METHODOLOGY   **Open Access**

Towards real-time probabilistic ash deposition forecasting for New Zealand

Rosa Trancoso[1] , Yannik Behr[2]* , Tony Hurst[2] and Natalia I Deligne[3]

Abstract

Volcanic ashfall forecasts are highly dependent on eruption source parameters (ESPs) and synoptic weather conditions at the time and location of the eruption. In New Zealand, MetService and GNS Science have been jointly developing an ashfall forecast system that incorporates four-dimensional high-resolution numerical weather prediction (NWP) and ESPs into the HYSPLIT model, a state-of-the art hybrid Eulerian and Lagrangian dispersion model widely used for volcanic ash. However, these forecasts are based on discrete ESPs combined with a deterministic weather forecast and thus provide no information on output uncertainty. This shortcoming hinders stakeholder decision making, particularly near the geographical margin of forecasted ashfall and in areas with large gradients in forecasted ash deposition. Our study presents a new approach that incorporates uncertainty from both eruptive and meteorological inputs to deliver uncertainty in the model output. To this end, we developed probability density functions (PDFs) for three key ESPs (plume height, mass eruption rate, eruption duration) tailored to New Zealand's volcanoes and combine them with NWP ensemble datasets to generate probabilistic ashfall forecasts using the HYSPLIT model. We show that the Latin Hypercube Sampling (LHS) technique can be used to representatively span this four-dimensional parameter space and allow us to add uncertainty quantification to rapid response forecast systems. For a case study of a hypothetical eruption at Tongariro, New Zealand we suggest that large parts of New Zealand's North Island would not receive adequate warning for potential ashfall if uncertainties were not included in the forecasts. We also propose new probabilistic summary products to support public information and emergency responders decision making.

Keywords: Volcanic hazard, Ashfall, Atmospheric models, HYSPLIT, Ensembles, Probabilistic forecast, Latin Hypercube, New Zealand, Near real-time forecasts

Introduction

Explosive volcanic eruptions produce ash (glass, crystals, and rock particles < 2 mm diameter) which can be carried hundreds of kilometres from the vent. It can disrupt and damage infrastructure and property, destroy crops, poison livestock, and pose serious risk for aviation and human health (Jenkins et al. 2015). Ash is dispersed in the atmosphere and it remains there until it settles onto Earth's surface (Carey and Bursik 2015). Besides tsunamis, volcanic eruptions are the hazard with the largest footprint and the greatest social and economic impact (Wilson et al. 2015). As a result, ashfall and ash dispersion forecasts are the most sought-after volcanic hazard information following an eruption (USGS 2022).

Where, when, and how much ash will be deposited depends on the style of eruption and the meteorological conditions. Only explosive eruptions will produce significant amounts of ash and these eruptions are typically characterised by the total of erupted mass and either the column height of the eruption plume or the radius of the umbrella cloud forming at the top of the plume (Constantinescu et al. 2021). Other important eruption source parameters (ESPs) are the duration of the eruption, the rate at which mass is ejected, the mass distribution along

*Correspondence: y.behr@gns.cri.nz

[2] GNS Science, Wellington, New Zealand
Full list of author information is available at the end of the article

Trancoso *et al. Journal of Applied Volcanology*   (2022) 11:13

Page 2 of 13

the eruption column, and the grain size distribution above the volcanic vent (Bonadonna et al. 2015). At most volcanoes in New Zealand it would take days to weeks to robustly quantify all these parameters following an eruption. To provide ashfall forecasts during the first hours after an eruption, prior assumptions must be made for some or all the ESPs.

To address this problem, MetService, New Zealand's national weather authority and host of the Wellington Volcanic Ash Advisory Centre (VAAC), and GNS Science, New Zealand's provider of geological hazard information, have jointly developed an ashfall forecasting system for New Zealand's most active volcanoes on or near the North Island (Fig. 1). This system (referred to as System 1 from hereon) combines a small number of eruption scenarios provided by GNS Science for each of the volcanoes with four-dimensional (4D) high-resolution numerical weather prediction (NWP) provided by MetService, to generate forecast maps of ash depth on the ground (Hurst and Davis 2017). The volcanic ash transport and deposition is computed with the HYSPLIT model, a state-of-the-art hybrid Eulerian and Lagrangian dispersion model (Stein et al. 2015a) used operationally by four of the nine VAACs (Prata and Rose 2015, Hort 2019). System 1 produces 24-h long forecasts updated every six hours using two hypothetical eruption scenarios (small and large), so that in case of an eruption a preliminary forecast is available.

However, because these forecasts are based on single-valued ESPs combined with a deterministic weather forecast, they cannot capture the breadth of possible outcomes and therefore provide no information on output uncertainty. Because uncertainty in the prior assumptions about eruption scenarios is large, it is critical to quantify how they affect the forecasts.

For example, a stakeholder (e.g., infrastructure manager) might take stronger mitigative actions if the forecast likelihood of exceeding a specified threshold is high, even if the expected value is low.

In this study, we aim to demonstrate a prototype system (referred to as System 2 from hereon) for near real-time probabilistic ashfall forecasts in New Zealand, that accounts for both uncertainties in ESPs and the current and future atmospheric state. As we do not attempt to optimise deposit modelling accuracy we adopt the HYSPLIT configuration used by Hurst and Davis (2017). Our focus lies on developing a methodology for uncertainty quantification of ashfall forecasts during rapid responses when information on ESPs is often incomplete and very uncertain. Forecasts can still be improved by, for example, using other or additional numerical models, higher-resolution weather models, or more extensive eruption databases. All these improvements can

be embedded in the same uncertainty quantification approach.

Methodology

Main sources of uncertainty in ashfall forecasts come from uncertainties in the ESPs, the NWP, and the choice of the numerical forecast model and its parameters. We focus here on the first two but briefly touch on the latter in the Discussion Section.

To quantify uncertainty in the ESPs in System 2 in the absence of any reliable ESP observations we compiled information and developed probability density functions (PDFs) for three key eruption source parameters (ESPs) – plume height, mass eruption rate, eruption duration – tailored to New Zealand's volcanoes. Next, we combined these PDFs with NWP ensembles, using Latin Hypercube Sampling (LHS) with Dependence, to generate probabilistic ashfall forecasts with the HYSPLIT model. Once some information on ESPs becomes available, new PDFs centred around this information can replace the prior ESP PDFs. We focus here, however, on the rapid response situation immediately after an eruption when typically very little is known about ESPs other than the onset of the eruption.

Uncertainty in ESPs

Quantifying uncertainty in the ESPs is challenging, particularly in the first hours of an eruption when reliable quantitative observations are often not available (Aubry et al. 2021). Three key ESPs for modelling volcanic ashfall transport are the eruption plume height (H), mass eruption rate (MER), and eruption duration (D). Depending on the situation (e.g., at night, bad weather, very large eruptions) H may not be trivial to reliably measure, can be variable on short timescales, and estimates may have large uncertainties. MER is rarely measured directly (e.g., Freret-Lorgeril et al. 2018) but usually derived from empirical correlations based on the observable plume height, in relationships that have confidence intervals spanning orders of magnitude (Aubry et al. 2021; Mastin et al. 2009). Finally, eruption duration is related to mass eruption rate and total volume erupted, and is difficult to forecast while an eruption is ongoing. Existing duration data consist of a mix of direct observations and calculated values, for example, using field data to estimate mass eruption rate, and from there calculating duration based on erupted mass, as done in Longchamp et al. 2011. What constitutes the 'end' of an eruption can also be difficult to ascertain.

Volcano dataset and ESP PDFs

We undertook a comprehensive study to develop, for the first time, PDFs for eruption plume height, the mass

Fig. 1 New Zealand 10 volcanic centres for which ashfall forecasts are routinely generated (black triangles in bottom panel) with local topography in the background.

Trancoso *et al. Journal of Applied Volcanology*   (2022) 11:13

Page 4 of 13
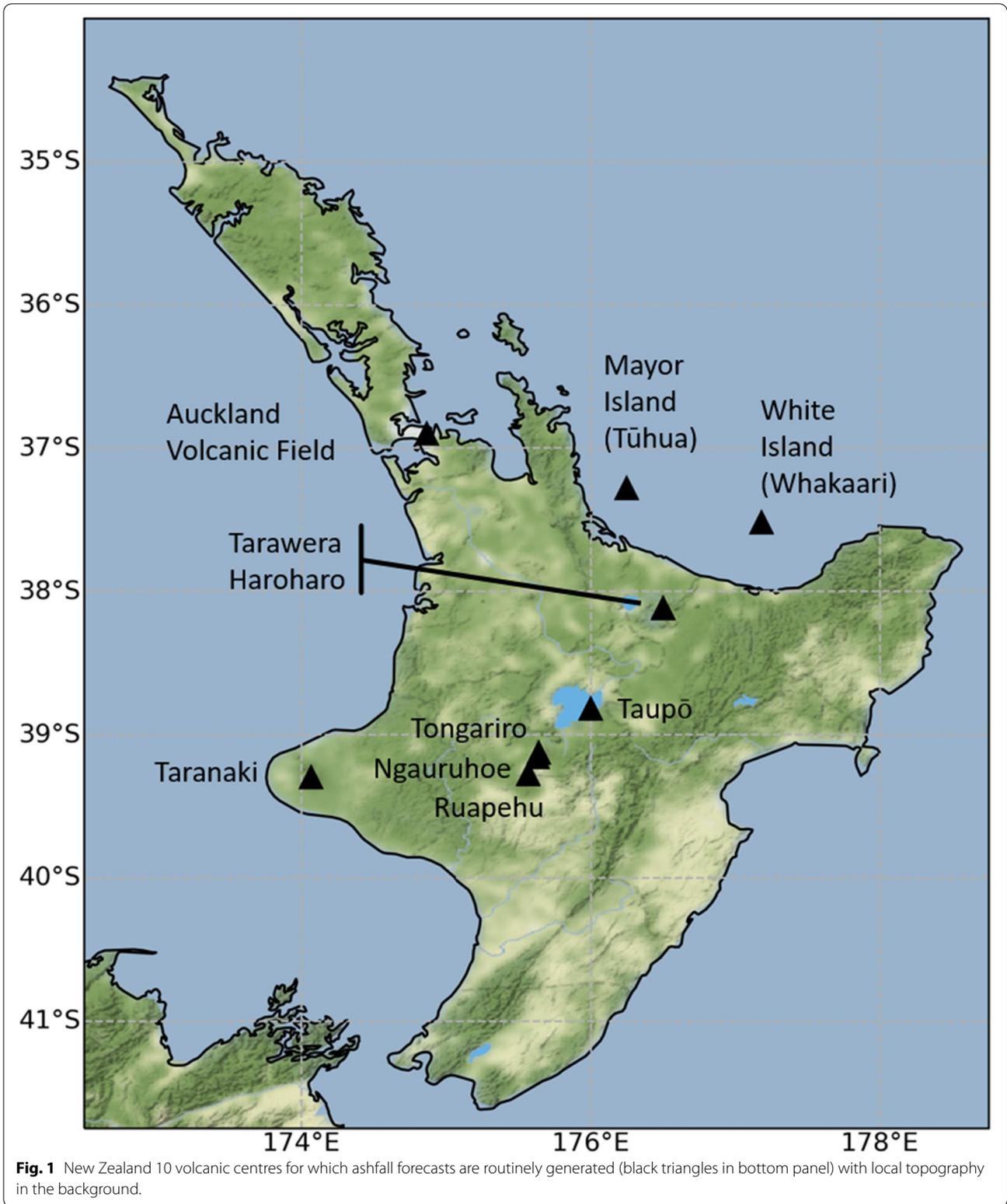
eruption rate, and eruption duration for the 10 volcanic centres for which forecasts are produced by System 1. The intention is to have a priori distributions ready to use if there is confirmation of an eruption. To achieve this goal, we compiled a large dataset, combining datasets from Deligne (2021) and Aubry et al. (2021). We briefly describe both below.

Deligne (2021) compiled measured, estimated, and calculated estimates for volcanic mass eruption rates, column heights, and durations from 213 events, a third of which are from New Zealand volcanoes. These include both prehistoric (estimates based on geologic record) and historical eruptions. Eruptions were categorized according to magma type (mafic, intermediate, silicic) and eruption style (steam-driven, magmatic small to moderate, magmatic large; see Deligne 2021 for details on the classification). Every entry was coded according to data type (e.g. instrumentally measured, observed, derived from geologic study) and assumptions made in determining values. In Deligne (2021), there was no attempt to evaluate the quality of the studies providing individual parameter values, nor an exclusion of derived values (e.g. mass eruption rate is often calculated based on erupted mass, duration, and/or other parameters, but is rarely independently measured). For this study, we removed ESP values compiled in Deligne (2021) derived from empirical or numerical models.

Aubry et al. (2021) independently compiled estimates of total erupted mass of fall deposits, duration, eruption column height, and atmospheric conditions from 134 eruptions from around the world from 1902 to 2016. Eruptions had to have independently measured/available values for all four parameters. Each eruption was independently evaluated by at least two experts, and uncertainty associated with each parameter was systematically evaluated, with estimates requiring a high degree of interpretation flagged. For this study, we used all ESP values compiled in Aubry et al. (2021), without consideration of the provided uncertainty.

The combined dataset includes all the Aubry et al. (2021) eruptions, and a subset of Deligne (2021) parameters as described above. In cases where an eruption is catalogued in both Aubry et al. (2021) and Deligne (2021), our default is to use the Deligne (2021) values, using the Aubry et al. (2021) ESP if none exists in the Deligne (2021) dataset. For the 85 eruptions that are only catalogued in Aubry et al. (2021), we attributed magma type and eruption style categories using the criteria documented in Deligne (2021). For 22 eruptions, there was no indication of eruption style apart from being classified as VEI 4 or smaller by the Smithsonian Global Volcanism Program: we assigned these to the 'Magmatic small to moderate' eruption style category. Table 1 summarizes

Table 1 Summary of eruptions data compiled to estimate ESP distributions

	Steam-driven	Magmatic small to moderate	Magmatic large	Total
Mafic	18 (7%)	78 (30%)	4 (2%)	100 (38%)
Intermediate	27 (10%)	85 (33%)	22 (8%)	134 (51%)
Silicic	1 (<1%)	15 (6%)	10 (4%)	26 (10%)
Total	46 (17%)	178 (69%)	36 (14%)	260 (100%)

the collected data we used drawn from Deligne (2021) and Aubry et al. (2021) and Fig. 2 shows the PDFs of ESPs by magma type.

Because we use the whole database to estimate ESP PDFs, uncertainties of ESPs in single studies are unlikely to have a strong influence on the PDFs. There are, however, most likely biases in the database. We expect the number of small eruptions is likely underestimated as many of them would have gone unnoticed or left no geological records. More importantly, the number of large eruptions is also likely to be underestimated due to their rare occurrence. To reduce these biases the database has to be continuously updated as new eruption studies are published.

Uncertainty in weather
Uncertainty in weather is nowadays quantified with NWP ensemble datasets (Cheung 2001). These datasets seek to capture forecast error growth by running many model realisations, with perturbed initial conditions that reflect uncertainty in the current atmospheric state, and perturbed model parameters that reflect uncertainty in physical processes. The non-linear nature of the equations governing atmospheric motion means that even with perfectly known initial conditions, the tiniest perturbations in atmospheric variables in the most accurate of NWP models may result in different outcomes after a finite time (Lorenz 1969; Zhang et al. 2019). The output of these ensembles is probabilistic, which more accurately reflects what is knowable about the future atmospheric state than the deterministic picture provided by a single model.

NWP ensembles
The NWP ensemble datasets used in this study are from the Global Ensemble Forecast System version 12 (GEFSv12) running at the National Oceanic and Atmospheric Administration (NOAA). The dataset is updated 4 times per day (cycles starting at 0, 6, 12 and 18:00 UTC) with 31 members (30 perturbed and 1 unperturbed/control) with approximately 25 km horizontal resolution and 64 vertical hybrid levels for atmospheric components,

Fig. 2 Logarithmic ESP distributions coloured by magma type. From top to bottom rows: Duration [h], Mass eruption rate [kg/s] and Column height above vent [km]. The diagonal panels show the empirical cumulative distribution functions; upper triangular panels show the distributions between pairs of ESPs; lower triangular panels show the same distributions interpolated using kernel density estimation (KDE). For the KDE gaussian kernels were used with the bandwidth computed using a heuristic described by Scott (1992). Contours are drawn at 0.1 intervals between 0 and 1

and out to 16 days of forecast at each cycle, except for 35 days at 0000 UTC (Stajner et al. 2020). The model output is public and available via NOMADS Server (2022) at a 0.5° resolution grid, 3 h temporal resolution, with surface and 12 pressure level fields up to 10 hPa, for the past 30 days.

Combining uncertainties and sampling with Latin Hypercube technique

Multiple approaches can be taken to quantify uncertainty both from the source parameters as well as from the weather conditions. In our rapid response context,

the biggest challenge of probabilistic forecasting is to optimize resources while ensuring the spread and natural variability of our multidimensional parameter space are well represented.

Monte Carlo sampling methods have been previously used in ash probabilistic forecasting (Hurst and Smith 2004; Magill et al. 2006) but they are only computationally feasible with simplifying assumptions that limit the sample size. For instance, only using a few eruption sizes, and/or only using wind profiles at the vent location. While these assumptions may be reasonable for long-term hazard studies, they would likely result in

Trancoso *et al. Journal of Applied Volcanology*   (2022) 11:13

Page 6 of 13

inaccurate forecasts for a single eruption and near real-time forecasting.

LHS is a stratified sampling approach that draws representative samples using a smaller sample size than Monte Carlo sampling, ensuring every section of the parameter space is sampled once (Mckay et al. 2000). LHS is a common statistical method for generating hazard maps with physics-based numerical models (e.g., Berger et al., 2011; Spiller et al. 2014). Sigg et al. (2018) and Prata et al. (2018) both used LHS techniques as an alternative to Monte Carlo in dispersion applications. Sigg et al. (2018) considered a known amount of pollutant and the other LHS parameters were associated with microscale boundary layer meteorology important for their specific small-scale application. Prata et al. (2018) considered large scale transport of airborne ash and considered various ESPs as independent parameters in their LHS.

For our purposes, we considered the GEFSv12 ensemble dataset to be reliable in a statistical sense (i.e., over time, ensemble member counts of some categorical weather event are consistent with their actual probability). The atmospheric evolution can be sampled by selecting an ensemble member, and as we considered the eruption and atmospheric state to be independent, the ensemble member label can be considered an independent parameter with uniform distribution (Stein et al. 2015b).

LHS with dependence

The nature of the sampling problem for ESPs is different because they are not independent of each other – an underlying assumption of standard LHS technique. To introduce correlation between ESPs, we assume that they approximately follow a multivariate Gaussian distribution (Packham and Schmidt 2008). We then first draw samples for each parameter, x, from a standard Gaussian distribution(mean = 0; variance = 1) and then match the mean μ and covariance matrix Σ of the prior eruption parameter distribution. Because we have three ESPs, x and μ are 3×1 vectors and Σ is a 3×3 matrix. A single LHS draw y can then be expressed as:

$$y = Lx + \mu \quad with \quad \Sigma = LL^T \tag{1}$$

where L is the Cholesky factorisation of Σ (Murphy 2022). It follows that y has the same mean as the original distribution. To see that y also has the same covariance matrix we can calculate the covariance of y:

$$Cov(y) = LCov(x)L^T = LIL^T = \Sigma \tag{2}$$

with I being the Identity matrix. This technique allows us to draw samples that preserve the correlation between, for example, MER and H (Mastin et al. 2009 and Fig. 2).

Figure 3 compares Monte Carlo and LHS techniques, demonstrating that LHS allows us to generate a representative sample of a given distribution with considerably fewer samples than Monte-Carlo sampling requires.

LHS modelling runs

As previously mentioned, we considered each perturbed atmospheric state to be independent of each other and of the ESPs. To optimize computational resources and maximize the weather variability information available, we constrained our ESP sample size to be the same as the number of NWP ensemble members, (30 for GEFSv12 forecasts). Figure 4 shows that a sample number of 30 suffices to have a good approximation to the PDFs derived from historical eruptions (Fig. 2).

We then used these "eruption scenarios" i.e., each point with a sampled MER, H, D and NWP ensemble member, to force independent ashfall dispersion model runs, constraining H to a maximum of 20 km above mean sea level to fit within the weather forcing data, and D to a maximum of 24 h which is the forecast length. Table 2 summarizes the run parameters used as input to the dispersion model.

HYSPLIT configuration

Ashfall dispersion was computed with the HYSPLIT model version v5.0.1 (April 2020), which is extensively used for volcanic ash transport and deposition by the international community (Rolph et al. 2017; Stein et al. 2015a).

The configuration options used were implemented in Hurst and Davis (2017) and a brief description follows. The particle size distribution used for andesite volcanoes such as Tongariro in our case study has a grain density of 1300 kg/m3, divided into 37 bins (see Fig. 5). We assumed a constant grain-size density of 1300 kg/m3 for all grain sizes consistent with Hurst and Davis (2017). This assumption could be refined in future work to reduce density for larger clasts which tend to contain more gas bubbles. The total mass erupted is distributed along the eruption column according to a Suzuki distribution with constant 4 and discretized in 10 levels. This distribution adds more mass to the upper levels of the eruption plume, as is typical for most eruptions (Suzuki 1983) and which approximates umbrella clouds seen in large eruptions. Model results were integrated to a grid of 1 km horizontal resolution spanning 14 degrees latitude and longitude, from 48.5°S to 33.5°S and 165.5°E to 180.5°E.

Case study

The hypothetical eruption studied here is located at Tongariro volcano (39.130°S, 175.642°E) with the main vent

Fig. 3 Performance of Monte Carlo and LHS techniques as a function of number of samples. **a** shows the 2D Gaussian distribution from which samples are drawn; **b** and **c** show the LHS and Monte Carlo estimate, respectively, from 20 random samples. **d** to **f** show the evolution of the estimates of first and second moment for LHS and Monte Carlo samples with respect to the number of samples

at 1978 m above sea level. It typically erupts andesite lavas (Leonard et al. 2021), which we classified as "intermediate magma" in our eruption database (Table 1 and Fig. 2) and thus draw ESP samples only from this subset of the eruption database. We could have used any of the other volcanic centres in New Zealand for this case study. We chose this one because it demonstrates well the effects of including uncertainty in ESPs and the NWP into ashfall forecasts.

Results

The proposed methodology was applied to a case study mimicking a real-time eruption forecast at Tongariro, making use of the full set (30 members) of GEFSv12 forecasts routinely made available by NOAA. We simulated a hypothetical eruption starting at 2021–07-15 00:00 UTC because during this day there was a wind shift near the volcano centre from south-westerly to westerly near Tongariro volcano that allowed us to illustrate the importance of weather dynamics.

Fig. 4 Comparison of original (blue) and LHS resampled (orange) distribution of ESP for 20 samples and intermediate magma (see Fig. 2)
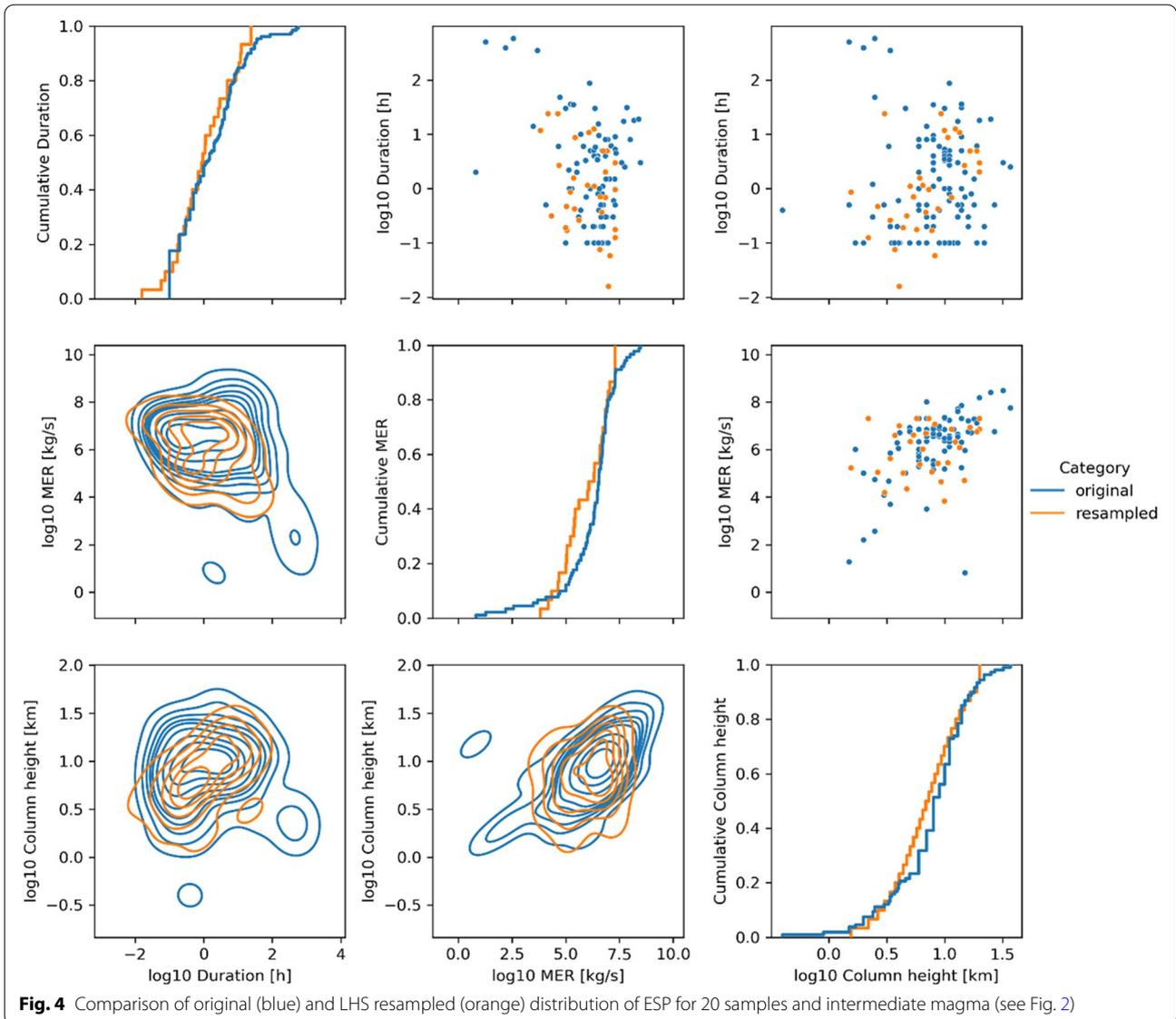
Table 2 HYSPLIT dispersion run parameters selected for each sample taken with LHS

Parameter	Sampling range
Plume height (H)	0–20 km above vent
Mass eruption rate (MER)	Unconstrained
Eruption duration (D)	0–24 h
Meteorological fields (Operational case study)	GEFSv12 members 1–30, 0.5° resolution, 3 hourly

Using LHS, 30 ESPs were drawn from the set of "intermediate magma" eruptions in the eruption database (Fig. 6). As mentioned before, each point of the sample is an "eruption scenario"—corresponding to discrete values of MER, H, D and ensemble member—used to configure a HYSPLIT run and then combined with one of the 30 NWP ensemble members. Each will produce a deterministic ashfall forecast that can vary significantly from other ensemble members as can be seen in the deterministic examples in Fig. 7. Here, eruptions longer than 6 h show more dispersion to the east because of a wind shift from south-westerlies to westerlies around 6 h after eruption. It is also interesting to note the clear split on the ashfall footprint due to wind patterns affected by the mountainous terrain that crosses the North Island. The ash plume disperses either north or south of the range, crossing it in its valleys if the wind is perpendicular (see terrain complexity on Fig. 1). This shows the importance of having a 4D NWP as it accounts for terrain features.

Fig. 5 Particle size distribution for andesitic ash (density 1300 kg/m3). Source: Hurst and Davis (2017)
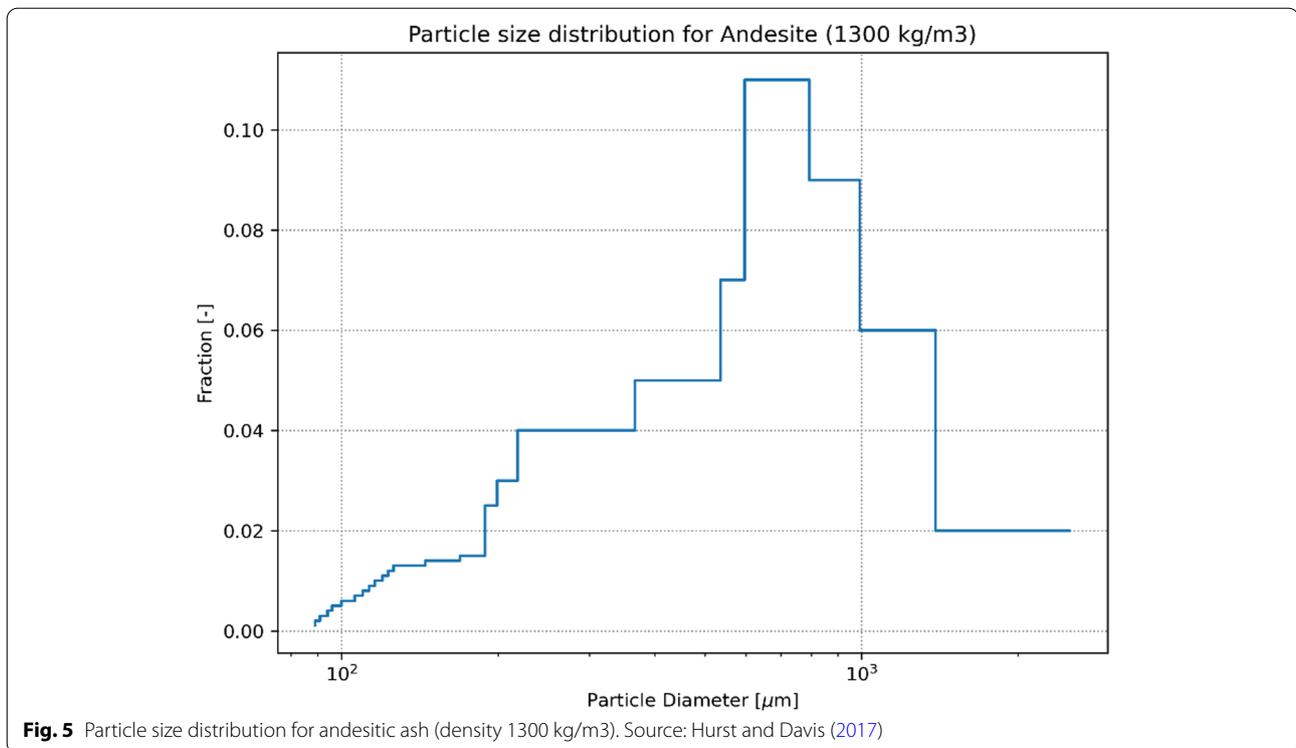
Ensemble results are then produced by combining the 30 deterministic scenario forecasts and summarized in products such as the ones presented below (Fig. 8). They allow us to see the probability of ashfall being present in a location, when and how much. One of the most useful products for emergencies is the hazard matrix (Fig. 8A; Prata et al. 2018) that clearly shows the areas that will be most likely affected by significant amounts of ash deposits. The comparison with a single ensemble member (blue contour in Fig. 8A) demonstrates the spread of the whole ensemble. For this case study, only areas close to the volcano are likely (> 50%) to experience ashfall greater than 1 mm or very likely (> 90%) to experience ashfall greater than 0.1 mm. A significant part of the ash deposits greater than 0.1 mm are likely to occur from the source to the north-east, and then offshore. Most importantly, by accounting for the uncertainties in the ESPs and the NWP, a much wider area is forecasted to possibly be affected by ashfall than if using only a single, deterministic forecast.

While Fig. 8A provides a concise summary of the whole ensemble forecast, Fig. 8B gives more detailed information on the probability of exceeding a given threshold of ash deposits.

It is also useful to look at the arrival time of ash given the ensemble spread. Figure 8C shows the arrival time of the median probability of accumulated ash thickness being higher than 0.01 mm (e.g. propagation in time of the 50% contour of the 0.01 mm exceedance probability). Here, the 0.01 mm threshold is low enough to be considered as ash presence. While the forecast was produced for 24 h, no more ash was deposited after 11 h. This is consistent with the 50% contour exceedance probability of 0.01 mm after 24 h, shown in Fig. 8B.

Discussion

Our case study demonstrates the large uncertainties that must be considered when issuing ashfall forecasts during a rapid response to a volcanic eruption. As a consequence, forecasts based on deterministic scenarios will likely miss large areas that are affected by ashfall and misestimate the amount of deposited ash. Our proposed approach lends itself to incrementally reduce uncertainty as new data become available. For example, if information on the eruption column height was available, a new LHS could be drawn where column height is constrained but other parameters are kept within their prior range. Currently, System 1 requires an expert to choose the most representative from the pre-computed scenarios. System 2 removes this need by having a prior distribution of ESPs rather than discrete scenarios. It comes however at the cost of significantly increased computational
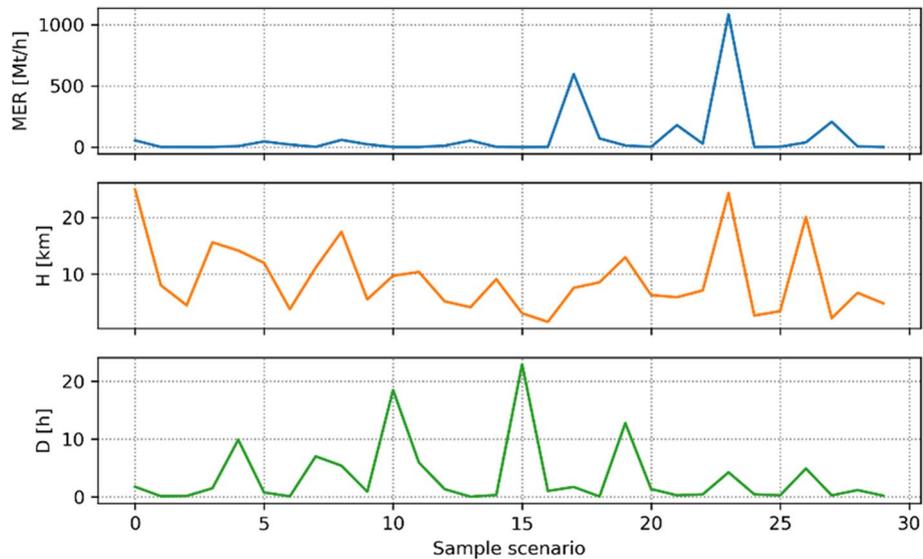
Fig. 6 LHS draw with 30 points used in the operational case study corresponding to 30 eruption scenarios. From top to bottom: mass eruption rate (MER), eruption column height above vent (H), and eruption duration (D)

resources. Using LHS is critical to keep the number of necessary forecast runs as low as possible in order to have a near real-time automated system. In the future, as computational resources and global ensembles become more accessible, the number of LHS samples could be increased and dispersion runs could be forced by high-resolution NWP ensembles derived from the global data-sets, using limited area models such as WRF which is already used extensively at MetService.

Other opportunities to improve uncertainty estimates include the variation of additional model parameters and the inclusion of additional forecast models. If we distinguish between input x and the parameters θ of the HYSPLIT model (M) we can write the probability Pr of ashfall at location i (af_i) as:

$$Pr(af_i) = Pr(af_i|M(\theta), x)Pr(x)\text{Pr}(M(\theta))$$

We focused here on ESPs that can be constrained by observations during a rapid response hence our model input x is comprised of the ESPs eruption plume height, mass eruption rate, and eruption duration as well as the NWP ensembles. Grain size distribution is part of the parameters θ, but given the importance of grain size distribution to ashfall forecasts, results from studies correlating column height and magma viscosity to grain size distribution could be used in the future (e.g. Costa et al. 2016) to include grain size distribution in x. Other parameters that could be included in the input are the mass profile along the plume or the umbrella cloud radius. The distinction between x and θ is mainly a practical one as it would be computationally very expensive if not infeasible to treat all θ as part of x even though this would provide the most accurate uncertainty estimate.

As ashfall forecasts depend on the forecast model, M, adding forecasts from different forecast models will also improve estimates of model uncertainty. Given K different forecast models one could then perform Bayesian model averaging (e.g. Bishop 2006) over such multi-model forecasts:

$$Pr(af_i) = \sum_{k=1}^{K} Pr(af_i|M_k(\theta_k), x)Pr(x)Pr(M_k(\theta_k))$$

Conveying probabilistic results to non-experts is challenging, especially with static maps (Thompson et al. 2015). Our proposed visual products lean on previous work on air-borne ash (Prata et al. 2018) and probabilistic hazard maps (Thompson et al. 2015) to present different ways to summarize the uncertainty information in a concise manner. The hazard matrix maps and probability of exceedance of cumulative ash fall provide a summary picture of the situation after a certain time, while the ash arrival time map conveys the dynamic evolution that leads to the previous products. However, these are simply a proposal and further stakeholder engagement will be required to find the best way of reporting probabilistic ashfall forecasts.

Fig. 7 Example of deterministic ashfall forecasts for 4 points of the LHS sample (each point is a set of ESPs and ensemble member). The points were chosen to show variability

Fig. 8 **A** Cumulative ashfall hazard map color-coding areas by the likelihood of ashfall exceeding a given threshold (see inset matrix for details) for a forecast 24 h after eruption. The blue lines show the contour of the 0.01 mm deposition for a single ensemble member. **B** Probability of cumulative ashfall exceeding 0.01 mm, for a forecast 24 h after eruption. The red line shows the median (50%) probability. **C** Propagation in time of the median probability of ashfall exceeding 0.01 mm in one-hour steps for a forecast 24 h after eruption. Most ash falls within 9 h after eruption

Conclusions

By quantifying uncertainty in ESPs combined with NWP ensemble datasets we were able to provide uncertainty estimation to ash dispersion forecasts during a rapid eruption response. This not only gives scientists a better tool to communicate ashfall hazards but also paves the way to impact-based forecasting.

We showed how the LHS technique can be successfully used to representatively span the high dimensional parameter space of combined uncertainty of ESPs and NWP, with a reduced sample size that allows for reasonable computational resource usage. This is a determining factor for a near real-time automated forecast system such as the one existing and developed in a joint effort between GNS Science and MetService.

We also explored and propose visual products that summarize probabilistic forecast, such as the hazard matrix, probability of exceedance of cumulative ash fall, and arrival time. These could form the bases to aid engagement with responding entities in selecting the most suitable way to communicate probabilistic ashfall. Future work could also include quantifying uncertainty in other relevant ESPs, such as grain size distribution and aggregation properties as they have a strong influence on ash deposition rates and location. Furthermore, the ash dispersion run times could be optimized by dynamically adapting each model configuration to the eruption size.

Finally, we would like to note that our methodology can also be applied to volcanoes in other regions or other particle transport problems such as air pollutants or radionuclides (Stein et al. 2015a) given the right database to quantify prior parameter distributions.

Abbreviations

D: Eruption Duration; ESPs: Eruption Source Parameters; GEFSv12: Global Ensemble Forecast System version 12; GENS-3: GEFSv12 reforecasts; H: Eruption Plume Height; LHS: Latin Hypercube Sampling; MER: Mass Eruption Rate; NOAA: National Oceanic and Atmospheric Administration, United States; NWP: Numerical Weather Prediction; PDF: Probability Density Function; VAAC: Volcanic Ash Advisory Centre.

Authors' contributions

TH, ND and YB compiled and built PDFs for the volcanic ESP. YB implemented LHS technique with dependence. RT collected and processed NWP data and ran the dispersion model experiments. RT and YB explored and processed the model output into the final summary maps. All authors drafted the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

The scripts and dataset(s) supporting the conclusions of this article are available in the GitHub repository, https://github.com/yannikbehr/ashfall_forecasts and Trancoso, R, Behr, Y, Hurst, T, & Deligne, N. (2022). HYSPLIT ensemble model output for case studies [Data set]. Zenodo.https://doi.org/10.5281/zenodo.6363619.
The HYSPLIT model used in this article was version 5.0.1 from https://www.ready.noaa.gov/index.php. The software is platform independent, based in Fortran programming language and requires external NetCDF libraries. The source code for version 5.0.1 is not public and was made available to MetService by the HYSPLIT developers. However, the most recent version of the model is available as pre-compiled binaries for multiple operating systems, to NOAA users and registered HYSPLIT users.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Metservice, Wellington, New Zealand. [2]GNS Science, Wellington, New Zealand. [3] Hawaiian Volcano Observatory, U.S. Geological Survey, Hilo, United States.

References

Aubry TJ, Engwell S, Bonadonna C, Carazzo G, Scollo S, Van Eaton AR et al (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

Berger J.O., Bayarri M.J., Calder E.S., Dalbey K., Lunagomez S., Patra A.K., Bruce E., Pitman E.T. and Wolpert R.L., 2011. Risk Assessment for Pyroclastic Flows: Combining Deterministic and Statistical Modeling. Proceedings of the 26th International Workshop on Statistical Modelling: Valencia, July 11–15, 2011, p.3 -- 9.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer Science+Business Media, LLC. https://doi.org/10.1198/tech.2007.s518BOMS. BOMS Analysis Chart Archive. corporateName=Bureau of Meteorology; 2022 [cited 2022 Mar 15]. http://www.bom.gov.au/australia/charts/archive/index.shtml Available from

Bonadonna C, Costa A, Folch A, Koyaguchi T. Chapter 33 - Tephra Dispersal and Sedimentation. In: Sigurdsson H, editor. The Encyclopedia of Volcanoes (Second Edition). Amsterdam: Academic Press; 2015 [cited 2022 Feb 2]. p. 587–97. https://www.sciencedirect.com/science/article/pii/B978012385938900033X Available from

Carey S, Bursik M. Chapter 32 - Volcanic Plumes. In: Sigurdsson H, editor. The Encyclopedia of Volcanoes (Second Edition). Amsterdam: Academic Press; 2015 [cited 2022 Feb 2]. p. 571–85. https://www.sciencedirect.com/science/article/pii/B9780123859389000328 Available from

Cheung KKW (2001) A review of ensemble forecasting techniques with a focus on tropical cyclone forecasting. Meteorol Appl 8(3):315–332

Constantinescu R et al (2021) The radius of the umbrella cloud helps characterize large explosive volcanic eruptions. Commun Earth Environ 2:1–8

Costa, A., Pioli, L., Bonadonna, C., 2016. Assessing tephra total grain-size distribution: Insights from field data analysis. Earth and Planetary Science Letters 443, 90–107. https://doi.org/10.1016/j.epsl.2016.02.040Deligne NI. Mass eruption rate, column height, and duration dataset for volcanic eruptions. Mass eruption rate, column height, and duration dataset for volcanic eruptions. GNS Science; 2021. https://shop.gns.cri.nz/sr_2021-12-pdf/ Available from

Deligne NI (2021) Mass eruption rate, column height, and duration dataset for volcanic eruptions. GNS Science Report. 10.21420/P18W-7674

Freret-Lorgeril, V. et al. Mass Eruption Rates of Tephra Plumes During the 2011–2015 Lava Fountain Paroxysms at Mt. Etna From Doppler Radar Retrievals. Frontiers in Earth Science 6, (2018).

Hort, M. VAAC Operational Dispersion Model Configuration Snap Shot. In Proceedings of the Conjoint 7th WMO VAAC Best Practices Workshop (VAAC BP/7) and 9th WMO/IUGG Volcanic Ash Scientific Advisory Group Meeting (VASAG/9), Washington, DC, USA, 21–22 November 2019)

Hurst T, Davis C (2017) Forecasting volcanic ash deposition using HYSPLIT. J Appl Volcanol 6(1):5

Hurst T, Smith W (2004) A Monte Carlo methodology for modelling ashfall hazards. J Volcanol Geoth Res 138(3):393–403

Jenkins S, Wilson T, Magill CR, Miller V, Stewart C, Blong R, et al. Volcanic ash fall hazard and risk. In: Loughlin S, Sparks S, Brown S, Jenkin S, Vye-Brown C, editors. Global Volcanic Hazards and Risk. Cambridge: Cambridge University Press; 2015 [cited 2022 Feb 2]. https://research-information.bris.ac.uk/en/publications/volcanic-ash-fall-hazard-and-risk-2 Available from

Leonard GS, Cole RP, Christenson BW, Conway CE, Cronin SJ, Gamble J, et al. Ruapehu and Tongariro stratovolcanoes: a review of current understanding. Open Access Te Herenga Waka-Victoria University of Wellington; 2021 Jan 1 [cited 2022 Mar 15]; https://openaccess.wgtn.ac.nz/articles/journal_contribution/Ruapehu_and_Tongariro_stratovolcanoes_a_review_of_current_understanding/17128517/1 Available from

Longchamp C, Bonadonna C, Bachmann O et al (2011) Characterization of tephra deposits with limited exposure: the example of the two largest explosive eruptions at Nisyros volcano (Greece). Bull Volcanol 73:1337–1352. https://doi.org/10.1007/s00445-011-0469-9

Lorenz EN (1969) The predictability of a flow which possesses many scales of motion. Tellus 21(3):289–307

Magill CR, Hurst AW, Hunter LJ, Blong RJ (2006) Probabilistic tephra fall simulation for the Auckland Region, New Zealand. J Volcanol Geoth Res 153(3):370–386

Mastin LG, Guffanti M, Servranckx R, Webley P, Barsotti S, Dean K et al (2009) A multidisciplinary effort to assign realistic source parameters to models of volcanic ash-cloud transport and dispersion during eruptions. J Volcanol Geoth Res 186(1):10–21

Mckay MD, Beckman RJ, Conover WJ (2000) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics. 42(1):55–61

Murphy KP. Probabilistic Machine Learning: An introduction. MIT Press; 2022 [cited 2022 Feb 3]. https://probml.github.io/pml-book/book1.html Available from:

NOMADS Server. NOMADS-NOAA Operational Model Archive and Distribution System. 2022 [cited 2022 Feb 14]. https://nomads.ncep.noaa.gov/ Available from

Packham N, Schmidt WM. Latin Hypercube Sampling with Dependence and Applications in Finance. Rochester, NY: Social Science Research Network; 2008 Oct. Report No.: ID 1269633. https://papers.ssrn.com/abstract=1269633 Available from

Prata AT, Dacre HF, Irvine EA, Mathieu E, Shine KP, Clarkson RJ (2018) Calculating and communicating ensemble-based volcanic ash dosage and concentration risk for aviation. Meteorol Appl 26(2):253–266

Prata F, Rose B. Chapter 52 - Volcanic Ash Hazards to Aviation. In: Sigurdsson H, editor. The Encyclopedia of Volcanoes (Second Edition). Amsterdam: Academic Press; 2015 [cited 2022 Feb 3]. p. 911–34. https://www.sciencedirect.com/science/article/pii/B9780123859389000523 Available from

Rolph G, Stein A, Stunder B (2017) Real-time Environmental Applications and Display sYstem: READY. Environ Model Softw 1(95):210–228

Scott DW (1992) Multivariate Density Estimation: Theory, Practice, and Visualization. John Wiley & Sons, New York, Chicester

Sigg R, Lindgren P, von Schoenberg P, Persson L, Burman J, Grahn H et al (2018) Hazmat risk area assessment by atmospheric dispersion modelling using Latin hypercube sampling with weather ensemble. Meteorol Appl 25(4):575–585

Spiller ET, Bayarri MJ, Berger JO, Calder ES, Patra AK, Pitman EB, Wolpert RL (2014) Automating emulator construction for geophysical hazard maps. SIAM/ASA J Uncertainty Quantification 2:126–152. https://doi.org/10.1137/120899285

Stajner I, Tallapragada V, Zhu Y, Alves H, McQueen J, Hamill T, et al. NOAA's unified forecast System for sub-seasonal predictions: Development and operational implementation plans of Global Ensemble Forecast System v12 (GEFSv12) at NCEP. Copernicus Meetings; 2020 Mar. Report No.: EGU2020–6212. https://meetingorganizer.copernicus.org/EGU2020/EGU2020-6212.html. Available from

Stein AF, Draxler RR, Rolph GD, Stunder BJB, Cohen MD, Ngan F (2015a) NOAA's HYSPLIT atmospheric transport and dispersion modeling system. Bull Am Meteorol Soc 96(12):2059–77

Stein AF, Ngan F, Draxler RR, Chai T (2015b) Potential use of transport and dispersion model ensembles for forecasting applications. Weather and Forecasting American Meteorological Society. 30(3):639–55

Suzuki T. A Theoretical Model for Dispersion of Tephra. Arc volcanism: physics and tectonics. Shimozuru, D., Yokohama, I. Tokyo: Terra Scientific Publishing; 1983. p. 95–113. https://pages.mtu.edu/~raman/Ashfall/Syllabus/Entries/2009/6/21_Ash_Blankets_files/Suzuki83.pdf Available from

Thompson MA, Lindsay JM, Gaillard J (2015) The influence of probabilistic volcanic hazard map properties on hazard communication. J Appl Volcanol 4:6. https://doi.org/10.1186/s13617-015-0023-0

USGS. Volcanic Ash Impacts & Mitigation - Volcanic Ash Impacts & Mitigation. 2022 [cited 2022 Feb 27]. https://volcanoes.usgs.gov/volcanic_ash/ Available from

Wilson TM, Jenkins S, Stewart C. Chapter 3 - Impacts from Volcanic Ash Fall. In: Shroder JF, Papale P, editors. Volcanic Hazards, Risks and Disasters. Boston: Elsevier; 2015 [cited 2022 Feb 2]. p. 47–86. https://www.sciencedirect.com/science/article/pii/B9780123964533000034 Available from

Zhang F, Sun YQ, Magnusson L, Buizza R, Lin S-J, Chen J-H et al (2019) What Is the Predictability Limit of Midlatitude Weather? Journal of the Atmospheric Sciences American Meteorological Society. 76(4):1077–91

Publisher's Note