters of the model. As a result, the total erupted mass is responsible for most of the variability of the model predictions implying that there is very good chance that this input factor can be estimated using field data of fallout deposits.

[45] We also observe that the value of S_{Ti} for lithic density (*DL*) and pumice density (*DP*) is quite small for all the 35 locations for Etna and at all the 114 locations for Ruapehu. Consequently, the two densities are not influential and can be fixed to a nominal value within their range of uncertainty without any significant loss in accuracy of results computed by the model.

[46] A detailed analysis of interactions between input factors shows that the Fall Time Threshold (*FTT*) has high values of the total order index and low values of the first order for a number of Etna locations (Figure 2). This means that it has several interactions with the other factors. For example, for the location about 7 km away from the volcanic vent there are strong interactions between *FFT* and *TM* (at Etna location 18).

[47] Sensitivity analysis is an important tool to check the reliability of a model. For example, in the case of Ruapehu, we find that one of the total order indices is largely negative (−1.07 at Ruapehu location 89) in the upwind position about 2.5 km far from the vent (Figure 3). Given that the total sensitivity index is theoretically greater than or equal to zero, a large negative quantity indicates that the model is not able to accurately describe the upwind sedimentation from plumes bent over by strong winds. Finally, the sensitivity analysis shows that far from the volcanic vent (>70 km) the

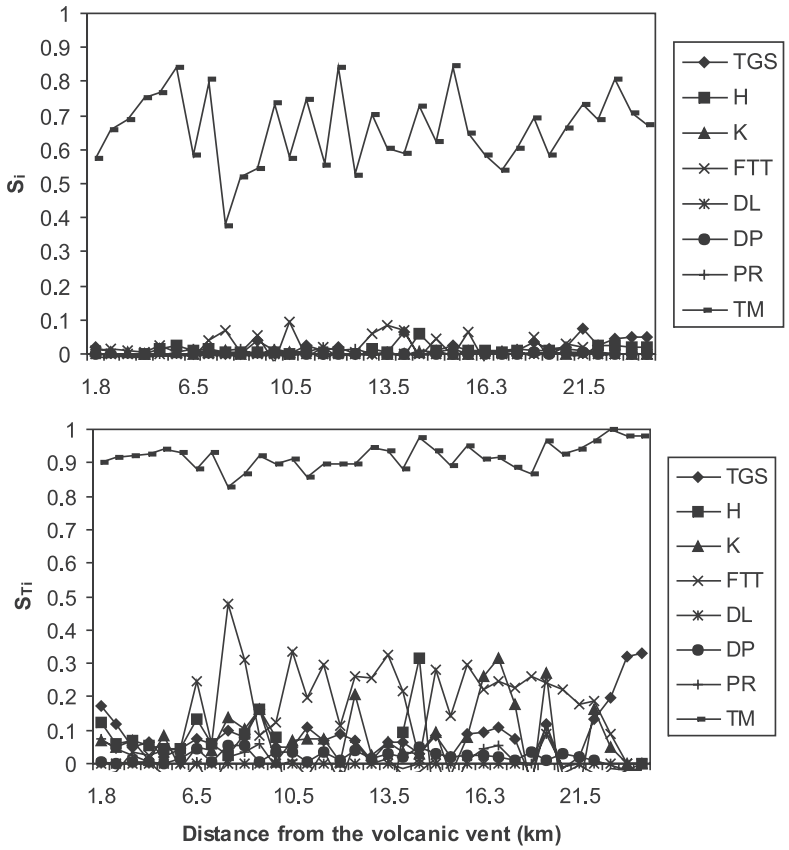

Figure 2. Plots showing the first and the total indices S_i and S_{Ti} obtained by the analysis with Simlab software of 4068 simulations of the TEPHRA model for the 22 July 1998 eruption of Etna volcano. Indices of each input factor are plotted in function of the distance from the volcanic vent (km). In the legend TGS is the total grain-size distribution, H is the plume height, K is the diffusion coefficient, FTT is the fall time threshold, DL is the density of lithics, DP is the density of pumices, PR is the plume ratio and TM is the total mass.

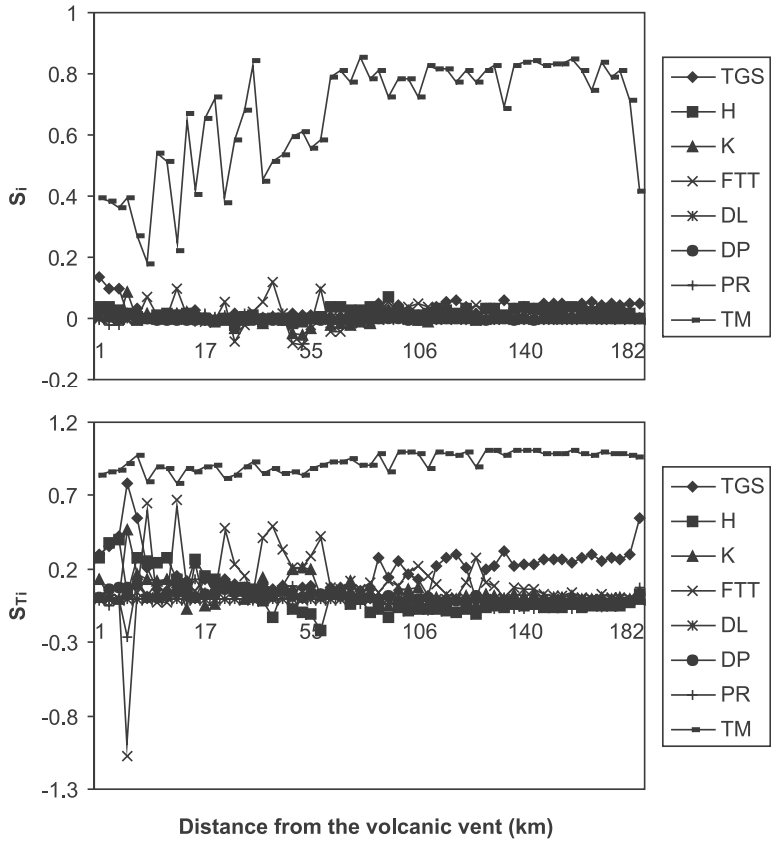

Figure 3. Plots showing the first and the total indices S_i and S_{Ti} obtained by the analysis with Simlab software of 4068 simulations of the TEPHRA model for the 17 June 1996 eruption of Ruapehu volcano. Indices of each input factor are plotted in function of the distance from the volcanic vent (km). Refer to Figure 2 for TGS, H, K, FTT, DL, DP, PR and TM.

total erupted mass does not interact significantly with the other factors.

[48] The next step of the analysis is the estimation of the total erupted mass using the field data for both volcanoes.

5.3. Total Erupted Mass Estimation

[49] The total erupted mass was estimated by the best fit formula (10) comparing the field data and the model outputs of the previous 4608 simulations. Figure 4 shows the posterior distribution of the total erupted mass for Etna, i.e., the pdf (probability distribution function) obtained after applying the best fit function. Higher values of the best fit function correspond to the combination of input factors that give the best agreement with the field data. The posterior distribution looks like a normal distribution slightly skewed on the right side (on a log scale) with the mean value at 1.5×10^9 kg. The result is very similar to the total erupted mass estimated using standard curve-fitting techniques (Table 1). Note that the range of the posterior distribution is much narrower than that of the prior distribution with uncertainty ranging between 10^8 and 10^{13} kg (Table 2).

[50] Figure 4 shows also the posterior distribution of the total erupted mass for Ruapehu. Although the distribution is more dispersed than Etna, by fitting it with a Gaussian distribution we obtained a mean value equal to 4.0×10^9 kg, which is also close to the value obtained using curve-fitting of field data (Table 1).

5.4. Iteration of Sensitivity Analysis

[51] We iterated the sensitivity tests using the estimated pdfs of the total erupted mass and setting the density of lithic and pumice at constant values, given that these two factors are non-influential. With this analysis we expected to obtain more accurate posterior distribution of mass per unit area computed by the model to see whether some further estimations of the total erupted mass is feasible.

[52] In this second test, pumice and lithic densities were fixed at 1080 and 2500 kg m^{-3} for Etna, and at 1000 and 2350 kg m^{-3} for Ruapehu. A new sample of 3584 points was generated by selecting $k = 6$ (the actual number of input factors) and $N = 256$.

[53] First and total sensitivity indices have been again calculated for Etna (Figure 5). Unlike the previous tests, in the Etna simulations all input factors have similar first order indices at all measurement locations (Figure 5). Contrary to the previous test in which the total erupted mass had a first order around 0.8 for many locations (Figure 2), the first order indices are all below 0.4, meaning that there is no chance of estimating any order factor in the model.

[54] Total sensitivity indices show that the factor *PR* is almost non-influential (associated total indices are below 0.1), and could be fixed in the subsequent analysis.

[55] On the other hand, *FTT* shows some peaks at the locations between 10 and 16 km from the vent, meaning

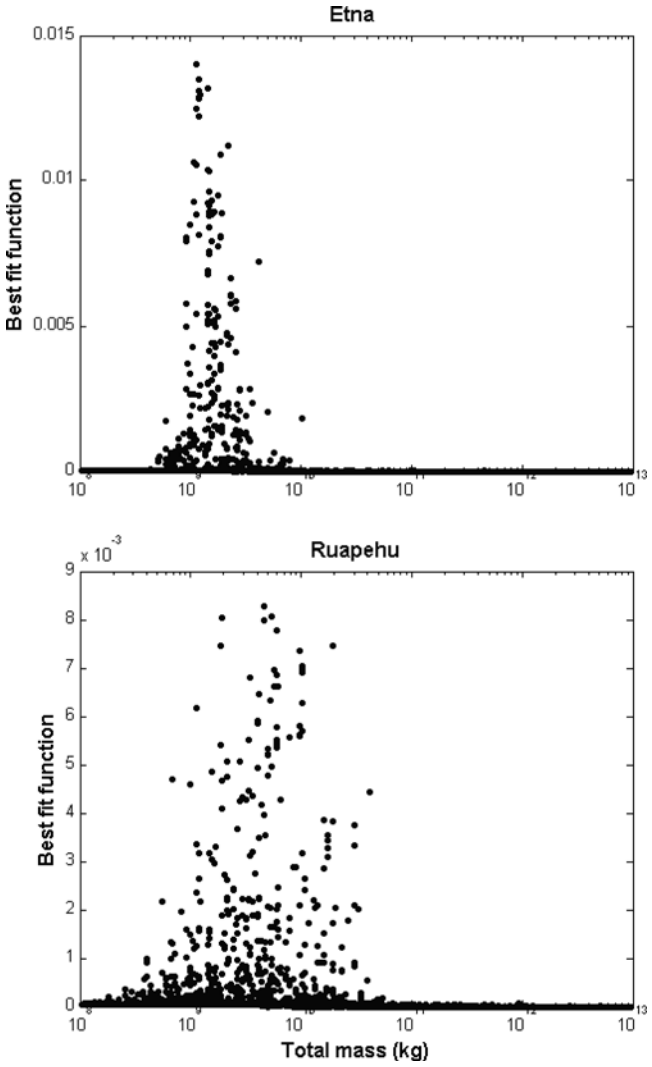

Figure 4. Plot of the best fit function (equation (10)) versus the total erupted mass ranging between 10^8 and 10^{13} kg obtained by comparing simulations of the TEPHRA model and field data of 22 July 1998 Etna deposit and of 17 June 1996 Ruapehu deposit.

that it contributes to uncertainty in model prediction within this range.

[56] The factor *TGS* gains importance with distance from the vent, meaning that at large distance it can be accurately predicted.

[57] First and total sensitivity indices have also been calculated for Ruapehu (Figure 6). Here, a relatively high value of the first order index for *TM* shows that we could, in principle, make an additional estimation of the total erupted mass using the existing data. Factors *K* and *PR* have low values of total sensitivity index, hence, they have minimal influence on model outputs. Instead, the factor *TGS* becomes more important beyond 100 km.

[58] We have also calculated the best fit function (10) for each of the 3584 model predictions and shown the scatterplots of the best fit function versus each input factor (Figures 7 and 8). For both volcanoes, the new posterior distributions of the total erupted mass are slightly updated and have mean values at 1.3×10^9 kg (for Etna) and 4.16×10^9 kg (for Ruapehu). The estimated mean values are closer to the estimates obtained with the classic approach (Table 1).

[59] A new test was carried out for a better calibration of the total erupted mass for the Ruapehu eruption. We used the posterior distribution of the total erupted mass obtained by the last test (Figure 8), while all the other factors remained the same. 3584 runs were performed; however, sensitivity analysis still showed high values of S_i and S_{Ti} for the total erupted mass at almost all locations. It means that the total erupted mass could be calibrated better because both the total and the first index are still high. Nevertheless, using the formula (10) by further iteration, we were not able to reduce more the value of the total sensitivity index S_{Ti}, and, hence, to improve the calibration of the total erupted mass.

5.5. Uncertainty Analysis

[60] A Monte-Carlo-based uncertainty analysis was performed by simulating the total mass per unit area on the ground over different locations from the volcanic vents of both Etna and Ruapehu using the sample of 3584 points obtained from the posterior distributions of the factors.

[61] For Etna case, we found that the mass per unit area measured at the 35 locations falls inside the 5% and the 95% percentiles of the model predictions.

[62] For the Ruapehu test, almost all the mass per unit area collected at the 114 locations falls inside the 5% and the 95% percentiles of the model predictions, excluding some locations very close to the vent or located at distances between 100 km and 150 km far from the vent.

[63] In addition, the mean value and the standard deviation of the simulations have been evaluated from the 3584 simulations available over different locations. For Etna (Figure 9), the predictive uncertainty in terms of coefficient variation (i.e., standard deviation/mean value) of the mass per unit area is about 60% within 16 km from the vent and increases with distance, reaching up to 80% at about 50 km from the vent. On the basis of these data, a given uncertainty can be found for each location. For example, if 500 g m^{-2} of tephra deposit is computed at Acireale we would expect 800 g m^{-2} as the worst case scenario. Similarly, if 1000 g m^{-2} of tephra is computed for Nicolosi and 300 g m^{-2} for Catania, we would expect a maximum accumulation of 1600 g m^{-2} and 500 g m^{-2} respectively for Nicolosi and Catania.

[64] The predictive uncertainty associated with the Ruapehu data set is instead higher than for Etna, being 150% beyond 60 km from the vent, 130% at 100 km, decreasing down to 120% at 200 km. Predictive uncertainty (standard deviation/mean value) is plotted in Figure 10.

6. Discussion and Conclusions

[65] In this paper, a new approach is presented to perform sensitivity analysis of input factors and uncertainty estimation of critical parameters used in analytical and numerical models. As an example, we present the results of sensitivity analyses carried out on specific input factors and uncertainty estimation of the total erupted mass computed on the ground

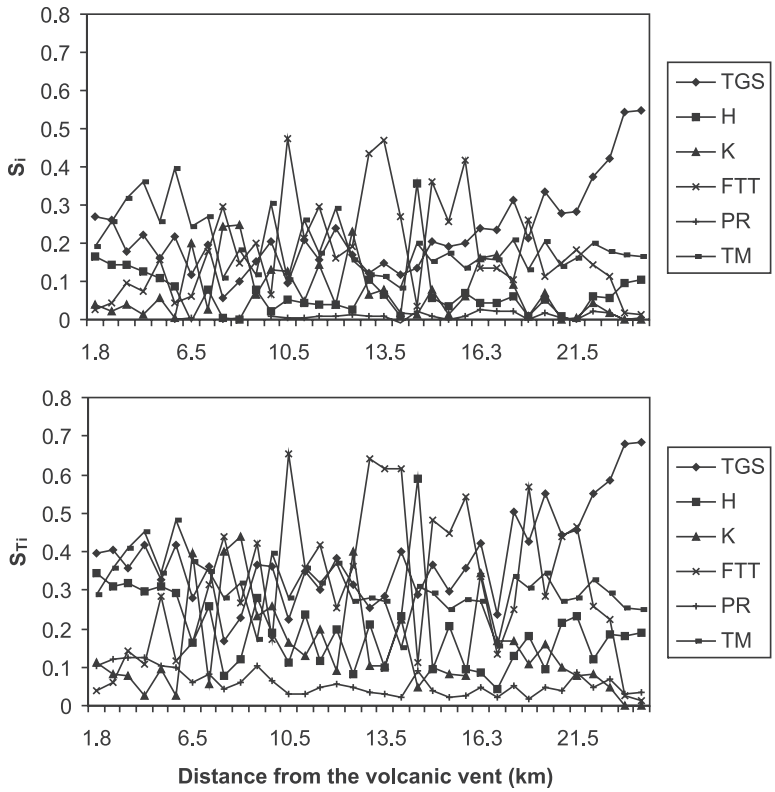

Figure 5. Plots showing the first and the total indices S_i and S_{Ti} resulting from the Simlab analysis of 3584 simulations obtained running TEPHRA for the 22 July 1998 eruption of Etna volcano. Indices of each input factor are plotted as a function of the distance from the volcanic vent (km).

using TEPHRA, an advection-diffusion model that describes particle sedimentation from volcanic plumes [*Bonadonna et al.*, 2005a]. This approach is based on a combination of sensitivity analysis and uncertainty estimation techniques that can be applied to all numerical models in any scientific field [*Crosetto and Tarantola*, 2001].

6.1. Sensitivity Analysis

[66] Sensitivity analysis enables us to classify model input factors according to these three categories:

[67] (1) Factors influencing the model predictions significantly, but not through the interaction with the other input factors.

[68] (2) Factors influencing the model predictions mainly through the interaction with the other factors.

[69] (3) Factors having overall negligible effect on the model predictions.

[70] The first category includes factors which have a high likelihood of being calibrated accurately. The second category includes factors that help us improve our knowledge of the physical model. Factors of the third category can be fixed at a nominal value within their range of uncertainty, thus simplifying the model through the reduction of the number of input factors.

[71] Sensitivity analysis carried out applying TEPHRA to two sub-plinian deposits have shown that the total erupted mass *(TM)* can be accurately estimated and falls in the first category. As a result, sensitivity analyses represent a valuable alternative for the calculation of the total erupted mass, whereas curve-fitting techniques of sparse field data can produce very large errors [*Bonadonna and Houghton*, 2005]. Our evaluation of the total erupted mass calculated by the second iteration of the sensitivity analysis is in good agreement with the total mass calculated by standard technique obtained by field data of the 22 July 1998 Etna eruption and of the 17 June 1996 Ruapehu eruption. In fact, for Etna we obtained the same value (1.3×10^9 kg), and for Ruapehu we found a relative difference of only 9.6% (4.16×10^9 kg respect to 4.6×10^9 kg obtained from field data). It is important to stress here that the total erupted mass calculated from field data both for Etna and Ruapehu is to be considered close to the true value because of the large area covered by the isomass maps and because the deposit was sampled soon after the eruption allowing for a good coverage of distal areas [*Andronico et al.*, 1999; *Bonadonna et al.*, 2005b]. Finally the approach proposed in this paper is more comprehensive and fast than the classic approach based on curve fitting of field data, in that we can obtain uncertainty bounds for the estimates, and we use an automated process that does not involve manual work of the operator. Similar promising results are also presented by *Connor and Connor* [2006] with the application of inversion techniques.

[72] We made some additional tests to check whether the estimation of the total erupted mass could be influenced by the number of field data. The results showed that the estimation of the total erupted mass can be reached with at least 10 locations in which the mass deposited is measured without a particular correlation with the distance of the locations from the volcanic vent (indeed, the sensitivity analysis showed that the total erupted mass is the most

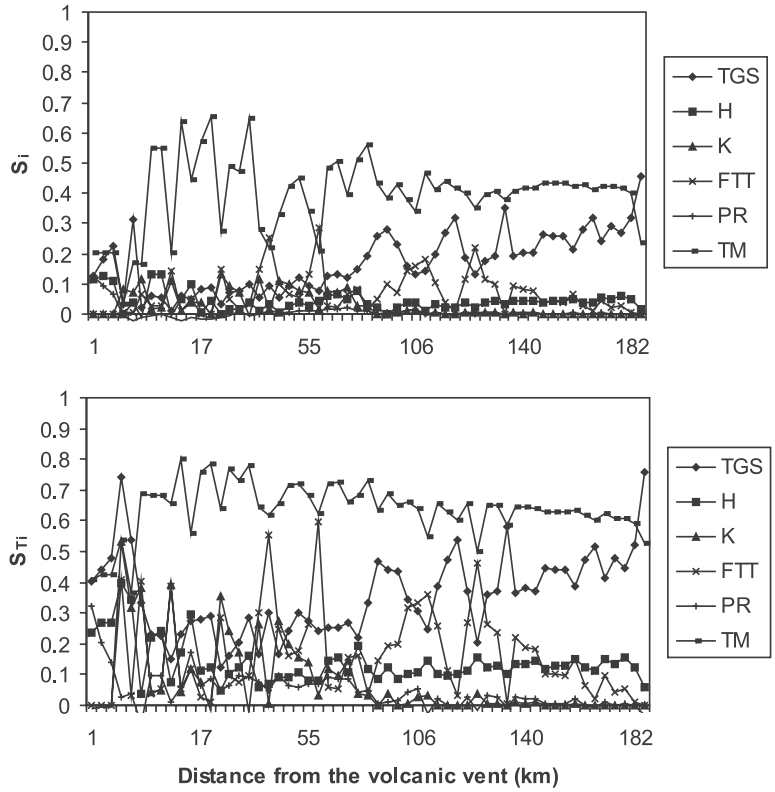

Figure 6. Plots showing the first and the total indices S_i and S_{Ti} resulting from the Simlab analysis of 3584 simulations obtained running TEPHRA for the 17 June 1996 eruption of Ruapehu volcano. Indices of each input factor are plotted as a function of the distance from the volcanic vent (km).

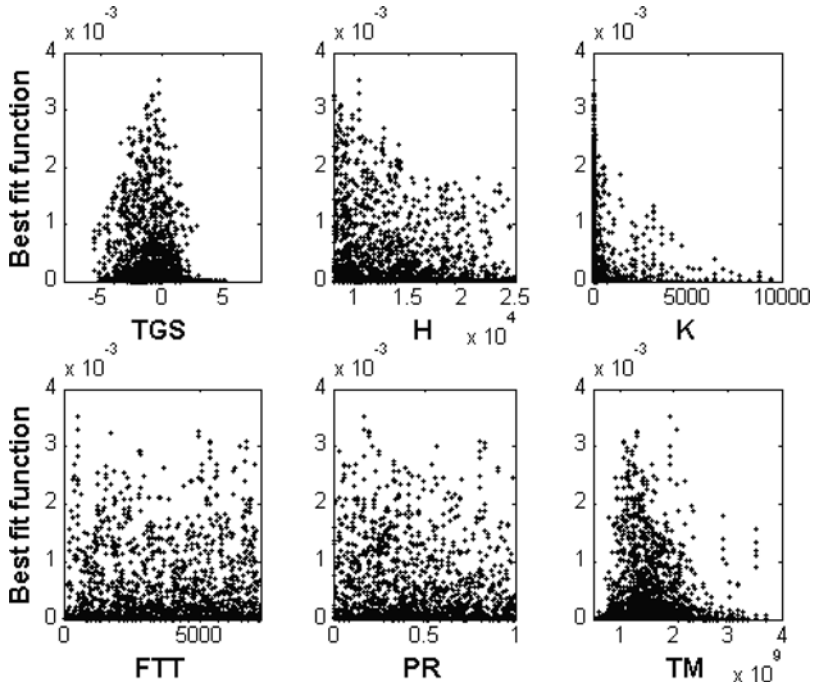

Figure 7. Scatterplots of the best fit function (equation 10) versus the total grain-size distribution *(TGS)*, the plume height *(H)*, the diffusion coefficient *(K)*, the fall time threshold *(FTT)*, the plume ratio *(PR)* and the total erupted mass *(TM)*. They are the result of comparisons between 3584 simulations carried by the TEPHRA model and the field data of 22 July 1998 Etna eruption.

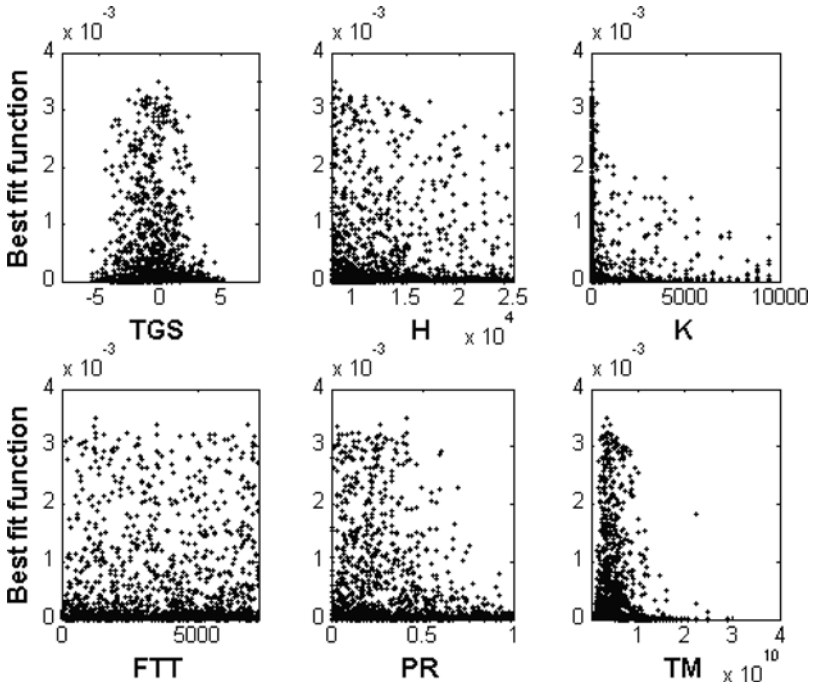

Figure 8. Scatterplots of the best fit function (equation 10) versus the total grain-size distribution *(TGS)*, the plume height *(H)*, the diffusion coefficient *(K)*, the fall time threshold *(FTT)*, the plume ratio *(PR)* and the total erupted mass *(TM)*. They are the result of comparisons between 3584 simulations carried by the TEPHRA model and the field data of 17 June 1996 Ruapehu eruption.

important factor at all locations). Hence this can be considered as the minimum number of locations required for a reliable estimation of the total erupted mass. However, sensitivity analyses on more case studies need to be analyzed to constrain the actual limits of applicability of the method.

[73] The input factors describing the density of lithics and pumices belong to the third category and can be fixed at some nominal values. This enables the modeler to reduce the dimensionality of the model to six input factors: total grain-size *(TGS)*, the plume height *(H)*, the diffusion coefficient *(K)*, the fall time threshold *(FTT)*, the plume ratio *(PR)*, the total erupted mass *(TM)*.

[74] Except for the total erupted mass and the particle density, the rest of the input factors belong to the second category, i.e., factors influencing the model prediction through the interaction with the other factors.

[75] In particular, the total grain-size distribution *(TGS)* has high values of both first and total sensitivity indices, especially at locations far from the volcanic vent (Figures 5 and 6). *TGS* shows also high values of best fit function that are peaked on $-1.14\ \Phi$ for Etna test (Figure 7) and $-0.58\ \Phi$ for Ruapehu test (Figure 8). This latter value is very close to the total grain-size distribution ($-0.8\ \Phi$) obtained applying the method of Voronoi to the same data set [*Bonadonna and Houghton*, 2005]. This is not the case of Etna, where this difference between the two approaches could be due to the availability of grain-size distributions data mainly for locations located far from the vent and therefore being characterized by a larger fraction of fines.

[76] Scatterplots of the best fit function versus plume Height (*H*), Diffusion Coefficient (*K*), Fall Time Threshold (*FTT*) and Plume Ratio (*PR*) show no pattern (Figures 7 and 8), confirming the results obtained by sensitivity analysis in which their first sensitivity indices are smaller than their total sensitivity indices and hence, they cannot be well calibrated (Figures 5 and 6). Some interactions between these input factors are expected due to the difference between the total sensitivity index and the first order index for the same factor. In addition, sensitivity analyses of these case studies (Figures 5 and 6) show that *H* controls the sedimentation near the volcanic vent (up to 10 and 20 km for Etna and Ruapehu respectively), whereas it shows strong interactions with other factors for sedimentation far from the vent. For Ruapehu, *K* is important up to 50 km from the vent; after that distance the associated first order index is below 0.1. This factor controls the diffusion of particles in the atmosphere combining the complex plume dynamics and atmospheric turbulent diffusion into a single factor [*Bonadonna et al.*, 2005a]. It could explain that beyond a certain distance from the vent, which value depends on the eruption, the effects of plume dynamics are negligible and the importance of this factor reduces (low total sensitivity order) compared to the other input factors. In both studied cases, *FTT* is an important factor. In fact it involves the terminal settling velocity of volcanic particles that has a first order effect on results of tephra dispersal models [*Pfeiffer et al.*, 2005] and influences strongly the volcanic particle deposition [*Riley et al.*, 2003].

[77] Sensitivity analyses show that *PR*, that describes the mass distribution along the eruptive column, does not significantly affect the model outputs. However, the simple parameterization used in this sensitivity analysis could oversimplify the description of the ascending eruption

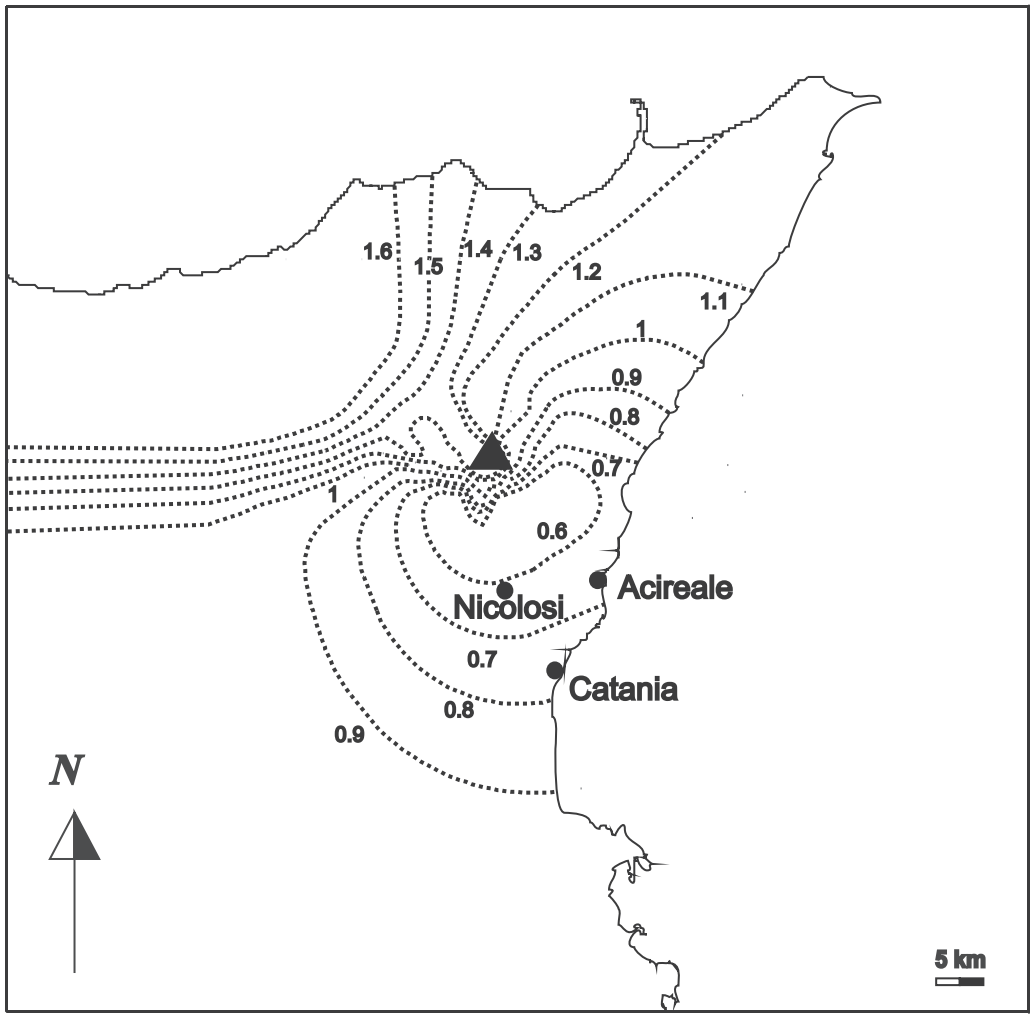

Figure 9. Predictive uncertainty (standard deviation/mean value) of the mass per unit area simulated for 22 July 1998 Etna eruption.

columns. In such case, other algorithms describing the complex mechanisms and thermodynamics of eruption columns [*Bursik*, 2001; *Woods*, 1988] should be implemented into the TEPHRA model and new sensitivity analysis should be performed. On the other hand, any modeling effort should be justified only if it yields better predictions.

6.2. Uncertainty Analysis

[78] We have also applied uncertainty analysis to evaluate uncertainty of the mass per unit area on the ground, due to the uncertainty on input factors. The results show that the coefficient of variation for Etna and Ruapehu ranges between 60% and 90% and between 100% and 150% respectively.

[79] For Ruapehu, we have noted that regions with larger uncertainty are located near the volcanic vent and between 100 and 150 km from the vent. Near the vent, the computed sedimentation clearly underestimates the sedimentation from bent-over plumes, whereas the region between 100 km and 150 km from the vent is characterized by a thickness double maximum and a sinusoidal pattern that cannot be described by simple advection-diffusion models [*Bonadonna et al.*, 2005b]. These features of the deposit are likely to be related to bent-over sedimentation structures and convective instabilities [*Bonadonna et al.*, 2005b].

6.3. Consideration on Our Case Studies (Etna and Ruapehu)

[80] We have considered Etna and Ruapehu (both sub-plinian eruptions) as case studies because they are both characterized by comprehensive data sets [*Andronico et al.*, 1999; *Bonadonna et al.*, 2005b]. *Bonadonna et al.* [2005a] have shown that TEPHRA can both describe sedimentation from sub-plinian and plinian plumes, and therefore we expect the results to be similar for the sedimentation from both strong and weak plumes. However, in order to test if the model results are a function of the eruption style, more tests should be run on smaller (e.g., Vulcanian) and larger eruptions (e.g., Plinian). In addition, *Bonadonna et al.* [2005a] had also shown that TEPHRA have difficulties describing proximal fallout from weak plumes, as confirmed by our results. Different physical approaches, used in models like FALL3D and VOLC-CALPUFF [*Costa et al.*, 2006; *Barsotti et al.*, 2008], could give better results in these cases. Sensitivity analysis and uncertainty estimations should also be carried out for these models in which the

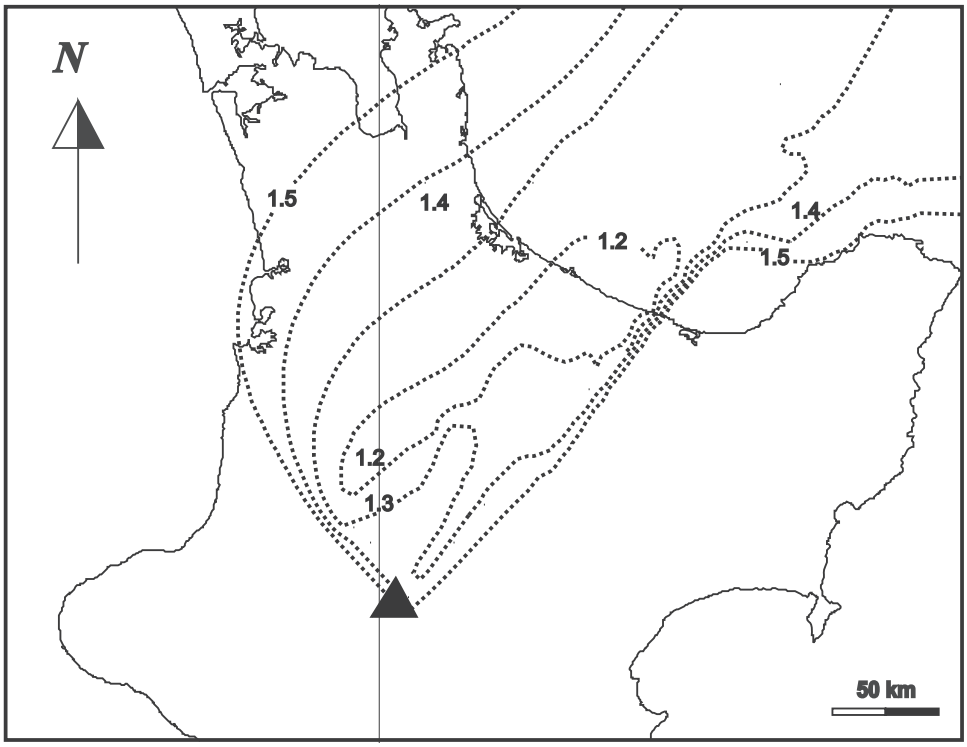

Figure 10. Predictive uncertainty (standard deviation/mean value) of the mass per unit area simulated for 17 June 1996 Ruapehu eruption.

meteorological data have an important effect on the model outputs [*Barsotti and Neri*, 2008].

6.4. Considerations on the Model Used (TEPHRA)

[81] The type of sensitivity analysis and uncertainty estimation presented in this paper can be applied to any analytical and numerical model. We have decided to test these analyses on the TEPHRA model as an example of a model that can describe the transport and sedimentation of volcanic particles. In fact, we were particularly interested in the comparison with ground field data. However, in order to fully understand the relevance of our results, it is also important to highlight some important limitations of the application of the TEPHRA model. First, TEPHRA is a proved powerful model for the compilation of hazard assessments but it cannot be used for the study of the dynamics of particle sedimentation. In fact, in order to allow for shorter computing times, the physical model is mainly based on simplified parameterizations of important processes, such as the spreading of the umbrella cloud, atmospheric diffusion and wind advection. As a result, its strength is on the probabilistic computation more than the physical model. Hence all our outcomes can only be considered as results related to this particular model for tephra dispersal and cannot be generalized to all models for particle sedimentation. Finally, TEPHRA does not account for particle aggregation, which has been shown to significantly affect sedimentation of fine particles [*Brazier et al.*, 1982; *Sparks et al.*, 1997], and consequently, TEPHRA should be used carefully when applied to fine deposits. As an example, *Bonadonna et al.* [2002], *Costa et al.* [2006] and *Scollo et al.* [2007] have assumed that aggregation occurred inside the eruption column and then they modified the input factor of the terminal settling velocity distribution.

[82] In addition, TEPHRA runs using an MPI interface on a Beowulf cluster [*Bonadonna et al.*, 2005a] and, therefore it allows for thousands of simulations to be made in a relative short time. Short computing time is crucial to sensitivity analysis and uncertainty estimations because they are based on probabilistic approaches. In fact, probabilistic approaches have been mainly used to compute hazard assessments [*Cioni et al.*, 2003; *Bonadonna et al.*, 2005a]. Moreover, we have shown that, thanks to the use of parallel processing, we can also use probabilistic techniques to perform sensitivity analysis and uncertainty estimations. Ongoing research focused on the use of sensitivity analysis with the aim to reduce the uncertainty of the mass per unit area computed from tephra dispersal models is crucial to the development of more reliable hazard maps.

Notation

Dimension of Each Term are Given in Brackets: L, Length; T, Time; M, Mass.

A dimensionless parameter that controls the shape of the mass distribution of volcanic particles in the plume.

C apparent eddy diffusivity empirically determined ($C = 0.04\ m^2\ s^{-1}$ [*Suzuki*, 1983]) [$L\ T^{-1}$]; equation (15).

CC computational cost of the simulation proportional to number of input factors ($CC = (2 \cdot k + 2) \cdot N$) where N is a value selected to guarantees statistically significant estimates of sensitivity indices

C_j particle concentration [M L^{-3}]; equation (11)

d particle diameter [L]

DL density of lithics [M L^{-3}]

DP density of pumices [M L^{-3}]

$E(Y)$ expected value of the output variable Y; equation (2)

$E(Y|x_i)$ expected value of the output variable Y conditional on x_i; equation (4)

$f(x)$ mathematical model used in the sensitivity analysis and uncertainty estimation; equation (1)

f_b best fit function used in the parameter estimation procedure and uncertainty estimation procedure; equation (10)

FTT fall time threshold [T]; diffusion of particles with fall times $\langle$ FTT is described by a linear law (equation (13)), diffusion of particles with fall times $\rangle$ FTT is described by a power law (equation (15))

i indices indicating the input factors of the model; indices of point sources along the eruptive plume

j indices of the sample points; indices of particles size

H total height of the eruption column [L]

k number of the model input factors

K horizontal diffusion coefficient ($K = K_x = K_y$) [L^2 T^{-1}]; equation (13)

K_x component (horizontal) of the diffusion coefficient [L^2 T^{-1}]; equation (11)

K_y component (horizontal) of the diffusion coefficient [L^2 T^{-1}]; equation (11)

K_z component (vertical) of the diffusion coefficient [L^2 T^{-1}]; equation (11)

$M(x, y)$ mass accumulated on the ground around a point of coordinates (x, y) [M L^{-2}]; equation (12)

M_{ij}^0 total erupted mass of a given grain-size fraction j released from a point source i along the erupting plume [M]; equation (13)

N_l number of horizontal layers in which the computation domain of the TEPHRA model is divided

$p(\phi)$ total grain-size distribution supposed in the sensitivity analysis; equation (16)

O_l vector of the observations compared with the results of simulations; equation (10)

PR adimensional parameter that describes the top and the bottom of the particle-source area in which the mass distribution is supposed uniform

$std(Y - O_l)$ standard deviation used in the best fit function; equation (10).

S_i first order sensitivity index; equation (6)

S_{ij} higher order sensitivity indices; equation (7)

S_{Ti} total sensitivity index; equation (8)

t time [T]

t_{ij} fall time of a particle of size j released from a point source i along the eruptive plume [T]

t'_i horizontal diffusion time in the volcanic plume at a point source i [T]; equation (14)

TGS Total grain-size distribution of volcanic particles inside the eruption column

TM total erupted mass [M]

$V(Y)$ variance of the model output Y; equation (3)

$V[E(Y|x_i)]$ top marginal variance that measures the importance of the factor x_i; equation (5)

$V_{j,i}$ terminal settling velocity of volcanic particles [L T^{-1}]

Y output of the model

Y_j output of the model related to the sample point j

δx_j wind transport of a particle of size j along the x axis within an atmospheric layer ($\delta x_j = w_x \delta t_j$) [L]

δy_j wind transport of a particle of size j along the x axis within an atmospheric layer ($\delta y_j = w_y \, \delta t_j$) [L]

δz thickness of each atmospheric layer [L]

δt_j time spent by a particle of size j within each atmospheric layer [T]

$x = [x_1, x_2, \ldots x_k]$ the uncertain input factors of the model in the domain $\Omega \in [0;1]^k$

$x_j = [x_{j1}, x_{j2}, \ldots x_{jk}]$ j = 1...N matrix with N row (sample point) and k column (the input factors); equation (9)

(x, y) coordinates of a point on the ground [L]

(x, y, z) coordinates of a point on the space [L].

μ median diameter of the total grain-size distribution of volcanic particles inside the eruption column

w_x component of the wind speed along the x axis [L T^{-1}]

w_y component of the wind speed along the y axis [L T^{-1}]

σ	standard deviation of the input factor *TGS*; equation (16)
σ_{ij}^2	variance of the Gaussian mass distribution on the ground of particles of size j released from a point source i [L^2]; equation (12)
Φ	granulometric unit: $\Phi = -\log_2(d)$ with d particle diameter in mm
Φ_{min}	minimum particle diameter
Φ_{max}	maximum particle diameter

[83] **Acknowledgments.** Authors are grateful to the JRC-Ispra and to the University of Hawaii for supporting the beginning of this project. Authors are also grateful to B.F. Houghton for the use of the field data from the 17 June 1996 eruption of Ruapehu and D. Andronico and P. Del Carlo for the use of field data of the 22 July 1998 eruption of Etna. Many thanks to L. Connor for inspiring advices on the use of TEPHRA. We greatly thank Gregg Bluth, an anonymous reviewer and the associate editor Anita L. Grunder for their valuable and constructive suggestions that improved the quality of the paper. We also thank M. Prestifilippo, G. Spata and M. Ratto for their suggestions and improvements on data analysis, M. Rosi for his encouragement on this research and A. Bonaccorso that allowed the purchase of the Beowulf cluster at INGV-CT where the simulations were run. This work was supported by INGV-GNV fellowship of one of the authors (S. Scollo) and by the FIRB Italian project ''Sviluppo Nuove Tecnologie per la Protezione e Difesa del Territorio dai Rischi Naturali'' funded by Italian Minister of University and Research.

References

Aloisi, M., M. D'Agostino, K. G. Dean, A. Mostaccio, and G. Neri (2002), Satellite analysis and PUFF simulation of the eruptive cloud generated by the Mount Etna paroxysm of 22 July 1998, *J. Geophys. Res.*, *107*(B12), 2373, doi:10.1029/2001JB000630.

Andronico, D., Del P. Carlo, and M. Coltelli (1999), The 22 July 1998 fire fountain episode at Voragine Crater (Mt Etna, Italy), Volcanic and Magmatic Studies Group, Annual Meeting, Birmingham, 5–6 January.

Armienti, P., G. Macedonio, and M. T. Pareschi (1988), A numerical-model for simulation of tephra transport and deposition: Applications to May 18, 1980, Mount-St-Helens eruption, *J. Geophys. Res.*, *93*(B6), 6463–6476.

Barsotti, S., and A. Neri (2008), The VOL-CALPUFF model for atmospheric ash dispersal: 2. Application to the weak Mount Etna plume of July 2001, *J. Geophys. Res.*, *113*, B03209, doi:10.1029/2006JB004624.

Barsotti, S., A. Neri, and J. S. Scire (2008), The VOL-CALPUFF model for atmospheric ash dispersal: 1. Approach and physical formulation, *J. Geophys. Res.*, *113*, B03208, doi:10.1029/2006JB004623.

Beven, K. J., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298.

Bonadonna, C., and B. F. Houghton (2005), Total grain-size distribution and volume of tephra-fall deposits, *Bull. Volcanol.*, *67*, 441–456.

Bonadonna, C., and J. C. Phillips (2003), Sedimentation from strong volcanic plumes, *J. Geophys. Res.*, *108*(B7), 2340, doi:10.1029/2002JB002034.

Bonadonna, C., G. G. J. Ernst, and R. S. J. Sparks (1998), Thickness variations and volume estimates of tephra fall deposits: The importance of particle Reynolds number, *J. Volcanol. Geotherm. Res.*, *81*, 173–187.

Bonadonna, C., G. Macedonio, and R. S. J. Sparks (2002), Numerical modelling of tephra fallout associated with dome collapses and Vulcanian explosions: Application to hazard assessment on Montserrat, in *The Eruption of Soufrière Hills Volcano, Montserrat, From 1995 to 1999*, edited by T. H. Druitt and B. P. Kokelaar, Geological Society, London.

Bonadonna, C., C. B. Connor, B. F. Houghton, L. Connor, M. Byrne, A. Laing, and T. K. Hincks (2005a), Probabilistic modeling of tephra dispersion: Hazard assessment of a multi-phase rhyolitic eruption at Tarawera, New Zealand, *J. Geophys. Res.*, *110*, B03203, doi:10.1029/2003JB002896.

Bonadonna, C., J. C. Phillips, and B. F. Houghton (2005b), Modeling tephra sedimentation from a Ruapehu weak plume eruption, *J. Geophys. Res.*, *110*, B08209, doi:10.1029/2004JB003515.

Brazier, S., A. N. Davis, H. Sigurdsson, and R. S. J. Sparks (1982), Fallout and deposition of volcanic ash during the 1979 explosive eruption of the Soufrière of St-Vincent, *J. Volcanol. Geotherm. Res.*, *14*, 335–359.

Bursik, M. I. (2001), Effect of wind on the rise height of volcanic plumes, *Geophys. Res. Lett.*, *28*, 3621–3624.

Bursik, M. I., S. N. Carey, and R. S. J. Sparks (1992a), A gravity current model for the May 18, 1980 Mount-St-Helens plume, *Geophys. Res. Lett.*, *19*, 1663–1666.

Bursik, M. I., R. S. J. Sparks, J. S. Gilbert, and S. N. Carey (1992b), Sedimentation of tephra by volcanic plumes: I. Theory and its comparison with a study of the Fogo A plinian deposit, Sao Miguel (Azores), *Bull. Volcanol.*, *54*, 329–344.

Carey, S. N., and H. Sigurdsson (1982), Influence of particle aggregation on deposition of distal tephra from the May 18, 1980, eruption of Mount St-Helens volcano, *J. Geophys. Res.*, *87*(B8), 7061–7072.

Carey, S. N., and R. S. J. Sparks (1986), Quantitative models of the fallout and dispersal of tephra from volcanic eruption columns, *Bull. Volcanol.*, *48*, 109–125.

Cioni, R., A. Longo, G. Macedonio, R. Santacroce, A. Sbrana, R. Sulpizio, and D. Andronico (2003), Assessing pyroclastic fall hazard through field data and numerical simulations: Example from Vesuvio, *J. Geophys. Res.*, *108*(B2), 2063, doi:10.1029/2001JB000642.

Connor, L. G., and C. B. Connor (2006), Inversion is the key to dispersion: Understanding eruption dynamics by inverting tephra fallout, in *Statistics in Volcanology, Society for Industrial and Applied Mathematics*, Special Publication of IAVCEI No. 1, edited by H. Mader, S. Cole, and C. B. Connor, pp. 231–242, Geological Society, London.

Connor, C. B., B. E. Hill, B. Winfrey, N. M. Franklin, and P. C. La Femina (2001), Estimation of volcanic hazards from tephra fallout, *Nat. Hazards Rev.*, *2*(1), 33–42.

Costa, A., G. Macedonio, and A. Folch (2006), A three dimensional Eulerian model for transport and deposition of volcanic ashes, *Earth Planet. Sci. Lett.*, *241*, 634–647.

Cronin, S. J., V. E. Neall, J. A. Lecointre, M. J. Hedley, and P. Loganathan (2003), Environmental hazards of fluoride in volcanic ash: a case study from Ruapehu volcano, New Zealand, *J. Volcanol. Geotherm. Res.*, *121*, 271–291.

Crosetto, M., and S. Tarantola (2001), Uncertainty and sensitivity analysis: Tool for GIS-based model implementation, *Int. J. Geogr. Inf. Sci.*, *15*(5), 415–437.

Ernst, G. G. J., R. S. J. Sparks, S. N. Carey, and M. I. Bursik (1996), Sedimentation from turbulent jets and plumes, *J. Geophys. Res.*, *101*, 5575–5589.

Fierstein, J., and M. Nathenson (1992), Another look at the calculation of fallout tephra volumes, *Bull. Volcanol.*, *54*, 156–167.

Froggatt, P. C. (1982), Review of methods estimating rhyolitic tephra volumes: Applications to the Taupo Volcanic Zone, New Zealand, *J. Volcanol. Geotherm. Res.*, *14*, 1–56.

Helton, J. C. (1993), Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal, *Reliab. Eng. Syst. Safety*, *42*, 327–367.

Holasek, R. E., and S. Self (1995), Goes weather-satellite observations and measurements of the May 18, 1980, Mount-St-Helens eruption, *J. Geophys. Res.*, *100*, 8469–8487.

Homma, T., and A. Saltelli (1996), Importance measures in global sensitivity analysis of model output, *Reliab. Eng. Syst. Safety*, *52*, 1–17.

Hurst, A. W., and R. Turner (1999), Performance of the program ASHFALL for forecasting ashfall during the 1995 and 1996 eruptions of Ruapehu Volcano, *N. Z. J. Geol. Geophys.*, *42*(4), 615–622.

Jansen, M. J. W., W. A. H. Rossing, and R. A. Daamen (1994), Monte Carlo estimation of uncertainty contributions from several independent multivariate sources, in *Predictability and Nonlinear Modelling in Natural Sciences and Economics*, edited by J. Gasman and G. Van Straten, pp. 334–343, Kluwer, Dordrecht.

Koyaguchi, T., and M. Ohno (2001a), Reconstruction of eruption column dynamics on the basis of grain-size of tephra fallout deposits. 1 Methods, *J. Geophys. Res.*, *106*, 6499–6512.

Koyaguchi, T., and M. Ohno (2001b), Reconstruction of eruption column dynamics on the basis of grain-size of tephra fallout deposits. 2. Application to the Pinatubo 1991 eruption, *J. Geophys. Res.*, *106*, 6513–6533.

Kunii, D. K., and O. Levenspiel (1969), *Fluidization Engineering*, John Wiley, New York.

Macedonio, G., M. T. Pareschi, and R. Santacroce (1988), A numerical simulation of the Plinian fall phase of 79 AD eruption of Vesuvius, *J. Geophys. Res.*, *93*, 14,817–14,827.

Morton, B., G. L. Taylor, and J. S. Turner (1956), Turbulent gravitational convection from maintained and instantaneous source, *Proc. R. Soc.*, *234*, 1–23.

Pfeiffer, T., A. Costa, and G. Macedonio (2005), A model for the numerical simulation of tephra fall deposits, *J. Volcanol. Geotherm. Res.*, *140*(4), 273–294.

Prata, A. J., and I. F. Grant (2001), Retrieval of microphysical and morphological properties of volcanic ash plumes from satellite data: Application to Mt Ruapehu, New Zealand, *Q. J. R. Meteorol. Soc.*, *127*, 2153–2180.

Pyle, D. M. (1989), The thickness, volume and grain-size of tephra fall deposits, *Bull. Volcanol.*, *51*, 1–15.

Riley, C. M., W. I. Rose, and G. J. S. Bluth (2003), Quantitative shape measurements of distal volcanic ash, *J. Geophys. Res.*, *108*(B10), 2504, doi:10.1029/2001JB000818.

Saltelli, A. (2002), Making best use of model valuations to compute sensitivity indices, *Comput. Phys. Commun.*, *145*, 280–297.

Saltelli, A., K. Chan, and M. Scott (Eds.) (2000), *Sensitivity Analysis*, John Wiley & Sons Publishers, New York.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola (2008), *Global Sensitivity Analysis the Primer*, John Wiley & Sons.

Scollo, S. (2006), Study of fallout processes of volcanic particles from explosive Etna eruptions, Ph.D. Alma Mater Studiorum, Università di Bologna, Italy.

Scollo, S., P. Del Carlo, and M. Coltelli (2007), Tephra fallout of 2001 Etna flank eruption: Analysis of the deposit and plume dispersion, *J. Volcanol. Geotherm. Res.*, *160*(1–2), 147–164.

Searcy, C., K. Dean, and W. Stringer (1998), PUFF: A high-resolution volcanic ash tracking model, *J. Volcanol. Geotherm. Res.*, *80*(1–2), 1–16.

Sobol', I. M. (1967), On the distribution of points in a cube and the approximate evaluation of integrals, *U.S.S.R. Comput. Math. Math. Phys.*, *7*, 86–112.

Sobol', I. M. (1990), Sensitivity estimates for nonlinear mathematical models, *Math. Modell. Comput. Exp.*, *1*, 407–414.

Sparks, R. S. J. (1986), The dimension and dynamics of volcanic columns, *Bull. Volcanol.*, *48*, 3–15.

Sparks, R. S. J., and L. Wilson (1982), Explosive volcanic eruptions: V. Observations of plume dynamics during the 1979 Soufrière eruption, St. Vincent, *Geophys. J. R. Astron. Soc.*, *69*, 551–570.

Sparks, R. S. J., M. I. Bursik, S. N. Carey, J. S. Gilbert, L. S. Glaze, H. Sigurdsson, and A. W. Woods (1997), *Volcanic Plumes*, John Wiley, New York.

Suzuki, T. (1983), A theoretical model for dispersion of tephra, in *Arc Volcanism, Physics and Tectonics*, edited by Terra Sci., Tokyo.

Walker, G. P. L. (1973), Explosive volcanic eruptions- A new scheme of classification, *Geol. Rundsch.*, *62*, 431–446.

Wilson, L., and G. P. L. Walker (1987), Explosive volcanic-eruptions, Ejecta dispersal in plinian eruptions-The control of eruption conditions and atmospheric properties, *Geophys. J. R. Astron. Soc.*, *89*, 657–679.

Woods, A. W. (1988), The fluid dynamics and thermodynamics of eruption columns, *Bull. Volcanol.*, *50*, 169–193.

Woods, A. W., R. E. Holasek, and S. Self (1995), Wind-driven dispersal of volcanic ash plumes and its control on the thermal structure of the plume-top, *Bull. Volcanol.*, *57*(5), 283–292.

C. Bonadonna, Section des Sciences de la Terre, Université de Genève, 13, rue des Maraîchers, CH-1205, Geneva, Switzerland. (costanza.bonadonna@terre.unige.ch)

M. Coltelli and S. Scollo, Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Catania, Piazza Roma 2, Catania 95123, Italy. (coltelli@ct.ingv.it; scollo@ct.ingv.it)

A. Saltelli and S. Tarantola, Joint Research Centre of the European Commission, Institute for the Protection and Security of the Citizen (IPSC), TP361, Ispra, VA I-21020, Italy. (andrea.saltelli@jrc.it; stefano.tarantola@jrc.it)