Journal of Volcanology and Geothermal Research 466 (2025) 108393

Contents lists available at ScienceDirect

Journal of Volcanology and Geothermal Research

journal homepage: www.journals.elsevier.com/journal-of-volcanology-and-geothermal-research

Global sensitivity analysis of models for volcanic ash forecasting

Emmy Scott *, Melody Whitehead, Stuart Mead, Mark Bebbington, Jonathan Procter

Volcanic Risk Solutions, Massey University, Private Bag 11222, Palmerston North 4442, New Zealand

ARTICLE INFO

Keywords:
Global sensitivity analysis
Tephra2
Fall3D
Tephra deposition
Forecasting

ABSTRACT

Volcanic ash is a widespread and destructive volcanic hazard. Timely and accurate forecasts for ash deposition and dispersal help mitigate the risks of volcanic hazards to society. Producing these forecasts requires numerous simulations with varying input parameters to encapsulate uncertainty and accurately capture the actual event to deliver a reliable forecast. However, exploring all possible combinations of input parameters is computationally infeasible in the lead up to an eruption. This research explores the input space of two volcanic ash transport and dispersion models, Tephra2, which is based on a simplified analytical solution, and Fall3D, which is a computational model based on more general assumptions, in the context of forecasting an unknown future eruption. We use the exemplar of Taranaki Mounga (Mount Taranaki), Aotearoa New Zealand, which has an estimated 30% to 50% chance of an explosive eruption in the next 50 years. We statistically determine how much each input parameter contributes to model output variance through a global sensitivity analysis via Sobol' indices and the extended Fourier Amplitude Sensitivity Test (eFAST). Our findings show that grain size distribution, diffusion, plume shape, and plume duration (Fall3D only) have a substantial first-order impact on model output variance. In contrast, mass, particle density, and plume height have minimal impact in the first-order but become influential when considering parameter-parameter inter-relationships (total-order). The results not only enhance our understanding of model sensitivities but also point to improved efficiency in forecasting efforts.

1. Introduction

Volcanic eruptions are powerful natural events that can endanger populations. Mitigating the societal impacts of eruption-related hazards can be accomplished through actions and plans based on hazard assessments and forecasts, which are reliant on the output of computational models (e.g., Folch, 2012; Selva et al., 2014; Sandri et al., 2016; Pardini et al., 2024). Advancements in computing power have enabled the development of increasingly complex volcanic ash transport and deposition models (VATDMs) (Folch, 2012). These models take input parameters – such as plume height, mass eruption rate (MER), and wind conditions – and simulate ash dispersal and deposition to produce outputs such as ashfall extent and thickness. More complex models can be developed by increasing the number of tuneable input parameters to capture the physical processes of ash transport and deposition (Höge et al., 2018; Malmborg et al., 2024). However, there is little evidence that more complex models produce more accurate forecasts of future events (Petropoulos et al., 2022). Additionally, more unconstrained input parameters may contribute to increases in uncertainty, and consequently, poorly constrained forecasts (Green and Armstrong, 2015; Saltelli et al., 2020).

Volcanic hazard model outputs rely on the accuracy of the input values (Selva et al., 2014; Engwell et al., 2024; Pardini et al., 2024), and are validated through simulating past eruptions (hindcasting) (Bonasia et al., 2012; Michaud-Dubuy et al., 2021). Successfully replicating a past event is a key step in model validation (Widiwijayanti et al., 2004; Connor and Connor, 2006; Scollo et al., 2007; Dietterich et al., 2017), based on the assumption that a model which accurately simulates one eruption provides a credible basis for forecasting future activity (Pyle, 2018). However, hindcasting often involves extensive tuning of input parameters, and performance on one eruption does not guarantee similar accuracy for others (Scollo et al., 2008a; Dietterich et al., 2017; Gueugneau et al., 2021).

During signs of heightened volcanic unrest, there is insufficient time to run hundreds of thousands of simulations of all possible input parameter combinations. Therefore, we want to know before an eruption, how many of the input parameter combinations produce statistically distinguishable results. In other words, we aim to identify which input parameters should be varied to explore worst-case scenarios and to capture the full range of physically plausible outcomes, including extreme but realistic combinations of conditions (i.e., end-member events) before an imminent eruption. Because ash transport and deposition is

* Corresponding author.
E-mail address: e.e.scott@massey.ac.nz (E. Scott).

https://doi.org/10.1016/j.jvolgeores.2025.108393
Received 28 February 2025; Received in revised form 30 May 2025; Accepted 6 June 2025
Available online 3 July 2025

Table 1
Sampled input parameter ranges for Tephra2 and Fall3D. All input parameters are sampled with a uniform distribution excluding spatial resolution, meteorological data, mass, and fall time threshold. The methods used to determine the number of simulations are described in Section 3.3.

Parameter Inputs	Tephra2	Fall3D	Reference
Spatial resolution	Specific points at 16 cardinal directions at 5 km, 10 km, and 30 km from the vent	0.125° × 0.125°	
Vertical resolution	37 Pressure levels, 0 to 32 km		Hersbach et al. (2020)
Meteorological data	ERA5 reanalysis		Hersbach et al. (2020)
Median grain size (phi)	−4 to 6		Pistolesi et al. (2015), Pioli et al. (2019)
Standard deviation grain size (phi)	0.3 to 4.5		Pioli et al. (2019)
Plume height (m)	1000 to 40000*		IVESPA database Aubry et al. (2021)
Plume shape (vertical distribution of ash)	Beta distribution Alpha$^{(T2)}$: 0.1 to 10 Beta: 0.1 to 10	Suzuki distribution Alpha$^{(F3D)}$: 3 to 5 Lambda: 1 to 10	Beta – University of South Florida (0000), Wiejaczka and Giachetti (2024) Suzuki – Sulpizio et al. (2012), Parra et al. (2016)
Density (kg/m^3)	Pumice: 450 to 2000 Lithic: 1000 to 4000	Minimum: 450 to 1000 Maximum: 2000 to 4000	IVESPA database Aubry et al. (2021), University of South Florida (0000)
Mass (kg)/MER (kg/s)	2.5× 10^9 to 1 × 10^{13}	Kept constant using Mastin et al. (2009) model	Min/max from Mastin et al. (2009) database
Diffusion	Diffusion coefficient: 100 to 100,000	Horizontal diffusion (m/s^2): 300 to 10,000	Diffusion coefficient: max in Volentik et al. (2010) Horizontal diffusion: range in Byun and Schere (2006)
Falltime threshold	100 to 150,000	N/A	
Plume duration (hours)	N/A	0.1 to 12	Most eruptions <10 h in Aubry et al. (2021)
Number of simulations	eFAST – 5000 × 12 winds = 60,000 Sobol' – 265,500 × 12 winds = 3.18 Million	eFAST – 1800 × 12 winds = 21,600 Sobol' – N/A	

strongly influenced by wind (Bursik, 2001; Bonadonna and Costa, 2013; Costa et al., 2016; Mulena et al., 2016), which varies in real-time, our focus remains on identifying key input parameters that contribute to output variability, rather than atmospheric conditions. Building on previous studies (Scollo et al., 2008b; Pardini et al., 2022), we examine input parameter ranges under complete uncertainty acknowledging that the future is fundamentally unknown.

This paper explores the input parameter space of two commonly used VATDMs Tephra2 (Connor et al., 2008) and Fall3D (Folch et al., 2009, 2020) for tephra deposition forecasting. We selected these two VATDMs because they differ in the number and style of input parameters, are widely used, and are open-source. Tephra2 is a semi-analytical model with fewer input parameters, making it computationally efficient. In contrast, Fall3D is a more complex Eulerian model that incorporates finer-scale physics and a greater number of input parameters, enabling higher-resolution simulation.

2. Volcanic ash transport and dispersion models

Various computational models have been developed to simulate the transport and deposition of volcanic ash. The following sections provide an overview of the two VATDM used in our study, Tephra2 and Fall3D.

2.1. Tephra2

Tephra2 is a numerical, two-dimensional VATDM that simulates deposited tephra mass. It is designed for fast computation, as the reference grid can be straightforward (e.g., one single coordinate, where one mass value is produced), and the wind is homogeneous over the whole reference grid.

Of the fourteen input parameters in Tephra2 (Connor and Hagdorn, 2018), four are kept constant throughout the simulations: the grain size range was kept at −7 to 7 phi (block/bomb to extremely fine ash) with 15 particle steps, resulting in one phi per bin; column steps (number of levels in the eruption plume) were kept constant at 100, as recommended by the Tephra2 developers (Connor et al., 2011); and the Eddy constant (diffusion of fine particles) is always fixed at 0.04 (Suzuki, 1983). The remaining ten input parameters are sampled from informed distributions (Table 1). We sample mass using log-uniform sampling and use the Mastin et al. (2009) equation (Eq. (1)) to ensure a realistic relationship between plume height and total erupted mass. As mass is needed in Tephra2, we converted volume to mass assuming a density of 2500 kg/m^3, which is a typical density for lithics (Carey and Sparks, 1986; Bonadonna et al., 1998). We also sample values around the regression line, where H is in km, and $Z \sim N(0,2)$, to obtain a range of plume heights that account for natural variability.

$$H = 25.9 + 6.64\text{x}log_{10}(\text{Volume}) + Z \quad (1)$$

We additionally sampled the fall time threshold input parameter dependent on the diffusion coefficient. During preliminary investigations, we found that certain combinations of fall time threshold (FTT) and diffusion coefficient produced unrealistic diffusion patterns (i.e., larger particles diffuse more than smaller particles) (supplementary material, Figure 14). To address this model-inherent issue, we used an equation derived from best-fit combinations of these input parameters, as identified in the literature (Eq. (2)):

$$FTT = 0.638\text{xDiffusion}^{(1.0792)} \quad (2)$$

2.2. Fall3D

Fall3D is a numerical, three-dimensional VATDM, that simulates tephra mass both in the atmosphere and on the ground. Compared to Tephra2, Fall3D is more computationally intensive, as it calculates particle transport and deposition across a three-dimensional domain.

Fall3D contains more than 20 input parameters, along with numerous sub-parameters (see Folch et al. (2023) for a full list of input

parameters). For this study, we sample nine key input parameters (Table 1). All the Fall3D input parameters were sampled uniformly. We excluded certain input parameters from our sampling, as some input parameters depended on multiple sub-parameter choices. For instance, one of the plume shape (i.e., the vertical distribution of ash in the eruption column) distribution options (named "PLUME"), which uses the one-dimensional buoyant plume theory model FPLUME-1.0 (Folch et al., 2016), requires an additional seventeen input parameter values or yes/no decisions. To facilitate insights into behaviour between the two VATDM, input parameter distributions were aligned. For example, we kept the total grain size distribution (TGSD) as a Gaussian distribution, which is the only option available in Tephra2, even though Fall3D provides five different input options (Gaussian, BiGaussian, Weibull, BiWeibull, Custom, and Estimate from column height and magma viscosity) (Folch et al., 2023).

Fall3D contains categorical variables that cannot be incorporated into global sensitivity analysis (GSA) like eFAST and Sobol' indices (Saltelli et al., 2008) (Section 3.1), as they are either non-numeric or do not have continuous ranges. Thus, a separate strategy was required to explore the influence of these categorical variables. Specifically, we examined distribution (categorised as a categorical variable because multiple distribution types are available), terminal velocity model, aggregation, flux limiter, and horizontal and vertical turbulence models. The definitions of each categorical variable can be found in Folch et al. (2009, 2020). To be able to compare plume shape distributions, specifically Suzuki and Top-Hat, we have normalised the two distributions i.e., the distributions have the same mean and standard deviation. Plume shapes are sampled at 5, 10, and 20 km plume heights. To avoid unnecessary redundancy, we reduced the sampling of terminal velocity models from seven to three distinct models. This decision was based on preliminary testing, where we observed that seven models could be grouped into three categories based on their similar outputs, allowing us to focus on the most representative models (supplementary material, Figure 14).

All the other input parameters, including simulation and plume duration, are kept constant. To avoid ignoring these potentially highly influential input parameters, we applied a local sensitivity analysis (LSA) (Fig. 1) to identify any individual or combinations of categorical variables that appeared to have a large effect on the output variance.

We excluded the MER models from the categorical sampling because, although MER is a key driver of the total amount of deposited mass, total erupted mass is a function of both MER and eruption duration. The correlation between MER and deposited mass is therefore only approximate unless duration is held constant. To reduce the influence of this correlation, we kept the MER input parameter consistent with the Mastin model in our GSA analysis based on findings from Dürig et al. (2023).

3. Methods

This section outlines our methodological approach, beginning with an introduction to global sensitivity analysis. We then provide an overview of the meteorological conditions examined, followed by the practical implementation of the methodology.

3.1. Global sensitivity analysis

Sensitivity analysis is a necessary step in understanding the relationship between input parameters and model outputs for complex computational models (Saltelli, 2002). Global sensitivity analysis (GSA) examines how much of the variation in model outputs can be attributed to different input parameters across an input parameter space range (Saltelli et al., 2004). GSA stems from the analysis of variance (ANOVA), which is used to determine if an input parameter includes values that provide statistically different outputs (Girden, 1992). While ANOVA statistically identifies significant relationships between individual input parameters and model outputs (Montgomery, 2017), GSA extends this concept by quantifying the contributions of each input variable – both individually and in combination – to the overall variability of the model output (Mara, 2009).

Several studies have conducted sensitivity analyses on VATDMs, focusing specifically on tephra deposition (Table 2). From these four studies, total erupted mass (TEM)/mass eruption rate (MER) and plume height emerge as influential input parameters, while plume shape and the total grain size distribution (TGSD) are influential in local sensitivity analysis (LSA), whereby one input parameter is changed at a time, but not as influential in GSA. These previous studies focused on limited subsets of input parameters. In contrast, this study applies GSA to the full input parameter space in a forecasting context (i.e., under no prior assumptions), capturing the effects of complex interactions that may significantly influence model output.

We apply the two main GSA methods: Sobol' indices (Sobol', 1990) and the extended Fourier Amplitude Sensitivity Test (eFAST) (Saltelli et al., 1999). Both Sobol' indices and eFAST provide first-order (S_i^{S} for Sobol', S_i^{F} for eFAST) and total-order (T_i^{S} for Sobol', T_i^{F} for eFAST) indices for each input parameter. Consider a model for a scalar $y = f(x)$, where x is a vector of k probabilistic input parameters. Sobol' indices are defined as:

$$S_i^{\mathrm{S}} = V[E(y|x_i)]/V(y)$$

which represents the proportion of the total variation attributable to variations in x_i. The total-order index is given by:

$$T_i^{\mathrm{S}} = 1 - V[E(y|x_{\sim i})]/V(y)$$

The eFAST method is analogous to Sobol' indices but variance is estimated as Z_i^2 via Fourier coefficients as described in the Appendix (Appendix A). The first-order index is expressed as:

$$S_i^{\mathrm{F}} = Z_i^2/Z_{total}^2$$

while the total-order index is given by:

$$T_i^{\mathrm{F}} = 1 - (Z_{\sim i}^2/Z_{total}^2)$$

where $Z_{\sim i}$ represents the summed complementary set of parameters (i.e., all parameters except x_i). See Appendix A for a more in-depth explanation.

For both Sobol' and eFAST, larger S_i and T_i values indicate a larger influence on model output variance. Given the high dimensionality and number of input parameters in the models, we use a 5% significance threshold for S_i rather than considering the relative number of input parameters per model (Puy et al., 2023). Sobol' indices provide a more comprehensive analysis by also providing second- and higher-order indices (Saltelli et al., 2008) (Appendix A). eFAST requires substantially fewer simulation runs to explore the entire input space compared to Sobol' indices, while still providing valuable sensitivity information (Saltelli et al., 2008). This meant that eFAST could be used for both our GSA of Tephra2 and Fall3D, while Sobol' indices is only applied to Tephra2.

3.2. Meteorological conditions

It is known that meteorological conditions play a major role in the deposition of ash (Bursik, 2001; Bonadonna and Costa, 2013; Mulena et al., 2016). No equivalent method exists for the stochastic sampling of wind, thus, we follow the advice of Phillips et al. (2023) and consider 12 meteorological conditions, corresponding to the 12 synoptic patterns in Aotearoa New Zealand Kidson (2000). These weather types are nationwide patterns that fall into three regimes: troughs (low-pressure systems), the trough group, high-pressure systems to the north and west-east flow in the south, the zonal group, and blocking patterns with high-pressure systems to the south, the blocking group.

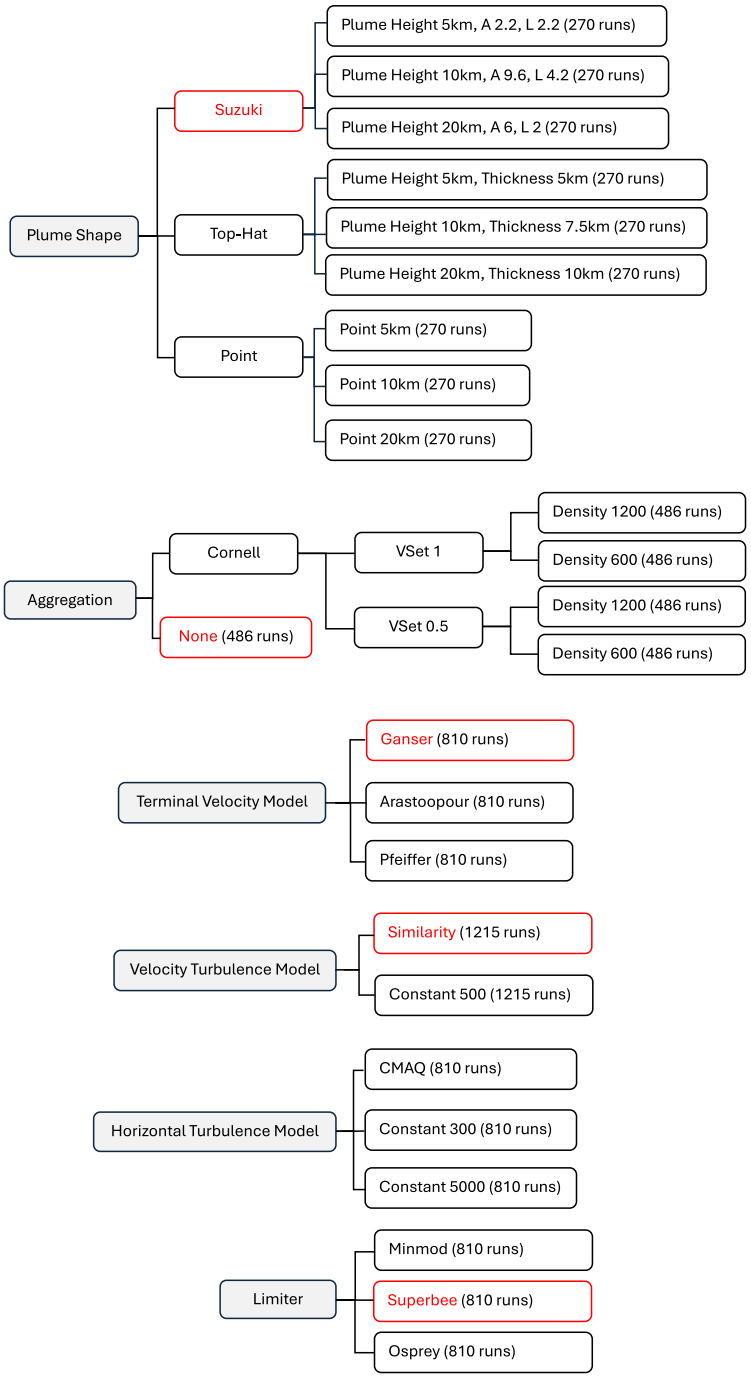


Fig. 1. Schematic of Fall3D categorical variables, totalling 2430 unique combinations. The red text indicates input variables that were kept constant in the GSA analyses as a result of the preliminary work.

Table 2
Influential and non-influential input parameters identified in sensitivity analyses of volcanic ash transport and dispersion models for tephra deposition.

Input Parameter	Scollo et al. (2008a)	Komorowski et al. (2008)	Scollo et al. (2008b)	Pardini et al. (2022)
Sensitivity Analysis Type	Local Sensitivity Analysis	Local Sensitivity Analysis	Global Sensitivity Analysis	Global Sensitivity Analysis
Number of Simulations	96 over three VATDM	93	7168	N/A
VATDM Explored	Fall3D, HAZMAP, TEPHRA	HAZMAP	TEPHRA	HYSPLIT
Total Erupted Mass/Mass Eruption Rate	Influential	Influential	Influential	N/A
Plume Height	Influential	Influential	Influential	Most Influential
Plume Shape	Influential	N/A	Non-influential	N/A
Total Grain Size Distribution	Influential	Influential	Influential	Least Influential
Particle Density	N/A	N/A	Non-influential	N/A
Wind Speed and Direction	N/A	Influential	N/A	Influential

Using stratified sampling (Cochran, 1977), we treated the 12 Kidson types as strata and selected one 24-hour period per Kidson type from February 2016 (Fig. 2). Each selected period was chosen to ensure that the Kidson type remained consistent throughout. We aggregate the GSA results from the 12 Kidson types to produce one overall result, as to provide a general overview of input parameter influence across different meteorological conditions. As different meteorological conditions may influence our GSA analysis, we also explore GSA outputs in both weak and strong wind fields (Fig. 2). Since the maximum eruption duration in our study is 12 h or less, we allowed an additional 12 h for finer grained particles to deposit. For Tephra2, the first hour in each Kidson type was used as the model does not consider temporal variability in weather.

Tephra2 uses minimal meteorological data, requiring only a text file of atmospheric height, wind speed, and direction that is assumed constant across the computational domain (Connor et al., 2011). Fall3D uses more complex meteorological data such as numerical weather prediction and reanalysis datasets, interpolating the required meteorological data (wind field, air temperature, friction velocity, atmospheric boundary layer height, and Monin–Obukhov length (turbulence height)) from the original grid driving the meteorological models to the computational domain (Folch et al., 2020). For this study, we use the European Centre for Medium-Range Weather Forecasts (ECMWF) hourly ERA5 reanalysis dataset, which spans from 1979 onwards with a vertical resolution of 37 pressure levels (Hersbach et al., 2020). For Fall3D, we use the ERA5 reanalysis data in NetCDF format with a spatial grid at half the resolution (0.125° × 0.125°) (Folch et al., 2012) (Fig. 3a). For Tephra2, we extract the necessary data from the ERA5 reanalysis data and run simulations on a spatial "grid" of the sampled locations and eruptive vent (Fig. 3b).

3.3. Practical implementation

We constrain our approach to an imminent eruption (i.e., meteorological conditions are known and assumed to be as accurate as possible) at Taranaki Mounga (Mount Taranaki), Aotearoa New Zealand where there is an estimated 30%–50% chance of an explosive eruption in the next 50 years (Damaschke et al., 2018). We sample deposit densities at locations across the region within 30 km of the vent, at 5 km, 10 km, and 30 km distances in 16 cardinal directions, with Fall3D sampling restricted to four cardinal directions at 5 km and no sampling at 10 km due to the grid size (Fig. 3). A similar approach of using 16 cardinal directions has been employed in previous studies (Macedonio et al., 2008). While Taranaki Mounga is used here as a case study, our study can be applied to any volcano worldwide, provided that meteorological data are available.

For our GSA, we use the `sensobol` package in R for the Sobol' indices (Puy et al., 2021), and investigate up to third-order interactions. The number of simulation runs generated for third-order Sobol' indices follows: $(N(k+2))+(k!/(2!(k-2)!)\ \mathrm{x}\ N)+(k!/(3!(k-3)!)\ \mathrm{x}\ N)$, where N is the initial sample size of the base sample matrix (typically varying from a few hundred to a few thousand Saltelli et al., 2008), and k is the number of input parameters. For Tephra2, we use $N = 1500$, resulting in 265,500 runs (per wind). We do not conduct Sobol' indices GSA on Fall3D due to computational limitations. Running the model a sufficient number of times to obtain a valid GSA result would have required several months worth of simulations.

We use the `SALib` package in Python to conduct our eFAST GSA (Herman and Usher, 2017). For eFAST, the number of simulations run is $N \times k$, where N is the number of samples to generate and k is the number of input parameters. For Tephra2 we use $N = 500$ ($k = 10$, resulting in 5000 runs per wind) and for Fall3D $N = 200$ ($k = 9$, resulting in 1800 runs per wind). The different values of N between eFAST and Sobol' for Tephra2 is due to Sobol' indices needing more repetitions to obtain second- and third-order indices. The different number of simulation runs between Tephra2 and Fall3D is based on the computational time it takes to run Fall3D. To conduct GSA on Fall3D, we used the Aotearoa New Zealand eScience Infrastructure (NeSI) High-Performance Computing (HPC) system.

4. Results

Here we show the results of the GSA analysis sampling tephra deposition at three distances and in 16 cardinal directions around Taranaki Mounga for Tephra2 (Fig. 3a), and two distance in four and 16 cardinal directions for Fall3D (Fig. 3b) based on 12 different wind scenarios (Fig. 2).

4.1. Weak vs strong wind

We first look at the Sobol' indices and eFAST outputs for two wind fields that convey a weak and strong wind (referred to here as a weak and strong wind scenario; Kidson type HE and T respectively (Fig. 2)).

For Tephra2 and Fall3D, we found that the output variance is concentrated over fewer input parameters in strong wind fields, while more input parameters become significant when the wind is weaker (Fig. 4,5,6). For Tephra2, only plume shape (alpha$^{(T2)}$ and beta) remains significant across both Sobol' and eFAST indices in the strong wind scenario (Fig. 4). For Fall3D, plume duration and horizontal diffusion are the only parameters deemed significant under the strong wind scenario (Fig. 6).

In the weak wind scenario for Tephra2, plume shape (alpha$^{(T2)}$) is consistently significant, especially off axis of the major direction of tephra deposition (ENE to ESE), while its influence decreases with distance (Fig. 4). The diffusion coefficient remains significant across all distances in eFAST but only in distal areas (30 km) for Sobol' indices. Plume height, fall time threshold (FTT), plume height, and plume shape (beta) are significant in proximal areas (5 km) in Sobol' indices, but lose significance distally. In eFAST the diffusion coefficient remains significant in distal areas.

For Fall3D, in weak wind conditions, median and standard deviation grain size are the only significant parameters in proximal areas, while the diffusion coefficient gains significance in distal regions (Fig. 6).

4.2. Tephra2

In the following sections, we look at the first-order sensitivity indices: S_i^S (Sobol') and S_i^F (eFAST), as well as the total-order indices: T_i^S (Sobol') and T_i^F (eFAST), over the whole GSA space, providing insights into the average trends over all meteorological conditions. Additionally, we also examine the S_i^S, S_i^F, T_i^S and T_i^F values in the major direction of tephra deposition, selecting values for the direction closest (out of 16 cardinal directions) to the predominant wind direction. It is important to investigate whether the results of our GSA vary in these regions of heightened tephra deposition.

4.2.1. First-order indices

The results of the GSA looking at the first-order indices (S_i) for Tephra2 for all sampling locations are shown in Fig. 7.

The most notable difference is in plume shape alpha$^{(T2)}$ (which controls how much mass is contributed to the upper part of the eruption column), which significantly influences output variance in 61%–82% of simulations in eFAST but only 33%–55% in Sobol' indices. eFAST shows a decline in the significance of median grain size and plume shape alpha$^{(T2)}$ with distance, whereas Sobol' indices reveal no consistent trend for plume shape alpha$^{(T2)}$.

Grain size standard deviation, lithic and pumice density, mass, plume height, and FTT are not significant first-order input parameters in either GSA method, suggesting their influence on tephra deposition is indirect or dependent on interactions with more dominant parameters like plume shape.

Examining the significant input parameters along the main tephra deposition axis, median grain size and plume shape (alpha$^{(T2)}$ and beta) remain the primary drivers of output variance.

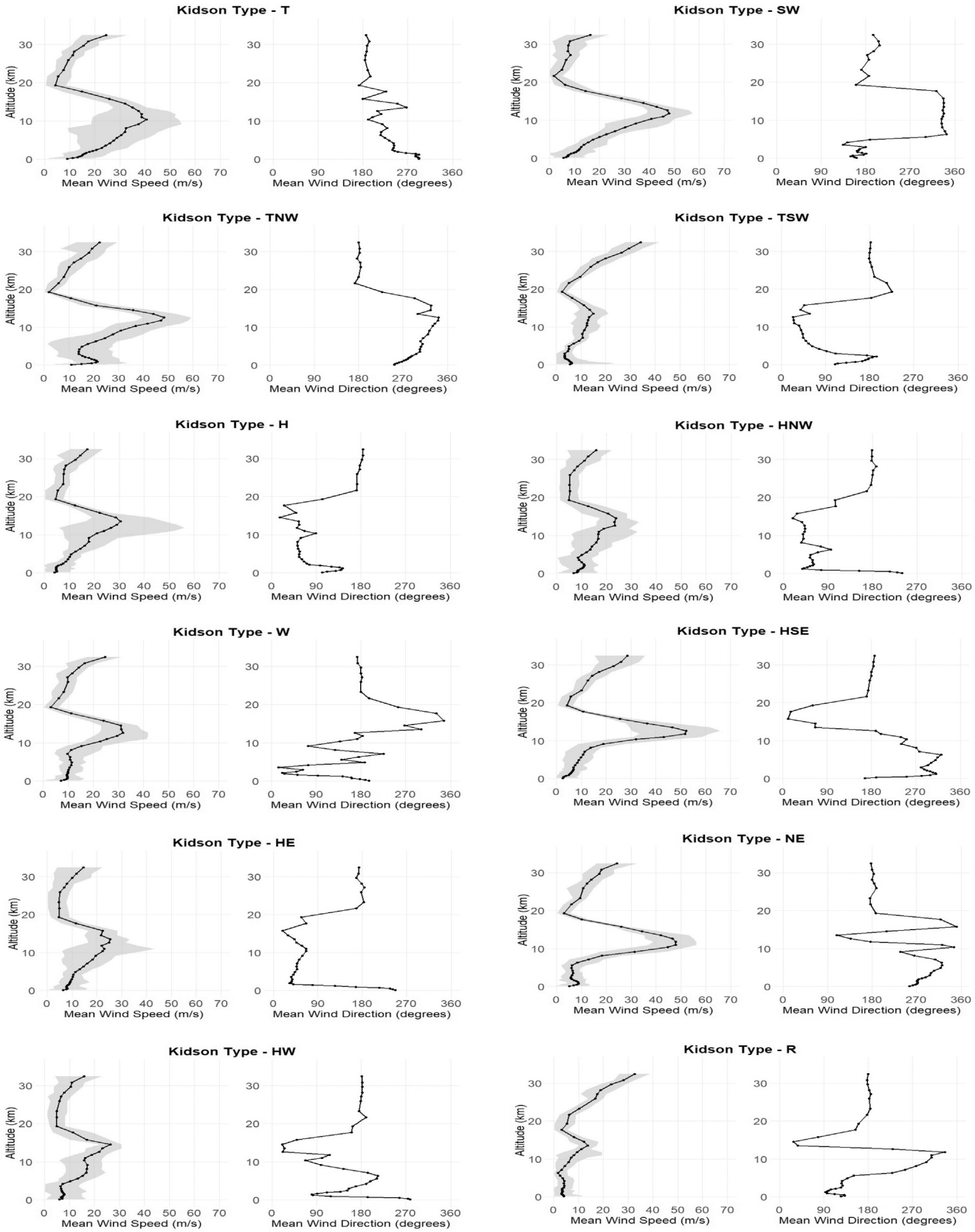
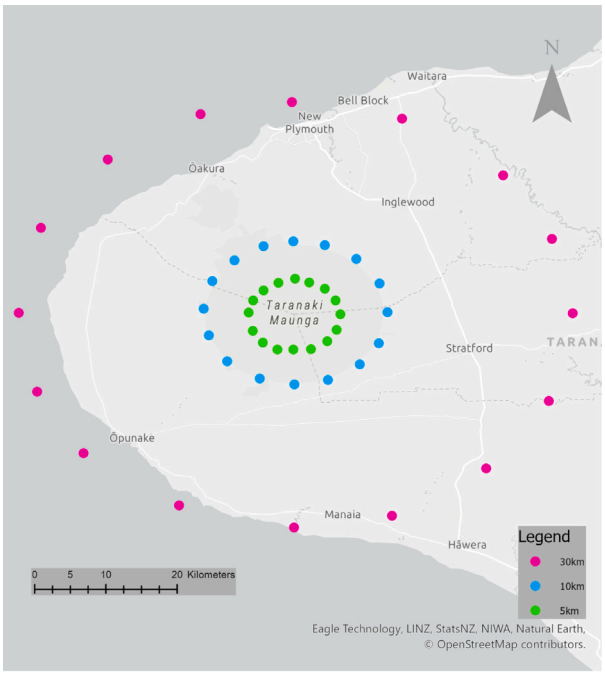
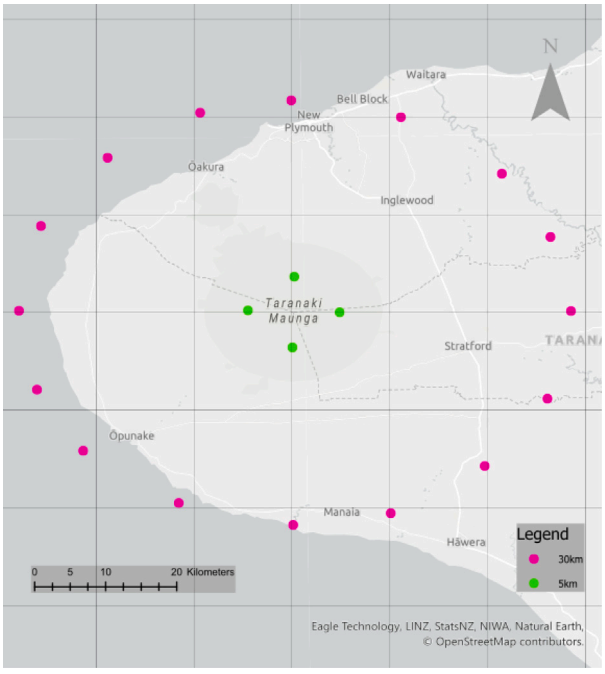


Fig. 2. Mean wind speeds (over all areas and time) and mean wind directions (from North) at different altitudes for the 12 24-hour long Kidson Types at Taranaki Mounga (39.2968° S, 174.0634° E), calculated over the selected 24-hour periods. Grey areas highlight the minimum and maximum wind speeds at different altitudes calculated over the selected 24-hour periods. Trough group: T, SW, TNW, TSW. Zonal group: H HNW, W. Blocking group: HSE, HE, NE, HW, R.



(a)



(b)

Fig. 3. Sampling locations at Taranaki Mounga (Mount Taranaki), Aotearoa New Zealand. (a) Tephra2 sampling locations at 5 km (green), 10 km (blue), and 30 km (purple) from the vent in 16 cardinal directions. (b) Fall3D sampling locations at 5 km in 4 cardinal directions (green) and 30 km in 16 cardinal directions (purple). The grey grid represents the ERA5 spatial resolution at 0.125° × 0.125°. Both models do not differentiate between tephra deposition on land or water.

4.2.2. Total-order indices

For the purpose of our analysis, the T_i are discussed in terms of rank, ranging from one (representing the lowest T_i value) to ten (representing the highest T_i value), as we investigate ten input parameters in Tephra2. Higher T_i indicates a greater influence on model output variance when considering interactions with all other input parameters.

Analysis of total-order indices (T_i) in Tephra2 revels consistent trends (Fig. 8). Median grain size and plume shape (alpha$^{(T2)}$) consistently rank highest.

Discrepancies between eFAST and Sobol' indices appear for grain size standard deviation, plume height, diffusion coefficient, FTT, and plume shape (beta). In eFAST, grain size standard deviation and diffusion coefficient rank relatively high (7-7.5 and 7.5-8, respectively) but drop in Sobol' indices (1–3 and 2–3.5). Conversely, FTT and plume shape (beta) increase in rank in Sobol' indices, shifting from 2–3.5 to 3–6.5 and from 1–3 to 4–8, respectively. These trends persist when considering T_i values specifically along the main axis of tephra deposition.

4.3. Fall3D

For Fall3D, we look at the first-order sensitivity indices S_i^F (eFAST) and total-order indices T_i^F, over the whole GSA space, providing insights into the average trends over all meteorological conditions. Additionally, we also examine the S_i^F and T_i^F values in the major direction of tephra deposition. This selection is based on four cardinal directions at a 5 km sampling distance or 16 cardinal directions at a 30 km sampling distance.

4.3.1. First and total-order indices

For Fall3D, first-order (S_i) outputs indicate that horizontal diffusion is the most significant input parameter, with its influence increasing with distance from the vent (Fig. 9). Median and standard deviation grain size, minimum and maximum density, and plume duration also contribute notably to output variance. In contrast, plume height and plume shape have minimal influence. These trends remain consistent when considering only sampling locations along the main axis of tephra deposition.

Total-order indices (T_i) for Fall3D also show that horizontal diffusion is the most significance input parameter (Fig. 9). Plume shape parameters alpha$^{(F3D)}$ and lambda gain more influence compared to their S_i values, with alpha$^{(F3D)}$ becoming a top-ranked input parameter, though lambda remains among the least influential, similar to minimum density.

When considering only the major axis of tephra deposition, ranking trends remain similar. However, plume height decreases in rank, alongside the low ranking of maximum density and plume shape lambda.

There is no overall variation in deposit density distribution when considering categorical variables in Fall3D when input parameters are perturbed along the major tephra deposition axis. However, plume shape distributions introduce slight variations in median deposit densities (Fig. 10).

Individual runs based on a baseline scenario, where only categorical input were altered, reveal that the largest variations in deposit density (red crosses) occur within the plume shape distribution category, particularly for plume heights exceeding 10 km. Minor variations also appear in individual runs for terminal velocity models (Fig. 10), while aggregation, vertical and horizontal turbulence models, and limiter categories show no significant density variation within their respective categories.

5. Discussion

This section discusses the key findings in our GSA analysis, including an exploration of higher-order Sobol' indices. We also examine the role of categorical variables, address the study's limitations, and consider the use of our VATDMs Tephra2 and Fall3D.

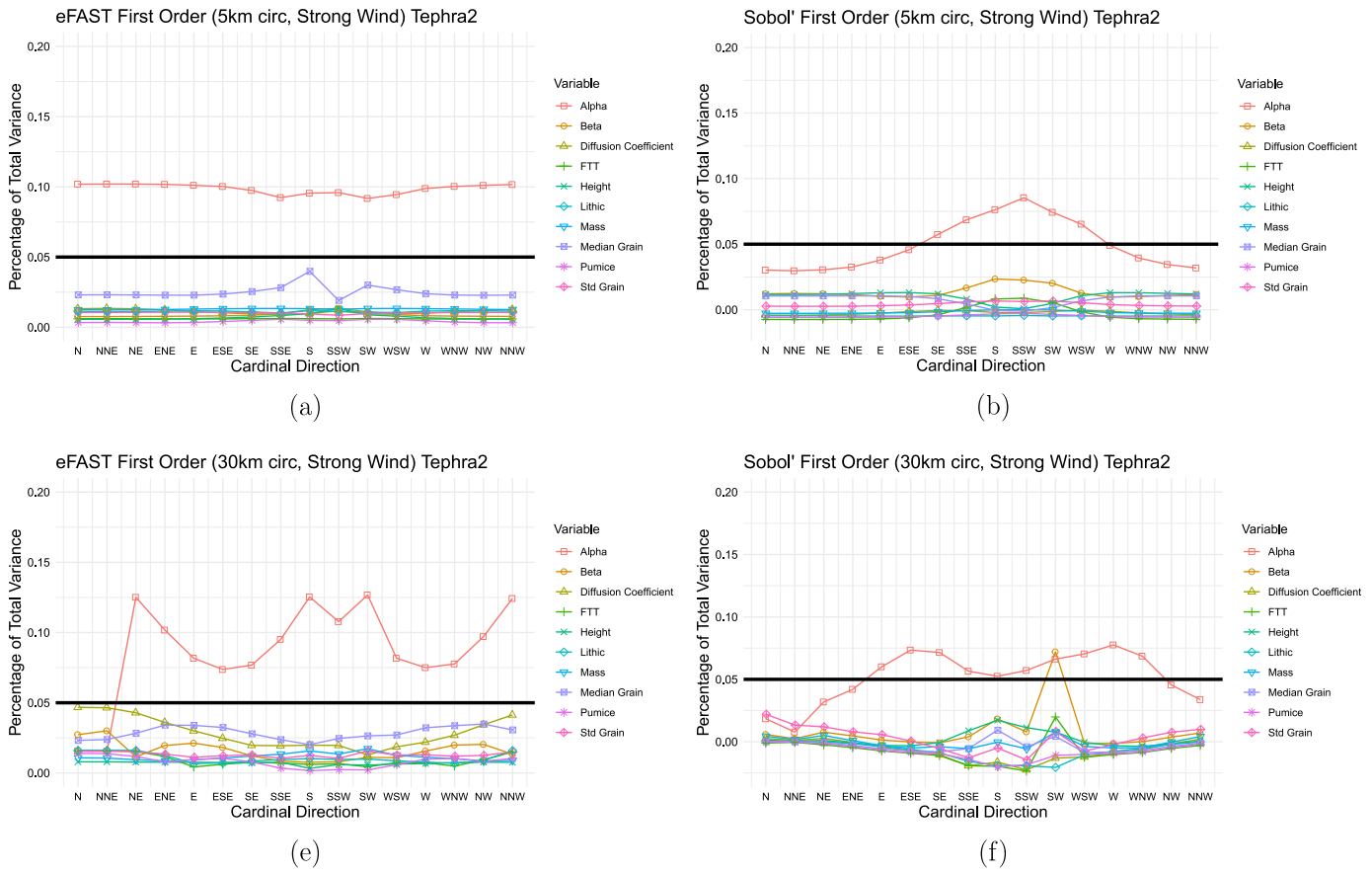


Fig. 4. S_i values from a strong wind scenario for Tephra2. Each coloured line and point shape represents an input parameter examined in the GSA analysis. Points indicate S_i values at each cardinal direction. Points above the solid black line represent significant S_i values (more than 5%). (a) 5 km Sampling distance via eFAST; (b) 5 km Sampling distance via Sobol' indices; (c) 10 km Sampling distance via eFAST; (d) 10 km Sampling distance via Sobol' indices; (e) 30 km Sampling distance via eFAST; (f) 30 km Sampling distance via Sobol' indices.

5.1. GSA

The objective of this study is to understand what input parameters contribute the most and least amount of variance in model outputs, which in turn can correctly represent uncertainty in forecasts, as well as reduce the number of simulations needed to produce probabilistic forecasts or ensembles. Our GSA analysis has shown that input parameters total grain size distribution, plume shape and diffusion (diffusion coefficient and horizontal diffusion) have significant first- and total-order influences on the output variance in both Tephra2 and Fall3D. In addition, plume duration (only an option in Fall3D) was also a significant input parameter. The other input parameters, mass (Tephra2 only), particle density, plume height, and FTT (Tephra2 only) were found to have only total-order influences on output variance for Tephra2 and Fall3D. These findings highlight that when designing forecasts, particular attention should be paid to significant input parameters as they contribute most to the variance in model outputs. These results differ from previous research, as Scollo et al. (2008b) (Table 2) found that in VATDM TEPHRA, using Sobol' indices, lithic and pumice densities and plume shape were not significant input parameters.

We conducted additional testing in our GSA by examining different sensitivity thresholds for S_i^S and S_i^F — 3% and 10% — compared to the 5% threshold originally used, to assess whether this would affect our conclusions. Results showed that even though the frequency of significant input parameters increased at 3% significance and decreased at 10%, the overall pattern was the same.

Our GSA results could inform the design of an optimal sampling or simulation strategy when the number of simulation runs is constrained. Both first- and total-order influences of input parameters can be utilised to identify which input parameters may be excluded from sampling designs. First-order indices, which correspond to the direct influence of individual input parameters, are particularly advantageous because their uncertainty is easier to quantify compared to the more complex and less well-understood interactions between input parameters.

5.2. Sobol' second and third-order indices

If the first-order interactions at a sampling point sum to one, it indicates no interaction between input parameters (Sobol', 1990). If they do not sum to one, second and third-order interactions account for the remaining proportion of the total-output variance. We found that the sum of first-order indices is higher in stronger wind fields and along the major deposition axis. This suggests that, in such scenarios, wind direction and speed dominate ash deposition patterns. Consequently, minimal simulations may be required in these cases, particularly when wind conditions are consistent along the major deposition axis.

Conversely, in weaker wind fields and opposite the major direction of tephra deposition, the sum of the first-order interaction terms is smaller, sometimes even negative, implying that other interactions, either between input parameters or outside of the selected input parameters used in the GSA, such as weak winds, may indirectly influence output variance. This pattern does not continue, however, for second- and third-order interactions. For both strong and weak wind fields, the sum of second-, and third-order interactions decreases with distance from the vent. This trend likely reflects the influence of input parameters, such as diffusion, which become increasingly more significant at further distance from the vent. At greater distances, the mass of ash

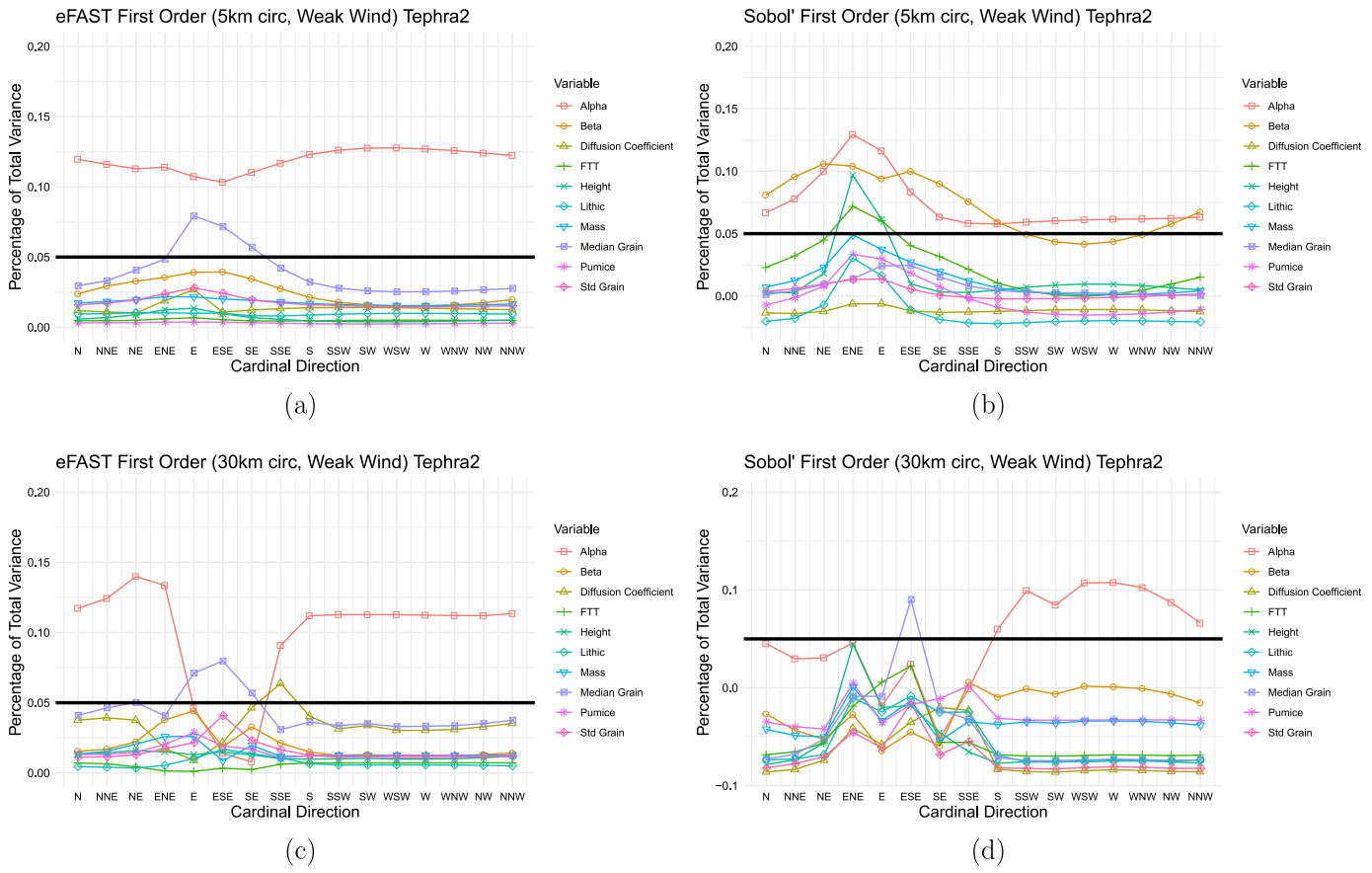


Fig. 5. S_i values from a weak wind scenario for Tephra2. Each coloured line and point shape represents an input parameter examined in the GSA analysis. Points indicate S_i values at each cardinal direction. Points above the solid black line represent significant S_i values (more than 5%). (a) 5 km Sampling distance via eFAST; (b) 5 km Sampling distance via Sobol' indices; (c) 30 km Sampling distance via eFAST; (d) 30 km Sampling distance via Sobol' indices.

and the prevalence of finer grained particles introduce more variability, potentially leading to complex interactions beyond the third order.

We also examined which input parameters and their interactions had the largest index values. By selecting the top five index values for each Kidson Type (12), by each sampling distance (3), and for each order (3), we identified a total of 540 high-value indices ($5 \times 12 \times 3 \times 3 = 540$). The first-order input parameter, plume shape alpha$^{(T2)}$, had the highest count among these, accounting for 14.6% (79/540). In comparison, 32% of the largest indices were related to median grain size, including 9.4% (51/540) as first-order contributions and the remainder from second- and third-order interactions with particle density, plume shape (alpha$^{(T2)}$ and beta), diffusion coefficient, plume height, FTT, standard deviation of grain size, and mass, ranging from 9% to 0.19%.

5.3. Categorical variables

We found that particle aggregation did not significantly affect the outputs of Fall3D. This differs from expectations, as most fine volcanic ash tends to settle out as particle aggregates (Carey and Sigurdsson, 1982), often leading to secondary thickening of deposits (e.g., Tsuji et al., 2020). Aggregation in Fall3D has been validated with the 17 September 1992 Mount Spurr and the 18 May 1980 Mount St. Helens eruptions, which showed strong agreement with observed deposit values (Folch et al., 2010). So for hindcasting, the use of aggregation models may improve model output accuracy. However, for a future unknown eruption, the extent of particle aggregation is uncertain, making it unclear whether particle aggregation model input parameter options will improve model output accuracy.

One possible explanation for the discrepancy between our results and previous studies is the consistency of the TGSD and particle density values used in our simulations. Research has shown that aggregation is sensitive to factors such as particle density, the TGSD of non-aggregated particles, and the specific aggregation model applied (Poret et al., 2017; Beckett et al., 2022).

Plume shape significantly influenced the outputs of Fall3D, with variations within individual plume shape distributions creating differences in deposit density. This finding aligns with the results of our GSA, which identified plume shape as a significant input parameter impacting output variance in both Tephra2 and Fall3D.

The other categorical input parameters: limiter, horizontal and vertical turbulence models, and terminal velocity models do not significantly affect output deposit variations. This suggests that their selection does not strongly influence the overall model output. Having multiple options within each categorical variable may still provide flexibility to achieve a closer fit during hindcasting by allowing adjustments to better match observed data for a specific eruption. However, while greater model expertise can help inform input parameter use, multiple options within categorical variables still introduce uncertainty for forecasting, as these choices may not accurately represent future conditions.

5.4. Limitations

We represented the variation in wind in the Taranaki region of Aotearoa New Zealand via Kidson types. However, the use of only 12 unique winds is unlikely to have captured the full range of meteorological conditions – which is impossible to do so, as wind is continuous – meaning that our GSA may not have fully captured input parameter interactions. The Kidson types are also based on country-wide patterns, and may not reflect all meteorological patterns seen at regional to local scales. A logical extension to this work would therefore be to

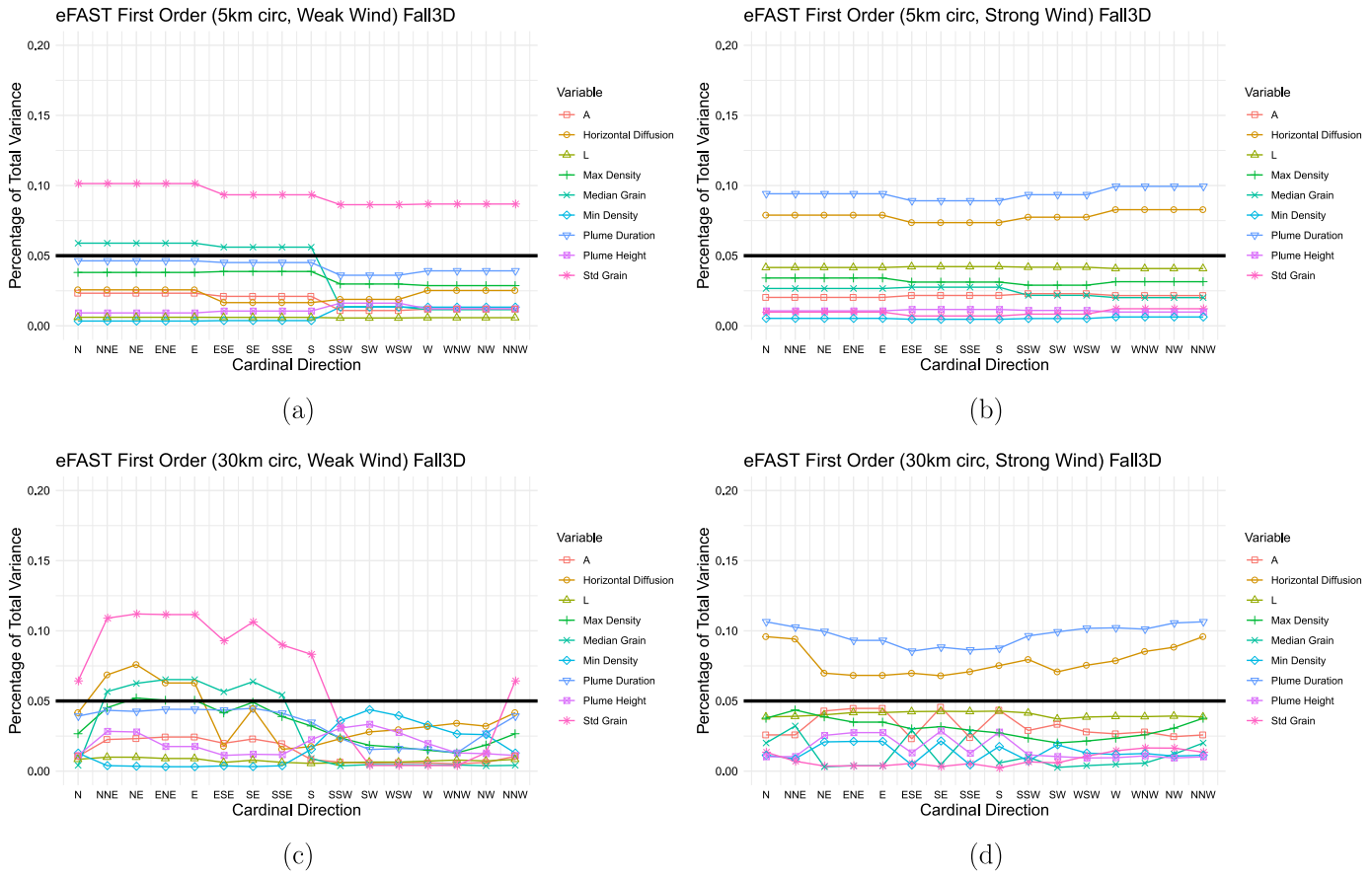


Fig. 6. S_i values via eFAST from a weak and strong wind scenario for Fall3D. Each coloured line and point shape represents an input parameter examined in the GSA analysis. Points indicate S_i values at each cardinal direction. Points above the solid black line represent significant S_i values (more than 5%). (a) Weak wind at 5 km sampling distance; (b) Strong wind at 5 km sampling distance; (c) Weak wind at 30 km sampling distance; (d) Strong wind at 30 km sampling distance.

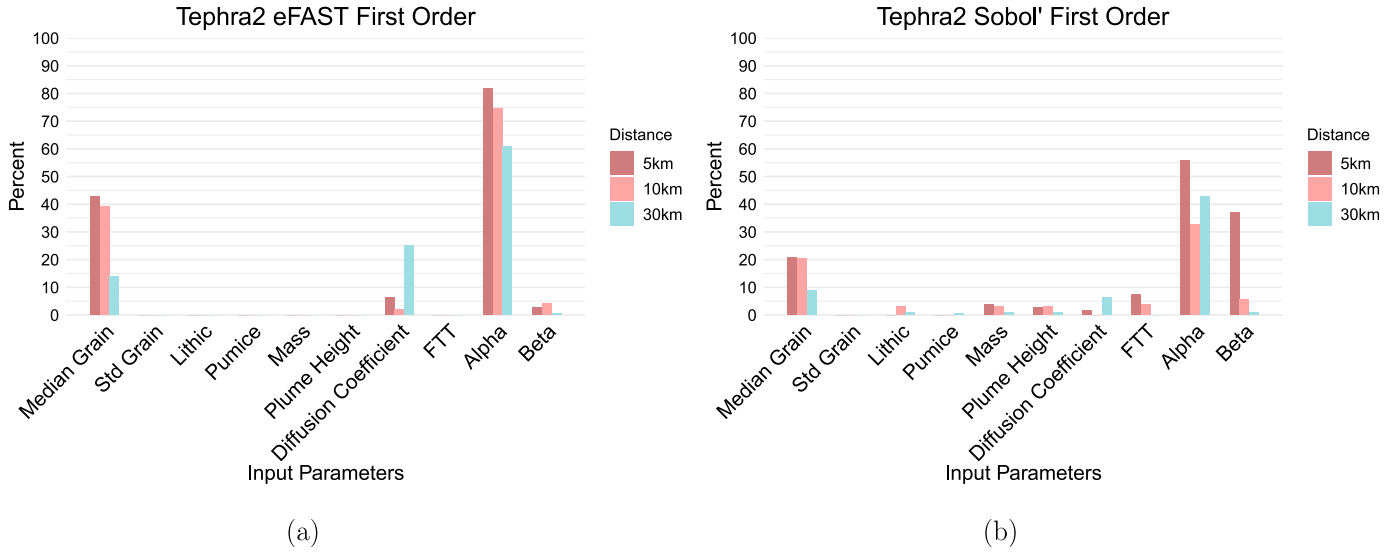


Fig. 7. Proportion of first-order S_i outputs for Tephra2 of input parameters that reach a 5% significance threshold. All winds are included. (a) S_i via eFAST incorporating 16 cardinal directions and sampling points. (b) S_i via Sobol' indices incorporating 16 cardinal directions and sampling points.

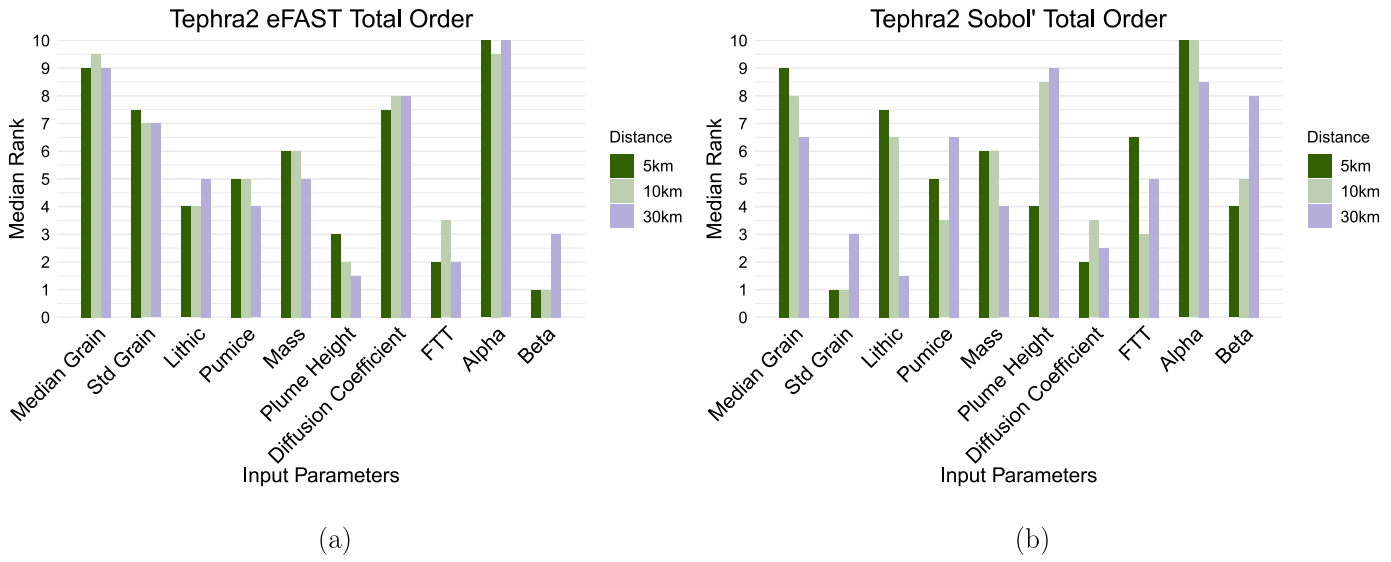


Fig. 8. Total-order T_i outputs for Tephra2. Fractional values result from averaging over multiple winds. All winds are included. (a) The median rank of T_i^{F} via eFAST incorporating 16 cardinal directions. (b) The median rank of T_i^{S} via Sobol' indices incorporating 16 cardinal directions.

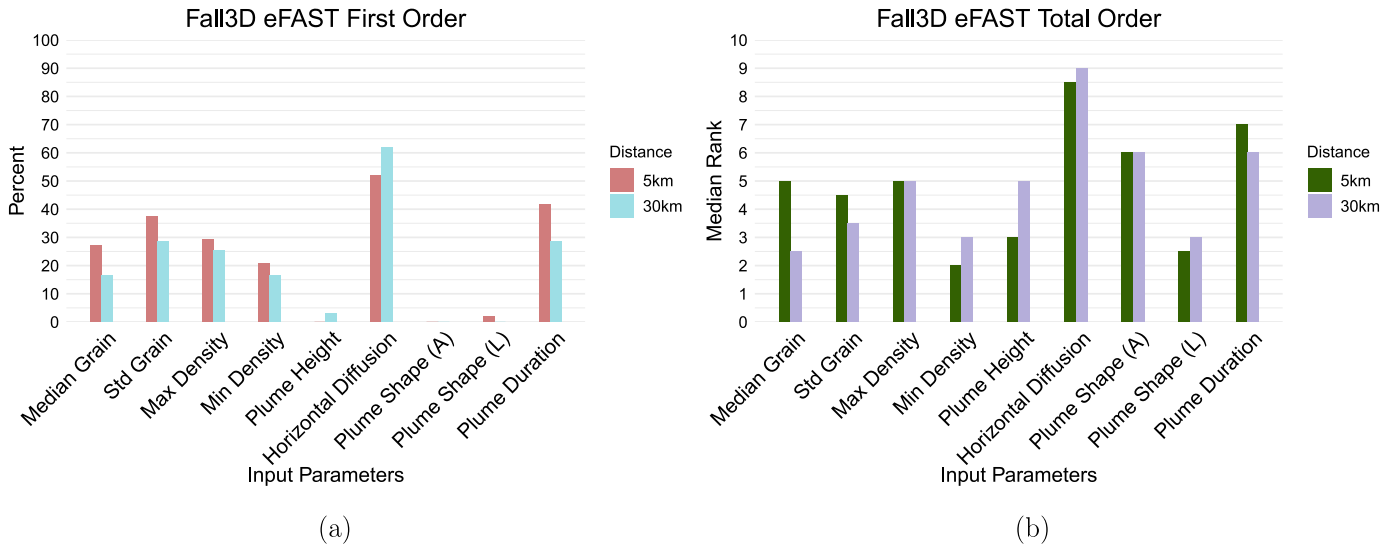


Fig. 9. Proportion of first-order S_i and total-order T_i outputs for Fall3D. Fractional T_i values result from averaging over multiple winds. All winds are included. (a) S_i^{F} via eFAST incorporating all cardinal directions and sampling points. (b) The median rank of T_i^{F} via eFAST summed in all cardinal directions.

expand the wind input variability, either through multiple days under each Kidson type or by stratifying wind sampling with location-specific patterns.

Both eFAST and Sobol' indices analyses identified the same key input parameters that significantly influence output variance in both Tephra2 and Fall3D. eFAST was used for both Tephra2 and Fall3D due to its computational efficiency. While Sobol' indices can provide more detailed insights into higher-order interactions, the occurrence of negative Sobol' indices suggests that eFAST is a more robust approach, particularly when analysing models with a high-dimensional input parameter space.

Sobol' indices produced negative S_i^{S} values (55% of the time). Negative S_i^{S} values can occur due to sampling variability due to an insufficient sample size (Saltelli et al., 2010). Even with over 260,000 simulations per wind field in Tephra2, negative S_i^{S} indicate that the number of samples to generate should be increased as S_i^{S} by definition should not be negative. However, increasing the number of simulations further may not get rid of negative values entirely, due to the high-dimensional input parameter space. Scollo et al. (2008b) also encountered this limitation, reporting negative Sobol' indices with only 4608 simulation runs.

5.5. Considerations of the models used (Tephra2 and Fall3D)

The type of GSA analysis presented in this paper can be applied to any volcanic hazard model. We decided to test Tephra2 and Fall3D as they have different benefits and limitations, especially in forecasting. Tephra2 is computationally efficient, with a single simulation typically taking less than a second, allowing thousands of simulations to be run in minutes on a standard computer. In contrast, Fall3D is more computationally intensive; a single simulation took between 10 and 18 min for short-duration eruptions (less than 12 h) with coarse grid sizes, even when using batch processing on a supercomputer system (NeSI). These differences in simulation time are largely due to the broader application of Fall3D, which simulates both ash deposition and atmospheric concentrations and incorporates large meteorological datasets.

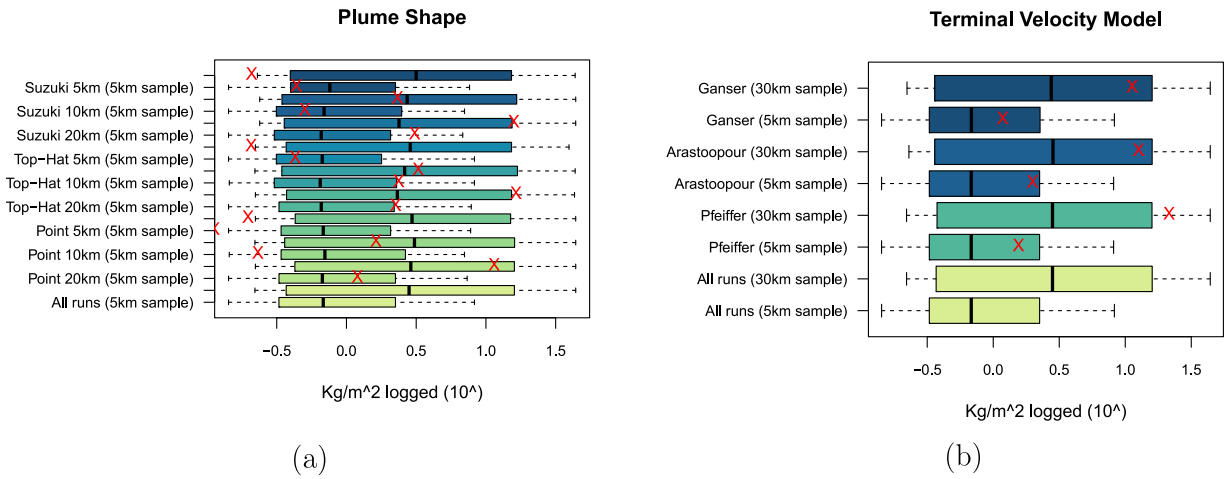


Fig. 10. Categorical simulation outputs with logged deposit density in the major direction of tephra deposition. Red crosses are individual runs on a base run where only the categorical input is changed.

Tephra2 simplifies the user experience by directly embedding different physical processes, such as turbulence and diffusion calculations within the code. For example, atmospheric diffusion is assumed to be Gaussian over the whole computational domain (Connor et al., 2008). Other notable assumptions of the model include a Gaussian particle size distribution, particles of every size are well-mixed throughout the entire eruption plume, and a layered atmosphere where windspeed and direction are constant within horizontal layers but allowed to vary between them (Connor et al., 2008). In addition, Tephra2 considers a consistent wind profile, which means that over large domains the winds cannot be accurately represented by a single profile. These simplifications enable the model to achieve short computing times, allowing thousands of simulations to be run in just a few minutes on standard computers.

Fall3D, by contrast, provides users the flexibility to define and manipulate more aspects of the model compared to Tephra2, such as particle shape and limiters. Additionally, Fall3D is usually used for simulating atmospheric concentrations of volcanic ash – a major concern for Volcanic Ash Advisory Centers (VAACs) (Beckett et al., 2024) – in addition to deposition.

6. Conclusions

While VATDM models such as Tephra2 and Fall3D can retrospectively simulate past eruption behaviours, identifying the key input parameters that significantly influence output variance in the context of forecasting is crucial for improving hazard forecast reliability. During or immediately before an eruption, VATDM users are constrained in the number of simulations that can be run in a fast-evolving situation. Our GSA revealed that grain size distribution, diffusion, and plume duration input parameters have the most significant impact on model output variance. Contrary to previous studies, our analysis demonstrated that the plume shape input parameters also significantly affect model output variance.

Therefore, when time is limited, VATDM users should focus on running simulations that prioritise varying these influential input parameters. Identifying these key input parameters allows users to direct efforts towards quantifying their variability, including understanding the correlations between them, which will improve the overall accuracy of the simulations. By more accurately quantifying distributions of significant input parameters, and leveraging limited simulation resources to best exploit these, more reliable hazard forecasts and better preparedness strategies can be created during volcanic eruptions. Additionally, this sensitivity information can also underpin the development of long-term tephra hazard assessments, where pre-established input parameter ranges are established in advance. As thousands of simulations are required to conduct a GSA, alternative strategies for assessing the importance of VATDM input parameters may include the use of emulators, which can be particularly useful for models with large computational demands (e.g., Harvey et al., 2018).

CRediT authorship contribution statement

Emmy Scott: Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Melody Whitehead:** Writing – review & editing, Methodology, Conceptualization. **Stuart Mead:** Writing – review & editing, Methodology, Conceptualization. **Mark Bebbington:** Writing – review & editing, Methodology. **Jonathan Procter:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Resilience to Nature's Challenges (RNC) Volcano Program, grant GNS-RNC047, and the Dr Eileen Fair Doctoral Scholarship in Earth Science. The author(s) wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high performance computing facilities, consulting support and/or training services as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme. URL https://www.nesi.org.nz. The authors would also like to thank the two anonymous reviewers for their insightful comments, which helped strengthen the manuscript.

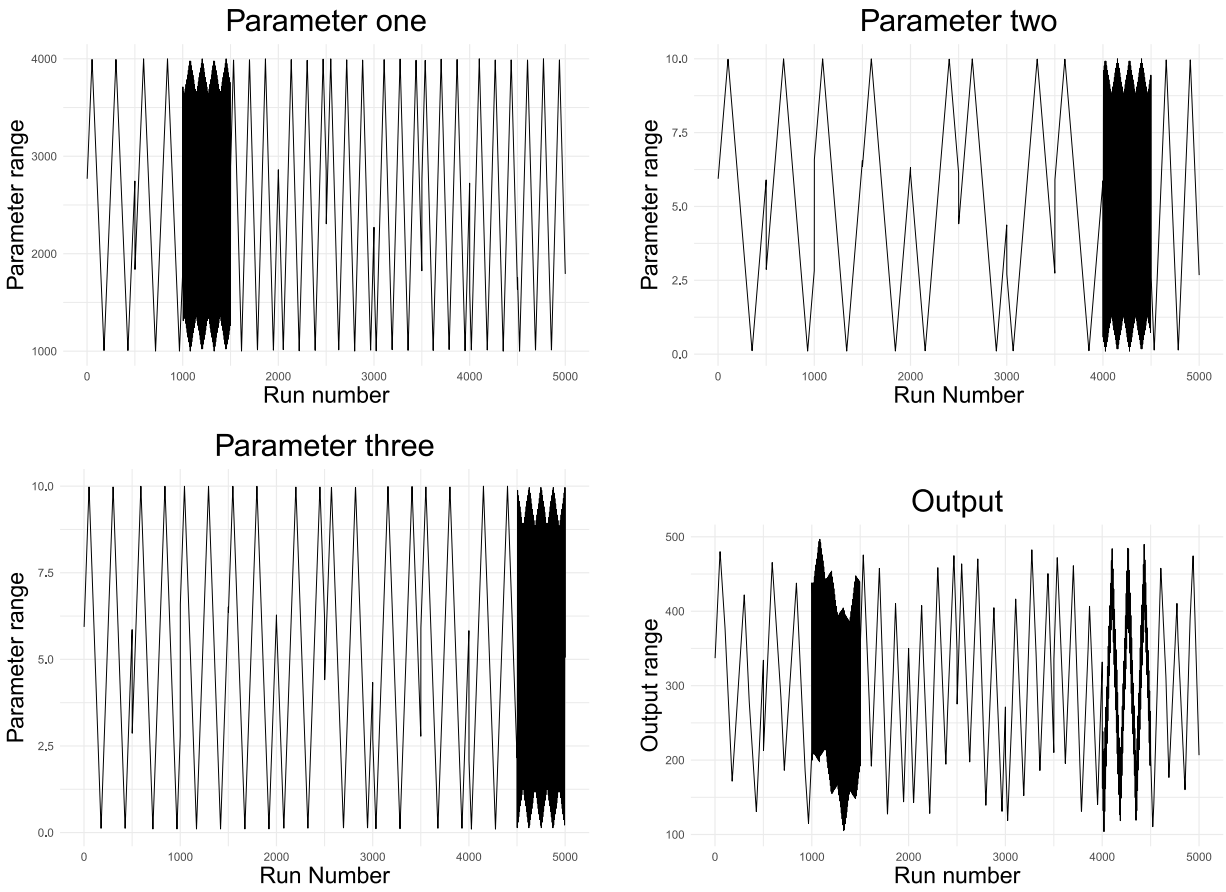


Fig. 11. Visualisation of the eFAST search curve, which shows the decomposition of output variance based on the spectral analysis of input parameters. Results show that input parameter one strongly contributes to the output variance through high spectral amplitude, input parameter two contributes moderately with a smaller amplitude, and input parameter three shows no contribution, as its spectral amplitude remains negligible.

Appendix A

A.1. Sobol' indices

Sobol' indices (Sobol', 1990) is a Monte Carlo-based calculation that quantifies the contribution of each input variable to the output variance of a model. Given a model in the form:

$$y = f(x), x = (x_1, x_2, \dots, x_k) \in R^k \tag{3}$$

where y is a scalar output and $x_1, \dots, x_k$ are k independent input parameters described by probability distributions, we can calculate the proportion of total variance conveyed to y by each input parameter x_i, known as the first-order effect (S_i^S).

$$S_i^S = V[E(y|x_i)]/V(y) \tag{4}$$

where

$$V[y] = \sum_{i=1}^{N}(y_i - \bar{y})^2/(N-1) \tag{5}$$

where N is the sample size, y_i is the output from the ith combination of inputs, and $\bar{y}$ is the mean output across all combinations of input parameters, and $E(y|x_i)$ is the expected value of y given a fixed value of x_i

We can also extract the proportion of variance contributed by the interaction between pairs of inputs (second-order effect, S_{ij}), triplets of input parameters (third-order effect, S_{ijl}), etc., up to the kth-order interaction, e.g.,:

$$S_{ij} = V[E(y|x_i, x_j)] - V[E(y|x_i)] - V[E(y|x_j)]/V(y) \tag{6}$$

For a model with three input parameters, the total the variance can be expressed as, for example: $S_1 + S_2 + S_3 + S_{1,2} + S_{1,3} + S_{2,3} + S_{1,2,3} = 1$.

S_i^S can then be used to rank input parameters by their relative first-order contribution to model output variation. When $\sum_{i=1}^{k} S_i^S = 1$, the model is additive, i.e., the total variance (y) can be fully decomposed as the sum of first-order effects, indicating that there is no interaction between input parameters. However, this is rarely the case in complex models (Puy et al., 2023), and the sum of the first-order indices usually falls well short of total model output variance.

Total-order indices T_i^S, which measure the first-order effects of x_i together with its interactions with all the other input parameters provide information on the non-additive features of a model (Saltelli et al., 2010).

$$T_i^S = 1 - V[E(y|x_{\sim i})]/V(y) \tag{7}$$

where $V[E(y|x_{\sim i})]$ is the portion of variance that is explained given all input parameters except x_i. For a three-parameter model, $T_1 = S_1 + S_{1,2} + S_{1,3} + S_{1,2,3}$.

A.2. eFAST - extended fourier amplitude sensitivity test

eFAST is a variance decomposition method, analogous to ANOVA (Cukier et al., 1978). A transformation (sinusoidal) function, $x = f(j), j = 1, 2, \dots, M$, is used to convert input parameter values to values along a search curve, based on the sample number M (number of repetitions per input combination) (Fig. 11) — this essentially turns a probability density function into a signal that can then be traced back to the output variance. Using Fourier coefficients, eFAST decomposes the variance of the model output, determining what fraction of the variance can be explained by the variation in each input parameter.

The variance contribution of each input parameter (Eq. (5)) (i) is estimated from the Fourier coefficients and its unique frequency (and harmonics of that frequency):

$$Z_i^2 = 2(A_j^2 + B_j^2) \tag{8}$$

where A_j and B_j are the Fourier coefficients corresponding to the input parameter's frequency.

The total variance of the model output is then computed as the sum of these variance estimates across all input parameters:

$$Z_{total}^2 = \sum_i Z_i^2 \tag{9}$$

Finally, the first-order sensitivity index (S_i^{F}) for a given input parameter are calculated as the variance attributed to that input parameter divided by the total variance of the model output:

$$S_i^{\mathrm{F}} = Z_i^2 / Z_{total}^2 \tag{10}$$

To calculate the total-order sensitivity index (T_i^{F}) of a given input parameter, eFAST calculates the summed sensitivity index contributions of the entire complementary set of input parameters (i.e., all input parameters except x_i). T_i^{F} is then derived as the proportion of variance remaining after accounting for the contribution of this complementary set:

$$T_i^{\mathrm{F}} = 1 - Z_{\sim i} \tag{11}$$

where $Z_{\sim i}$ represents the contribution of the complementary set to the total variance:

$$Z_{\sim i} = Z_{total} - Z_i^2 \tag{12}$$

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jvolgeores.2025.108393.

Data availability

Data will be made available on request.

References

Aubry, T., Engwell, S., Bonadonna, C., Carazzo, G., Scollo, S., Van Eaton, A.R., Taylor, I.A., Jessop, D., Eychenne, J., Gouhier, M., Mastin, L.G., Wallace, K.L., Biass, S., Bursik, M., Grainger, R.G., Jellinek, A.M., Schmidt, A., 2021. The Independent Volcanic Eruption Source Parameter Archive (IVESPA, version 1.0): A new observational database to support explosive eruptive column model validation and development. J. Volcanol. Geotherm. Res. 417, 107295. http://dx.doi.org/10.1016/j.jvolgeores.2021.107295.

Beckett, F.M., Bensimon, D., Crawford, A., Deslandes, M., Guidard, V., Hort, M.C., Jeoffrion, M., Kristiansen, N., Lucas, C., Nishijo, A., Osores, S., Renard, E., Servranckx, G., Snee, E., Trancoso, R., Vazquez, E., 2024. VAAC Model Setup Tables 2023. Technical Report, http://dx.doi.org/10.5281/zenodo.10671098.

Beckett, F.M., Rossi, E., Devenish, B., Witham, C., Bonadonna, C., 2022. Modelling the size distribution of aggregated volcanic ash and implications for operational atmospheric dispersion modelling. Atmos. Chem. Phys. 22 (5), 3409–3431. http://dx.doi.org/10.5194/acp-22-3409-2022.

Bonadonna, C., Costa, A., 2013. Modeling tephra sedimentation from volcanic plumes. Model. Volcan. Process. 173–202. http://dx.doi.org/10.1017/CBO9781139021562.009.

Bonadonna, C., Ernst, G.G.J., Sparks, R.S.J., 1998. Thickness variations and volume estimates of tephra fall deposits: the importance of particle Reynolds number. J. Volcanol. Geotherm. Res. 81 (3), 173–187. http://dx.doi.org/10.1016/S0377-0273(98)00007-9.

Bonasia, R., Costa, A., Folch, A., Macedonio, G., Capra, L., 2012. Numerical simulation of tephra transport and deposition of the 1982 El Chichón eruption and implications for hazard assessment. J. Volcanol. Geotherm. Res. 231–232, 39–49. http://dx.doi.org/10.1016/j.jvolgeores.2012.04.006.

Bursik, M., 2001. Effect of wind on the rise height of volcanic plumes. Geophys. Res. Lett. 28 (18), 3621–3624. http://dx.doi.org/10.1029/2001GL013393.

Byun, D., Schere, K.L., 2006. Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. Appl. Mech. Rev. 59 (2), 51–77. http://dx.doi.org/10.1115/1.2128636.

Carey, S.N., Sigurdsson, H., 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. http://dx.doi.org/10.1029/JB087iB08p07061.

Carey, S., Sparks, R.S.J., 1986. Quantitative models of the fallout and dispersal of tephra from volcanic eruption columns. Bull. Volcanol. 48 (2), 109–125. http://dx.doi.org/10.1007/BF01046546.

Cochran, W.G., 1977. Sampling Techniques, third ed. John Wiley & Sons, New York.

Connor, L.J., Connor, C.B., 2006. Inversion is the key to dispersion: understanding eruption dynamics by inverting tephra fallout. Stat. Volcanol. 1, http://dx.doi.org/10.1144/IAVCEI001.18.

Connor, L.J., Connor, C.B., Bonadonna, C., 2008. Forecasting tephra dispersion using TEPHRA2. URL: https://vhub.org/resources/574/download/Tephra2_Manual.pdf.

Connor, L.J., Connor, C.B., Saballos, A., 2011. Tephra2 Users Manual. University of South Florida, Tampa, FL, URL: https://theghub.org/resources/756/download/Tephra2_Users_Manual.pdf. (Accessed 24 September 2017).

Connor, L.J., Hagdorn, M., 2018. Tephra2. URL: https://github.com/geoscience-community-codes/tephra2.git.

Costa, A., Pioli, L., Bonadonna, C., 2016. Assessing tephra total grain-size distribution: Insights from field data analysis. Earth Planet. Sci. Lett. 443, 90–107. http://dx.doi.org/10.1016/j.epsl.2016.02.040.

Cukier, R.I., Levine, H.B., Shuler, K.E., 1978. Nonlinear sensitivity analysis of multiparameter model systems. J. Comput. Phys. 26 (1), 1–42.

Damaschke, M., Cronin, S.J., Bebbington, M.S., 2018. A volcanic event forecasting model for multiple tephra records, demonstrated on Mt. Taranaki, New Zealand. Bull. Volcanol. 80, 1–14. http://dx.doi.org/10.1007/s00445-017-1184-y.

Dietterich, H.R., Lev, E., Chen, J., Richardson, J.A., Cashman, K.V., 2017. Benchmarking computational fluid dynamics models of lava flow simulation for hazard assessment, forecasting, and risk management. J. Appl. Volcanol. 6 (1), 9. http://dx.doi.org/10.1186/s13617-017-0061-x.

Dürig, T., Schmidt, L.S., Dioguardi, F., 2023. Optimizing mass eruption rate estimates by combining simple plume models. Front. Earth Sci. 11, 1250686. http://dx.doi.org/10.3389/feart.2023.1250686.

Engwell, S., Mastin, L.G., Bonadonna, C., Barsotti, S., Deligne, N.I., Oladottir, B.A., 2024. Characterising, quantifying, and accessing eruption source parameters of explosive volcanic eruptions for operational simulation of tephra dispersion: a current view and future perspectives. Bull. Volcanol. 86 (7), 67. http://dx.doi.org/10.1007/s00445-024-01706-y.

Folch, A., 2012. A review of tephra transport and dispersal models: Evolution, current status, and future perspectives. J. Volcanol. Geotherm. Res. 235–236, 96–115. http://dx.doi.org/10.1016/j.jvolgeores.2012.05.020.

Folch, A., Costa, A., Basart, S., 2012. Validation of the FALL3D ash dispersion model using observations of the 2010 Eyjafjallajökull volcanic ash clouds. Atmos. Environ. 48, 165–183. http://dx.doi.org/10.1016/j.atmosenv.2011.06.072.

Folch, A., Costa, A., Durant, A., Macedonio, G., 2010. A model for wet aggregation of ash particles in volcanic plumes and clouds: 2. Model application. J. Geophys. Res.: Solid Earth 115 (B9), http://dx.doi.org/10.1029/2009JB007176.

Folch, A., Costa, A., Macedonio, G., 2009. FALL3D: A computational model for transport and deposition of volcanic ash. Comput. Geosci. 35 (6), 1334–1342. http://dx.doi.org/10.1016/j.cageo.2008.08.008.

Folch, A., Costa, A., Macedonio, G., 2016. FPLUME-1.0: An integral volcanic plume model accounting for ash aggregation. Geosci. Model. Dev. 9 (1), 431–450. http://dx.doi.org/10.5194/gmd-9-431-2016.

Folch, A., Costa, A., Macedonio, G., Mingari, L., 2023. FALL3D. Zenodo, http://dx.doi.org/10.5281/zenodo.7802290.

Folch, A., Mingari, L., Gutierrez, N., Hanzich, M., Macedonio, G., Costa, A., 2020. FALL3D-8.0: a computational model for atmospheric transport and deposition of particles, aerosols and radionuclides – Part 1: Model physics and numerics. Geosci. Model. Dev. 13 (3), 1431–1458. http://dx.doi.org/10.5194/gmd-13-1431-2020.

Girden, E.R., 1992. ANOVA: Repeated measures. (84), Sage, Newbury Park, http://dx.doi.org/10.4135/9781412983419.

Green, K.C., Armstrong, J.S., 2015. Simple versus complex forecasting: The evidence. J. Bus. Res. 68 (8), 1678–1685. http://dx.doi.org/10.1016/j.jbusres.2015.03.026.

Gueugneau, V., Charbonnier, S., Esposti Ongaro, T., de' Michieli Vitturi, M., Peruzzetto, M., Mangeney, A., Bouchut, F., Patra, A., Kelfoun, K., 2021. Synthetic benchmarking of concentrated pyroclastic current models. Bull. Volcanol. 83 (11), 75. http://dx.doi.org/10.1007/s00445-021-01491-y.

Harvey, N.J., Huntley, N., Dacre, H.F., Goldstein, M., Thomson, D., Webster, H., 2018. Multi-level emulation of a volcanic ash transport and dispersion model to quantify sensitivity to uncertain parameters. Nat. Hazards Earth Syst. Sci. 18 (1), 41–63. http://dx.doi.org/10.5194/nhess-18-41-2018.

Herman, J., Usher, W., 2017. SALib: An open-source Python library for sensitivity analysis. J. Open Source Softw. 2 (9), 97. http://dx.doi.org/10.21105/joss.00097.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049. http://dx.doi.org/10.1002/qj.3803.

Höge, M., Wöhling, T., Nowak, W., 2018. A Primer for Model Selection: The Decisive Role of Model Complexity. Water Resour. Res. 54 (3), 1688–1715. http://dx.doi.org/10.1002/2017WR021902.

Kidson, J.W., 2000. An analysis of New Zealand synoptic types and their use in defining weather regimes. Int. J. Climatol. 20 (3), 299–316. http://dx.doi.org/10.1002/(SICI)1097-0088(20000315)20:3<299::AID-JOC474>3.0.CO;2-B.

Komorowski, J.-C., Legendre, Y., Caron, B., Boudon, G., 2008. Reconstruction and analysis of sub-plinian tephra dispersal during the 1530 A.D. Soufrière (Guadeloupe) eruption: Implications for scenario definition and hazards assessment. J. Volcanol. Geotherm. Res. 178 (3), 491–515. http://dx.doi.org/10.1016/j.jvolgeores.2007.11.022.

Macedonio, G., Costa, A., Folch, A., 2008. Ash fallout scenarios at Vesuvius: Numerical simulations and implications for hazard assessment. J. Volcanol. Geotherm. Res. 178 (3), 366–377. http://dx.doi.org/10.1016/j.jvolgeores.2008.08.014.

Malmborg, C.A., Willson, A.M., Bradley, L.M., Beatty, M.A., Klinges, D.H., Koren, G., Lewis, A.S.L., Oshinubi, K., Woelmer, W.M., 2024. Defining model complexity: An ecological perspective. Meteorol. Appl. 31 (3), http://dx.doi.org/10.1002/met.2202.

Mara, T., 2009. Extension of the RBD-FAST method to the computation of global sensitivity indices. Reliab. Eng. Syst. Saf. 94, 1274–1281. http://dx.doi.org/10.1016/j.ress.2009.01.012.

Mastin, L.G., Guffanti, M., Servranckx, R., Webley, P., Barsotti, S., Dean, K., Durant, A., Ewert, J.W., Neri, A., Rose, W.I., Schneider, D., Siebert, L., Stunder, B., Swanson, G., Tupper, A., Volentik, A., Waythomas, C.F., 2009. A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions. J. Volcanol. Geotherm. Res. 186 (1), 10–21. http://dx.doi.org/10.1016/j.jvolgeores.2009.01.008.

Michaud-Dubuy, A., Carazzo, G., Kaminski, E., 2021. Volcanic hazard assessment for tephra fallout in Martinique. J. Appl. Volcanol. 10 (1), 8. http://dx.doi.org/10.1186/s13617-021-00106-7.

Montgomery, D.C., 2017. Design and Analysis of Experiments. John Wiley & Sons, Hoboken, NJ.

Mulena, G.C., Allende, D.G., Puliafito, S.E., Lakkis, S.G., Cremades, P.G., Ulke, A.G., 2016. Examining the influence of meteorological simulations forced by different initial and boundary conditions in volcanic ash dispersion modelling. Atmos. Res. 176–177, 29–42. http://dx.doi.org/10.1016/j.atmosres.2016.02.009.

Pardini, F., Barsotti, S., Bonadonna, C., de'Michieli Vitturi, M., Folch, A., Mastin, L., Osores, S., Prata, A.T., 2024. Dynamics, Monitoring, and Forecasting of Tephra in the Atmosphere. Rev. Geophys. 62 (4), http://dx.doi.org/10.1029/2023RG000808.

Pardini, F., de' Michieli Vitturi, M., Andronico, D., Esposti Ongaro, T., Cristaldi, A., Neri, A., 2022. Real-time probabilistic assessment of volcanic hazard for tephra dispersal and fallout at Mt. Etna: the 2021 lava fountain episodes. Bull. Volcanol. 85 (1), 6. http://dx.doi.org/10.1007/s00445-022-01614-z.

Parra, R., Bernard, B., Narváez, D., Le Pennec, J.-L., Hasselle, N., Folch, A., 2016. Eruption Source Parameters for forecasting ash dispersion and deposition from vulcanian eruptions at Tungurahua volcano: Insights from field data from the July 2013 eruption. J. Volcanol. Geotherm. Res. 309, 1–13. http://dx.doi.org/10.1016/j.jvolgeores.2015.11.001.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M.Z., Barrow, D.K., Ben Taieb, S., Bergmeir, C., Bessa, R.J., Bijak, J., Boylan, J.E., Browell, J., Carnevale, C., Castle, J.L., Cirillo, P., Clements, M.P., Cordeiro, C., Cyrino Oliveira, F.L., De Baets, S., Dokumentov, A., Ellison, J., Fiszeder, P., Franses, P.H., Frazier, D.T., Gilliland, M., Gönül, M.S., Goodwin, P., Grossi, L., Grushka-Cockayne, Y., Guidolin, M., Guidolin, M., Gunter, U., Guo, X., Guseo, R., Harvey, N., Hendry, D.F., Hollyman, R., Januschowski, T., Jeon, J., Jose, V.R.R., Kang, Y., Koehler, A.B., Kolassa, S., Kourentzes, N., Leva, S., Li, F., Litsiou, K., Makridakis, S., Martin, G.M., Martinez, A.B., Meeran, S., Modis, T., Nikolopoulos, K., Önkal, D., Paccagnini, A., Panagiotelis, A., Panapakidis, I., Pavía, J.M., Pedio, M., Pedregal, D.J., Pinson, P., Ramos, P., Rapach, D.E., Reade, J.J., Rostami-Tabar, B., Rubaszek, M., Sermpinis, G., Shang, H.L., Spiliotis, E., Syntetos, A.A., Talagala, P.D., Talagala, T.S., Tashman, L., Thomakos, D., Thorarinsdottir, T., Todini, E., Trapero Arenas, J.R., Wang, X., Winkler, R.L., Yusupova, A., Ziel, F., 2022. Forecasting: theory and practice. Int. J. Forecast. 38 (3), 705–871. http://dx.doi.org/10.1016/j.ijforecast.2021.11.001.

Phillips, J., Williams, S., Lee, A., Jenkins, S., 2023. Quantifying uncertainty in probabilistic volcanic ash hazard forecasts, with an application to weather pattern based wind field sampling. Bull. Volcanol. 85 (11), 68. http://dx.doi.org/10.1007/s00445-023-01664-x.

Pioli, L., Bonadonna, C., Pistolesi, M., 2019. Reliability of total grain-size distribution of tephra deposits. Sci. Rep. 9 (1), 10006. http://dx.doi.org/10.1038/s41598-019-46125-8.

Pistolesi, M., Cioni, R., Bonadonna, C., Elissondo, M., Baumann, V., Bertagnini, A., Chiari, L., Gonzales, R., Rosi, M., Francalanci, L., 2015. Complex dynamics of small-moderate volcanic events: the example of the 2011 rhyolitic Cordón Caulle eruption, Chile. Bull. Volcanol. 77 (1), 3. http://dx.doi.org/10.1007/s00445-014-0898-3.

Poret, M., Costa, A., Folch, A., Martí, A., 2017. Modelling tephra dispersal and ash aggregation: The 26th April 1979 eruption, La Soufrière St. Vincent. J. Volcanol. Geotherm. Res. 347, 207–220. http://dx.doi.org/10.1016/j.jvolgeores.2017.09.012.

Puy, A., Beneventano, P., Levin, S.A., Lo Piano, S., Portaluri, T., Saltelli, A, 2023. Models with higher effective dimensions tend to produce more uncertain estimates. Sci. Adv. 8 (42), http://dx.doi.org/10.1126/sciadv.abn9450.

Puy, A., Lo Piano, S., Saltelli, A., Levin, S., 2021. sensobol: an R package to compute variance-based sensitivity indices. URL: https://cran.r-project.org/package=sensobol.

Pyle, D.M., 2018. What can we learn from records of past eruptions to better prepare for the future? In: Fearnley, Carina J, Bird, Deanne K, Haynes, Katharine, McGuire, William J, Jolly, Gill (Eds.), Observing the Volcano World: Volcano Crisis Communication. Springer International Publishing, Cham, pp. 445–462. http://dx.doi.org/10.1007/11157{_}2017{_}5.

Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. Comput. Phys. Comm. 145 (2), 280–297. http://dx.doi.org/10.1016/S0010-4655(02)00280-1.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput. Phys. Comm. 181 (2), 259–270. http://dx.doi.org/10.1016/j.cpc.2009.09.018.

Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Nelson Espeland, W., Kay, J., Lo Piano, S., Mayo, D., 2020. Five ways to ensure that models serve society: a manifesto. Nature 582, 482–484. http://dx.doi.org/10.1038/d41586-020-01812-9.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis: the Primer. John Wiley & Sons, http://dx.doi.org/10.1002/97804707251 8.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity analysis in practice: a guide to assessing scientific models, vol. 1, Wiley Online Library, http://dx.doi.org/10.1002/0470870958.

Saltelli, A., Tarantola, S., Chan, K.P.-S., 1999. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics 41 (1), 39–56. http://dx.doi.org/10.1080/00401706.1999.10485594.

Sandri, L., Costa, A., Selva, J., Tonini, R., Macedonio, G., Folch, A., Sulpizio, R., 2016. Beyond eruptive scenarios: Assessing tephra fallout hazard from Neapolitan volcanoes. Sci. Rep. 6, http://dx.doi.org/10.1038/srep24271.

Scollo, S., Del Carlo, P., Coltelli, M., 2007. Tephra fallout of 2001 Etna flank eruption: Analysis of the deposit and plume dispersion. J. Volcanol. Geotherm. Res. 160 (1), 147–164. http://dx.doi.org/10.1016/j.jvolgeores.2006.09.007.

Scollo, S., Folch, A., Costa, A., 2008a. A parametric and comparative study of different tephra fallout models. J. Volcanol. Geotherm. Res. 176 (2), 199–211. http://dx.doi.org/10.1016/j.jvolgeores.2008.04.002.

Scollo, S., Tarantola, S., Bonadonna, C., Coltelli, M., Saltelli, A., 2008b. Sensitivity analysis and uncertainty estimation for tephra dispersal models. J. Geophys. Res.: Solid Earth 113 (B6), http://dx.doi.org/10.1029/2006JB004864.

Selva, J., Costa, A., Sandri, L., Macedonio, G., Marzocchi, W., 2014. Probabilistic short-term volcanic hazard in phases of unrest: A case study for tephra fallout. J. Geophys. Res.: Solid Earth 119 (12), 8805–8826. http://dx.doi.org/10.1002/2014JB011252.

Sobol', I.M., 1990. On sensitivity estimation for nonlinear mathematical models. Mat. Model. 2 (1), 112–118.

Sulpizio, R., Folch, A., Costa, A., Scaini, C., Dellino, P., 2012. Hazard assessment of far-range volcanic ash dispersal from a violent Strombolian eruption at Somma-Vesuvius volcano, Naples, Italy: implications on civil aviation. Bull. Volcanol. 74 (9), 2205–2218. http://dx.doi.org/10.1007/s00445-012-0656-3.

Suzuki, T., 1983. A Theoretical Model for Dispersion of Tephra. Arc Volcanism: Phys. Tectonics 95–113.

Tsuji, T., Nishizaka, N., Ohnishi, K., 2020. Influence of particle aggregation on the tephra dispersal and sedimentation from the October 8, 2016, eruption of Aso volcano. Earth, Planets Space 72 (1), 104. http://dx.doi.org/10.1186/s40623-020-01233-y.

University of South Florida, 0000. Tephra2 Forward Modeling Tool. Geo-science Community Codes URL: https://gscommunitycodes.usf.edu/geoscicommunitycodes/public/tephra2/forward-model.php.

Volentik, A.C.M., Bonadonna, C., Connor, C.B., Connor, L.J., Rosi, M., 2010. Modeling tephra dispersal in absence of wind: Insights from the climactic phase of the 2450BP Plinian eruption of Pululagua volcano (Ecuador). J. Volcanol. Geotherm. Res. 193 (1), 117–136. http://dx.doi.org/10.1016/j.jvolgeores.2010.03.011.

Widiwijayanti, C., Voight, B., Hidayat, D., Patra, A., Pitman, E.B., 2004. Validation of TITAN2D flow model code for pyroclastic flows and debris avalanches at Soufrière Hills Volcano, Montserrat, BWI. In: AGU Fall Meeting Abstracts, vol. -1, p. 1428.

Wiejaczka, J., Giachetti, T., 2024. Determining the umbrella cloud geometry of unwitnessed silicic explosive eruptions: A case study from Mount Mazama (Oregon, United States). J. Volcanol. Geotherm. Res. http://dx.doi.org/10.1016/j.jvolgeores.2024.108015.