https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Towards the European Tephra Hazard Map: Methodology and Catalogue of Footprints

Arnau Folch[1], Laura Sandri[2], Sara Barsotti[3], Simone Aguiar[4], Beatriz Martínez Montesinos[1,2], Alejandra Guerrero[1], Pablo Tierz[1], Heribert Pascual[1], Eva Hernandez[1], Leonardo Mingari[1], and Antonio Costa[5]

[1]Geociencias Barcelona (GEO3BCN-CSIC), Barcelona, Spain
[2]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Italy
[3]Icelandic Meteorological Office (IMO), Iceland
[4]Instituto de Investigação em Vulcanologia e Avaliação de Riscos (IVAR), Ponta Delgada, Portugal
[5]Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Sede di Perugia, Italy

Correspondence: Arnau Folch (afolch@geo3bcn.csic.es)

Abstract. Many parts of Europe are exposed to tephra hazards that jeopardise regions around volcanoes and the continental airspace. Quantifying and mapping the impacts from future eruptions is relevant to aviation stakeholders and to national and regional civil protection administrations and agencies. This paper presents a novel methodology for long- and short-term Probabilistic Volcano Hazard Assessment (PVHA) from tephra based on a large catalogue of footprints and associated metadata, each footprint representing a single-scenario model realisation. The methodology is used to generate a first version of the European Tephra Hazard Map (ETHM), a collaborative initiative aiming at providing single- and multi-volcano mappings consistent and homogeneous across spatio-temporal scales and volcanoes. More than 53,000 footprints from 12 European volcanoes were computed by running the FALL3D atmospheric dispersal model on the accelerated partition of the MareNostrum-5 supercomputer. In regional model nests (2 km model grid resolution), footprints identify areas prone to hazardous tephra fallout. At European scale (10 km model grid resolution), footprints delineate hazardous airspace regions in terms of ash concentration at critical flight-level layers. The catalogue of footprints is stored in a Simulation Data Lake (SDL@CINECA) to ensure accessibility and re-usability, as well as to ease the addition of footprints from other volcanoes in the future. In parallel, a lightweight service facilitates downstream hazard mapping and dynamic update of hazard based on user-given Probability Density Functions (PDFs), decoupling the computational workload of simulating the physical process from the actual hazard quantification.

1 Introduction

Volcanic hazard assessment is a major topic of modern volcanology and the scientific pillar underpinning increased awareness of future events and the mitigation of eruption impacts on individuals, society and the environment. Among the many volcano-related phenomena, tephra dispersal and fallout is the most-widely spread volcanic hazard, with potential impacts spanning from regional-scale deposition to continental-scale airspace disruptions punctuated in time. Ash clouds jeopardise civil aviation (e.g., Prata and Rose, 2015; Aydar et al., 2024), while tephra fallout can trigger a variety of impacts on build-

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

ings, infrastructures, transport networks, energy lines, airports, power plants, water supply systems, or agriculture (e.g., Wilson et al., 2014). Based on tephra hazard mappings, civil protections can implement evidence-based emergency plans and evacuation routes, administrations can plan safer critical infrastructures (e.g. McBirney et al., 2003), and civil aviation stakeholders can optimise engine maintenance costs or ban airspace regions in the presence of volcanic clouds jeopardising aircrafts (e.g. Folch and Sulpizio, 2010). The key question in any volcano hazard assessment is how to capture the inherent uncertainty regarding the characteristics of future volcanic activity, including its timing and frequency. Physics-based Probabilistic Volcano Hazard Assessment from Tephra (PVHA-T or simply PVHA) is the standard methodology to consistently integrate through physical and statistical modelling the uncertainties from both the volcanic system, the so-called Eruption Source Parameters (ESPs), and the atmospheric conditions that will drive the dispersal of tephra during a future eruption. The standard PVHA workflow entails a series of steps that include:

1. Definition of the range of eruptive scenarios, in which volcanologists constrain possible scenarios depending on a combination of ESPs. For example, ranges for the Volcanic Explosivity Index (VEI; Newhall and Self, 1986) or for the total Erupted Mass (M) or for the Eruption Magnitude (M_g; Mason et al., 2004) can be used to bound uncertainties on the erupted mass/volume of the tephra. Similarly, ranges for Eruption Duration (D) or for Mass Eruption Rate (MER or simply $\dot{M}$) or for plume height (H_c) can be used as proxies to bound the Eruption Intensity (I). The definition of scenario ranges for a specific volcano is underpinned by the geological record and the eruption catalogues informing on its past behaviour or, as an alternative for poorly-characterized systems, by using sets of analogue volcanoes (Tierz et al., 2019).

2. Random sampling of ESPs within the scenario range in order to create a sufficiently large number of "what-if" samples covering all plausible outcomes. Early PVHAs relied on the concept of "single eruption scenario", assuming that the next eruption from a given volcano will be similar to a fixed "representative scenario" and neglecting the natural variability and the knowledge-related epistemic uncertainties. Today, modern PVHAs (e.g. Sandri et al., 2016) circumvent this limitation by using Probability Density Functions (PDFs) to explore a continuum of ESPs and their likelihood within the considered range, together with the unknown meteorological conditions. A volcano PDF represents the relative likelihood that a continuous random variable (a certain ESP) takes on a specific value, with the area under the curve giving the probability of the ESP falling within a range of values.

3. For every sampled scenario and meteorological conditions, run a physics-based model for the transport and deposition of tephra. Given the amount of samples involved (from many hundreds to thousands) and the resulting computational cost, early PVHAs performed during the 90s and 2000s (e.g. Barberi et al., 1990) relied on wind profiles and semi-analytical Gaussian models such as HAZMAP (Pfeiffer et al., 2005; Costa et al., 2009) or TEPHRA2 (Bonadonna et al., 2005). Nowadays, capacity High-Performance Computing (HPC) allows to perform PVHAs with fully-numeric models that are more generic, account for real 4D meteorological patterns and, moreover, can compute airborne concentrations and timestamps in addition to the fallout deposit. Several examples of fully-numeric PVHA exist after the pioneer works of Folch and Sulpizio (2010) and Sulpizio et al. (2012) originally focused on Vesuvius, *e.g.* for Popocatépetl (Bonasia

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

et al., 2014), Icelandic volcanoes (Biass et al., 2014), Neapolitan volcanoes (Sandri et al., 2016), Tacaná (Vázquez et al., 2019), Jan Mayen (Norway) (Titos et al., 2022) or Campi Flegrei (Martínez Montesinos et al., 2022) among others.

4. Finally, quantification of tephra hazard by statistically combining single-scenario simulations to produce exceedance probability maps, showing the probability that a specific hazard intensity measure (*e.g.*, deposition of 10 cm of ash or concentration above 2 $\mathrm{mg/m^3}$) is exceeded at a given point or atmospheric level; isomass hazard maps, showing the expected load for a fixed probability level (*e.g.*, the "1-in-100-year" fallout deposit), etc. Many types of output maps are possible but, in any case, the final goal of any PVHA is to move from the PDF describing what a volcano might do to curves and maps showing, respectively, the likelihood of impacts at specific points or over a regular grid.

The time frame or "time window" is an important aspect of volcanic hazard, with a distinction often made between long-term, short-term and real-time frames (e.g. Marzocchi and Bebbington, 2012). The so-called long-term PVHA is commonly done during "peace time" and considers long (years to decades) time scales, running models with many years of past weather data (re-analyses datasets like ERA-5) and assuming that the past atmospheric conditions are statistically representative of future weather patterns (Cioni et al., 2003). Short-term (weeks to months) PVHA is performed during volcanic crises or unrest episodes and can include information from volcano monitoring to further constrain uncertainties on ESPs together with seasonal weather forecasts (Selva et al., 2014). Finally, real-time (hours to days) PVHA is done during ongoing eruptions driving dispersal models with operational weather forecasts and, eventually, involving assimilation of observations into models through ensemble-based forecasts (Pardini et al., 2023). Note that, from a conceptual point of view and in the context of a probabilistic framework, these different time frames can be viewed as parts of a continuum in which PDFs evolve in time, with uncertainties growing from the better-constrained real-time to the poorer-constrained long-term time frames. In fact, a purely deterministic forecast can be considered as an end-member situation in which a PDF collapses to a Dirac delta, reducing ESPs and weather uncertainties to zero. On the other hand, the time window is also relevant to multi-volcano PVHA, used to quantify the collective hazard posed by several active volcanoes to a specific region. Unlike traditional single-volcano mapping, this approach "stacks" several PVHAs to account for the overlapping impact of multiple eruptive centres considering their relative eruption frequency during the period of interest. Examples of multi-volcano PVHA using semi-analytical tephra dispersal models include applications to New Zealand (Hurst and Smith, 2010) or the Asia-Pacific region (Jenkins et al., 2012, 2018; Bear-Crozier et al., 2016). Only more recently, Massaro et al. (2023) carried out the first multi-volcano PVHA using a fully numerical model for tephra fallout in southern Italy. However, similar studies have not been conducted yet over wider regions or for airborne hazard.

The standard PVHA workflow described above presents an important caveat, namely, that the sampling of scenarios by means of PDFs occurs upstream of the modelling component (see Fig. 1). As a result, new sets of simulations are required each time a PDF varies, something that can happen for several reasons. First of all, PDFs are built upon volcano catalogues that can be very irregular in spatio-temporal coverage and populated with non-standardized and heterogeneous records, often with insufficient metadata characterization. On the other hand, significant gaps exist between the actual number of past eruptions and the number of records in volcano databases (this is known as the completeness problem, see *e.g.* Jenkins et al., 2012;

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Burgos et al., 2022). All these aspects introduce subjective biases when defining a PDF. On the other hand, PDFs may be updated whenever new field studies unveil uncatalogued events or simply redefine some records in a catalogue, *e.g.* when a new stratigraphic study is able to discern different eruptive events within a previous single record. Finally, PDFs can vary in time reflecting our better knowledge of the system, *e.g.* if during an unrest episode volcano monitoring can constrain the range of expected scenarios. A key question in all these situations is weather an existing hazard mapping can be updated without rerunning the scenario modelling step of the PVHA workflow. Note that the answer to this question relates to decoupling the simulated scenarios from their probabilities of occurrence, a "philosophical approach" underpinning statistical emulation methods for geophysical hazard mapping (Bayarri et al., 2009; Spiller et al., 2014).

The objective of this paper is twofold. First, it presents a novel methodology for long- and short-term PVHA based on the catalogue of footprints concept (see Fig. 1). The innovative idea here is not just to sample over a "full-range" of scenarios but to create and store a massive and homogeneous collection of footprints. A footprint can be defined as a single model realization together with associated metadata regarding the scenario and the model inputs, which presents benefits in terms of storage optimisation and workflow orchestration. Compared to the classical "PDF-driven" scenario sampling, much narrower in range, this novel approach pays the penalty of needing a larger number of runs (footprints span low or even negligible probability scenarios laying outside the bounds of a PDF). However, the "catalogue-based" approach presents a tremendous advantage: footprints are computed only once and stay independent of the PDFs, which are used downstream in the workflow to weight all (or a subset of) the footprints in the catalogue. This allows, for example, to update/re-compute hazard mappings through a lightweight post-process service at almost no computational cost. Using this methodology and as a second objective, this paper also introduces the European Tephra Hazard Map (ETHM) initiative. Many parts of Europe are exposed to hazardous phenomena that precede, accompany and follow volcanic eruptions (Gasparini, 1993), and this manuscript presents the first step towards a continental-scale tephra hazard mapping consistent across spatio-temporal scales and volcanoes. Section 2 introduces the design requirements and the basic concepts behind the ETHM project, which has been developed by the EuroHPC Center of Excellence for Exascale in Solid Earth (ChEESE; Folch et al., 2023). Section 3 presents the methodology of the ETHM and the software stack to run each step of the workflow, including the scenario sampling, the footprint generation and cataloguing, and the single- and multi-volcano mapping. As detailed in Section 4, the off-line part of the ETHM workflow, which serves to populate the catalogue of footprints, can be run using a Common Workflow Language (CWL) file that orchestrates and automates its execution. Finally, Section 5 draws some conclusions and discusses possible next steps for this collaborative initiative. The subsequent hazard map creation service constitutes the bulk of a second paper (Guerrero et al., 2026b) and is not discussed in depth here.

2 Requirements and Overview of the ETHM

The methodology and characteristics of the ETHM initiative were designed considering the following a priori requirements:

1. The ETHM should be user-oriented, meaning that end-users must have the option to furnish their own PDFs to build customized mappings from a vast catalogue of footprints and by means of a lightweight map-creation service. The

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

only limitation of the methodology is that the lower/upper limits of a user-given PDF should lay within the sampled ranges of ESPs. As detailed in Section 3.3, footprints for the ETHM were generated running the FALL3D model (Folch et al., 2020) over 30 years of ERA-5 re-analysis data (Hersbach et al., 2020) but, in principle, the rest of the workflow and the mapping methodology are model-agnostic. Beyond the FALL3D model specificities, the idea is to separate the volcanological expertise, assumed to be on the user-end, from the computational steps that require different sets of skills, regular access to supercomputing resources, and data storage, preservation, and curation capabilities (see Sec. 3.4).

2. The ETHM should allow for single and multi-volcano hazard mapping. End-users must have the option to perform both single- and multi-volcano mappings by blending contributions from several volcanoes in a selected region (*e.g.* Italy) or even from the whole continent. In addition, the mapping service must be able to deliver either conditional or absolute probability maps (Marzocchi et al., 2010). Note that conditional maps are possible only for single-volcano mapping because a time frame is needed in order to blend contributions from two or more volcanoes on a single map.

3. The ETHM should embrace both long- and short-term time scales, the later consisting on mapping with PDFs informed by volcano monitoring and/or on filtering footprints by season or sets of months (*e.g.* use only winds from winter months, use only October winds, etc).

4. The ETHM should be multi-scale, covering both short-range tephra fallout and long-range ash cloud dispersal at relevant atmospheric levels, ensuring map consistency across continental and regional scales. To this end, all simulations follow a nested approach with footprints sharing a same continental-scale computational domain but with different regional higher-resolution nests. Here, the only limitation to both single- and multi-volcano mapping, is the implicit assumption that tephra fallout hazards are limited to the regional nests (few hundreds of kilometres around a volcano). In contrast, collective ash cloud hazards can span over the continental domain (several thousand of kilometres).

5. The ETHM should be open source and as collaborative as possible. The methodology and the accompanying software stack must ease the future addition of footprints from other volcanoes not included in this first version. In this sense, the ETHM project is intended as a multi-step and long-term collaborative initiative, where the scientific community must be able to progressively populate the catalogue of footprints by following the methodology introduced here.

6. Finally, the ETHM methodology should be exportable and replicable to other parts of the world. As a first step, an on-going collaboration is already replicating the methodology for different Mexican volcanoes.

With the above premises in mind, this first version of the ETHM populates the catalogue with footprints from 12 volcanoes located in 4 different regions (Azores, Canaries, Iceland and Italy). These pioneer volcanoes, listed in Table 1, were selected based on their current state of unrest and/or their high-potential risk in case of reactivation. The computational domains used to run the footprints are shown in Figure 2 and consist on a unique European-scale domain at 10 km grid resolution and four regional nests at 2 km grid resolution. According to the European Catalogue of Volcanoes and Volcanic Areas (ECV) of the International Association of Volcanology and Chemistry of the Earth's Interior (IAVCEI), there are about 50 active volcanoes in

Europe, but this count number slightly varies depending on the definition of "active" and on the geographic span (Barsotti et al., 2024). It is clear that a catalogue including footprints from "all volcanoes" will require more time and resources. However, and as a starting point, we consider that the selected subset of 12 volcanoes already provides a first valuable assessment for Europe. The unprecedented number of high-resolution footprints should enable the quantification of hazard and uncertainty even for low-probability but high-impact events (end-member scenarios), both on the ground (ash fallout hazard on roofs, roads, assets) and at relevant flight levels (hazard on aviation routes and airport operations).

3 Methodology

Figure 3 sketches the ETHM workflow, divided in a series of off-line (HPC upstream) and on-line (HPC downstream) workflow steps. The off-line steps need to be run only when a new volcano is added to the catalogue, and yield to the storage and cataloguing of footprints and associated metadata in a Simulation Data Lake. The methodology underpinning the different ETHM workflow steps is the main topic of this paper and is explained in detail along the following sub-sections. Additionally, Section 4 explains how to execute the off-line steps of the ETHM workflow in order to contribute to enlarge the catalogue of footprints consistently. The on-line workflow steps orchestrate the final single- and multi-volcano mappings and can be run by users in-house or, preferentially, through a post-process on-demand service described in a second paper (Guerrero et al., 2026b).

3.1 Scenario Sampling

For each volcano, the first step is to sample the ESPs (*i.e.* the model inputs defining the tephra emission terms and the particle granulometry) within all plausible scenarios. Several choices exist considering that not all ESPs are independent and that, in addition, empirical relationships can be used to infer some ESPs from other(s) in a consistent way. We opted for a six-dimensional sampling space for the following set of ESPs:

1. Eruption start timestamp (T_o), *i.e.* eruption start day and hour with dates comprised between 01/01/1995 and 31/12/2024 in order to span 30 years of ERA-5 re-analysis data.

2. Erupted mass (M). When sampling, erupted mass is actually converted to Eruption Magnitude (M_g), defined as $M_g = log_{10} M - 7$ (Pyle, 2015). This conversion is done in order to have a linear sampling scale, more suitable for parameters which, like mass, can vary across several orders of magnitude. Note that, in the context of the ETHM, the erupted mass refers only to tephra and not to lavas, pyroclastic flows, or any other mass contribution. In the case of mixed effusive-explosive activity, only the former contribution should be included. The erupted volume of tephra and the Volcanic Explosivity Index (VEI; Newhall and Self, 1986) can be obtained from mass by assuming a Dense Rock Equivalent (DRE) and a bulk density respectively.

3. Eruption Duration (D), up to a maximum of 7 days (168 hours) for the sake of limiting the computational time of footprints. It is obvious that longer-lasting eruptions occur in nature, often alternating or combining phases of effusive

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

and explosive activity (e.g. Jenkins et al., 2007). In this case, footprints must be understood as representing a single time window or eruptive phase of a long-lasting eruption. The combination of total erupted mass M and eruption duration D yields to the time-averaged mass eruption rate ($\dot{M} = M/D$) and to the eruption intensity (I), defined as $I = log_{10}\dot{M} + 3$ (Pyle, 2015). Furthermore, a time-averaged column height (H_c) can be inferred from $\dot{M}$ using an empirical power-law (e.g. Mastin et al., 2009). However, the limitations and large error bounds associated with of these type of empirical relationships are well-known (see Costa et al., 2026, for a detailed review). This drawback is circumvented here by introducing a fourth sampling parameter that perturbates the "Mastin-like" column height (H_M).

4. Standard column height deviation (σ_c). Following Aubry et al. (2023), we consider a value of $\sigma_c = 3.5$ km, meaning that the central "Mastin-like" value is randomly perturbed between $H_M \pm 3.5$ km. Moreover, a minimum floor value of 2 km above vent level is imposed to H_c, so that $H_c = \max(2, H_M \pm 3.5)$. Minor eruptions with column heights below this threshold are not included in the footprints.

5. Magma viscosity (μ). This ESP is used together with the inferred column height H_c to estimate the Total Grain Size Distribution (TGSD) following Costa et al. (2016), which correlated bi-lognormal distributions (bi-Gaussian in ϕ-units) to a variety of eruptive styles and magma compositions.

6. Finally, the the so-called Suzuki parameter (A), commonly used to describe how tephra and volcanic gases are distributed vertically within an eruption column (e.g. Suzuki, 1983; Pfeiffer et al., 2005).

Sampling these 6 ESPs and the derivatives thereof provides all the volcanological inputs necessary to run the FALL3D model. However, given that some ESPs are not independent, an important aspect is how to define the sampling range along each dimension in order to avoid unrealistic combinations of values. For example, erupted mass and duration are typically bind, with higher and lower magnitude eruptions displaying shorter and longer durations respectively. These types of dependencies can be included when sampling the ESPs by defining the ESPs ranges over a small number of "representative classes", typically up to three (*e.g.* small-, medium- and large-size scenarios). Table 2 shows the sampled ranges for the 12 volcanoes, based on the following rationale:

– **Sete Cidades**: two classes covering small- and medium-magnitude explosive activity. Eruption durations were adjusted to obtain $\dot{M}$ and H_c consistent with available field observations (Cole et al., 2008; Queiroz et al., 2008, 2015). For the three Azores volcanoes, the viscosity range was adopted based on trachytic magmas from Santa Barbara eruption (Sete Cidades) (Laeger et al., 2019) and from Furnas and Pico Alto volcanoes (Jeffery et al., 2016, 2017).

– **Furnas**: three classes representing small, medium and large explosive eruptions. Eruption durations were adjusted to obtain $\dot{M}$ and H_c consistent with field-based estimates (Cole et al., 1995, 1999; Guest et al., 1999, 2015).

– **Fogo**: three classes corresponding to small-, medium- and large-explosive eruptions identified in the geological record. Eruption durations were adjusted to obtain $\dot{M}$ and H_c consistent with field-based estimates (Walker and Croasdale, 1971; Wallenstein et al., 2005; Aguiar, 2019).

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

- **Teide**: two classes corresponding to Strombolian to violent Strombolian eruptions (Martí et al., 2008) and to sub-Plinian phonolitic events (Martí et al., 2012) respectively. The magmatic parameters were derived from the petrological analyses (Dorado et al., 2023).
- **La Palma**: a single class reflecting the historical activity (Longpré and Felpeto, 2021) and incorporating the magnitude and phenomenology of the latest 2021 eruption (Bonadonna et al., 2023).
- **Reykjanes**: a single class with eruption intensity calibrated in light of the Reykjanes fires (1210-1240 CE), when an explosive eruption occurred offshore the SW of Reykjanes (Larsen et al., 2013). No larger ash-rich eruptions are known for this volcanic system (Sigurgeirsson and Einarsson, 2019).
- **Hekla**: three classes (small, moderate and large) based on the explosive phases of eruptions for which tephra volumes are adequately known (Larsen and Thordarson, 2019). Hekla is the third most active volcano in Iceland during the last millennium, with 23 recorded eruptions (VEI 0-5). About 100 eruptions are known in the last 9000 years, including purely explosive, purely effusive and mixed eruptions. The largest explosive eruptions (VEI 6) occurred 4300 and 3000 years ago. Around 28% to class 1, 60% belong to class 2, and 11% class 3 (Larsen et al., 2013).
- **Katla**: three classes based on the Catalogue of Icelandic Volcanoes (Larsen and Guðmundsson, 2019). Characterised by explosive basaltic eruptions occurre at the central volcano, with tephra volumes (bulk volume) ranging from 0.02 to over 2 km^3 accompanied by jökulhlaups. Looking to the last 20 historical eruptions occurred over a period of 1100 years, class 1 is the most frequent scenario (up to 45%), whereas the classes 2 and 3 have equally occurred with a frequency of 25%. Larger eruptions are considered possible but quite rare.
- **Campi Flegrei**: three classes corresponding to small-, medium- and large-explosive eruptions (Orsi et al., 2009) or, respectively, the classes 1, 2 and 3 considered in Sandri et al. (2016) in terms of total erupted mass. Include the reference explosive eruptions of Averno (class 1), Astroni (class 2) and Agnano-MonteSpina (class 3), that, together with the effusive class (Monte Nuovo eruption as reference) cover the range of sizes observed in the last 15ky (*i.e.* after the caldera-forming Neapolitan yellow tuff event). The range in durations is justified to provide values of $\dot{M}$ and H_c consistent with these three scenarios.
- **Vesuvius**: three classes corresponding to violent Strombolian, sub-Plinian and Plinian activity respectively (Sandri et al., 2016). The three size classes include the reference eruptions for Somma-Vesuvius activity (Macedonio et al., 2008), *i.e.* 1944 AD (violent Strombolian, class 1), 1631 AD and 472 AD (sub-Plinian, class 2) and 79 AD (Plinian, class 3). The range in durations is justified to provide values of $\dot{M}$ and H_c consistent with these three scenarios.
- **Ischia**: a single class exploring the largest type of eruption occurred in the last 10ky (Massaro et al., 2023) that is the Cretaio Eruption (Orsi et al., 1996; Selva et al., 2019). This class characterizes the only eruptive activity at Ischia, whose eruptive products are able to affect Italy mainland, beyond the island itself. The mass associated with eruption is almost entirely due to a fallout phase (Primerano et al., 2021).

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

– **Etna**: a single class with footprints considering paroxysmal eruptions for which the duration (from 0.5 to 4 hours) must be understood as an "effective" duration during which most of the tephra is emitted.

Once the sampling ranges were defined for each ESP, we sampled our 6-dimensional distribution using an Optimised Latin Hypercube Sampling (OLHS; Tang, 1993). Sampling was performed using the quasi-Monte Carlo module in the SciPy's library (Virtanen et al., 2020) with a random-cd optimization, which applies repeated random permutations of coordinate strata to iteratively minimize the centred discrepancy of the sample. Cantered-discrepancy minimization is known to provide greater robustness in 2D and 3D sub-projections compared to other discrepancy metrics, and this behaviour extends well to higher-dimensional spaces such as our 6D parameter domain. The only limitation of the OLHS is that the number of sampled points must be equal to the square of a prime number. Here we considered $47^2 = 2209$ samples. In other words, we generated 4418 footprints per volcano, 2209 for the European domain and 2209 for the regional ones, summing a total of 53,016 footprints (12 volcanoes x 4418). The specific number of OLHS samples was selected as a compromise between having a sufficiently large number of scenarios per volcano and the computational time and resources needed to run and to store more than 53,000 FALL3D simulations. The next possible number compatible with the OLHS constrains would have been $53^2 = 2809$ samples per volcano, which implies an increase in computational and storage resources by almost 30%. As an illustrative example, Figure 4 shows the sampling results for the particular case of Vesuvius volcano (Italy), for which 3 classes were considered (see Table 2).

3.2 Model Inputs

The next ETHM workflow step is to map the sampled values from the OLHS unit space to the actual range of ESP values. To this end, a python script automates the assignment of values for each sampled footprint by using a template for the main FALL3D input file. In addition, the script creates all the ETHM folder tree structure, with one folder per footprint and ordering footprints by region and volcano ID. Table 3 summaries the FALL3D model setup including domains, resolution and particle ranges. Note that a particle cut-off (*i.e.* an effective granulometry) is used in both types of domains; only fine and coarse particles are considered for long- and short-range dispersal respectively. Consequently, simulations in the European-scale domain exclude particles larger than 64 μm (4ϕ) and simulations in the regional nests neglect particles smaller than 128 μm (3ϕ). These two cut-off are justified by the different target hazards of the respective footprints. On the other hand and for the sake of simplicity, ash aggregation was disabled in all simulations. This can entail some over- and under-estimation of hazard for the long- and short-range cases respectively.

The preparation of FALL3D inputs also includes the downloading and cropping of 30 years of ERA-5 meteorological data over the European domain. In general, this can be done by using a second and independent python script that retrieves the required subset of surface and pressure level variables from the Copernicus Climate Data Store server (Hersbach et al., 2023). After data downloading, the script concatenates the fields in "monthly" netCDF files of up to 41 days each. For each month, the 10 extra days are added in order to cover cases in which the eruption start time is sampled at the last days of a month, with a maximum duration of 7 days and leaving 3 extra days for cloud dispersal after the eruption end. As a result, about a 30%

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

overlap exists among two meteorological files from consecutive months. This makes a total of 360 (12 months x 30 years) netCDF files, each of about 13 GB and overall disk space of 4.7 TB. Given the limitations in the Copernicus service queuing request and bandwidth, several weeks were needed to retrieve 4.7 TB of data and upload it to the HPC facilities.

3.3 Footprint Generation

For each target volcano and domain, the 2209 scenarios were simulated running the FALL3D dispersal model version 9.1 (Folch et al., 2020). Simulations were done on the accelerated partition of the MareNostrum-5 supercomputer (MN5-ACC) hosted at the Barcelona Supercomputing Center (BSC). The accelerated partition of MN5 has 1120 computing nodes, each equipped with 4 NVIDIA Hopper (H100) GPUs and 2 Intel Sapphire rapids (80 cores per node). Computational resources to run the ETHM were obtained through an EuroHPC regular access mode project granted with 32,000 node hours on MN5-ACC. The 53,016 FALL3D model executions were launched using 1 node (4 H100 GPUs) per footprint and making use of the slurm array utility to automate the queuing of workloads with common job pattern. The simulation time per footprint is variable, from from one to ten minutes wall clock time depending on factors such as domain size or the number of simulated hours (eruption duration plus 3 extra days for dispersal).

The FALL3D model outputs variables of interest in a netCDF file with a user-defined period, typically every one hour. This standard model setup can produce substantially large output files, something difficult to manage in our case given that the total output size of footprints would exceed 50 TB of disk space. As an alternative, FALL3D can also write the so-called "footprint files", in which only a reduced set of variables relevant to hazard is written in a compressed netCDF file (see Table 4). This alternative output option presents two important advantages. Firstly, the time dimension is eliminated from output and the number of vertical model layers reduced, resulting on output file sizes very much reduced with respect to a standard model output, typically by two or more orders of magnitude. Secondly, footprint output files furnish precomputed hazard results for fixed hazard intensity measures (thresholds of concentration and deposit load), reducing the subsequent blending of footprints to simple matrix-vector product operations. As indicated in Table 4, three threshold values of concentration (0.2, 2 and 10 $\mathrm{mg/m^3}$) were set at relevant flight level layers (surface to FL050, FL050 to FL250 and FL250 to FL400). These three concentration thresholds agree with the new Quantitative Volcanic Ash (QVA) products adopted by the International Civil aviation Organization Meteorology Panel (ICAO) and correspond to low, medium-to-high and very high ash concentrations respectively. Similarly, three thresholds for deposit load hazard (1, 10 and 100 $\mathrm{kg/m^2}$) were set as representative of ground accumulations of 1 mm, 1 cm and 10 cm respectively and accepted as thresholds for tephra-related damage on infrastructures and buildings (e.g. Wilson et al., 2014). For the European domain, the resulting footprint files contain: i) the concentration footprint (*i.e.* a binary value of 1 or 0 indicating if every concentration threshold has been exceeded or not at each point), ii) the arrival time (*i.e.* time spanned between the eruption start and the superation of the threshold, if any), iii) the persistence (*i.e.* total time over the run with concentration above the threshold) and, iv) the maximum of concentration over the simulation time. This variable is stored just in case end-users want to perform a hazard assessment with threshold values (intensity measures) different from the precomputed ones and also to post-process hazard maps maps (see Sec. 3.5). Similarly, for each regional domain, the FALL3D footprint files contain: i) the deposit load footprint (*i.e.* a binary value of 1 or 0 indicating if every deposit

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

load threshold has been exceeded or not at each point), ii) the accumulation time (*i.e.* time needed to accumulate the threshold loads) and, iii) the maximum of load over time. For both types of runs, European and regional, a JSON file with all relevant metadata is attached in order to ease the subsequent blending of footprints and the catalogue discovery. For illustrative purposes Fig. 5 and Fig. 6 show some results for a randomly picked footprint from Katla volcano for the European and the Icelandic domains respectively.

3.4 Footprint Aggregation and Storage

The last two off-line steps of the workflow (see Fig 3) are the aggregation and the storage of footprints to populate the catalogue. For each volcano, single footprint files were aggregated by natural months (footprints from all Januaries winds, from all Februaries, etc) resulting on only 24 netCDF files per volcano, 12 for the European domain plus 12 for the respective regional nest. Considering the 12 volcanoes, the aggregation step reduces the total number of footprint files from 53,016 to 288 (24 files x 12 volcanoes) with the objectives of speeding-up the blending of footprints during the subsequent map generation operations and facilitating the management and storage of the catalogue. Furthermore, single footprint metadata characterizing the eruption scenario (*e.g.* specific values for M, M_g, I, VEI, H_c, etc) is also embedded in the aggregated files.

Aggregated footprint files were stored in the CINECA Simulation Data Lake (SDL) to populate a first version of the ETHM catalogue of footprints. The SDL, developed and implemented by the Geosphere Infrastructures for Questions into Integrated Research (Geo-INQUIRE) project, provides a cloud-based storage and management platform, making large-scale simulation datasets discoverable, reusable, and repeatable for the scientific community, particularly in the Solid Earth and geophysical domains. Data in the SDL is openly available under the CC-BY-4.0 license. The SDL is hierarchically organized in three main entities, namely, experiments, simulations and datasets. Experiments are the highest rank entity and consist on collections of files that can include 1-to-N simulation runs and datasets. A simulation is a realization of a workflow, the life cycle of how data is produced and is made of files that ideally allow to re-run the simulation. Finally, datasets are files or parameter values that are used by the simulations. The SDL infrastructure provides also a Digital Object Identifier (DOI) minting and management service, so that a persistent identifier can be assigned at the experiment level. In order to adapt these SDL entities to our catalogue of footprints we grouped all the volcanoes from a same region in a single experiment and considered the monthly-aggregated footprint files as dataset entities. Table 5 summarizes the characteristics of the first version of the catalogue of footprints hosted in the SDL. The 53,016 footprints and associated metadata from the 12 volcanoes are grouped in 288 files occupying a total of about 65 GB of disk space.

3.5 Hazard Mapping

The on-line part of the ETHM workflow (*i.e.* HPC downstream, see Fig. 3) consists on the generation, retrieval and visualization of European-scale maps at the pre-defined atmospheric layers and of regional-scale maps at ground level. As detailed in Guerrero et al. (2026b), pre-defined PDFs for the 12 volcanoes will be used to compute and distribute static maps through the European Plate Observing System (EPOS) portal. In parallel, a lightweight post-process service allows to compute single- and multi-volcano mappings from user-given PDFs. The tephra hazard mapping service allows user flexibility to provide PDFs in

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

terms of different ESPs for each volcano (*e.g.* give a PDF in terms of VEI, in terms of mass M, in terms of column height H_c, etc) and with any arbitrary binning.

Let's consider that $f(\Phi)$ is the PDF of a random variable Φ, here representing any of the ESPs characterizing a footprint (VEI, H_c, M, etc). Assume also that the continuous PDF $f(\Phi)$ is discretized using n bins with central bin values $(\Phi_1, \Phi_2, ..., \Phi_n)$ and bin widths $(\Delta\Phi_1, \Delta\Phi_2, ..., \Delta\Phi_n)$, in which $f(\Phi)$ takes constant values $(f_1, f_2, ..., f_n)$ with $f_i \geq 0$. Then, by definition, the probability of the random variable Φ taking its value within the i-th discrete bin is given by:

$$P(\Phi_i - \Delta\Phi_i/2 \leq \Phi \leq \Phi_i + \Delta\Phi_i/2) \equiv f_i \Delta\Phi_i = w_i \tag{1}$$

where w_i is the i-th bin area (also referred to as the "bin weight"). From a mathematical perspective, a true PDF is always normalised but in Bayesian statistics and in the context of hazard it is also common to work with "unnormalized densities" that, by abuse of language, are often also called PDFs. For any normalized PDF, here denoted by $f^N(\Phi)$, it holds that:

$$P(\forall\Phi) = \sum_{i=1}^{n} f_i^N \Delta\Phi_i = \sum_{i=1}^{n} w_i^N = 1 \tag{2}$$

that is, the probability of Φ taking any value is exactly 1. In the context of probabilistic hazard, the so-called "conditional maps" are built upon normalized (canonical) PDFs f^N and, therefore, maps are conditional to the materialization of the event triggering the hazard. In our case, conditional maps implicitly assume that the probability of having an eruption (with Φ taking any possible value) in a given time frame is equal to 1. In contrast, the so-called "absolute maps" result from "unnormalized densities" or "absolute PDFs", here denoted by $f^A(\Phi)$, for which the probability P of having an eruption within the time frame $\Delta T = [t, t + \Delta t]$ is:

$$P(\forall\Phi, \Delta T) = \sum_{i=1}^{n} f_i^A \Delta\Phi_i = \sum_{i=1}^{n} w_i^A = \sum_{i=1}^{n} P(\Phi_i, \Delta T) \leq 1 \tag{3}$$

where $P(\Phi_i, \Delta T)$ is the probability of having an eruption within the i-th Φ-bin during the time frame ΔT. Note that the normalization of an absolute PDF f^A implies that:

$$w_i^N \equiv \frac{w_i^A}{\sum_{i=1}^{n} w_i^A} = \frac{w_i^A}{P(\forall\Phi, \Delta T)} \longrightarrow w_i^A = P(\forall\Phi, \Delta T)\, w_i^N \tag{4}$$

which can be seen as a conditional probability explicitly separating the probability along time (when) from the normalized weight of the scenarios (how). Furthermore, as $P(\forall\Phi, \Delta T \to \infty) \to 1$, conditional maps can be viewed as a particular case of absolute maps in which the time frame is made arbitrarily large.

3.5.1 Single-volcano Mapping

Consider a group of m footprints $(\psi_1, \psi_2, ..., \psi_m)$ for a generic quantity ψ. In our particular case, each single footprint $\psi_j(\boldsymbol{x})$ is defined on a 2D regular grid at points $\boldsymbol{x} \in (x, y)$, and can represent a single-scenario probability of exceedance map, an arrival time map, an accumulation map, or a map of disruption persistence. In any case of the cases, it is important to note that,

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

by construction, each single footprint ψ_j is associated to a unique value of the underpinning ESP. As a result, given a PDF in terms of an ESP, the m footprints can be grouped in the Φ-bins so that:

$$\sum_{i=1}^{n} m_i = m \tag{5}$$

where m_i is the number of footprints belonging to the i-th bin of the PDF $f(\Phi)$ and n is the number of Φ-bins as in (2). Different typologies of maps are possible by blending footprints or subsets of footprints from the catalogue (see Table 6 for a summary), including:

1. Probability maps (Ψ_{PB}), giving the probability to exceed a given hazard intensity measure, and computed at each point $\boldsymbol{x}$ as:

$$\Psi_{PB}(\boldsymbol{x}) = P(\forall\Phi, \Delta T) \sum_{i=1}^{n} w_i^N \left(\frac{1}{m_i} \sum_{j=1}^{m_i} [\psi_{PB}]_j^i(\boldsymbol{x}) \right) \tag{6}$$

   where $[\psi_{PB}]_j^i$ are the m_i concentration (or load) single-scenario footprints laying in the i-th bin. In other words, $[\psi_{PB}]_j^i$ are the single-footprint probability maps filled with 0/1 depending on the exceedance of the threshold (intensity measure), and Ψ_{PB} are the final maps resulting from weighting all footprints along the Φ-bins. Note that there is one probability map $\Psi_{PB}(\boldsymbol{x})$ for each threshold of concentration (and atmospheric level) and for each threshold of deposit load. Note also that probability maps can be either absolute maps or conditional maps in case $P(\forall\Phi, \Delta T) = 1$ is assumed in (6).

2. Hazard maps (Ψ_{PR}), giving the intensity of a hazard (airspace contamination or ground accumulation) at a given exceedance probability $y\%$ (conditional to eruption occurrence), and computed at each point $\boldsymbol{x}$ as:

$$\Psi_{PR}(\boldsymbol{x}) = \Pi_{100-y} \left(\sum_{i=1}^{n} w_i^N \left[[\psi_{max}]^i(\boldsymbol{x}) \right] \right) \tag{7}$$

   where:

   – The sum in brackets represents the so-called weighted "statistical mixture" of the maximum of concentration (maximum of load) over time among all the n Φ-bins. The weighted statistical mixing is the weighted average of the respective PDFs from the n bins, that is, results from statistically mixing the n distributions from each Φ-bin (e.g., Selva et al., 2014). It implies that the n bins are considered as n distinct alternative statistical models, and a sample of size N_{tot} (e.g., 1000) for the mixed distribution can be obtained by mixing n samples from the n bins, each of size $w_i^N \times N_{tot}$. In this way, each sample of the mixed distribution comes from one of the distributions relative to the n bins, and single samples are not averaged. This is well explained in Figure 4 of Selva et al. (2014). On the other hand, we consider co-monotonic sampling, meaning that the same set of sampled footprints is used at all grid points, ensuring the samples from each sub-population move in perfect unison.

   – Π_{100-y} represents the $(100 - y)$-th percentile of the maximum of concentration (maximum of load) over time for all single-scenario footprints sampled in the mixed distribution from the n Φ-bins (see point above). By definition, Π_{100-y} has a probability of $y\%$ of being exceeded.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

These hazard maps represent the minimum impact expected with a probability of $y\%$ over a given time window. Note that, by definition, these maps are conditional maps but, nonetheless, time dependency is implicit because the normalized wights w_i^N in (7) can actually evolve with time. For example, if w_2^N and w_4^N represent the probabilities of a VEI=2 and of a VEI=4 scenarios, their relative weights may depend on the time window considered (10 years, 100 years, etc). Regarding percentiles, it is common to use the median (50-th percentile) plus other values like 85-th (15% probability of exceedance) or 98-th (2% probability of exceedance), which correspond to values above $\mu + 1\sigma$ and $\mu + 2\sigma$ being μ and σ the mean and standard deviation of a Gaussian distribution (even if the volcano PDF is not).

3. Arrival time maps (Ψ_{AR}), giving the averaged time needed to exceed a concentration threshold for each atmospheric layer, and computed as:

$$\{\forall \psi \subset \psi_{AR}(\boldsymbol{x}) > 0\}_{\boldsymbol{x} \in (x,y)} : \quad \Psi_{AR}(\boldsymbol{x}) = \sum_{i=1}^{n} w_i^N \left(\frac{1}{m_i} \sum_{j=1}^{m_i} [\psi_{AR}]_j^i(\boldsymbol{x}) \right) \tag{8}$$

where $[\psi_{AR}]_j^i$ are the m_i single-scenario footprint arrival time maps in the i-th Φ-bin. Note that, for each point $\boldsymbol{x} \in (x,y)$, the footprint averaging is done only for the subset of footprints for which $[\psi_{AR}]_j^i > 0$.

4. Similarly, ground accumulation time maps (Ψ_{AC}) are computed in the same way but using $\psi_{AC}(\boldsymbol{x})$ instead of $\psi_{AR}(\boldsymbol{x})$ in (8), *i.e.*,

$$\{\forall \psi \subset \psi_{AC}(\boldsymbol{x}) > 0\}_{\boldsymbol{x} \in (x,y)} : \quad \Psi_{AC}(\boldsymbol{x}) = \sum_{i=1}^{n} w_i^N \left(\frac{1}{m_i} \sum_{j=1}^{m_i} [\psi_{AC}]_j^i(\boldsymbol{x}) \right) \tag{9}$$

5. Persistence maps (Ψ_{PE}), giving the averaged time during which a concentration threshold has been exceeded, and computed as:

$$\{\forall \psi \subset \psi_{AR}(\boldsymbol{x}) > 0\}_{\boldsymbol{x} \in (x,y)} : \quad \Psi_{PE}(\boldsymbol{x}) = \sum_{i=1}^{n} w_i^N \left(\frac{1}{m_i} \sum_{j=1}^{m_i} [\psi_{PE}]_j^i(\boldsymbol{x}) \right) \tag{10}$$

where $[\psi_{PE}]_j^i$ are the m_i single-scenario persistence footprints laying in the i-th Φ-bin. As in (8), this is done only for subset of footprints for which $[\psi_{AR}]_j^i > 0$ at each point of the map.

As an illustrative example, single-volcano mappings are shown for the particular case of Vesuvius in Figures 7 and 8 for the European and regional domains respectively. Illustrative single hazard maps for Vesuvius were obtained assuming a PDF in terms of VEI and with only 3 bins of VEI 3, 4 and 5, with normalized probabilities of 0.65, 0.24 and 0.11 respectively (Marzocchi et al., 2004). However, the lightweight map production service in the ETHM allows to recalculate maps using other user-given PDFs at low extra computational cost.

3.5.2 Multi-volcano mapping

The composite hazard from multiple volcanoes can be derived from single-volcano mappings assuming that the volcanoes erupt independently. Consider a set of K volcanoes, each with their single-volcano mapping $(\Psi^1, \Psi^2, ..., \Psi^K)$ computed as

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

explained in the previous section. Then, the multi-volcano probability maps can be obtained as (Sandri et al., 2018):

$$\Psi_{PB}(\boldsymbol{x}) = 1 - \prod_{k=1}^{K} \left[1 - P_k(\Delta T)\ \Psi_{PB}^{k}(\boldsymbol{x})\right] \tag{11}$$

where each map $\Psi_{PB}^{k}(\boldsymbol{x})$ is given by (6) and $P_k(\Delta T)$ are the probabilities of having (at least one) eruption for the K volcanoes blended, all referred to the same time frame $\Delta T = [t, t+\Delta t]$. In other words, Ψ_{PB} represent the probability that at least one of the K volcanoes exceed the hazard intensity measure (threshold) at every grid point and within the time window considered. At each grid point, equation (11) allows to compute multi-volcano hazard curves for each intensity measure. From these, hazard (percentile) maps $\Psi_{PR}(\boldsymbol{x})$ can be estimated by interpolation of exceedance probabilities. On the other hand, multi-volcano arrival (accumulation) maps, and multi-volcano persistence maps are easily computed as probability-weighted averages:

$$\{\forall \psi \subset \Psi_{AR}(\boldsymbol{x}) > 0\}_{\boldsymbol{x}\in(x,y)}: \quad \Psi_{AR}(\boldsymbol{x}) = \sum_{k=1}^{K} P_k^{N}(\Delta T)\ \Psi_{AR}^{k}(\boldsymbol{x}) \tag{12a}$$

$$\{\forall \psi \subset \Psi_{AC}(\boldsymbol{x}) > 0\}_{\boldsymbol{x}\in(x,y)}: \quad \Psi_{AC}(\boldsymbol{x}) = \sum_{k=1}^{K} P_k^{N}(\Delta T)\ \Psi_{AC}^{k}(\boldsymbol{x}) \tag{12b}$$

$$\{\forall \psi \subset \Psi_{AR}(\boldsymbol{x}) > 0\}_{\boldsymbol{x}\in(x,y)}: \quad \Psi_{PE}(\boldsymbol{x}) = \sum_{k=1}^{K} P_k^{N}(\Delta T)\ \Psi_{PE}^{k}(\boldsymbol{x}) \tag{12c}$$

where, for every volcano, Ψ_{AR}^{k}, Ψ_{AC}^{k} and Ψ_{PE}^{k} are given by (8), (9) and (10) respectively. Note that, in the above expressions, the superscript N in $P_k^{N}(\Delta T)$ explicitly indicates that time probabilities need to be normalized to reflect the fact that these types of maps are conditional.

4 Running the ETHM workflow

In order to enlarge the catalogue with footprints from new volcanoes, it is recommended to execute the ETHM workflow (off-line part) with the provided Python scripts and the Common Workflow Language (CWL) file that orchestrates and automates, as much as possible, the execution of different steps (see Table 7). The CWL (Crusoe et al., 2022) is an open standard to connect and run command line tools and their dependencies to create workflows portable across a variety of platforms. In our case, this includes a series of Python scripts for footprint sampling and preparation of the FALL3D input files, creation of a hierarchic folder tree structure for European-scale and regional-scale footprints, and pre-configured slurm files to submit arrays of batch jobs to high-performance computing (HPC) clusters. The CWL workflow does not integrate the ERA-5 data download due the complexity generated by the data size. The slurm batch files are tailored by default to the MN5-ACC queue system, but can be easily adapted to the specificity of other machines and cluster management and job scheduling systems. The different software components are distributed through an in-house GitLab instance (see Code and data availability) and can be used to consistently add new volcanoes in a future. The reader is referred to the online documentation for details on how to install, configure and execute the different components. Note that the execution of some workflow steps, in particular the

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

model execution on HPC premises and the uploading of monthly aggregated files to the SDL (red boxes in Fig. 3), cannot be automated and must be done manually.

5 Conclusions and next steps

We implemented a first version of the European Tephra Hazard Map (ETHM), an initiative from the EuroHPC ChEESE Center of Excellence to facilitate single- and multi-volcano hazard mappings consistent across scales and volcanoes. The novel underpinning concept is that of the catalogue of footprints, a massive and open dataset of scenarios hosted in a Simulation Data Lake at CINECA. This methodology allows to decouple the modelling part of the PVHA workflow, which normally requires dedicated HPC resources and skills, from the downstream PDF-driven mapping. End users like volcanologists, academia or hazard assessment stakeholders can focus their expertise on the volcanological aspects, making use of a lightweight post-process service to create or update different typologies of maps by blending subsets of footprints in the catalogue (Guerrero et al., 2026a). The first version of the ETHM has already populated the catalogue with more than 53,000 footprints from 12 volcanoes located in 4 different European regions. However, the ETHM initiative has an incremental and collaborative ambition, making the necessary software stack open in order to facilitate future engagement and contributions from the scientific community. In the short term, the plan is to export the methodology to other regions of the world, including Mexico and New Zealand.

Beyond the immediate purpose of the ETHM, the unprecedented synthetic dataset of simulations is also of interest in the current era of data-driven modelling. Neural networks are already being used to accelerate, emulate, and improve atmospheric transport modelling, replacing expensive numerical simulations with fast, differentiable, and often physics-informed AI models. For example, deep learning models like SwinTransformer and UNet are being used to predict the global distribution of inert trace gases, specifically CO2 (e.g. Benson et al., 2025). Other foundation models like Aurora (Bodnar et al., 2025) can be fine-tuned for diverse tasks, including air pollution monitoring. In volcanology, machine learning techniques are boosting many applications like cloud detection and estimation of critical parameters needed for dispersal simulations (e.g. Torrisi et al., 2024) or ensemble augmentation (Mingari et al., 2025). However, AI-driven models for tephra transport and deposition are almost untapped and the synthetic catalogue of footprints generated here offers, for the first time, a training dataset to start exploring this path.

Code and data availability. FALL3D is available under the version 3 of the GNU General Public License (GPL) at https://gitlab.com/fall3d-suite/fall3d/. The catalogue of footprints is available under the CC-BY-4.0 license at https://sdl.hpc.cineca.it (DOI numbers 10.82544/sdl-372, 10.82544/sdl-344, 10.82544/sdl-368 and 10.82544/sdl-365 for the Azores, Canary, Iceland and Italy volcanoes respectively). The CWL and Python scripts to run the ETHM workflow are available at https://gitlab.geo3bcn.csic.es/cheese2p/ETHM.

Author contributions. Conceptualization, AF, LS, SB, AG, PT, AC; Data curation, HP, EH; Funding Acquisition, AF, AC; Methodology, AF, LS, SB, AG, PT, AC; Project administration, AF; Software, AF, AG, HP, LM; Resources and execution, AF, LS, SB, SA, BM, AG, PT; Supervision, AF, AC, LS; Writing—original draft, AF; Writing—review and editing, AF, LS, SB, AC, PT. All authors have read and approved the final version of the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) ChEESE-2P Center of Excellence (CoE) under Grant Agreement No 101093038, and the Spanish project PCI2022-134973-2 funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR. The SDL is hosted at the CINECA supercomputing system and has been developed and implemented by the Geo-INQUIRE project under Grant Agreement No 101058518. ERA-5 data was retrieved from the Copernicus Climate Change Service, Climate Data Store. Computational resources to run the ETHM were provided by the MareNostrum-5 supercomputer (ACC partition) through the EuroHPC Regular Access project EHPC-REG-2024R02-128. MareNostrum-5 is hosted at the Barcelona Supercomputing Center (BSC). Figures have been made using QGIS version 3.44.7-Solothurn. We thank Olaya Dorado for guidance on ranges of ESPs used for Canaries volcanoes.

References

Aguiar, S. C.: Contribuição para o estudo da estratigrafia dos vulcões Fogo e Furnas (S. Miguel, Açores) : implicações para a sua história eruptiva recente, Ph.D. thesis, University of Azores, http://hdl.handle.net/10400.3/5079, 2019.

Aubry, T. J., Engwell, S. L., Bonadonna, C., Mastin, L. G., Carazzo, G., Van Eaton, A. R., Jessop, D. E., Grainger, R. G., Scollo, S., Taylor, I. A., Jellinek, A. M., Schmidt, A., Biass, S., and Gouhier, M.: New Insights Into the Relationship Between Mass Eruption Rate and Volcanic Column Height Based On the IVESPA Data Set, Geophys. Res. Lett., 50, e2022GL102 633, https://doi.org/10.1029/2022GL102633, 2023.

Aydar, E., Cubukcu, H., Bal, C., Cluzel, N., Çağdas, ., Aladağ, H., Ersoy, O., and Laporte, D.: Volcanic jets to commercial jets: synopsis and diagnosis, Bull. Volcanol., 86, 1–15, https://doi.org/10.1007/s00445-024-01759-z, 2024.

Barberi, F., Macedonio, G., Pareschi, M. T., and Santacroce, R.: Mapping the tephra fallout risk: an example from Vesuvius, Italy, Nature, 344, 142–144, https://doi.org/10.1038/344142a0, 1990.

Barsotti, S., Scollo, S., Macedonio, G., Felpeto, A., Peltier, A., Vougioukalakis, G., de Zeeuw van Dalfsen, E., Ottemöller, L., Pimentel, A., Komorowski, J.-C., Loughlin, S., Carmo, R., Coltelli, M., Corbeau, J., Vye-Brown, C., Di Vito, M., de Chabalier, J.-B., Ferreira, T., Fontaine, F. R., Lemarchand, A., Marques, R., Medeiros, J., Moretti, R., Pfeffer, M. A., Saurel, J.-M., Vlastelic, I., Vogfjörd, K., Engwell, S., and Salerno, G.: The European Volcano Observatories and their use of the aviation colour code system, Bull. Volcanol., 86, 23, https://doi.org/10.1007/s00445-024-01712-0, 2024.

Bayarri, M., Berger, J., and Liu, F.: Modularization in Bayesian analysis, with emphasis on analysis of computer models, Bayesian Analysis, 4, https://doi.org/10.1214/09-BA404, 2009.

Bear-Crozier, A. N., Miller, V., Newey, V., Horspool, N., and Weber, R.: Probabilistic Volcanic Ash Hazard Analysis (PVAHA) I: development of the VAPAH tool for emulating multi-scale volcanic ash fall analysis, J. Appl. Volc., p. 1:3, https://doi.org/10.1186/s13617-016-0043-4, 2016.

Benson, V., Bastos, A., Reimers, C., Winkler, A. J., Yang, F., and Reichstein, M.: Atmospheric Transport Modeling of CO2 With Neural Networks, Journal of Advances in Modeling Earth Systems, 17, e2024MS004 655, https://doi.org/10.1029/2024MS004655, 2025.

Biass, S., Scaini, C., Bonadonna, C., Folch, A., Smith, K., and Höskuldsson, A.: A multi-scale risk assessment for tephra fallout and airborne concentration from multiple Icelandic volcanoes; Part 1: Hazard assessment, Nat. Hazards Earth Syst. Sci., pp. 2265–2287, https://doi.org/10.5194/nhess-14-2265-2014, 2014.

Bodnar, C., Bruinsma, W. P., Lucic, A., Stanley, M., Allen, A., Brandstetter, J., Garvan, P., Riechert, M., Weyn, J. A., Dong, H., Gupta, J. K., Thambiratnam, K., Archibald, A. T., Wu, C.-C., Heider, E., Welling, M., Turner, R. E., and Perdikaris, P.: A foundation model for the Earth system, Nature, 641, 1180–1187, https://doi.org/10.1038/s41586-025-09005-y, 2025.

Bonadonna, C., Connor, C. B., Houghton, B. F., Connor, L., Byrne, M., Laing, A., and Hincks, T. K.: Probabilistic modeling of tephra dispersal: Hazard assessment of a multiphase rhyolitic eruption at Tarawera, New Zealand, J. Geophys. Res.-Sol. Ea., 110, https://doi.org/10.1029/2003JB002896, 2005.

Bonadonna, C., Pistolesi, M., Dominguez, L., Freret-Lorgeril, V., Rossi, E., Fries, A., Biass, S., Voloschina, M., Lemus, J., Romero, J. E., Zanon, V., Pastore, C., Reyes Hardy, M.-P., Di Maio, L. S., Gabellini, P., Martin-Lorenzo, A., Rodriguez, F., and Perez, N. M.: Tephra sedimentation and grainsize associated with pulsatory activity: the 2021 Tajogaite eruption of Cumbre Vieja (La Palma, Canary Islands, Spain), Frontiers in Earth Science, 11, https://doi.org/10.3389/feart.2023.1166073, 2023.

Bonasia, R., Scaini, C., Capra, L., Nathenson, M., Siebe, C., Arana-Salinas, L., and Folch, A.: Long-range hazard assessment of volcanic ash dispersal for a Plinian eruptive scenario at Popocatépetl volcano (Mexico): implications for civil aviation safety, Bull. Volcanol., p. 76:789, https://doi.org/10.1007/s00445-013-0789-z, 2014.

Burgos, V., Jenkins, S. F., Bebbington, M., Newhall, C., and Taisne, B.: A new perspective on eruption data completeness: insights from the First Recorded EruptionS in the Holocene (FRESH) database, J. Volcanol. Geoth. Res., 431, 107 648, https://doi.org/j.jvolgeores.2022.107648, 2022.

Cioni, R., Longo, A., Macedonio, G., Santacroce, R., Sbrana, A., Sulpizio, R., and Andronico, D.: Assessing pyroclastic fall hazard through field data and numerical simulations: Example from Vesuvius, J. Geophys. Res.-Sol. Ea., p. 108:B2, https://doi.org/10.1029/2001JB000642, 2003.

Cole, P., Queiroz, G., Wallenstein, N., Gaspar, J., Duncan, A., and Guest, J.: An historic subplinian/phreatomagmatic eruption: the 1630 AD eruption of Furnas volcano, Sa˜o Miguel, Azores, J. Volcanol. Geoth. Res., 69, 117–135, https://doi.org/10.1016/0377-0273(95)00033-X, 1995.

Cole, P., Guest, J., Queiroz, G., Wallenstein, N., Pacheco, J., Gaspar, J., Ferreira, T., and Duncan, A.: Styles of volcanism and volcanic hazards on Furnas volcano, São Miguel, Azores, J. Volcanol. Geoth. Res., 92, 39–53, https://doi.org/10.1016/S0377-0273(99)00066-9, 1999.

Cole, P., Pacheco, J., Gunasekera, R., Queiroz, G., Gonçalves, P., and Gaspar, J.: Contrasting styles of explosive eruption at Sete Cidades, São Miguel, Azores, in the last 5000 years: Hazard implications from modelling, Journal of Volcanology and Geothermal Research, 178, 574–591, https://doi.org/10.1016/j.jvolgeores.2008.01.008, 2008.

Costa, A., Dell'Erba, F., Di Vito, M. A., Isaia, R., Macedonio, G., Orsi, G., and Pfeiffer, T.: Tephra fallout hazard assessment at the Campi Flegrei caldera (Italy), Bull. Volcanol., 71, 259–273, https://doi.org/10.1007/s00445-008-0220-3, 2009.

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

Costa, A., Folch, A., Suzuki, Y. J., Mingari, L., and Macedonio, G.: Eruption Source Parameters in Volcanic Plume Modeling: Advances, Challenges, and Future Directions, Reviews of Geophysics, 64, e2025RG000 897, https://doi.org/10.1029/2025RG000897, 2026.

Crusoe, M. R., Abeln, S., Iosup, A., Amstutz, P., Chilton, J., Tijanić, N., Ménager, H., Soiland-Reyes, S., Gavrilović, B., Goble, C., and Community, T. C.: Methods included: standardizing computational reuse and portability with the Common Workflow Language, Commun. ACM, 65, 54–63, https://doi.org/10.1145/3486897, 2022.

Dorado, O., Wolff, J. A., Ramos, F. C., and Martí, J.: Ba, Sr, and Rb feldspar/melt partitioning in recent eruptions from Teide-Pico Viejo volcanic complex, Tenerife: New insights into pre-eruptive processes, Frontiers in Earth Science, 112023, https://doi.org/10.3389/feart.2023.1155724, 2023.

Folch, A. and Sulpizio, R.: Evaluating long-range volcanic ash hazard using supercomputing facilities: application to Somma-Vesuvius (Italy), and consequences for civil aviation over the Central Mediterranean Area, Bull. Volcanol., pp. 1039–1059, https://doi.org/10.1007/s00445-010-0386-3, 2010.

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

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Folch, A., Abril, C., Afanasiev, M., Amati, G., Bader, M., Badia, R. M., Bayraktar, H. B., Barsotti, S., Basili, R., Bernardi, F., et al.: The EU Center of Excellence for Exascale in Solid Earth (ChEESE): Implementation, results, and roadmap for the second phase, Future Gener. Comp. Syst., pp. 47–61, https://doi.org/j.future.2023.04.006, 2023.

Gasparini, P.: Research on Volcanic Hazards in Europe, Science, pp. 1759–1760, https://doi.org/10.1126/science.260.5115.1759, 1993.

Guerrero, A., Tierz, P., and Folch, A.: Volcano Energy MOdelling and Research Analysis (VEMORA): A tool for probabilistic density functions of the energy eruptions, Volcanica, p. submitted, https://doi.org/xxx, 2026a.

Guerrero, A., Tierz, P., Folch, A., and team, E.: Title goes here, Volcanica, p. submitted, https://doi.org/xxx, 2026b.

Guest, J., Gaspar, J., Cole, P., Queiroz, G., Duncan, A., Wallenstein, N., Ferreira, T., and Pacheco, J.-M.: Volcanic geology of Furnas Volcano, São Miguel, Azores, J. Volcanol. Geoth. Res., 92, 1–29, https://doi.org/10.1016/S0377-0273(99)00064-5, 1999.

Guest, J. E., Pacheco, J. M., Cole, P. D., Duncan, A. M., Wallenstein, N., Queiroz, G., Gaspar, J. L., and Ferreira, T.: Chapter 9 The volcanic history of Furnas Volcano, São Miguel, Azores, Geological Society, London, Memoirs, 44, 125–134, https://doi.org/10.1144/M44.9, 2015.

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., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present, https://doi.org/10.24381/cds.adbb2d47, 2023.

Hurst, T. and Smith, W.: Volcanic ashfall in New Zealand – probabilistic hazard modelling for multiple sources, New Zealand J. Geol. Geophys., pp. 1–14, https://doi.org/10.1080/0028830100363112, 2010.

Jeffery, A. J., Gertisser, R., O'Driscoll, B., Pacheco, J. M., Whitley, S., Pimentel, A., and Self, S.: Temporal evolution of a post-caldera, mildly peralkaline magmatic system: Furnas volcano, São Miguel, Azores, Contributions to Mineralogy and Petrology, 171, 42, https://doi.org/10.1007/s00410-016-1235-y, 2016.

Jeffery, A. J., Gertisser, R., Self, S., Pimentel, A., O'Driscoll, B., and Pacheco, J. M.: Petrogenesis of the Peralkaline Ignimbrites of Terceira, Azores, Journal of Petrology, 58, 2365–2402, https://doi.org/10.1093/petrology/egy012, 2017.

Jenkins, S., Magill, C., and Mcaneney, J.: Multi-stage volcanic events: A statistical investigation, J. Volcanol. Geoth. Res., pp. 275–288, https://doi.org/10.1016/j.jvolgeores.2006.12.005, 2007.

Jenkins, S., Magill, C., McAneney, J., and Blong, R.: Regional ash fall hazard I: a probabilistic assessment methodology, Bull. Volcanol., 74, 1699–1712, https://doi.org/10.1007/s00445-012-0627-8, 2012.

Jenkins, S. F., Magill, C. R., and Blong, R. J.: Evaluating relative tephra fall hazard and risk in the Asia-Pacific region, Geosphere, pp. 492–509, https://doi.org/10.1130/GES01549.1, 2018.

Laeger, K., Petrelli, M., Morgavi, D., Lustrino, M., Pimentel, A., Paredes-Mariño, J., Astbury, R. L., Kueppers, U., Porreca, M., and Perugini, D.: Pre-eruptive conditions and triggering mechanism of the 16 ka Santa Bárbara explosive eruption of Sete Cidades Volcano (São Miguel, Azores), Contributions to Mineralogy and Petrology, 174, 11, https://doi.org/10.1007/s00410-019-1545-y, 2019.

Larsen, G. and Guðmundsson, M.: Catalogue of Icelandic Volcanoes, https://icelandicvolcanoes.is/?volcano=KAT, iMO, UI and CPD-NCIP, 2019.

Larsen, G. and Thordarson, T.: Catalogue of Icelandic Volcanoes, https://icelandicvolcanoes.is/?volcano=HEK, iMO, UI and CPD-NCIP, 2019.

Larsen, G., Sverrisdottir, G., Hjartarson, Á., Einarsson, P., and Jóhannesson, H.: HEKLA, pp. 189–209, Háskólaútgáfan, Ísland, 2013.

Longpré, M.-A. and Felpeto, A.: Historical volcanism in the Canary Islands; part 1: A review of precursory and eruptive activity, eruption parameter estimates, and implications for hazard assessment, J. Volcanol. Geoth. Res., 419, 107 363, https://doi.org/10.1016/j.jvolgeores.2021.107363, 2021.

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

Martí, J., Aspinall, W., Sobradelo, R., Felpeto, A., Geyer, A., Ortiz, R., Baxter, P., Cole, P., Pacheco, J., Blanco, M., and Lopez, C.: A long-term volcanic hazard event tree for Teide-Pico Viejo stratovolcanoes (Tenerife, Canary Islands), J. Volcanol. Geoth. Res., 178, 543–552, https://doi.org/10.1016/j.jvolgeores.2008.09.023, 2008.

Martí, J., Sobradelo, R., Felpeto, A., and García, O.: Eruptive scenarios of phonolitic volcanism at Teide–Pico Viejo volcanic complex (Tenerife, Canary Islands), Bull. Volcanol., 74, 767–782, https://doi.org/10.1007/s00445-011-0569-6, 2012.

Martínez Montesinos, B., Titos, M., Sandri, L., Rudyy, O., Cheptsov, A., Macedonio, G., Folch, A., Barsotti, S., Selva, J., and Costa, A.: On the feasibility and usefulness of high-performance computing in probabilistic volcanic hazard assessment: An application to tephra hazard from Campi Flegrei, Front. Earth Sci., pp. 10–2022, https://doi.org/10.3389/feart.2022.941789, 2022.

Marzocchi, W. and Bebbington, M. S.: Probabilistic eruption forecasting at short and long time scales, Bull. Volcanol., pp. 1777–1805, https://doi.org/10.1007/s00445-012-0633-x, 2012.

Marzocchi, W., Sandri, L., Gasparini, P., Newhall, C., and Boschi, E.: Quantifying probabilities of volcanic events: The example of volcanic hazard at Mount Vesuvius, J. Geophys. Res.-Sol. Ea., 109, https://doi.org/10.1029/2004JB003155, 2004.

Marzocchi, W., Sandri, L., and Selva, J.: BET_VH: a probabilistic tool for long-term volcanic hazard assessment, Bull. Volcanol., 72, 705–716, https://doi.org/10.1007/s00445-010-0357-8, 2010.

Mason, B. G., Pyle, D. M., and Oppenheimer, C.: The size and frequency of the largest explosive eruptions on Earth, Bull. Volcanol., pp. 735–748, https://doi.org/10.1007/s00445-004-0355-9, 2004.

Massaro, S., Stocchi, M., Martínez Montesinos, B., Sandri, L., Selva, J., Sulpizio, R., Giaccio, B., Moscatelli, M., Peronace, E., Nocentini, M., Isaia, R., Titos Luzón, M., Dellino, P., Naso, G., and Costa, A.: Assessing long-term tephra fallout hazard in southern Italy from Neapolitan volcanoes, Nat. Hazards Earth Syst. Sci., pp. 2289–2311, https://doi.org/10.5194/nhess-23-2289-2023, 2023.

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

McBirney, A. R., Serva, L., Guerra, M., and Connor, C. B.: Volcanic and seismic hazards at a proposed nuclear power site in central Java, J. Volcanol. Geoth. Res., 126, 11–30, https://doi.org/10.1016/S0377-0273(03)00114-8, 2003.

Mingari, L., Folch, A., Pascual, H., and Titos, M.: Expanding ensemble forecasts with generative AI: application to volcanic clouds, EGUsphere, https://doi.org/10.5194/egusphere-2025-6097, 2025.

Newhall, C. G. and Self, S.: The Volcanic Explosivity Index (VEI): An Estimate of Explosive Magnitude for Historical Volcanism, pp. 143–150, American Geophysical Union (AGU), ISBN 9781118665244, https://doi.org/10.1029/HG002p0143, 1986.

Orsi, G., De Vita, S., and di Vito, M.: The restless, resurgent Campi Flegrei nested caldera (Italy): constraints on its evolution and configuration, J. Volcanol. Geoth. Res., 74, 179–214, https://doi.org/10.1016/S0377-0273(96)00063-7, 1996.

Orsi, G., Di Vito, M. A., Selva, J., and Marzocchi, W.: Long-term forecast of eruption style and size at Campi Flegrei caldera (Italy), Earth Planet. Sc. Lett., 287, 265–276, https://doi.org/10.1016/j.epsl.2009.08.013, 2009.

Pardini, F., de'Michieli Vitturi, M., Andronico, D., Esposti Ongaro, T., Cristaldi, A., and Neri, A.: Real-time probabilistic assessment of volcanic hazard for tephra dispersal and fallout at Mt. Etna: the 2021 lava fountain episodes, Bull. Volcanol., p. 1:6, https://doi.org/10.1007/s00445-022-01614-z, 2023.

Pfeiffer, T., Costa, A., and Macedonio, G.: A model for the numerical simulation of tephra fall deposits, J. Volcanol. Geoth. Res., 140, 273–294, https://doi.org/10.1016/j.jvolgeores.2004.09.001, 2005.

Prata, F. and Rose, B.: Chapter 52 - Volcanic Ash Hazards to Aviation, in: The Encyclopedia of Volcanoes (Second Edition), edited by Sigurdsson, H., pp. 911–934, Academic Press, Amsterdam, second edition edn., ISBN 978-0-12-385938-9, https://doi.org/10.1016/B978-0-12-385938-9.00052-3, 2015.

Primerano, P., Giordano, G., Costa, A., de Vita, S., and Di Vito, M. A.: Reconstructing fallout features and dispersal of Cretaio Tephra (Ischia Island, Italy) through field data analysis and numerical modelling: Implications for hazard assessment, J. Volcanol. Geoth. Res., 415, 107 248, https://doi.org/10.1016/j.jvolgeores.2021.107248, 2021.

Pyle, D. M.: Chapter 13 - Sizes of Volcanic Eruptions, in: The Encyclopedia of Volcanoes (Second Edition), edited by Sigurdsson, H., pp. 257–264, Academic Press, Amsterdam, second edition edn., ISBN 978-0-12-385938-9, https://doi.org/10.1016/B978-0-12-385938-9.00013-4, 2015.

Queiroz, G., Pacheco, J., Gaspar, J., Aspinall, W., Guest, J., and Ferreira, T.: The last 5000 years of activity at Sete Cidades volcano (São Miguel Island, Azores): Implications for hazard assessment, J. Volcanol. Geoth. Res., 178, 562–573, https://doi.org/10.1016/j.jvolgeores.2008.03.001, 2008.

Queiroz, G., Gaspar, J. L., Guest, J. E., Gomes, A., and Almeida, M. H.: Chapter 7 Eruptive history and evolution of Sete Cidades Volcano, São Miguel Island, Azores, Geological Society, London, Memoirs, 44, 87–104, https://doi.org/10.1144/M44.7, 2015.

Sandri, L., Costa, A., Selva, J., Tonini, R., Macedonio, G., Folch, A., and Sulpizio, R.: Beyond eruptive scenarios: assessing tephra fallout hazard from Neapolitan volcanoes, Sci. Rep., p. 1:24271, https://doi.org/10.1038/srep24271, 2016.

Sandri, L., Tierz, P., Costa, A., and Marzocchi, W.: Probabilistic Hazard From Pyroclastic Density Currents in the Neapolitan Area (Southern Italy), Journal of Geophysical Research: Solid Earth, 123, 3474–3500, https://doi.org/10.1002/2017JB014890, 2018.

Selva, J., Costa, A., Sandri, L., Macedonio, G., and Marzocchi, W.: Probabilistic short-term volcanic hazard in phases of unrest: A case study for tephra fallout, J. Geophys. Res.-Sol. Ea., 119, 8805–8826, https://doi.org/10.1002/2014JB011252, 2014.

Selva, J., Acocella, V., Bisson, M., Caliro, S., Costa, A., Della Seta, M., De Martino, P., de Vita, S., Federico, C., Giordano, G., Martino, S., and Cardaci, C.: Multiple natural hazards at volcanic islands: a review for the Ischia volcano (Italy), J. Appl. Volc., p. 5, https://doi.org/10.1186/s13617-019-0086-4, 2019.

Sigurgeirsson, M. and Einarsson, S.: Catalogue of Icelandic Volcanoes, https://icelandicvolcanoes.is/?volcano=REY, iMO, UI and CPD-NCIP, 2019.

Spiller, E. T., Bayarri, M. J., Berger, J. O., Calder, E. S., Patra, A. K., Pitman, E. B., and Wolpert, R. L.: Automating Emulator Construction for Geophysical Hazard Maps, SIAM/ASA Journal on Uncertainty Quantification, 2, 126–152, https://doi.org/10.1137/120899285, 2014.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Sulpizio, R., Folch, A., Costa, A., Scaini, C., and Dellino, P.: 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., pp. 2205–2218, https://doi.org/10.1007/s00445-012-0656-3, 2012.

Suzuki, T.: A theoretical model for dispersion of tephra, in: D. Shimozuru and I. Yokoyama (eds) Arc Volcanism: Physics and Tectonics, Terra Scientific Publishing, Tokyo, pp. 95–116, 1983.

Tang, B.: Orthogonal Array-Based Latin Hypercubes, Journal of the American Statistical Association, 88, 1392–1397, https://doi.org/10.1080/01621459.1993.10476423, 1993.

Tierz, P., Loughlin, S. C., and Calder, E. S.: VOLCANS: an objective, structured and reproducible method for identifying sets of analogue volcanoes, Bull. Volcanol., p. 76, https://doi.org/10.1007/s00445-019-1336-3, 2019.

Titos, M., Martínez Montesinos, B., Barsotti, S., Sandri, L., Folch, A., Mingari, L., Macedonio, G., and Costa, A.: Long-term hazard assessment of explosive eruptions at Jan Mayen (Norway) and implications for air traffic in the North Atlantic, Nat. Hazards Earth Syst. Sci., pp. 139–163, https://doi.org/10.5194/nhess-22-139-2022, 2022.

Torrisi, F., Corradino, C., Cariello, S., and Del Negro, C.: Enhancing detection of volcanic ash clouds from space with convolutional neural networks, J. Volcanol. Geoth. Res., 448, 108 046, https://doi.org/10.1016/j.jvolgeores.2024.108046, 2024.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al.: Fundamental algorithms for scientific computing in Python, Nat. Methods, 17, 261–272, 2020.

Vázquez, R., Bonasia, R., Folch, A., Arce, J. L., and Macías, J. L.: Tephra fallout hazard assessment at Tacaná volcano (Mexico), J. South Am. Earth. Sci., pp. 253–259, https://doi.org/10.1016/j.jsames.2019.02.013, 2019.

Walker, G. L. and Croasdale, R.: Two Plinian-type eruptions in the Azores, Journal of the Geological Society, 127, 17–55, https://doi.org/10.1144/gsjgs.127.1.0017, 1971.

Wallenstein, N., Delgada, P., Chester, D., and Duncan, A.: Methodological implications of volcanic hazard evaluation and risk assessment: Fogo Volcano, São Miguel, Azores, Zeitschrift fur Geomorphologie, Supplementband, 140, 129–149, 2005.

Wilson, G., Wilson, T., Deligne, N., and Cole, J.: Volcanic hazard impacts to critical infrastructure: A review, J. Volcanol. Geoth. Res., pp. 148–182, https://doi.org/10.1016/j.jvolgeores.2014.08.030, 2014.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

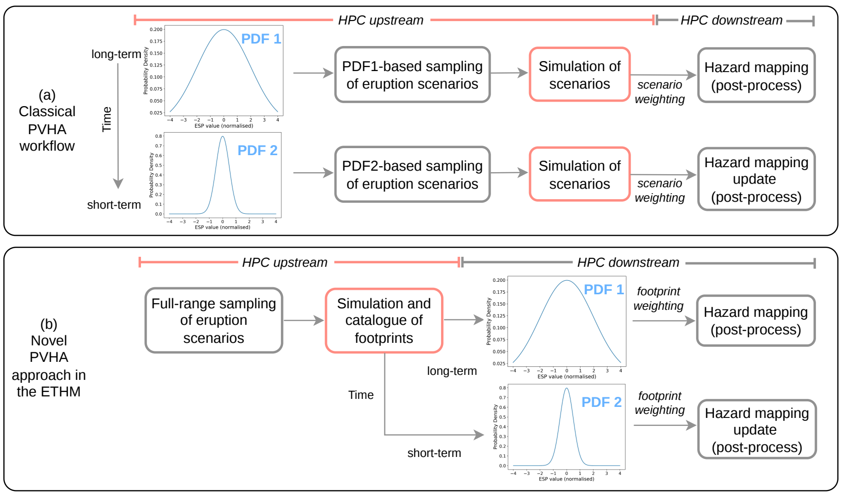

Figure 1. Conceptual comparison between: (a) the classical PVHA workflow and, (b) the novel workflow strategy adopted in the ETHM. The red boxes indicate the PVHA workflow steps that typically run of HPC premises and divide the HPC upstream from the HPC upstream components. In the novel catalogue-based approach, the time-dependent PDFs (illustrated by PDF 1 and PDF 2 in the sketch) are needed only by the upstream components; see text for details.

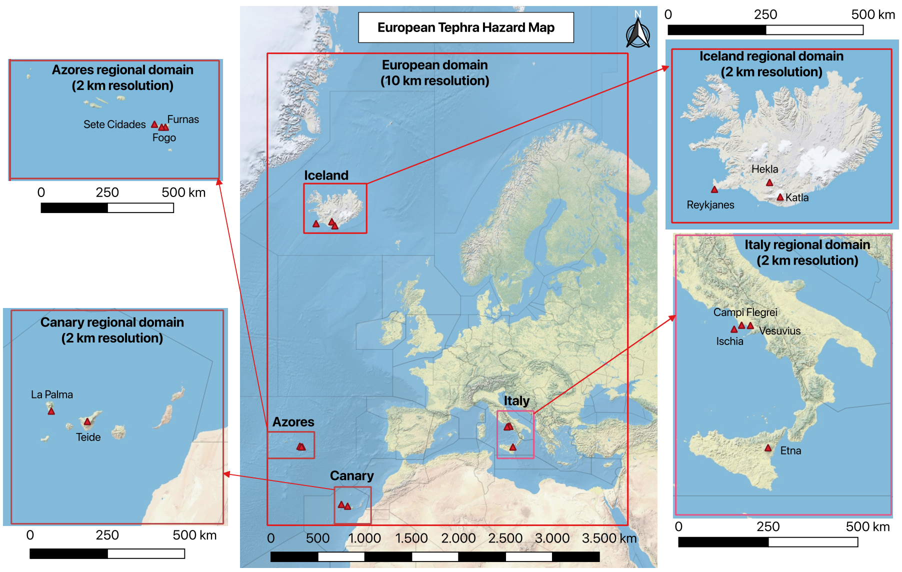

Figure 2. FALL3D computational domains for the ETHM. Footprints from all 12 volcanoes share a same continental-scale domain at 10-km grid resolution (central panel) whereas 4 regional nests are considered for deposit footprints from Iceland, Italy, Azores and Canary archipelagos. Volcano locations are indicated by the red triangles. For reference, the light gray lines in the central map show the International Civil Aviation Organization (ICAO) Flight Information Regions (FIRs). Map created with QGIS 3.44. Base map source: Esri Physical. FIR contours extracted from https://github.com/jaluebbe/FlightMapEuropeSimple (last access: 24 September 2026).

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

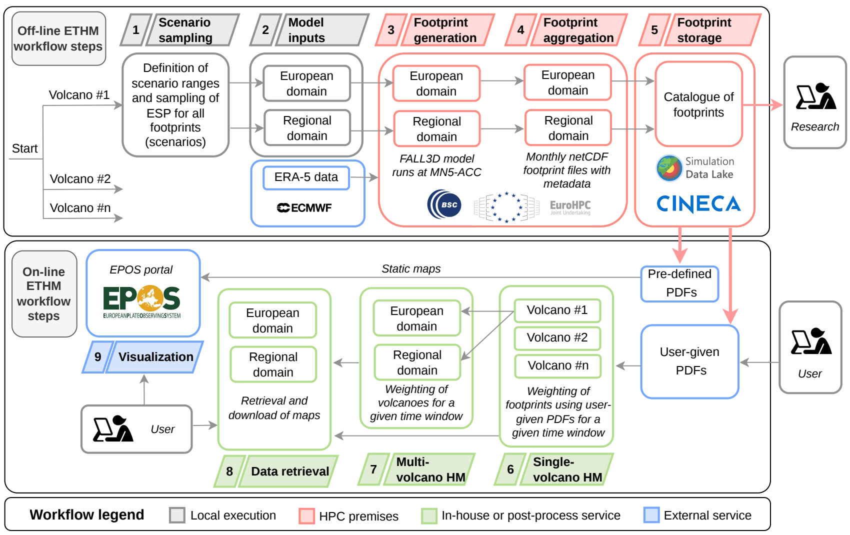

Figure 3. Workflow steps of the ETHM, divided into off-line (top, HPC upstream) and on-line (bottom, HPC downstream) components. The off-line steps 1 to 5 are the main subject of this paper, which addresses also the methodology for the on-line steps 6 to 9. The lightweight service for hazard mapping/update and related results constitute the subject of a second paper (Guerrero et al., 2026a).

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

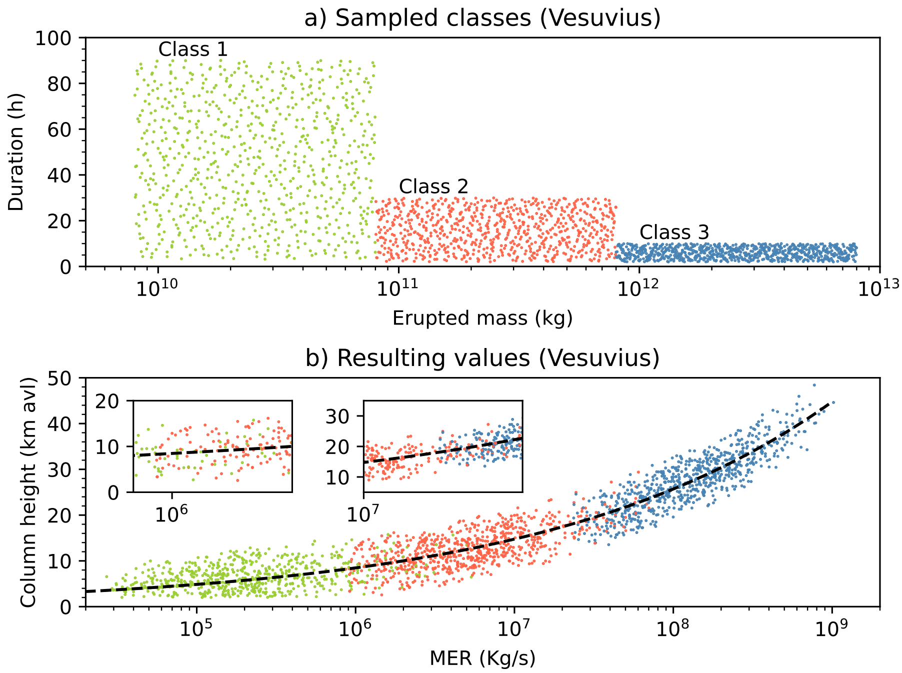

Figure 4. Scenario sampling results for Vesuvius volcano, where each dot in the plots represents one of the 2209 footprints. (a) Assumed ranges for erupted mass of tephra (M) and eruption duration (D) considering 3 eruption classes: Strombolian (class 1, green dots), sub-Plinian (class2, red dots) and Plinian (class 3, blue dots); see Table 2 for details. (b) Resulting Mass Eruption Rates (MER) and eruption column heights (H_c) above vent level. Note the 2 km floor value discussed in the text. The dashed line shows the "Mastin-like" power law for reference. The two inset zoom plots illustrate how the 3 eruption classes can mix when considered in terms of MER and height, even if they are disjoint in terms of mass and duration. This is a consequence of the non-linear relationships among different ESPs.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

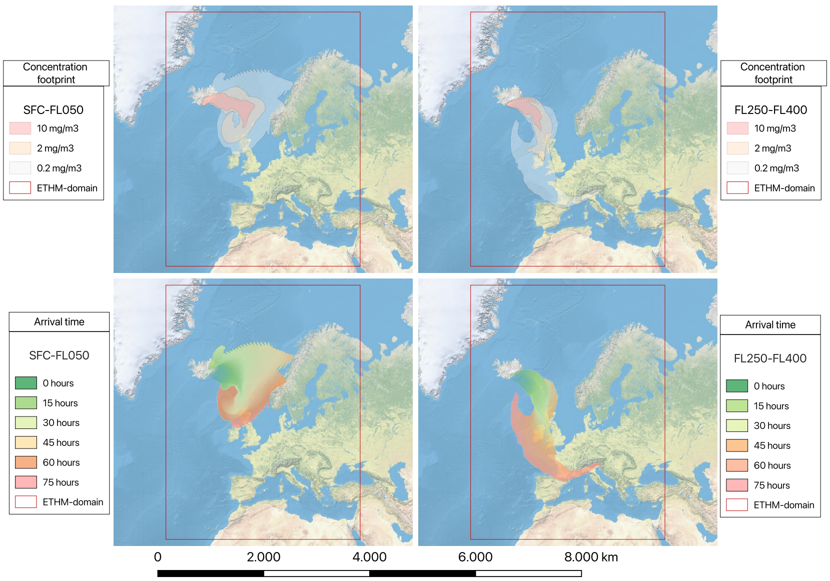

Figure 5. Example of a single footprint from Katla volcano over the European domain. Top panels: concentration footprints for three thresholds of 0.2 (blue), 2 (brownish) and 10 (reddish) $\mathrm{mg/m^3}$ at flight level layers of surface to FL050 (left) and FL250 to FL400 (right). Bottom panels: arrival times (in h) for the 0.2 $\mathrm{mg/m^3}$ concentration threshold at flight level layers of surface to FL050 (left) and FL250 to FL400 (right). The scenario of this particular Katla footprint corresponds to an eruption starting on 24 Jan 2006 at 05:00 UTC, eruption duration D =2 hour, column height H_c =22.6 km a.v.l., eruption magnitude M_g =4.2, eruption intensity I =10.4, time-averaged mass flow rate $\dot{M} = 2.8 \times 10^7$ kg/s and VEI=3. For reference, the light gray lines show the International Civil Aviation Organization (ICAO) Flight Information Regions (FIRs). Map created with QGIS 3.44. Base map source: Esri Physical. FIR contours extracted from https://github.com/jaluebbe/FlightMapEuropeSimple (last access: 24 September 2026).

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

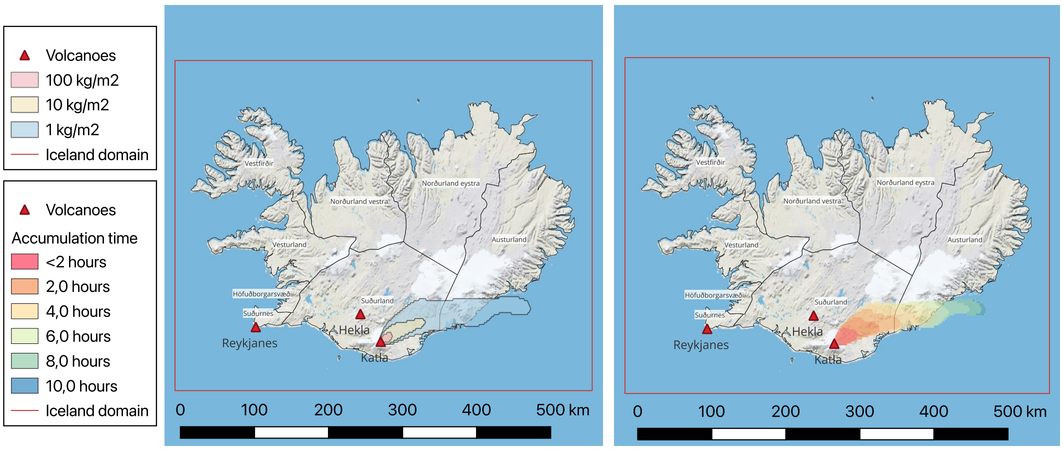

Figure 6. Same Katla footprint (scenario) than in Figure 5 but for the regional domain. Left: ground load footprint for three thresholds of 1 (blue), 10 (brownish) and 100 (reddish) $\mathrm{kg/m^2}$. Right: deposit accumulation times (in h) for the 1 $\mathrm{kg/m^2}$ load threshold. For reference, black lines show the Icelandic level 1 administrative levels (8 regions), with names highlighted in white. Map created with QGIS 3.44. Base map source: Esri Physical. Iceland administrative levels extracted from https://simplemaps.com/gis/country/is (last access: 24 September 2026).

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

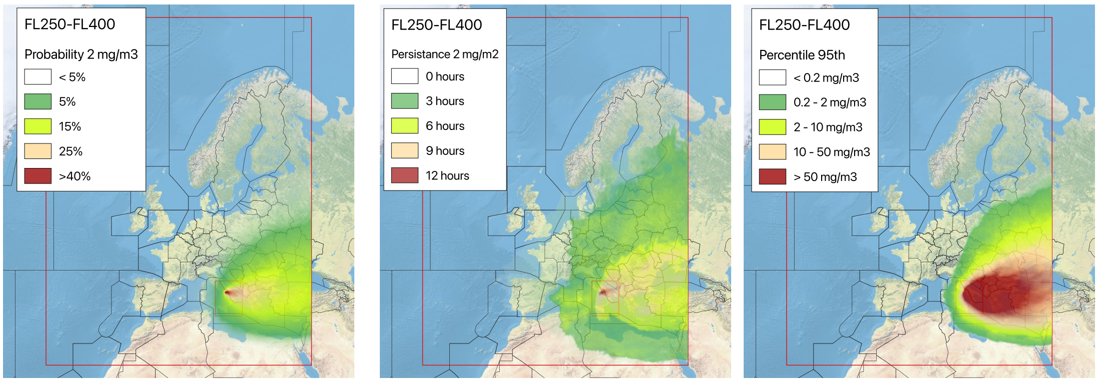

Figure 7. Example of mapping for Vesuvius on the European domain and flight level layer from FL205 to FL400. Left: probability of exceedance map for concentration threshold (hazard intensity measure) of 2 $\mathrm{mg/m^3}$. Center: persistence map for the same concentration threshold (in h). Right: hazard map showing concentration contours (in $\mathrm{mg/m^3}$) for the 95-th percentile (5% probability of exceedance). See text for details on the underlaying PDF used to draw the maps from the catalogue of footprints. Map created with QGIS 3.44. Base map source: Esri Physical. FIR contours extracted from https://github.com/jaluebbe/FlightMapEuropeSimple (last access: 24 September 2026).

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

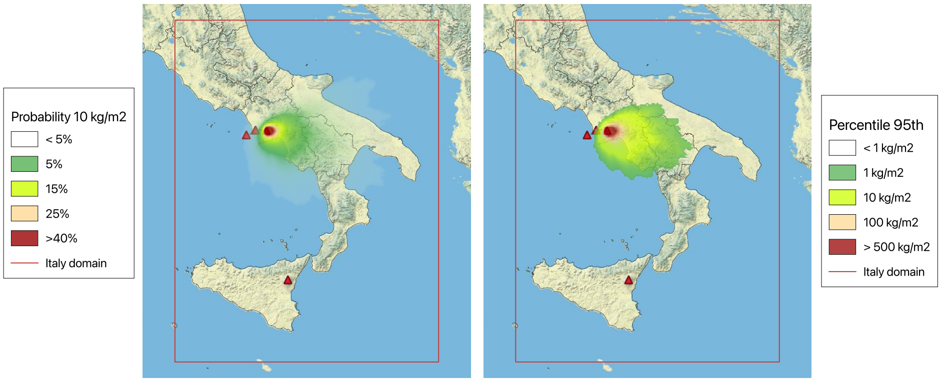

Figure 8. Example of ground deposit mapping for Vesuvius on the Italian regional domain (red box). Left: exceedance probability map for a deposit load of 10 $\mathrm{kg/m2}$. Right: hazard map showing isomass contours (in $\mathrm{kg/m2}$) for the 95-th percentile (5% probability of exceedance). Results using the same PDF than in Figure 7. Map created with QGIS 3.44. Base map source: Esri Physical. Italian administrative levels extracted from https://simplemaps.com/gis/country/it (last access: 24 September 2026).

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Table 1. List of the 4 regions and 12 volcanoes included in the first version of the ETHM. The volcano ID follows the Smithsonian Institution's Global Volcanism Program (GVP) that uses a hierarchical 6-digit number to uniquely identify any volcano in the world.

Region	Volcano (shortname)	Volcano ID (GVP)	longitude (^{o}E)	latitude (^{o}N)
Azores	Sete Cidades	382080	-25.78	37.87
	Furnas	382100	-25.31	37.77
	Fogo	382090	-25.47	37.76
Canary	Teide	383030	-16.64	28.27
	La Palma (†)	383010	-17.83	28.57
Iceland	Reykjanes (†)	371020	-22.67	63.82
	Hekla	372070	-19.66	63.98
	Katla	372030	-19.08	63.63
Italy	Campi Flegrei	211010	14.14	40.83
	Vesuvius	211020	14.42	40.82
	Ischia	211030	13.89	40.73
	Etna	211060	15.00	37.75

(†) For fissure-like systems, vent coordinates are given as reference only. For continental-scale simulations small vent position variations have no impact on the simulation results. For regional-scale nests, the FALL3D model allows perturbations of the vent position.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Table 2. Range of sampled ESPs (eruption mass M; eruption duration D; magma viscosity) and of derived ESPs (eruption intensity I; column height H_c; eruption volume; and VEI). Note that for volcanoes with more than one eruption class the reported values of derived ESPs span all the classes. Ranges for the timestamp T_o (ranging from 01/01/1995 to 31/12/2024), Suzuki parameter-A (from 2 to 5) and standard column height deviation ($\sigma_c = \pm 3.5$ km) are the same for all volcanoes. See text for details.

		Range of sampled ESPs				Range of derived ESPs		
Region	Volcano	Number of classes	Mass M (kg)	Duration D (h)	Viscosity (Pa s)	Col height (km avl)	Volume (km^3)	VEI (-)
Azores	Sete Cidades	1	$1.0\,10^{10}$ - $8.0\,10^{10}$	5 - 20	$5.0\,10^{2}$ - $3.2\,10^{4}$	2 - 33	0.005 - 2.5	1 - 5
		2	$8.0\,10^{10}$ - $5.0\,10^{12}$	5 - 100	-			
	Furnas	1	$1.0\,10^{10}$ - $8.0\,10^{10}$	5 - 20	$5.0\,10^{2}$ - $3.2\,10^{4}$	2 - 52	0.005 - 15	1 - 6
		2	$8.0\,10^{10}$ - $5.0\,10^{11}$	10 - 20	-			
		3	$5.0\,10^{11}$ - $3.0\,10^{13}$	5 - 20	-			
	Fogo	1	$1.0\,10^{10}$ - $8.0\,10^{10}$	5 - 20	$5.0\,10^{2}$ - $3.2\,10^{4}$	2 - 55	0.005 - 15	1 - 6
		2	$8.0\,10^{10}$ - $5.0\,10^{11}$	2 - 30	-			
		3	$5.0\,10^{11}$ - $3.0\,10^{13}$	3 - 20	-			
Canary	Teide	1	$1.0\,10^{9}$ - $5.0\,10^{10}$	2 - 96	$1.0\,10^{3}$ - $3.1\,10^{4}$	2 - 46	0.005 - 5	1 - 4
		2	$5.0\,10^{10}$ - $1.0\,10^{13}$	2 - 48	-			
	La Palma	1	$1.0\,10^{9}$ - $5.0\,10^{11}$	2 - 96	$1.0\,10^{2}$ - $1.0\,10^{4}$	2 - 24	0.005 - 0.25	1 - 3
Iceland	Reykjanes	1	$2.7\,10^{10}$ - $2.7\,10^{11}$	1 - 154	$1.0\,10^{2}$ - $1.0\,10^{6}$	2 - 22	0.01 - 0.013	2 - 3
	Hekla	1	$2.7\,10^{10}$ - $2.7\,10^{11}$	1 - 35	$1.0\,10^{3}$ - $1.0\,10^{6}$	2 - 43	0.01 - 1.2	2 - 4
		2	$2.7\,10^{11}$ - $1.3\,10^{12}$	2 - 32	-			
		3	$1.2\,10^{12}$ - $2.4\,10^{12}$	1 - 19	-			
	Katla	1	$2.8\,10^{10}$ - $2.8\,10^{11}$	1 - 35	$1.0\,10^{3}$ - $1.0\,10^{6}$	2 - 56	0.01 - 7.0	2 - 5
		2	$2.8\,10^{11}$ - $1.4\,10^{12}$	2 - 32	-			
		3	$1.4\,10^{12}$ - $1.4\,10^{13}$	1 - 98	-			
Italy	Flegrei	1	$2.5\,10^{9}$ - $2.5\,10^{10}$	2 - 30	$2.0\,10^{5}$ - $6.3\,10^{6}$	2 - 67	0.001 - 12.5	1 - 6
		2	$2.5\,10^{10}$ - $2.5\,10^{11}$	1 - 15	$7.9\,10^{4}$ - $4.0\,10^{7}$			
		3	$2.5\,10^{11}$ - $2.5\,10^{13}$	1 - 10	-			
	Vesuvius	1	$8.0\,10^{9}$ - $8.0\,10^{10}$	3 - 90	$2.5\,10^{3}$ - $3.1\,10^{4}$	2 - 48	0.004 - 4	1 - 5
		2	$8.0\,10^{10}$ - $8.0\,10^{11}$	2 - 30	$1.3\,10^{4}$ - $3.1\,10^{6}$			
		3	$8.0\,10^{11}$ - $8.0\,10^{12}$	2 - 10	-			
	Ischia	1	$1.0\,10^{11}$ - $3.0\,10^{11}$	3 - 168	$7.9\,10^{4}$ - $2.0\,10^{6}$	2 - 25	0.05 - 0.15	3 - 4
	Etna	1	$3.0\,10^{5}$ - $2.8\,10^{10}$	0.5 - 4	$1.0\,10^{2}$ - $1.0\,10^{3}$	2 - 20	$<$ 0.001 - 0.01	0 - 3

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Table 3. Configuration of the FALL3D model for the European domain and the regional nests.

		European Domain	Regional Nests
Target hazard		airborne ash	tephra fallout
	Number of footprints	2209	2209
	Number of volcanoes	12	12
FALL3D domains	Longitude range	(32ºW, 37ºE)	region dependent
	Latitude range	(25ºN, 75ºN)	region dependent
	Altitude (top)	footprint dependent	footprint dependent
Model resolution	Horizontal	10 km	2.5 km
	Vertical	250 to 500 m	250 to 500 m
Granulometry	Number of bins	4	13
	Particle size range	8 to 64 μm	128 μm to 8 mm
	ϕ range (†)	[4,7]	[-3,3]
	ϕ increment	1.0	0.5

(†) The ϕ scale is a standard logarithmic transformation used to describe particle grain size distributions of tephra deposits and is defined as $d = 2^{-\phi}$ where d is the particle diameter expressed in mm.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Table 4. Contents of a netCDF footprint file for the European and regional domains.

	Dimensions	Comments
Dimension variables		
Longitude	lon	domain dependent
Latitude	lat	domain dependent
Ground load intensity measure	n_l	3 pre-defined values of 1, 10 and 100 $\mathrm{kg/m^2}$
Concentration intensity measure	n_c	3 pre-defined values of 0.2, 2 and 10 $\mathrm{mg/m^3}$
Flight levels (†)	fl	3 pre-defined values of FL050, FL250 and FL400
European domain		
Maximum of concentration	(lon, lat, fl)	maximum value of concentration over time
Concentration footprint	(lon, lat, fl, n_c)	discrete value (0 or 1) for each concentration threshold
Arrival time	(lon, lat, fl, n_c)	time from eruption start to disruption (given in h)
Persistence	(lon, lat, fl, n_c)	disruption persistence (given in h)
Regional domains		
Maximum of load	(lon, lat)	maximum value of ground load over time
Load footprint	(lon, lat, n_l)	discrete value (0 or 1) for each load threshold
Accumulation time	(lon, lat, n_l)	time from eruption start to disruption (given in h)

(†) Altitude expressed in hundreds of feet of nominal standard atmosphere pressure. Concentration values are integrated in vertical, *i.e.* from surface to FL050, from FL050 to FL250 and from FL250 to FL400 respectively.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Table 5. Characteristics of the catalogue of footprints stored in the CINECA Simulation Data Lake (SDL).

Experiment name	Experiment ID	Number of volcanoes	Number of objects (†)	Number of footprints	Data volume	DOI
ETHM_Azores_Region	exp372	3	72	13,254	18.83 GB	10.82544/sdl-372
ETHM_Canary_Region	exp344	2	48	8,836	9.95 GB	10.82544/sdl-344
ETHM_Iceland_Region	exp368	3	72	13,254	19.79 GB	10.82544/sdl-368
ETHM_Italy_Region	exp365	4	96	17,672	16.33 GB	10.82544/sdl-365
TOTAL	-	12	288	53,016	64.90 GB	-

(†) Aggregated footprint files, includes European and regional domains in 12 monthly files. See text for details.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Table 6. Types of maps in the ETHM and how these are computed from footprints for the single- and multi-volcano cases.

Type of map	Map symbol	Domain $^{(\dagger)}$	Single-volcano	Multi-volcano	Comments $^{(\dagger\dagger)}$
Probability maps	$\Psi_{PM}(\mathbf{x})$	E/R	Eq. (6)	Eq. (11)	A/C $^{(*)}$
Hazard (percentile) maps	$\Psi_{PR}(\mathbf{x})$	E/R	Eq. (7)	see text	C $^{(**)}$
Arrival time maps	$\Psi_{AR}(\mathbf{x})$	E	Eq. (8)	Eq. (12a)	C$^{(*)}$
Persistence maps	$\Psi_{PE}(\mathbf{x})$	E	Eq. (10)	Eq. (12c)	C$^{(*)}$
Accumulation time maps	$\Psi_{AC}(\mathbf{x})$	R	Eq. (8)	Eq. (12b)	C$^{(*)}$

(†) European (E) or Regional (R). (††) Absolute (A) or Conditional (C) maps. (*) one map for each hazard intensity level (threshold). (**) One map for each percentile value.

https://doi.org/10.5194/egusphere-2026-4485
Preprint. Discussion started: 25 September 2026

Table 7. Software stack tools available to run the off-line part of the ETHM workflow. See Fig. 3 and text for details on the workflow steps.

Component	step	Comments
`ETHM_workflow.cwl`	-	Runs workflow steps 1 and 2 as a CWL pipeline
`olhs_sampling.cwl`	1	Generates eruption scenarios using OLHS and maps sampled values to physically meaningful ESPs
`footprint_generator.cwl`	2	Creates folder tree and replaces variables inside FALL3D templates
`meteo.py`	2	Downloads and concatenates ERA-5 data in monthly files
`FTP_merge`	4	Aggregates footprints and JSON metadata into monthly netCDF files