Computer Applications in Science and Engineering (*CASE*) Department
Barcelona Supercomputing Center (BSC, CNS)

Modeling and Simulation as a Service (MSaaS): a tool for the study of volcanic phenomena.

The final project for Informatics Engineering Master's degree

Universitat Politècnica de Catalunya

Facultat d'Informàtica de Barcelona

Author:

Diana Alejandra Guerrero López

Main Advisor:

Arnau Folch Duran [1]

Advisor:

Jordi Torres Viñals [2]

Barcelona, 28th of october 2021

[1] Barcelona Supercomputing Center (BSC-CNS)
[2] Computers Architecture Department. UPC, FIB.

Modeling and Simulation as a Service (MSaaS): a tool for the study of volcanic phenomena.

As the EaaS (Everything as a Service) paradigm is taking power in the computational environment, other fields of science have started to apply virtualization as a solution to the infrastructure problems. Additionally, the large capacity of new machines allows researchers to develop new models of reality simulation that control a huge amount of variables converted into data. The main goal of this project is to provide a Modeling and Simulation as a Service tool flexible, scalable, and with a trustworthy engine for the simulation of specific events or the construction of probabilistic hazard maps for volcanic ash fallout. The development process of a Modelling and Simulation as a Services (MSaaS) tool for volcanic ash fallout has been built upon microservices and Dockers architecture and over the HAZMAP simulation tool for Linux. The four main phases of a Simulation process (problem characterization, tool design, service development, and evaluation of the model) are detailed and documented in the present project.

Index

1. INTRODUCTION

More than 500 million people live in areas under the influence of potentially active volcanoes (Organización Panamericana de la Salud, 2002) . Volcanic ash fallout resulting from atmospheric dispersal of ash particles (Poulidis et al., 2018) threatens populations and infrastructures laying around volcanoes and up to hundreds of kilometers away. Centimeter to millimeter thickness deposits of ash on the ground encompass impacts on buildings, infrastructures, water and energy supplies, roads and traffic, agriculture, farming, and overall worsening of air quality (Wilson et al., 2012). In the proximal, effects can be even more harmful and provoke roof collapses in cases where the thickness of accumulated ash exceeds tens of centimeters (Hampton et al., 2015).

The quantification of this natural hazard and its associated risk is therefore an important aspect for society, territorial planning, and civil protection managers. However, the underlying analysis is a process surrounded by multiple uncertainties due to the impossibility to predict when the next eruption will happen and how it will be (Madankan et al., 2014). First of all, geologists study past events to infer the future but, because the deposits from past events blur with time, the number and quality of records available for a given volcano often impedes developing trustworthy predictive models (Loughlin et al., 2015). Even with information on the latest events at hand, our knowledge about the past volcano activity is necessarily limited. On the other hand, it is impossible to know how the atmospheric conditions, which ultimately dictate the dispersal pattern of ash, will be at the uncertain time of the next eruption (Denlinger et al., 2012) . For these reasons, probabilistic approaches and methodologies have been proposed over the last years in which thousands of physics-based simulations of single scenarios (type of eruption and atmospheric conditions) are combined according to their likelihood of occurrence (Bear-Crozier et al., 2016; Madankan et al., 2014). Traditionally, the atmospheric dispersion models had been used only to reproduce and understand past events (single and constrained scenarios) but, with the advance of computer capacity and related technology, a new application emerged: the physics-based hazard mapping (Denlinger et al., 2012).

Hazard assessment defines which zones are potentially exposed to ash fallout and which intensity is expected. Hazard maps actually delineate areas enclosing probabilities of given affectation by ash loads (Vázquez et al., 2021). The main source of information for hazard map generation is the geological record of past events (historical and beyond). However, the incompleteness of available datasets, derived from the life cycle of some volcanoes (thousands of years), the old recording tool’s limitations, and the loss of the deposits with time causes many uncertainties about past events. The probabilistic approach aims at integrating to the model the uncertainty generated by our ignorance about the key meteorological and eruption conditions of the next event (Leonard et al., 2014), namely, wind intensity and direction, column height and particle granulometry by the synthetic generation of hundreds or thousand events. In this way, a set of scenarios and their corresponding simulations is developed based on the analysis of past events to obtain a probabilistic characterization, resulting in a large amount of events to be simulated.

The general workflow for probabilistic hazard mapping is as follows (Figure 1):

- First, the historical data from past eruptions is analysed and processed to obtain a Probability Density Function (PDF) that relates the eruption magnitude of the volcano under consideration with its probability of occurrence.
- Second, wind and meteorological conditions over the last few years are retrieved from reanalysis datasets to get a climatological representation of the present (last decades) atmospheric conditions.
- Third, synthetic eruption scenarios are created based on the data obtained. This process is known as sampling and helps defining the initial conditions of the event. Every sampled scenario is assigned randomly to a wind (meteorological) condition.
- Forth, with the volcano and wind conditions the multiple (thousands) scenarios can be simulated using an atmospheric dispersal model.
- Finally, single results are combined and processed to generate hazard maps.

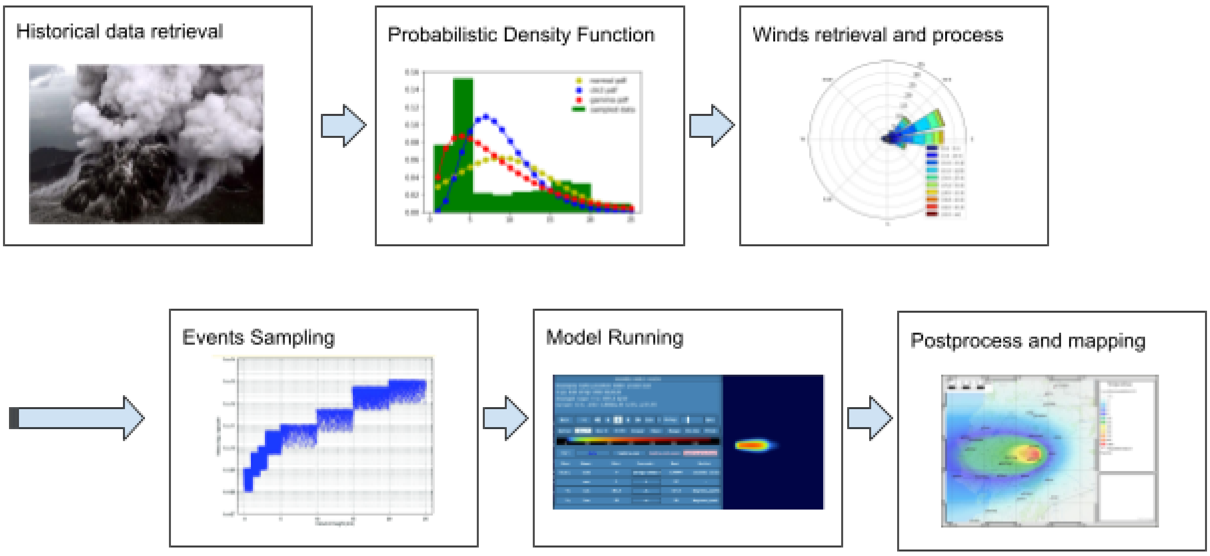


Figure 1. Hazard mapping workflow

In addition to hazard mapping, this work also considers single-scenario simulations to reproduce known past events or forecasting the impacts of an on-going eruption. Note that this "well-constrained" scenario can actually be seen as an end-member in which uncertainties do not exist. The single run workflow is obviously less complex than in the multiple scenario hazard mapping (Figure 2). In this case, the initial conditions are defined by the user and the wind retrieval is made on the basis of a specific (known) date. Then, the dispersal model run provides deterministic results of the simulated event.

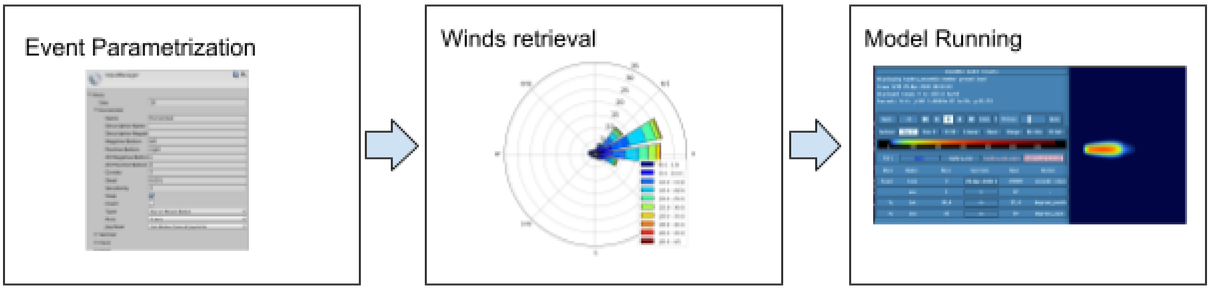


Figure 2. Single event (deterministic) simulation workflow

The amount of computational resources required by the probabilistic approach can be substantial and become a limiting factor for users (Leonard et al., 2014). Additionally, and very importantly, the workflow in Figure 1 must be coded and requires skills from computer science, volcanology, atmospheric dispersion modelling, parallel programming, and statistics. Furthermore, access to the required computing resources can also be an issue for which a Modelling and Simulation as a Service (MSaaS) paradigm supported by a Cloud Computing environment provides a solution.

The main objective of this project is to implement a tool for both single eruption simulation and probabilistic volcanic ashfall hazard mapping. The goal of this tool for Modelling and Simulation as a Service (MSaaS) for ash fallout is to aid users with a piece of basic knowledge in modelling and computer science. Additionally, the workflow for probabilistic hazard mapping has been also developed and coded as a Service applying different know-how lines such as statistics, maths, computer science, and volcanism.

This document contains the development of the MaaS according to the workflow presented before. First, the simulation and modelling services are briefly explained as an introduction to cloud services, including the sources of data and simulation tools. Then, the methodology is developed step by step, with the workflow defined to solve the problems of uncertainties through techniques suggested in a volcanic context. After that, the architecture and microservices design are introduced. Additionally, in order to illustrate the capabilities of the project, a specific study case is shown: the ashfall hazard map for Popocatepetl volcano in Mexico. Finally, the conclusions of the project are presented based on the expert experience and the case of study developed.

1.1. OBJECTIVES

1.1.1. General Objective

To develop and implement a computational tool of SMaaS focused on the volcanic ash fallout analysis

1.1.2. Specific Objectives

1. To model the physics of the problem and establish the main requirements of the tool.
2. To design the underlying workflow, modules, microservices, and architecture of the tool and its deployment in the Cloud.
3. To implement a user interface for the execution of the service.
4. To evaluate the resulting SMaaS over concrete cases (deterministic single scenario and probabilistic approach accounting for uncertainties).

2. PROJECT BACKGROUND

Virtualization and Cloud Computing have attracted many researchers in recent years (Cayirci, 2013). The convenience that Cloud Computing offers in terms of self-demand service, access, flexibility, elasticity, resource pooling, and measured services has increased the need for services based on simulation and modelling environments. Also, the requirements of infrastructure with high capacity make the Modeling and Simulation as a Service (MSaaS) a proper solution for risk management researchers.

According to Caycirci (2013), the main objective of the MSaaS paradigm is to supply the environment required to parameterize and run a simulation on self-demand. In the last years, a joint community effort to develop the state-of-the-art and the standardization of this new approach has yielded to the reference architecture of MSaaS. The hierarchy of M&S services (Figure 3), presented by Siegfried and Van Den Berg (2019), shows three subcategories of MSaaS explained in the next subsections.

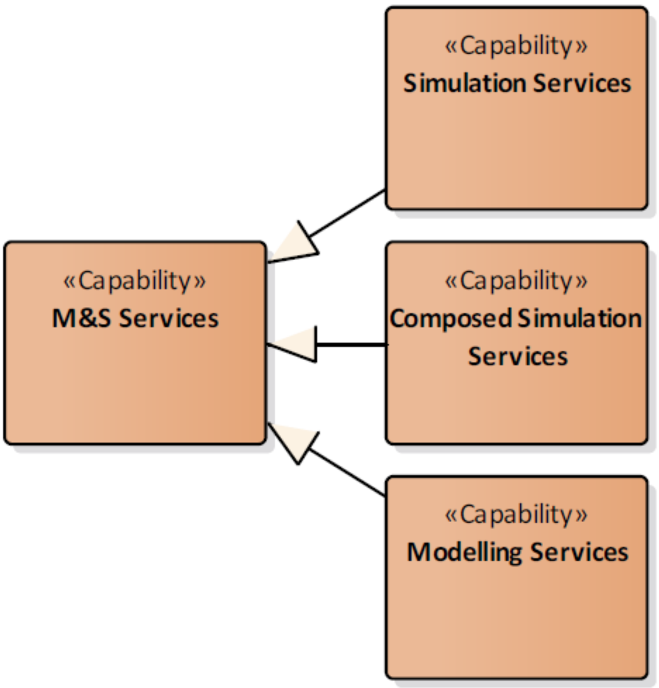


Figure 3. M&S Services Hierarchy. (Van Den Ber, 2019)

2.1 Simulation Services

The simulation Services are the functionalities that allow to define, run, and process a code that emulates a part of the real world. In contrast, the modelling services are the tools required to construct a Simulation Service. The aim of the MSaaS in this work is to provide the environment to simulate single ashfall scenarios and the generation of hazard maps through a probabilistic approach simulation.

The different simulation services interact between them providing information in every step of the workflow. Figure 4 describes the workflow from the simulation services perspective. The vertical boxes are the services explained forward and the small icons represent the microservices embedded. The microservices are explained in detail in the chapter Microservices Design.

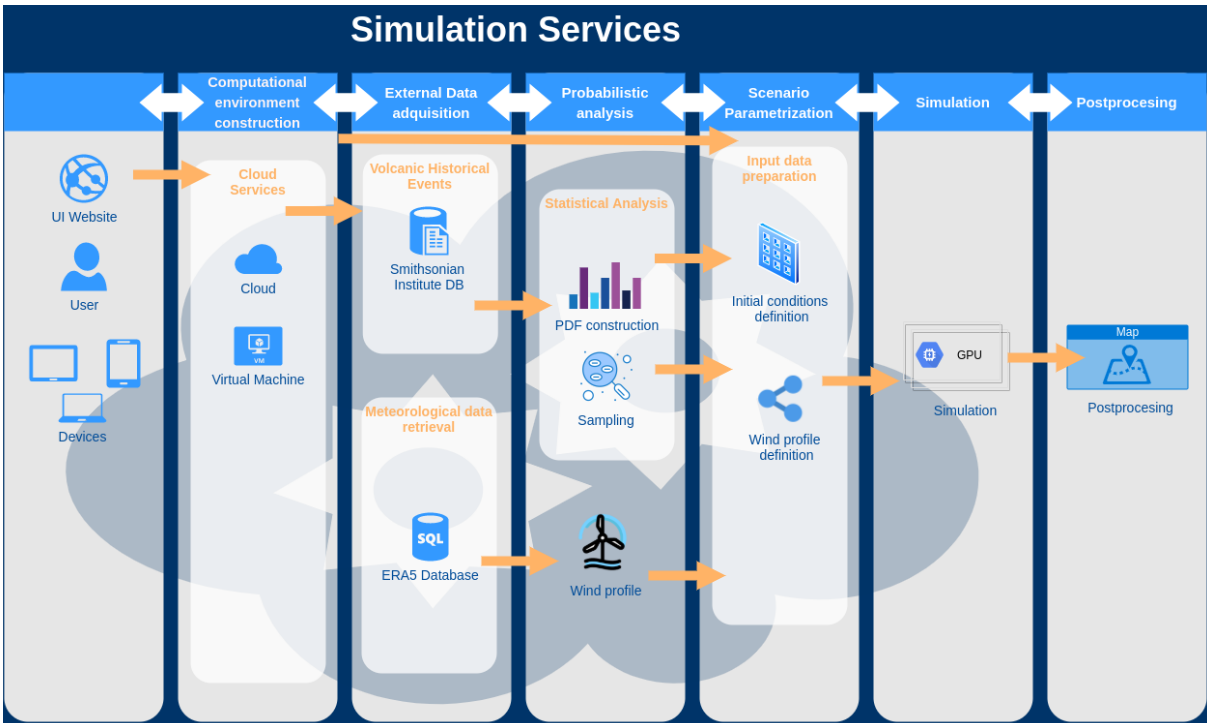


Figure 4. Simulation services

These services are supported by a group of functionalities listed below.

- **Computational environment construction**. This service is responsible for creating all the bases necessary to run the model in terms of hardware (capabilities) and software (simulation code, libraries) required for the service.

- **External data acquisition**. Service in charge of downloading and processing meteorological data from forecasts and reanalysis**. The dispersal model needs a wind vertical profile, which must be downloaded and processed in order to feed the simulation tool. In the case of a single event, running only during a given (known) time slab, specific times are downloaded. Otherwise, 6-hourly wind profiles at the volcano location are processed for the last 10 years. Additionally, the process retrieves the eruption parameters of the historical events of the volcano. For a single event, eruption data is furnished manually by the user through a form. In contrast, the probabilistic approach considers catalog data with all the events recorded by the Smithsonian Institute.
- **Probabilistic analysis**.** Defines the relation magnitude-temporal of the volcano. This process is composed of the following successive steps:
  - *Evaluation of data quality:* that means that only those events in the catalog with a magnitude associated will be taken into account.
  - *Data adjust:* where the eruption magnitude of each event is translated in terms of ash-plume height by a simple sampling technique.
  - *Magnitude probability density function*: calculates the probability that the next event has a certain column height value.
  - *Events recurrence:* that defines the behavior of the volcano in terms of active/not active year. We also analyze the magnitude recurrence to find a recurrence between events with the same magnitude. Due events with low magnitude are more frequent than higher magnitude ones.
  - *Probability Density Function (PFD) construction:* Recurrence and magnitude is converted into a function of probabilities.

  Additionally to the PDF construction, the sampling and the wind profiles are also part of this simulation service. The sampling consists of retrieving a unique set of values for the model inputs; one for each simulation. For the dispersal model considered here (see the HAZMAP Modelling tool in Chapter 2.2), samples are composed by a combination of 4 inputs: eruption column height, total erupted mass, particle granulometry and wind profile. For column height, sampling is based on the PDF. As a result, more likely column heights (with more probability in the PDF) have more samples associated. On the other hand, total erupted mass and granulometry are calculated by equations depending on the column height. Finally, wind profiles are obtained as described previously. The complete catalog is converted into daily-hourly wind profiles.

- **Scenario parametrization**. Entails a characterization of the initial conditions of a given scenario (or scenarios). It means that the sampled inputs are given to the model input files with the format required by the simulation tool. Also, a random function assigns a wind profile to every sample (senario).

- **Simulation**. Implies the simulation of scenarios or the group of scenarios.
- **Postprocessing**[**3]**:** the simulation results from all the scenarios are merged to build the final hazard map(s) of the volcano. Every map represents the distribution of the probabilities to reach a value equal or higher than an ash load (thickness is also valid) specified by the user. The resulting number of maps depends on the number of threshold load values.

2.2 Modeling Services

The previous Simulation Services involve a series of architectures, repositories, data providers, and tools (development, modelling, visual, instrumentation) to be developed. The modeling services aim to integrate all the services mentioned. The tools that support modelling services (see Figure 5) are explained next, with more detailed information given in the subsequent microservices design section.

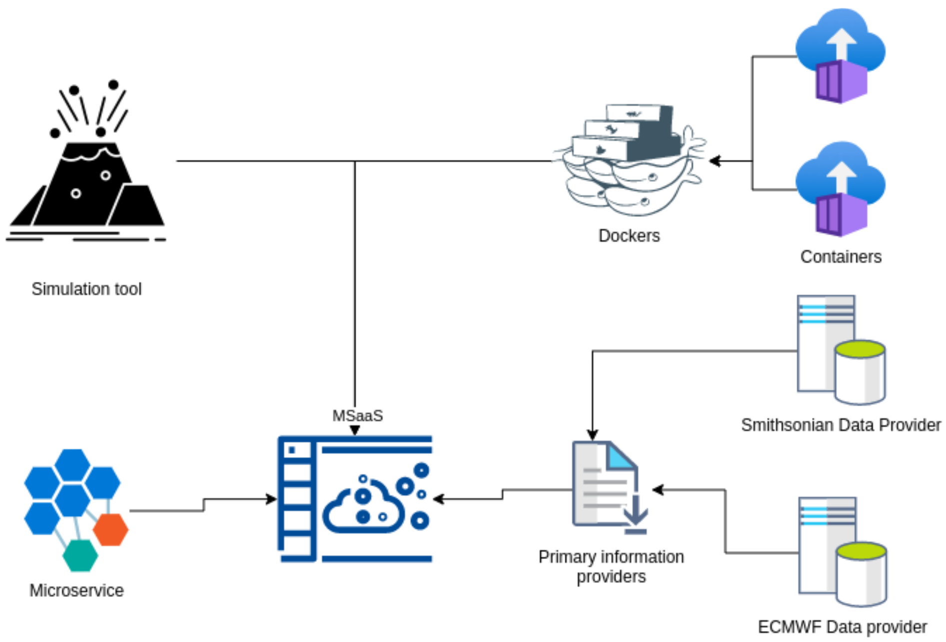


Figure 5. Modelling Services

- **Primary Information providers.** The meteorological information is divided in the individual scenario and probabilistic hazard map approach. However, the ERA-5 reanalysis is, in both cases, the source of 4-times daily wind profiles (direction and velocity). ERA-5 is a project developed by the European Centre for Medium-Range Weather Forecasts and offers a catalog of historical records of meteorological variables daily. The available data covers the Earth in a grid and implements the atmospheric layer using 137 levels from the surface up to a height of 80km (ECMWF, 2021).

[3] The services marked with ** are part of the probabilistic approach simulation only.

On the other hand, in the case of volcano historical activity, the Smithsonian Institution offers the information of volcano activity during the last 14.000 years. The Holocene Eruptions database provides the date and VEI (Volcanic Explosivity Index) of every registered event (Global Volcanism Program, 2013). The unique volcano ID is required and the service is accessible by Web Feature Service (WFS).

- **Modelling Tool.** HAZMAP is a Fortran-90 code which allows modeling the processes of diffusion, transport and sedimentation of ash particles in the atmosphere. The model simplifies the three dimensional system to a two dimensional system, providing mass accumulation at ground on a regular lon-lat grid (Macedonio et al., 2005). HAZMAP runs over Linux based SO and requires certain inputs related to the eruptive event:
  - the wind profile (velocity and direction) at the volcano location,
  - the plume parameters (column height, altitude, mass eruption rate),
  - the ash characteristic (essentially particle granulometry and properties) and,
  - the study zone limits or bounds and resolution of the computational domain (grid size, grid center, cell size).
- **Microservices.** This architecture emerged as an alternative to the monolithic code paradigm (i.e. one single code including all the functions) (Zheng and Wei, 2018). Microservices implements small functionalities that could be used for the user or other services because they offer autonomy, easy integration, small size, and decentralized gestion (Pachghare,2016).
- **Containers.** Standard software units that pack code and necessary libraries to run a service (IBM Cloud Education, 2021). The containers provide many advantages for developing such as portability, scalability, possibility to run multiple containers at the same time, and efficiency in terms of machine configuration (Mouat, 2015).
- **Dockers.** Can be described as a container's technology for virtualization (Anderson, 2015). This platform offers two components: Docker Engine, that allows the creation of containers and its running process, and Docker Hub, the cloud service for container distribution (Mouat, 2015). One of the biggest benefits of Dockers is that they provide solutions for container standardization on IT(Mouat, 2015).

2.3 Methodology

Before establishing a methodology for the project it is worth clarifying the concepts of simulation and modelling process. In terms of MSaaS, simulation is understood as the representation of reality. In contrast, modelling describes all the systems and methodologies that aim to deal with the simulation process (Kelton et al, 2014). The development of this project considers the following steps, adapted from Kelton et al. (2014), that describe the basics for the modelling.

- *Understand the system*, as well as we exposed in the first chapter, the problematic and the system behaviour have been analysed. Then, some problems surged, such as uncertainties, high computational requirements, variate knowledge and others. Also, some solutions are presented and deling in the Cloud Computing environment.
- *Formulate the model representation,* that is explained in the following chapter (Tool Design). Here we explain the workflow for probabilistic hazard mapping. As well as the model is implemented by Cloud, also the Service design and architecture.
- *Design the experiments, run the experiments* and *analyze your results.* After the design and the software development a case of study was displayed. The case of study for Popocatepetl volcano is presented in the third chapter. Here an ash fallout hazard map is developed and analyzed taking into account the volcano's past.

3. TOOL DESIGN

This section describes the schematization of the system. The design process must provide a solution to the problem through a model construction (Budgen, 1994). The first step to answer the three questions of design (what, how and who) is to define a workflow of the process (Kavis, 2014). The workflow not only solves the sequence and data dependencies, it also defines the methods for issues like probabilistic analysis and sampling.

Due to the complexity of the probabilistic approach, the workflow is explained more in detail. In the case of the single simulation the workflow is only overviewed because the processes involved are actually detailed in the probabilistic approach.

After this, the definition of the microservices is based on the workflow tasks. This includes not only the design of the workflow itself but also the definition of the communication between services and the virtualization process and requirements.

3.1 Workflow

The workflow describes the methodology adopted for ash fallout hazard assessment. Figure 6 shows a representation of the workflow for probabilistic hazard mapping. The assessment involves climatically representative wind profiles required by the tephra dispersal model, the generation of simulations giving the fallout deposit resulting from each eruption scenario, and the resulting hazard probability that results from simulation (footprints) aggregation.

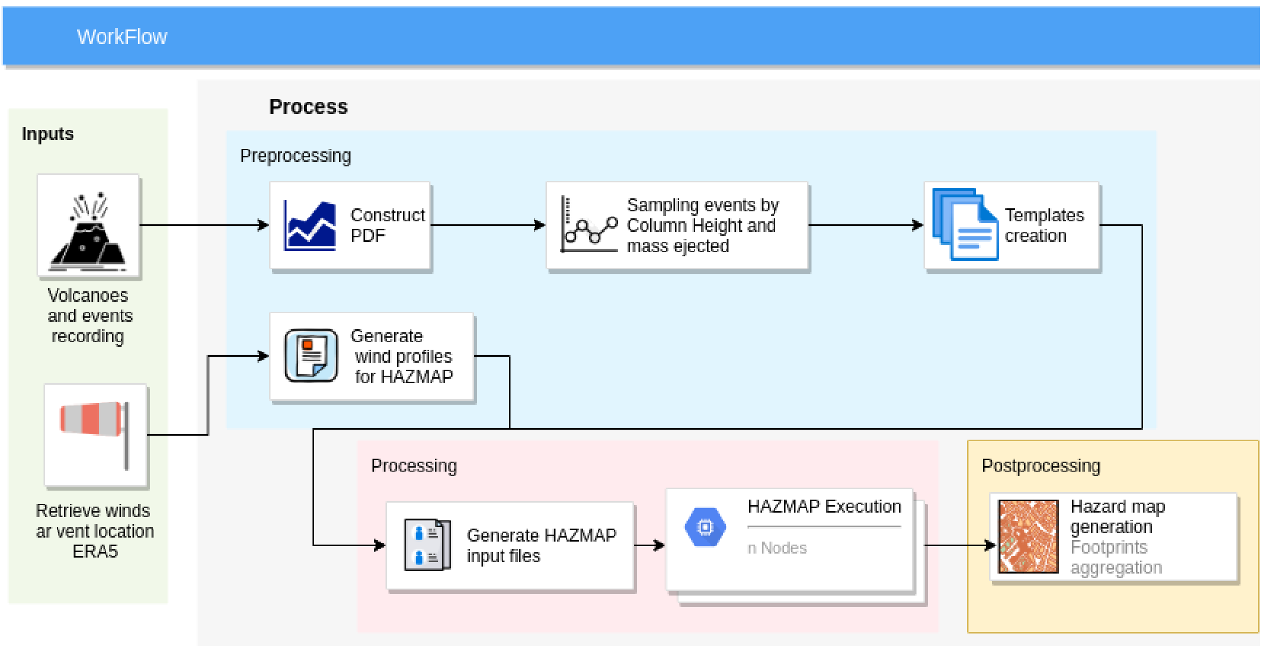


Figure 6. Probabilistic hazard assessment workflow.

The workflow also includes some application decisions. In this way the methodologies and models used to integrate all the systems are adjusted to the particular phenomena. The selection of the mathematical and statistical procedures was based in the literature and works related to past volcanology studies. As well as the project is modular, the changes into mathematical or statistical approaches proposed don't carry a total change in the system.

3.1.1. Volcanoes and event recording

The first step in the Hazard Map construction is the obtention and selection of data. The Smithsonian Institution supports the Global Volcanism Program (GVP), devoted to a better understanding of Earth's active volcanoes and their eruptions during the last 10.000 years (Holocene period). The Global Volcanism Program maintains a database that currently contains 1432 volcanoes with eruptions during the Holocene and more than 7.000 reports on volcanic activity (Global Volcanism Program, 2013). The data includes the (approximate) date of the event, the uncertainty status of the event, the method of data obtention (could be direct observation or geological deposits datation) and the magnitude traduced into the VEI (Volcano Explosivity Index), which is related to volume of the eruption and column height reached by the plume. The data is provided by the WFS service as we describe:

Smithsonian DB: WFS service

a. DB obtained WFS
   i. Link: https://webservices.volcano.si.edu/geoserver/GVP-VOTW/ows?version=1.0.0&typeName=GVP-VOTW:Smithsonian_VOTW_Holocene_Eruptions
   ii. Tables:
      1. Smithsonian_VOTW_Holocene_Eruptions, Smithsonian_VOTW_Holocene_Volcanoes

   In both tables the relation between volcano and volcano eruption are related through the volcano ID assigned in the Volcanoes table.

3.1.2. Retrieve winds at vent location

The HAZMAP model simulations use vertical wind profiles above the vent to advect the volcanic ash and determine the deposits on the ground. The model considers the winds as an static input, and does not consider changes on the profiles along the horizontal axis. Here we simulate postmortem and synthetic events, and assess them using wind profiles of the last 10 years. ERA-5 reanalysis dataset has been selected to this end, and the API from the

Copernicus Climate Data Store (CDS) is the tool used to download the data. Because the data have to be downloaded in a regular lat-lon grid, and the number of fields per file are limited, the downloads are configured in packages of 6 months and the grid is a box with 37 pressure vertical levels and four points per level around the vent.

The process is carried on by a python script that reads the list of the targeted volcanoes and their location from a csv file:

- Retrieve_ERA5_volcanoes.py (it needs a registered user credentials)
    - Input: Reads the volcanoes list and gets id, name, lat and lon
    - Workflow: Carry on a set of calls to the Copernicus CDS
        - Call - wait - download - call .....
    - Variables retrieved: U and V wind components and the geopotential (z)
    - Timing: To complete a volcano 10 years download can take from 3 hours to a day depending on the CDS workload (lower during nights and weekends).
    - Outputs: A netCDF file for each six months period and per volcano. The weight of each file if the time frequency is 6 hours is approx. 640 KB.

3.1.3. Probability Density Functions

The probabilistic approach requires a number of synthetic events defined using the PDF result. The samples represent a given eruption scenario, characterized by a given set of eruption source parameters and fixed meteorological conditions (wind profile). Given these, tephra dispersal models simulate the resulting fallout deposit (load of deposited material in kg/m2 or, equivalently, deposit thickness). For each volcano, results for the 14.608 simulations or footprints (one per wind profile) can be combined to obtain hazard maps and/or to perform cost analysis.

Values of eruption column height for each footprint are sampled following the corresponding PDF distribution. From these, total erupted mass was sampled within the range of values given in Table 1 that relates CH, VEI, volume and mass values. The number of samples per volcano has been defined by the number of wind profiles.

The Probability Density Functions tells which is the probability of having a certain column height (CH) in the case of an eruption in a determined volcano. Typically, eruption databases contain information on the Volcanic Explosivity Index (VEI) rather than on column heights (this happens because VEI is related to deposit volume, which is mostly the “natural” parameter that geologists obtain from field studies). As a result, the construction of PDFs in terms of VEI is straightforward, and simply consists of counting the number of events in each VEI category within the complete part of the database (i.e. from the turning point to present). These PDFs expressed in terms of VEI can then be formulated in terms of CH using the well-established relation between VEI and height. However, the problem with this conversion is that the VEI-height relationship is not univocal; a wide range of possible heights is associated to each VEI bin (see Figure 7). For this reason, the conversion deserves several comments:

1. The bias generated by the uncertain VEI-height relationship can be minimized by sampling randomly for column heights within the VEI-height range over many events. To this purpose, we generated an expanded "synthetic catalogue" with 1000 compatible column heights per each record. For example, if a volcano has a record of 20 eruptions with VEI = 2, the expanded database contains 20.000 possible values of column height, obtained by sampling randomly between 1 and 5 km (the possible range of heights compatible with VEI=2).

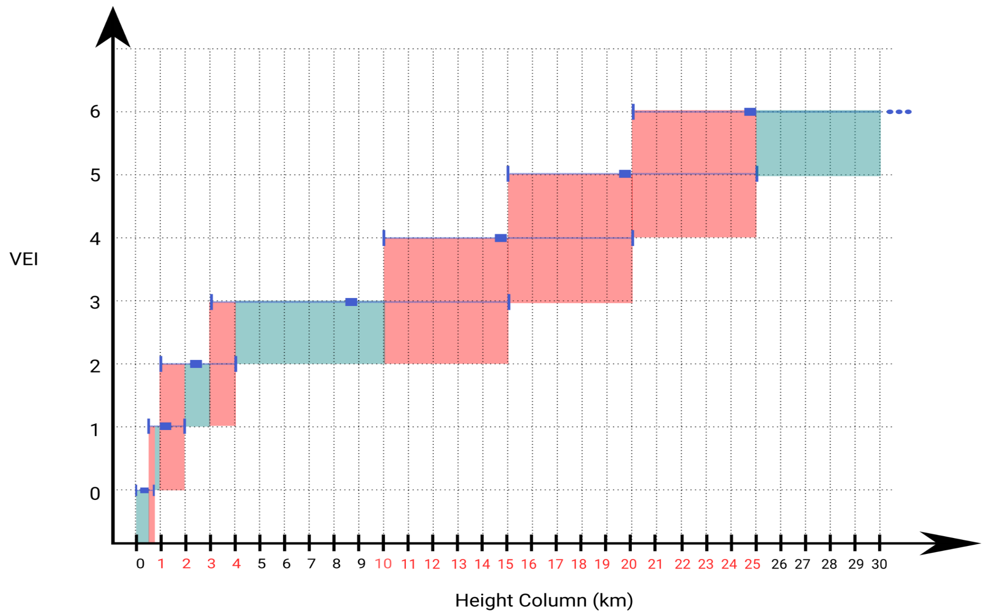


Figure 7. Column Height range related to VEI. *The red blocks correspond to Column heights related to two or more VEIs.*

2. With this expanded database one could, in principle, compute the discrete PDFs in terms of height by simple counting and binning. However, with no other corrections, over and oversampling situations would exist due to the overlapping of height intervals (note that some values of column height are related to two or more VEIs, whereas others are related to just one VEI value; see Figure 7). This drawback can be circumvented by finding an analytical PDF that represents the global behavior of the discrete data and decreases the variability in overlapping bins.

3. Three different analytical PDFs are typically used in volcanology, Normal, Chi-squared, and Gamma. Among these 3 possibilities, the best-fit option was selected for each volcano based on the chi-squared test and the p-values.

This process was coded in a python script and required the results of the Volcanoes recording task (3.1.1.)

a. Script: EruptionDB_PDFs_auto.py ➜ search the volcanoes data and generate the PDF image and file.

b. Establish the volcanoes selection criteria and the period studied:

    i. volcano_ID="85123"

c. Outputs: PDF image, PDF file, file with volcano information: ID, lat, lon, and elevation.

   Example: PDF_285020_Hokkaido-Komagatake_1y.dat (Column_Height, Probability %)

```
1.00 9.88
2.00 9.08
3.00 8.02
4.00 7.03
5.00 6.13
6.00 5.33
7.00 4.63
8.00 4.02
...
```

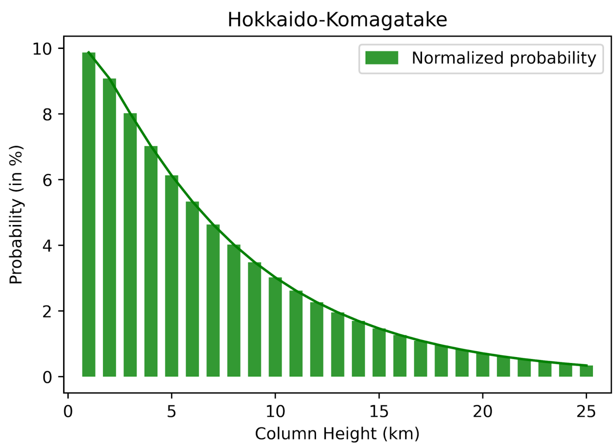


Figure 8. Example of PDF result. The results (left) could be interpreted as the graphic shows (right)

3.1.4. Wind profiles generation

4-daily wind profiles above each volcano were extracted from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-5 reanalysis dataset during the last 10 years of data available. The total number of wind profiles is 14.608 per volcano, each one representing a possible climatologically representative meteorological condition. Figure 9 shows wind roses for Merapi volcano at 3 different atmospheric levels of 5, 10, and 15 km above the sea level. However every wind profile contains the data of the meteorological conditions (velocity and direction) in a total of 37 atmospheric levels.

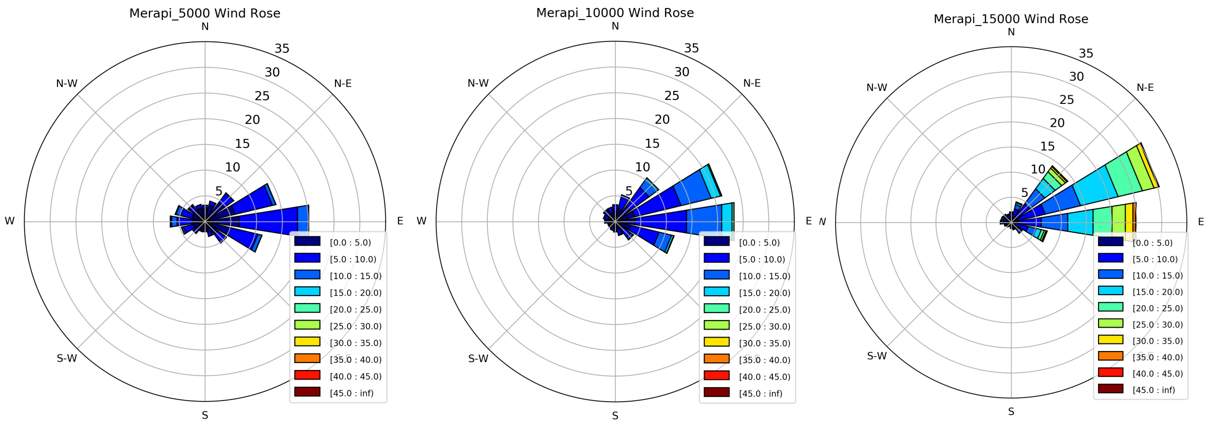


Figure 9. Wind roses for Merapi volcano for altitudes of 5, 10 and 15 km above the volcano summit.

This process comprehends several steps carried on by a python script called concatenate.py:

1. Input: Reads the volcano id and gets altitude, lat and lon
2. Workflow:
    a. Concatenates the netCDF files from ERA5 into one volcano 10 years file
        i. Checks if the whole set of files is available
        ii. Decompress each file
        iii. Concatenate them into one netCDF file
    b. Call ERA2dat Fortran90 program that:
        i. Interpolates to vent location
        ii. Interpolates to the defined vertical levels
3. Outputs:
        i. netCDF file per volcano and 10 years that weights 50 MB
        ii. One ASCII file with the vertical levels specified and the 10 years profiles, the file weights approx. 102 MB

iii. One ASCII file for each time step with the levels specified and the vertical wind profile. Each file weighs 12 K (or 256 K, depending on the minimum file weight defined by the system).

3.1.5. Sampling Events

Footprints represent tephra dispersal model results for a given eruption scenario, characterized by a given set of eruption source parameters and fixed meteorological conditions (wind profile). Given these, tephra dispersal models simulate the resulting fallout deposit (load of deposited material in kg/m2 or, equivalently, deposit thickness). For each volcano, results for the 14,608 footprints (one per wind profile) can be combined to obtain hazard maps and/or exceedance maps.

Values of eruption column height for each footprint were sampled following the corresponding PDF distribution. From these, total erupted mass was sampled within the range of values given in Table 1.

CH (km)	VEI	Minimum volume (m^3)	Minimum volume (m^3)	Minimum mass (kg)	Minimum mass (kg)
1	1.0	1e+4	1E+05	1E+07	1E+08
2	1.5	5E+05	5E+06	5E+08	5E+08
3	2.0	1E+06	1E+07	1E+09	1E+10
4	2.5	5E+06	5E+07	5E+09	5E+10
5	2.5	5E+06	5E+07	5E+09	5E+10
6	3.0	1E+07	1E+08	1E+10	1E+11
7	3.0	1E+07	1E+08	1E+10	1E+11
8	3.0	1E+07	1E+08	1E+10	1E+11
9	3.0	1E+07	1E+08	1E+10	1E+11
10	3.0	1E+07	1E+08	1E+10	1E+11
11	3.5	5E+07	5E+08	5E+10	5E+11
12	3.5	5E+07	5E+08	5E+10	5E+11
13	3.5	5E+07	5E+08	5E+10	5E+11
14	3.5	5E+07	5E+08	5E+10	5E+11

15	3.5	5E+07	5E+08	5E+10	5E+11
16	4.5	5E+08	5E+09	5E+11	5E+12
17	4.5	5E+08	5E+09	5E+11	5E+12
18	4.5	5E+08	5E+09	5E+11	5E+12
19	4.5	5E+08	5E+09	5E+11	5E+12
20	4.5	5E+08	5E+09	5E+11	5E+12
21	5.0	1E+09	1E+10	1E+12	1E+13
22	5.0	1E+09	1E+10	1E+12	1E+13
23	5.0	1E+09	1E+10	1E+12	1E+13
24	5.0	1E+09	1E+10	1E+12	1E+13
25	5.0	1E+09	1E+10	1E+12	1E+13

Table 1. The relation between eruption column height (CH in km above volcano), VEI and range of erupted volume and mass used to generate footprints.

The VEI “1.5, 2.5, 3.5 and 4.5” were defined by technical conditions and they denote Column Heights that belong with more than one VEI.

In this process, the PDF information is required to be joint to the relation between mass and column height (Table 1) . The final script generates an ESP file where the samples have the column height and mass values.

1. Script: gen_lhs.m (Octave)
2. Needs:
    a. lhs_desing.m (Octave) generates samples in Nx Dimensions from 0 to 1
    b. h_mass_rel_nh.dat ➔ relation between mass and column height. This file is mandatory.
    c. PDF file ➔ the file must content CH value and its probability
3. Input values: V volcano ID and the number of samples (nro_events=500)
4. Outputs: A file with Column Height (km) and mass (kg)

```
22.1      5.5000e+12
22.7      5.5796e+12
...
```

Example of sampling for 500 events with Cotopaxi PDF

3.1.6. Hazmap Input File templates creation

The HAZMAP model requires a Hazmap input file (see the example below) where some of the parameters are defined about the volcano, event information (like the column height that subsequently affects the number and dimension of the grid cells), and simulation configuration. The data of the volcano could be obtained from the Smithsonian Institution Web Service. Meanwhile, the event values, such as CH and total mass, are obtained from the ESP file. The simulation structure is mainly predefined, except for the number of grid cells and dimensions of the cells which depend on the column height value.

Example Hazmap input file for a Volcano Template_*volcanoID*_geninp.inp

```
0            # Mode (0=deposit; 1=hazmap)
2            # Grid_type (0=uniform in UTM, 1=from file, 2=uniform in l$
2            # Settling velocity model (2=Ganser)
1            # Settling velocity is a function of Z
0            # Write list of cell coordinates (file grid.out)
0            # Write particles spectra
0            # Write settling velocities (file vset.out)
```

```
3                  # Format of the output file (1=list points,2=ASCII,3=ASCII$
xxx  xxx           # Number of grid cells          (used if grid_type=0    or$
xxx  xxx           # Dimensions of the cells   (used if grid_type=0 in UTM or$
110.44 -7.54            # Center of the grid        (used if grid_type=0 in UTM or$
xxx                # Total Mass
1                  # Mass distribution model (0=from file, 1=Suzuki)
110.44 -7.54 2930       # Coordinates of the vent (X,Y,Z)
xxx                # Column height (a.s.l.)
100                # Number of vertical source points
4. 1.              # Suzuki coefficients
10000.             # Diffusion coefficient
1                  # Number of loading thresholds
```

* Values with XXX have to be changed automatically with a script based on the samples obtained before.

3.1.7. Footprints generation

This part of the process generates the HAZMAP input files and runs the model for each footprint. All the steps are carried on here by a bash script called Script-run-footprints that is called by another bash script manually set for each volcano. The latter only sends the previous script to the execution queue with the volcano name as an argument. Following two subsections summarize the process steps.

3.1.7.1. Hazmap inputs generation

The Hazmap input files are generated by a pre-processing program called geninp. In addition, another program called TGSD defines the granulometry used in the simulations. Following lines list the steps:

1. Sets the directories and copy the executables
2. To generate each Hazmap input file (footprint):
    a. Sets the counter and gets the mass and column height from the list of events sampled (Sampling Events).
    b. Runs the TGSD.x program to define the problem granulometry. The only argument is the column height.

c. Writes column height and mass using awk commands and adds granulometry into the geninp input file.

d. Runs geninp

3. Outputs:

   a. Hazmap input files (one per event/footprint) in ASCII format. Each file weights 12 K (or 256 K, depending on the minimum file weight defined by the system).

3.1.7.2. Run Hazmap

This last step only runs the Hazmap model for each event. For this purpose a code has been developed to automatically

1. Outputs:

   a. Hazmap footprints in netCDF format. The files weight 1.3 MB each, however this quantity is only orientative because it depends on the number of cells.

3.1.8. Hazard map generation

Once the footprints are simulated and stored into the machine, the post processing computes two different kinds of hazard maps.

- **Hazard maps** give, at any grid point, the probability of exceeding a given threshold value of the ground load. For example, maps for a load threshold of 1 kg/m^2 give the probability of tephra ground load exceeding this value. Maps are computed for thresholds of 1, 10, 100, and 500 kg/m^2, roughly corresponding to deposit thicknesses of 1 mm, 1 cm, 10 cm and 50 cm respectively.
- **Exceedance probability maps** give, at any grid point, which minimum deposit load (thickness) can be expected with a given probability. Exceedance probability maps are computed for 5, 50, and 95% percentiles. These maps are computed from footprints by ranking for column mass and finding the value of a given percentile.

Both processes are run by the Fortran script named footprints_2_HM. The final files correspond to one NetCDF for each Hazard or exceedance value. If all the listed values are developed we will obtain 4 Hazard maps and 3 Exceedance probability maps.

1. Inputs

   a. Volcano ID

   b. Footprints. The 14608 footprints are mandatory.

2. Outputs:
    a. Hazard maps NetCDF files
    b. Exceedance probability NetCDF files

3.2 Microservices design

The workflow defined a total of 6 services that are composed of one or more microservices. The development of the tool required the use of the following languages, tools, and frameworks.

- Dockers,
- Django,
- Python,
- Octave,
- Fortran

3.2.1. External Data Obtention

The external data is composed of volcanic activity information and a meteorological setup. Both of the services consume the web services of each institution, which are explained below.

3.2.1.1. Smithsonian Institution

The Smithsonian Institution provides a WFS. Web Feature Service (WFS) provides an interface standard that allows a client to get geographical feature data from an internet server using platform-independent requests (Global Volcanism Program, 2013). The following XML files show the structure of the data acceded. The information is identified by Volcano_Number key.

Volcano Data XML


```
<xsd:schema xmlns:GVP-VOTW="volcano.si.edu" xmlns:gml="http://www.opengis.net/gml"
xmlns:xsd="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified"
targetNamespace="volcano.si.edu">
  <xsd:import namespace="http://www.opengis.net/gml"
  schemaLocation="https://webservices.volcano.si.edu/geoserver/schemas/gml/2.1.2/feature
  .xsd"/>
  <xsd:complexType name="Smithsonian_VOTW_Holocene_VolcanoesType">
    <xsd:complexContent>
      <xsd:extension base="gml:AbstractFeatureType">
        <xsd:sequence>
          <xsd:element maxOccurs="1" minOccurs="1" name="Volcano_Number"
          nillable="false" type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Volcano_Name" nillable="true"
          type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Primary_Volcano_Type"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Last_Eruption_Year"
          nillable="true" type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Country" nillable="true"
          type="xsd:string"/>
```

```
          <xsd:element maxOccurs="1" minOccurs="0" name="Geological_Summary"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="1" name="Region" nillable="false"
          type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="1" name="Subregion" nillable="false"
          type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Latitude" nillable="true"
          type="xsd:decimal"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Longitude" nillable="true"
          type="xsd:decimal"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Elevation" nillable="true"
          type="xsd:short"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Tectonic_Setting"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Geologic_Epoch"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Evidence_Category"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Primary_Photo_Link"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Primary_Photo_Caption"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Primary_Photo_Credit"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Major_Rock_Type"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="GeoLocation" nillable="true"
          type="gml:GeometryPropertyType"/>
        </xsd:sequence>
      </xsd:extension>
    </xsd:complexContent>
  </xsd:complexType>
  <xsd:element name="Smithsonian_VOTW_Holocene_Volcanoes"
  substitutionGroup="gml:_Feature"
  type="GVP-VOTW:Smithsonian_VOTW_Holocene_VolcanoesType"/>
</xsd:schema>
```

Volcano Activity XML

```
<xsd:schema
xmlns:GVP-VOTW="volcano.si.edu" xmlns:gml="http://www.opengis.net/gml"
xmlns:xsd="http://www.w3.org/2001/XMLSchema" elementFormDefault="qualified"
targetNamespace="volcano.si.edu">
  <xsd:import namespace="http://www.opengis.net/gml"
  schemaLocation="https://webservices.volcano.si.edu/geoserver/schemas/gml/2.1.2/feature.x
  sd"/>
  <xsd:complexType name="Smithsonian_VOTW_Holocene_EruptionsType">
    <xsd:complexContent>
      <xsd:extension base="gml:AbstractFeatureType">
        <xsd:sequence>
          <xsd:element maxOccurs="1" minOccurs="1" name="Volcano_Number" nillable="false"
          type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="Volcano_Name" nillable="true"
          type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="1" name="Eruption_Number" nillable="false"
          type="xsd:int"/>
```

```
          <xsd:element maxOccurs="1" minOccurs="0" name="Activity_Type" nillable="true"
          type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="ExplosivityIndexMax"
          nillable="true" type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="ExplosivityIndexModifier"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="ActivityArea" nillable="true"
          type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="ActivityUnit" nillable="true"
          type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="StartEvidenceMethod"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="StartDateYearModifier"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="StartDateYear" nillable="true"
          type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="StartDateYearUncertainty"
          nillable="true" type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="StartDateDayModifier"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="StartDateMonth" nillable="true"
          type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="StartDateDay" nillable="true"
          type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="StartDateDayUncertainty"
          nillable="true" type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="EndDateYearModifier"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="EndDateYear" nillable="true"
          type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="EndDateYearUncertainty"
          nillable="true" type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="EndDateDayModifier"
          nillable="true" type="xsd:string"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="EndDateMonth" nillable="true"
          type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="EndDateDay" nillable="true"
          type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="EndDateDayUncertainty"
          nillable="true" type="xsd:int"/>
          <xsd:element maxOccurs="1" minOccurs="0" name="GeoLocation" nillable="true"
          type="gml:GeometryPropertyType"/>
        </xsd:sequence>
      </xsd:extension>
    </xsd:complexContent>
  </xsd:complexType>
  <xsd:element name="Smithsonian_VOTW_Holocene_Eruptions" substitutionGroup="gml:_Feature"
  type="GVP-VOTW:Smithsonian_VOTW_Holocene_EruptionsType"/>
</xsd:schema>
```

3.2.1.2. ERA5

The meteorological data, mainly wind data in this project, is obtained from the ECMWF Institution that provides the information of the historical setup through a Climate Data Store (CDS) Web API. The information needs to be accessed by a key parametrization and configuration of the API in the client machine (ECMWF, 2021). The data could be downloaded using Python. The following query corresponds to a specific wind profile for Popocatepetl volcano on the date/ time 01/07/2015 12:00h.

```
import cdsapi
c = cdsapi.Client()
c.retrieve("reanalysis-era5-pressure-levels",
{
"variable": "wind",
"pressure_level": "1000",
"product_type": "reanalysis",
"year": "2015",
"month": "07",
"day": "01",
"time": "12:00",
"latitude": "19.023",
"longitude": "-98.622",
"format": "NetCDF"

}, "volcano_10y_wind.nc")
```

3.2.2. PDF Generation*[4]

This process consists of the activity data analysis and statistical approach to obtain the Column Height recurrences. Three main distributions were applied based on the literature of volcanic behavior for continuous functions: LogNormal, ChiSquared and Gamma. The data was evaluated to correspond to the events defined by VEI, in which a total of 1000 column height values were sampled in order to represent the whole range of VEIs definition. The script was developed in Python and needs as an input the volcanic activity obtained from Smithsonian Institution in Json format. Finally, a file called PDF_<volcano_ID>.data i

s obtained and contains the information of Column height and probability of occurrence like the following example.

```
1.00 9.88
2.00 9.08
3.00 8.02
```

[4] The services marked with * are process that only concerns to probabilistic approach

```
4.00 7.03
5.00 6.13
6.00 5.33
7.00 4.63
8.00 4.02
...
```

3.2.3. Sampling*

The generation of the Column Height and mass combination is supported by the values defined in Table 1. The approach applied is LHS (Latin Hypercube Sampling) which needs less sample generation to construct a statistically valid dataset compared to Montecarlo and Random Sampling techniques (Matala, 2008). For this purpose, every Column Height was obtained separately and the number of samples was obtained by the relation between Column Height and recurrence versus the total number of footprints (predefined by winds profile in 14608). For example, if one bin has 10% of recurrence 1461 samples will be generated in this bin. The result is a .dat file with the assigned mass (in kg), the column height value (meters), the VEI, and footprint ID.

```
1.98e+11  14366.2  3.5  12234
6.57e+10     5882  3     4022
1.36e+12  22013.5  5    14375
7.62e+10  6670.71  3     6346
...
```

3.2.4. Event Parametrization

This service is charged with input files generation. The automatization demands the creation of several files and the conversion of the data obtained before into HAZMAP input files. The simulation code provides some tools to convert simple parameters such as mass and column height into granulometria layers. The dependencies and data flow of these files are detailed in the Appendix 3.

3.2.4.1. Volcano parametrization (location, altitude, resolution, and size of the grid)

This microservice takes the information from Smithsonian Volcanoes data for location (lon, lat) and altitude in meters. Meanwhile, the resolution is given by the size of the cell grid and the number of cells. The value is obtained by the maximum Volcanic Explosivity Index recorded, assuming that a volcano with more explosivity will need a grid bigger than in small events.

3.2.4.2.Event Parametrization (duration, date)

The event's initial conditions definition is given by ESP file with the combination of Column Height and mass. Additionally, the duration of the emission is mandatory and impacts the simulation. The value is obtained by the Column Height value and mass applying the formula given by Mastin (2014) for MER (Mass Eruption Rate), which is a relationship between mass and duration of the event.

3.2.4.3.Granulometry

Granulometry refers to total grain size distribution (TGSD). TGSD is directly related to the reaching ash, because fine ash tends to travel a larger distance than granular ash, and also defines the concentration in the land. For this purpose TGSD tool, taken from FALL3D code and written in Fortran, was adjusted to follow a certain volcano grain size distribution derived from plume height. These values are part of the Geninp.inp file such as the following example.

```
 27     # Number of particles types, Diameter, density, shape, wt%
0.320E-01   1200.0   0.900    0.04980
0.226E-01   1200.0   0.900    0.12314
0.160E-01   1200.0   0.900    0.35307
0.113E-01   1200.0   0.900    0.84677
0.800E-02   1200.0   0.900    1.72528
0.566E-02   1200.0   0.900    2.97133
0.400E-02   1200.0   0.900    4.32939
0.283E-02   1200.0   0.900    5.34861
0.200E-02   1200.0   0.900    5.61907
0.141E-02   1292.9   0.900    5.08156
0.100E-02   1385.7   0.900    4.11575
0.707E-03   1478.6   0.900    3.31883
0.500E-03   1571.4   0.900    3.17662
0.354E-03   1664.3   0.900    3.89824
0.250E-03   1757.1   0.900    5.30408
0.177E-03   1850.0   0.900    6.96723
0.125E-03   1942.9   0.900    8.34136
0.884E-04   2035.7   0.900    8.94584
...
```

3.2.4.4. ERA5 NetCDF to winds.dat

This step converts the ERA5 NetCDF file into small .dat files which contain the wind records in a specific time point (date and hour). The Script that generates these files is a bash file that creates the 14608 files with the Footprint ID.

3.2.4.5.Geninp2Hazmap.inp

Once the Geninp file has the volcano, event, and granulometry information the executable GENINP is recommended to use. The source distribution Geninp can read the initial simulation data directly from the file sources.inp) or, more simply, computes the mass along the vertical column in a discrete number of points. The input needed are the Geninp file and the Footprint ID and the result is a hazmap.inp file (For more information see the Appendix 4.

3.2.5.Run

The simulation process as well as the post-processing of the resulting information were developed by using the MareNostrum 4 capabilities and the personal authentication of the BSC CASE department. The process is launched by a batch script, shown below, which requests 48 nodes and a total of 11 hours for execution. The bash script Script-run-footprints takes the hazmap.inp file, wind.dat file and executes every footprint assigning an ID value.

```
#!/bin/bash
#SBATCH --job-name=volcano_run_372070
#SBATCH --workdir=.
#SBATCH --output=.%x_%j.out
#SBATCH --error=./%x_%j.err
#SBATCH --nodes=1
#SBATCH --cpus-per-task=1
#SBATCH --tasks-per-node=48
#SBATCH --ntasks=48
#SBATCH --time=11:00:00
#SBATCH --qos=bsc_case

date
./Script-run-footprints 372070
date
```

3.2.6. Post-processing

3.2.6.1. Probabilistic hazard mapping

The probabilistic hazard mapping process is developed by an executable Script that takes all the footprints listed in the input file and creates the maps according to the given loads and percentiles (load for probabilistic hazard map and percentiles for exceedance maps). Additionally, the number of footprints to be processed and the path of the footprints files need to be specified, like in the following example of the input file.

```
!
!    Example of input file for footprints_2_HM
!
THRESHOLDS
  LOADS_(KG/M2)   = 0.1 1 10 100
  PERCENTILES_(%) = 5 50 95

FOOTPRINTS
  NUMBER = 4
  PATH   = /Users/.../Footprints/Merapi$
FILE_LIST
 Merapi-footprint-00001.nc
 Merapi-footprint-00002.nc
 Merapi-footprint-00003.nc
 Merapi-footprint-00004.nc
...
```

4. CASE OF STUDY

4.1. Popocatepetl volcano hazard map

The Popocatepetl volcano is an andesitic composite volcano located in the southern end of the Sierra Nevada and centered in the Mexican Volcanic Belt (Arana-Salinas et al., 2010; Arciniega-Ceballos et al., 1999; De la Cruz-Reyna and Tilling, 2008). It is considered the potentially most dangerous volcano in Mexico (Arana-Salinas et al., 2010). The present cone is a huge edifice formed after the destruction of past similar structures by gravitational collapse (Siebe and Macías, 2003).

The volcanic activity in the last 800 years has been rated as VEI 1 – 3, with a prevalence of fumarolic and seismic activity in the last 70 years (Arciniega-Ceballos et al., 1999; Delgado-Granados et al., 2001). Geological studies show at least 4 Plinian or large Plinian eruptions occurred at 14.000, 4965, 2150, and 1100 yr BP. The eruptive record of this volcano consists of 64 eruptions since 7150 BCE according to the Smithsonian Institution. (Global Volcanism Program, 2013).

One of the most important characteristics of Popocatepetl volcano is the distance (less than 100 km) to the two principal Mexico's cities: 70 km southeast of Mexico city, with 25 million people, and 40 km west of Puebla, with a population of 1.5 million (Arana-Salinas et al., 2010; De la Cruz-Reyna & Tilling, 2008). A Plinian event with wind prevalence towards northwest could affect Mexico city infrastructure including electric network, sewer system and the international airport (Siebe & Macías, 2003).

4.1.1. Historical Events

The Global Volcanism Project from the Smithsonian Institution supplies data about the date and Volcanic Explosivity Index (VEI) of Holocene events on 1426 volcanoes (Global Volcanism Program, 2013). The main information of the volcano and data from the main activity of the volcano is recorded here. However, the Smithsonian Institution just provides the VEI, which means that the parameters of the event must be estimated. The following table shows an example of activity characterization of Popocatepetl volcano in the three oldest events. The events are sorted by date, with the most recent activity in the last row and the oldest Holocene register in the first. Each event consists of 9 fields or cells (tab sheet columns) detailed in Table 2. Field values can be introduced manually (if known) or, alternatively, estimated. The complete information of the events is recorded in Appendix 1.

Volcano	Popocatepetl
Summit Elevation	5393,00

Population within 30 km				6,34E+05				
Activity record								
Year	**Month**	**Day**	**VEI**	**Mass (kg)**	**Volume (m3)**	**Min Column height (km)**	**Max Column height (km)**	**Mean Column height (km)**
-7150	**	**	4	5,00E+11	5,00E+08	10,00	20,00	15,00
-6250	**	**	**	Volume is not specified	VEI is not specified	VEI is not specified	VEI is not specified	VEI is not specified
-5150	**	**	4	5,00E+11	5,00E+08	10,00	20,00	15,00

Table 2. Example of volcano activity obtention and main parameters estimation.

4.1.2. Probability Density Function

Probability Density Functions (PDFs) are functions giving the probability of a given variable (in this case eruption column height) falling within a range of values. PDFs can be relative (i.e. total probability normalized to 1) or absolute (i.e. total probability less than one). In our context, absolute PDFs give the probability distribution of Column Height assuming that the eruption occurred, then, the probability of having an eruption of any column height is 1. This is illustrated in Figure 10 for the particular case of Popocatepetl volcano, where different approaches are applied. The approach selected corresponds to the best fit curve, which was evaluated by the root mean squared error and taking into account previous studies of volcanoes behavior. The figures are automatically generated.

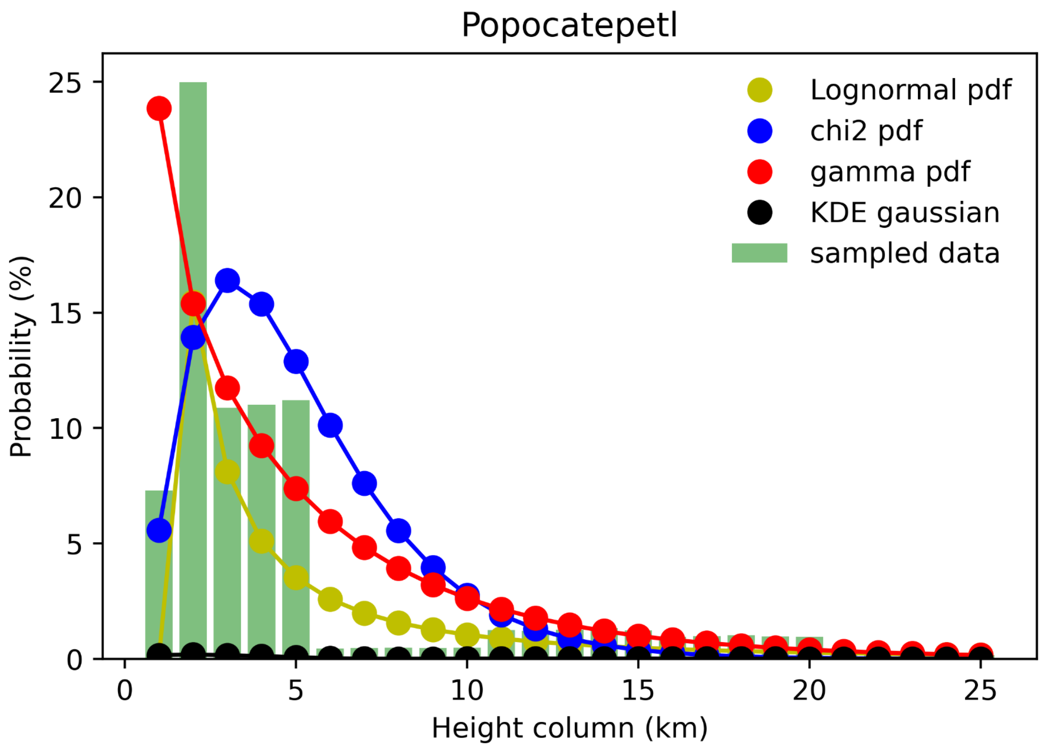


Table 3 shows the values obtained for Popocatepetl volcano. The first column represents the plume height (Column Height or CH) over the volcano summit (in km), and the second one refers to the probability of occurrence expressed in percentage.

CH	P(%)	CH	P(%)	CH	P(%)	CH	P(%)	CH	P(%)
1.00	23.86	**6.00**	5.96	**11.00**	2.16	**16.00**	0.83	**21.00**	0.34
2.00	15.40	**7.00**	4.83	**12.00**	1.78	**17.00**	0.69	**22.00**	0.28
3.00	11.75	**8.00**	3.93	**13.00**	1.47	**18.00**	0.58	**23.00**	0.24
4.00	9.25	**9.00**	3.21	**14.00**	1.21	**19.00**	0.48	**24.00**	0.20
5.00	7.39	**10.00**	2.63	**15.00**	1.00	**20.00**	0.40	**25.00**	0.17

Table 3. PDF values for Popocatepetl volcano

4.1.3. Winds Profile Obtention

We obtained 4-hourly wind profiles above each volcano from ERA5 reanalysis dataset during the period 2011-2020 (10 years). The total number of profiles is 14608 per volcano, each representing one possible meteorological condition. Every wind profile contains the wind speed (in the components X, Y) on a total of 37 atmospheric pressure levels, which could be converted into altitudes. Appendix 2 contains an example of the wind profiles structure. Figure 11 shows the wind roses in the Popocatepetl summit location for 5000, 10.000, and 15.000 m.a.s.l., that resume the wind behavior (speed and direction) in the last 10 years.

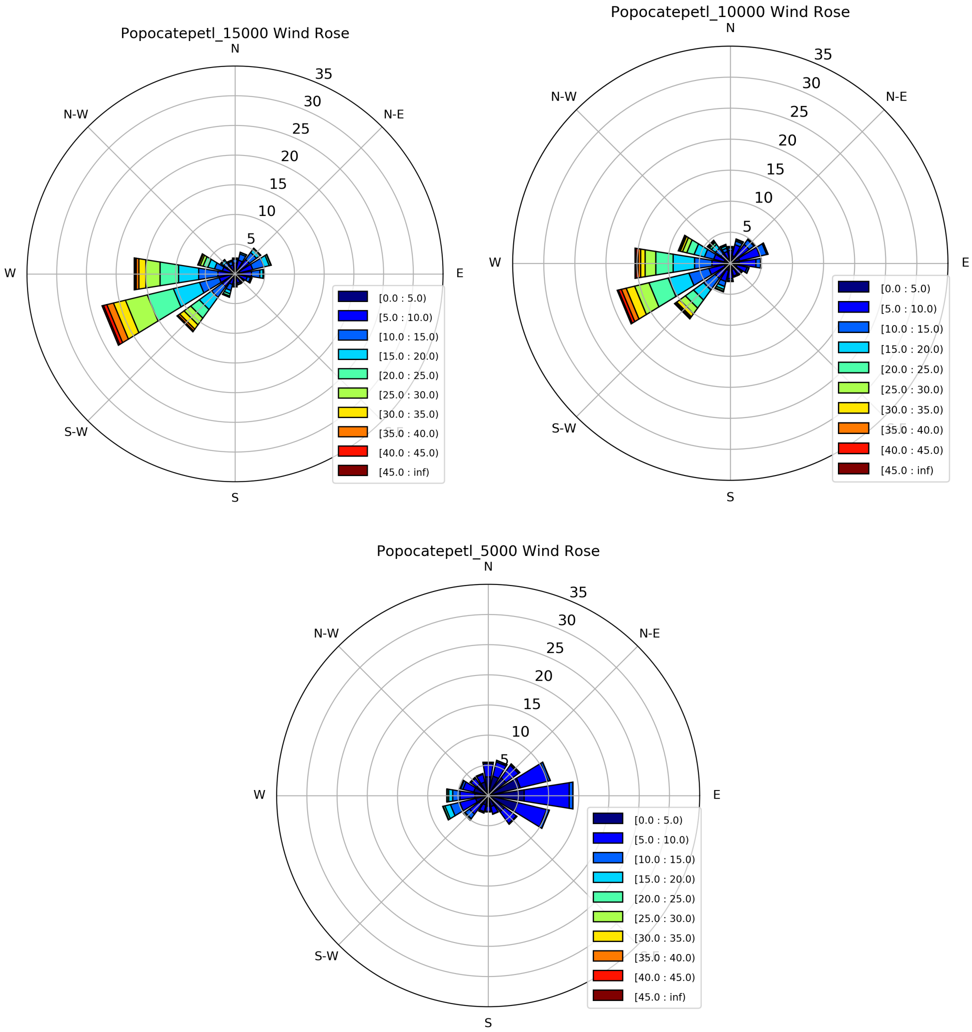


Figure 11. Wind roses 5000, 10000, and 15000 m.a.s.l. in the Popocatepetl volcano location.

4.1.4. Synthetic Catalog Sampling

Footprints represent tephra dispersal model results for a given eruption scenario, characterized by a given set of eruption source parameters and fixed meteorological conditions (wind profile). Given these, tephra dispersal models simulate the resulting fallout deposit (a load of deposited material in kg/m2 or, equivalently, deposit thickness). Values of eruption column height for each footprint were sampled following the corresponding PDF distribution.

For each bin, footprint parameters were sampled using the Latin Hypercube Sampling technique, which allows a better distribution of samples across the sampling space (in this case, delimited by column height and mass ranges). Figure 12 shows the resulting footprint parameters for Popocatepetl volcano. Each 'X' represents a combination of Column Height and Mass values to be simulated.

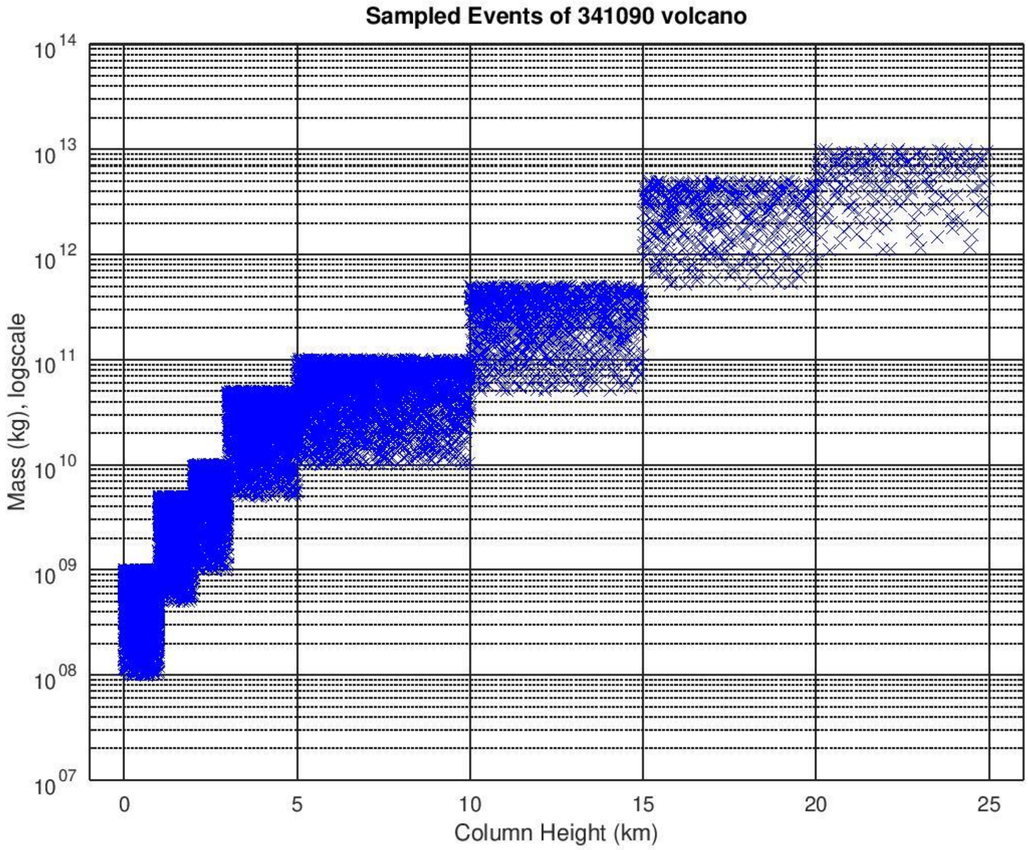


Figure 12. Synthetic events obtention by Sampling technique.

4.1.5. Template Creation.

Through the ESP file, basic volcano data, Geninp, TGSD granulometry script, and wind profiles, each of the templates with the initial conditions were generated automatically. The following is an example of an input data file for Popocatepetl volcano. For more information about the input file generation and dependencies see Appendix 3.

```
#hazmap.inp
0            # Mode (0=deposit, 1=hazard map, 2=barycenters)
2           # Grid_type (0=uniform, 1=from file)
```

```
2            # Settling velocity model
1           # Settling velocity is a function of Z
0            # Write list of cells coord. (file grid.out)
0            # Flag: write particles spectra
0            # Flag: write settling velocity profile
3           # Format of the output file
300    300 # Number of grid cells (used if grid_type=0)
0.20000E-01  0.20000E-01  # Dimensions of the cells (used if grid_type=0)
-98.6      19.0        # Center of the grid (used if grid_type=0)
1           # Number of thresholds (used if mode=1)
100.00       # Loading threshold (used if mode=1)   1
10000.       # Diffusion coefficient
27        # Number of particles types
0.32000E-01  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   1
0.22600E-01  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   2
0.16000E-01  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   3
0.11300E-01  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   4
...
```

4.1.6. Simulation and Postprocessing.

The simulation process for this purpose was developed through the MareNostrum 4 credentials for CASE department users. A total of 14608 footprints were obtained in a time-lapse of around 8 hours. Finally, hazard and exceedance probability maps were produced to facilitate the analysis of the footprints for each volcano. The results are given into NetCDF files.

Exceedance probability maps give, at any grid point, which minimum deposit load (thickness) can be expected with a given probability. Exceedance probability maps were computed for 5, 50, and 95% percentiles. These maps are computed from footprints by ranking for column mass and finding the value of a given percentile.

For illustrative purposes, Figure 13 shows the worst 5% exceedance probability map (in kg/m2), processed and formatted in the QGIS program.

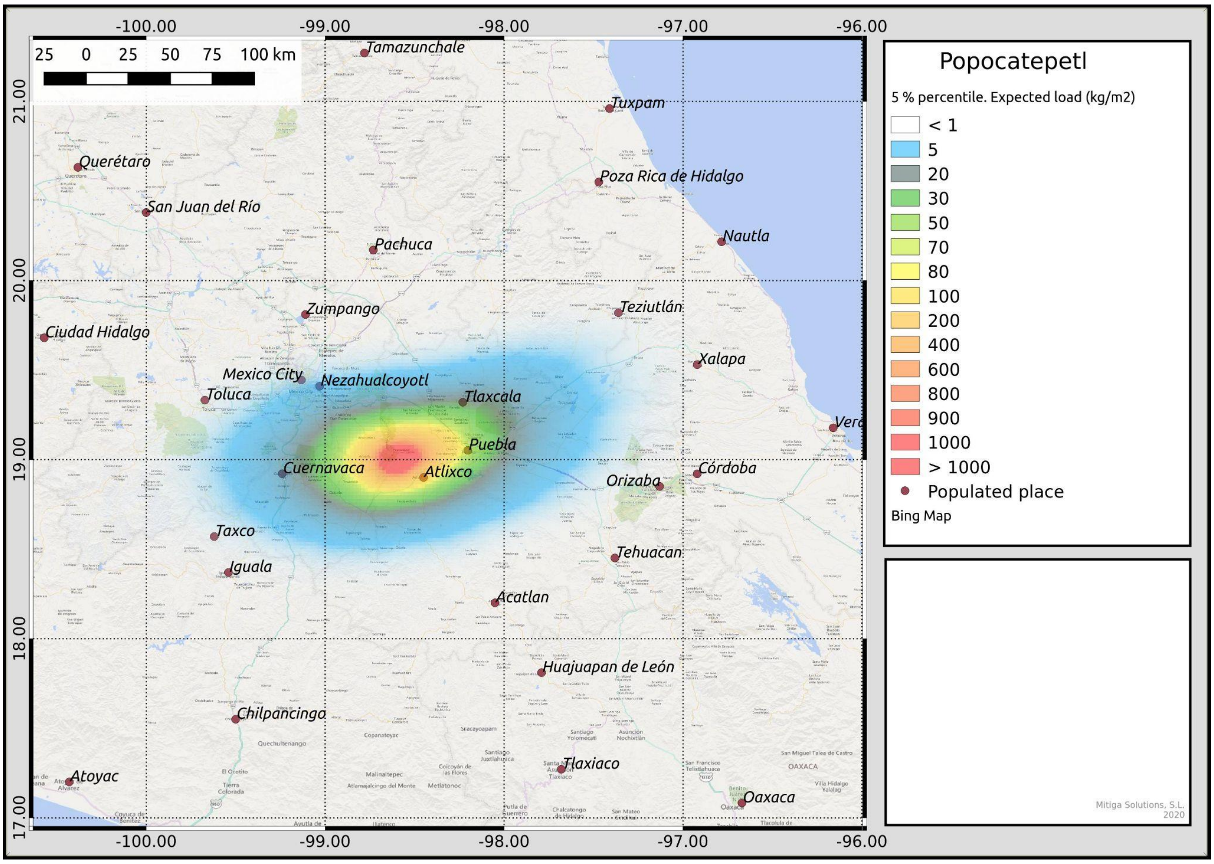


Figure 13. Popocatepetl volcano exceedance probability map (5% percentile).

Hazard maps give, at any grid point, the probability of exceeding a given threshold value of the ground load. For example, maps for a load threshold of 1 kg/m2 give the probability of tephra ground load exceeding this value. Maps were computed for thresholds of 1, 10, 100, and 500 kg/m2, roughly corresponding to deposit thicknesses of 1 mm, 1 cm, 10 cm, and 50 cm. Figures 14 and 15 two hazard maps for thresholds on 1 mm and 10 cm (in %) for the case of Popocatepetl volcano.

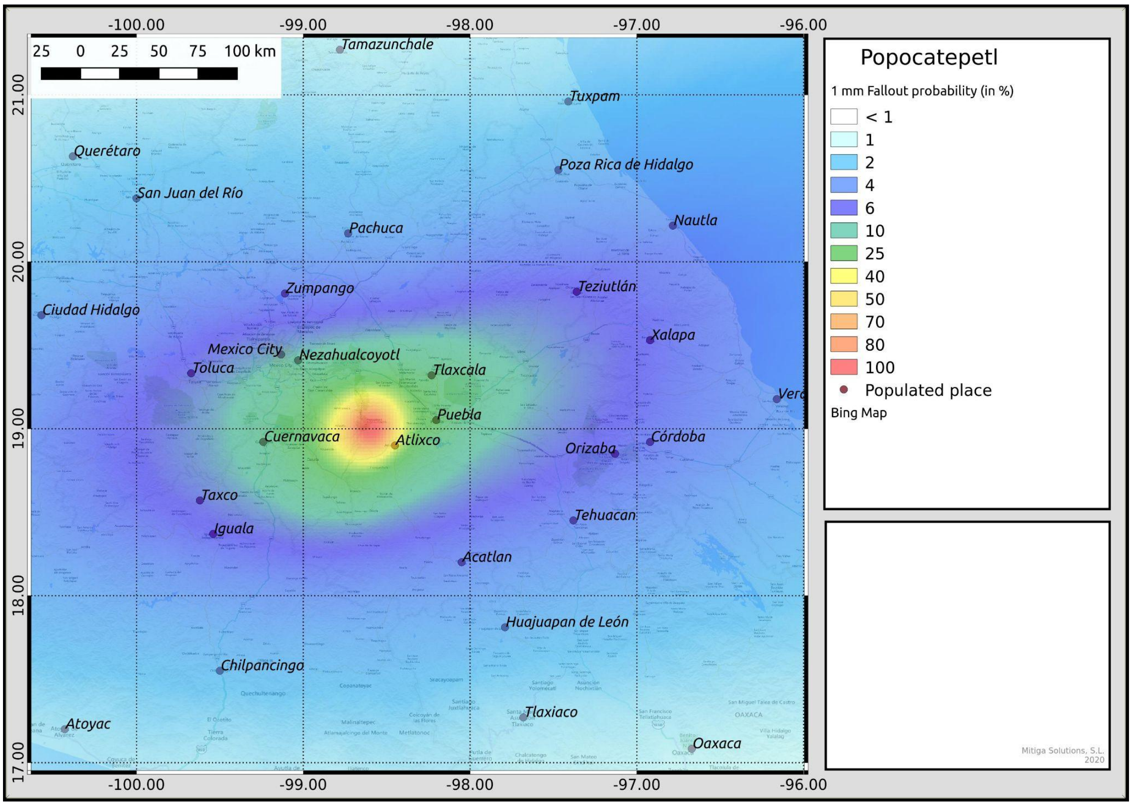


Figure 14. Probabilistic hazard map for 1mm fallout in Popocatepetl volcano surroundings.

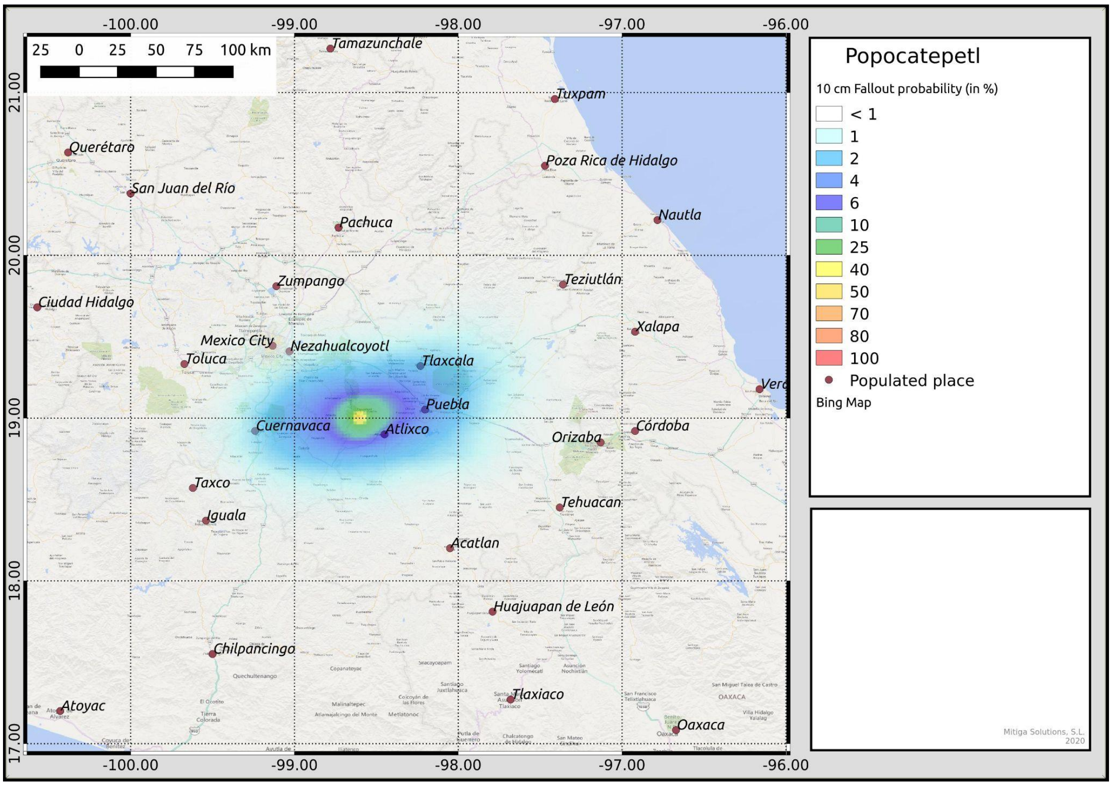


Figure 15. Probabilistic hazard map for 10 cm fallout in Popocatepetl volcano surroundings.

5. CONCLUSIONS

- The project developed the main structure of a MSaaS applying the supercomputing capabilities of the BSC to develop ash fallout probabilistic hazard maps.
- The development of MSaaS is crucial in environments where different knowledge areas are applied because this decreases the multi-area expertise requirements for researchers.
- The proposed methodologies and approach to develop the probabilistic hazard map could be updated without affecting the whole system because of the modular design of microservices architecture.
- The proposed methodology is able to reproduce the volcano behaviour. Although the official hazard map shows a larger area affected, the winds behaviour is similar and the difference is given by ash dispersion which is not considered in this project.
- The simulation of real phenomena by computational codes demands high resources and infrastructure capabilities making the Supercomputation a requirement in these areas.
- The present project was focused on the construction of the methodology and the backend coding. The front end and the customization of services must be developed in future phases.

6. NOTES

6.1. Software de Terceros

Para la implementación del proyecto se utilizó el software de modelamiento de cenizas HAZMAP de openaccess.

6.2. Dirección del proyecto

El proyecto fue dirigido por el ENVIRONMENTAL SIMULATIONS GROUP MANAGER del Departamento CASE de BSC, Arnau Folch, bajo la modalidad de contrato.

7. REFERENCES

Anderson, C. (2015). Docker [Software engineering]. *IEEE Software*, *32*(3), 102-c3. https://doi.org/10.1109/ms.2015.62

Arana-Salinas, L., Siebe, C., & Macías, J. L. (2010). Dynamics of the ca. 4965yr 14C BP "Ochre Pumice" Plinian eruption of Popocatépetl volcano, México. Journal of Volcanology and Geothermal Research, 192, 212–231. https://doi.org/10.1016/j.jvolgeores.2010.02.022

Arciniega-Ceballos, A., Chouet, B. A., & Dawson, P. (1999). Very long-period signals associated with vulcanian explosions at Popocatepetl volcano, Mexico. Geophysical Research Letters, 26(19), 3013–3016. https://doi.org/10.1029/1999GL005390

Bear-Crozier, A. N., Miller, V., Newey, V., Horspool, N., & Weber, R. (2016). Probabilistic Volcanic Ash Hazard Analysis (PVAHA) I: Development of the VAPAH tool for emulating multi-scale volcanic ash fall analysis. *Journal of Applied Volcanology*, *5*(1). https://doi.org/10.1186/s13617-016-0043-4

Budgen, D. (1994). *Software design*. Addison-Wesley.

Capra, L., Norini, G., Groppelli, G., Macías, J. L., & Arce, J. L. (2008). Volcanic hazard zonation of the Nevado de Toluca volcano, México. *Journal of Volcanology and Geothermal Research*, *176*(4), 469–484. https://doi.org/10.1016/j.jvolgeores.2008.04.016

Cayirci, E. (2013). Modeling and simulation as a cloud service: A survey. *2013 Winter Simulations Conference (WSC)*. http://dx.doi.org/10.1109/wsc.2013.6721436

De la Cruz-Reyna, S., & Tilling, R. I. (2008). Scientific and public responses to the ongoing volcanic crisis at Popocatépetl Volcano, Mexico: Importance of an effective hazards warning system. Journal of Volcanology and Geothermal Research, 170, 121–134. https://doi.org/10.1016/j.jvolgeores.2007.09.002

Delgado-Granados, H., Cárdenas Gonzales, L., & Piedad Sánchez, N. (2001). Sulfur dioxide emissions from Popocatepetl volcano (México): case study of a high-emission rate passively degassing erupting volcano. Journal of Volcanology and Geothermal Research, 108, 107–120.

Denlinger, R. P., Pavolonis, M., & Sieglaff, J. (2012). A robust method to forecast volcanic ash clouds. *Journal of Geophysical Research: Atmospheres*, *117*(D13), n/a-n/a. https://doi.org/10.1029/2012jd017732

ECMWF. (2021). *ERA5: Data documentation - Copernicus Knowledge Base - ECMWF Confluence Wiki*. ERA5: Data Documentation . https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation

Global Volcanism Program, 2013. Volcanoes of the World, v. 4.10.1. Venzke, E (ed.). Smithsonian Institution. Downloaded 18 Aug 2021. https://doi.org/10.5479/si.GVP.VOTW4-2013

Hampton, S. J., Cole, J. W., Wilson, G., Wilson, T. M., & Broom, S. (2015). Volcanic ashfall accumulation and loading on gutters and pitched roofs from laboratory empirical experiments: Implications for risk assessment. *Journal of Volcanology and Geothermal Research*, *304*, 237–252. https://doi.org/10.1016/j.jvolgeores.2015.08.012

IBM Cloud Education. (2021). *Containerization*. IBM Cloud. https://www.ibm.com/cloud/learn/containerization

Kavis, M. J. (2014). *Architecting the cloud: Design decisions for cloud computing service models (saas, paas, and iaas)*. John Wiley & Sons.

Kelton, W. D., Sadowski, R. P., Swets, N. B., & Zupick, N. B. (2014). *Simulation with arena*. College le Overruns.

Keramat, M., & Kielbasa, R. (1997). Latin hypercube sampling monte carlo estimation of average quality index for integrated circuits. In *Analog Design Issues in Digital VLSI Circuits and Systems* (pp. 131–142). Springer US. http://dx.doi.org/10.1007/978-1-4615-6101-9_11

Leonard, G. S., Stewart, C., Wilson, T. M., Procter, J. N., Scott, B. J., Keys, H. J., Jolly, G. E., Wardman, J. B., Cronin, S. J., & McBride, S. K. (2014). Integrating multidisciplinary science, modelling and impact data into evolving, syn-event volcanic hazard mapping and communication: A case study from the 2012 Tongariro eruption crisis, New Zealand. *Journal of Volcanology and Geothermal Research*, *286*, 208–232. https://doi.org/10.1016/j.jvolgeores.2014.08.018

Loughlin, S. C., Sparks, R. S. J., Sparks, S., Brown, S. K., Jenkins, S. F., & Vye-Brown, C. (2015). Global volcanic hazards and risk. Cambridge University Press.

Macedonio, G., Costa, A., & Longo, A. (2005). A computer model for volcanic ash fallout and assessment of subsequent hazards. *Computers & Geosciences*, *31*(7), 837–845. https://doi.org/10.1016/j.cageo.2005.01.013

Madankan, R., Pouget, S., Singla, P., Bursik, M., Dehn, J., Jones, M., Patra, A., Pavolonis, M., Pitman, E. B., Singh, T., & Webley, P. (2014). Computation of probabilistic hazard maps and source parameter estimation for volcanic ash transport and dispersion. *Journal of Computational Physics*, *271*, 39–59. https://doi.org/10.1016/j.jcp.2013.11.032

Mastin, L. G. (2014). Testing the accuracy of a 1-D volcanic plume model in estimating mass eruption rate. *Journal of Geophysical Research: Atmospheres*, 119, 2474– 2495. https://doi.org/10.1002/2013JD020604

Matala, A. (2008). Sample Size Requirement for Monte Carlo simulations using Latin Hypercube Sampling. Independent Research Projects in Applied Mathematics. Retrieved from http://www.sal.tkk.fi/vanhat_sivut/Opinnot/Mat-2.4108/pdf-files/emat08.pdf

Menčík, J. (2016). Latin hypercube sampling. In *Concise Reliability for Engineers*. InTech. http://dx.doi.org/10.5772/62370

Mouat, A. (2015). *Using docker: Developing and deploying software with containers*. "O'Reilly Media, Inc."

Organización Panamericana de la Salud. (2002). Los Volcanes y la protección de la salud. Organización Panamericana de la Salud.

Pachghare, V. (2016). Microservices Architecture for Cloud Computing. *Journal of Information Technology and Sciences* , *2*(1). https://doi.org/10.3403/30268907u

Poulidis, A. P., Phillips, J. C., Renfrew, I. A., Barclay, J., Hogg, A., Jenkins, S. F., Robertson, R., & Pyle, D. M. (2018). Meteorological controls on local and regional volcanic ash dispersal. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-24651-1

Siegfried, R., & Van Den Berg, T. (2019). *Modelling and Simulation as a Service, Volume 1: MSaaS Technical Reference Architecture (La modelisation et simulation en tant que service, Volume 1: Architecture de reference technique).* NATO SCIENCE AND TECHNOLOGY ORGANIZATION NEUILLY-SUR-SEINE (FRANCE) NEUILLY-SUR-SEINE France.

Siebe, C., & Macías, J. L. (2003). Volcanic hazards in the Mexico City metropolitan area from eruptions at Popocatépetl, Nevado de Toluca, and Jocotitlán stratovolcanoes and monogenetic scoria in the Sierra Chichinautzin Volcanic Field. GSA Penrose Conference, 77. Boulder, Colorado: Geological Society of America.

Vázquez, R., Macías, J. L., & Arce, J. L. (2021). Integrated hazards maps of the Tacaná Volcanic complex, Mexico-Guatemala: Ashfall, block-and-ash flows, and lahars. *Journal of South American Earth Sciences*, *107*, 103146. https://doi.org/10.1016/j.jsames.2020.103146

Wilson, T. M., Stewart, C., Sword-Daniels, V., Leonard, G. S., Johnston, D. M., Cole, J. W., Wardman, J., Wilson, G., & Barnard, S. T. (2012). Volcanic ash impacts on critical infrastructure. *Physics and Chemistry of the Earth, Parts A/B/C*, *45–46*, 5–23. https://doi.org/10.1016/j.pce.2011.06.006

Zheng, L., & Wei, B. (2018). Application of microservice architecture in cloud environment project development. *MATEC Web of Conferences*, *189*, 03023. https://doi.org/10.1051/matecconf/201818903023

APPENDIX 1. Popocatepetl volcano historical events

The values given in black font correspond to primary information. Meanwhile, the red ones correspond to approximations obtained in the present project.

Volcano	Popocatepetl							
Summit Elevation	5393,00							
Population within 30 km	6,34E+05							
Activity record								
Year	**Month**	**Day**	**VEI**	**Mass (kg)**	**Volume (m3)**	**Min Column height (km)**	**Max Column height (km)**	**Mean Column height (km)**
-7150	**	**	4	5,00E+11	5,00E+08	10,00	20,00	15,00
-6250	**	**	**	Volume is not specified	VEI is not specified	VEI is not specified	VEI is not specified	VEI is not specified
-5150	**	**	4	5,00E+11	5,00E+08	10,00	20,00	15,00
-3700	**	**	5	5,00E+12	5,00E+09	12,00	25,00	18,50
-2370	**	**	**	Volume is not specified	VEI is not specified	VEI is not specified	VEI is not specified	VEI is not specified
-1890	**	**	**	Volume is not specified	VEI is not specified	VEI is not specified	VEI is not specified	VEI is not specified
200	**	**	5	5,00E+12	3,20E+09	12,00	25,00	18,50
250	**	**	**	Volume is not specified	VEI is not specified	VEI is not specified	VEI is not specified	VEI is not specified

823	3	1	4	1,58E+12	1,42E+09	10,00	20,00	15,00
1345	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1354	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1363	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1488	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1504	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1509	**	**	**	Volume is not specified	VEI is not specified	VEI is not specified	VEI is not specified	VEI is not specified
1512	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1518	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1519	9	**	3	5,00E+10	5,00E+07	3,00	15,00	9,00
1528	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1530	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1539	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1542	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00

1548	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1571	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1580	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1590	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1592	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1642	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1663	10	13	3	5,00E+10	5,00E+07	3,00	15,00	9,00
1666	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1697	10	20	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1720	**	**	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1802	**	**	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1827	**	**	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1834	**	**	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1852	**	**	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1919	2	19	1	5,00E+08	5,00E+05	0,50	2,00	1,25

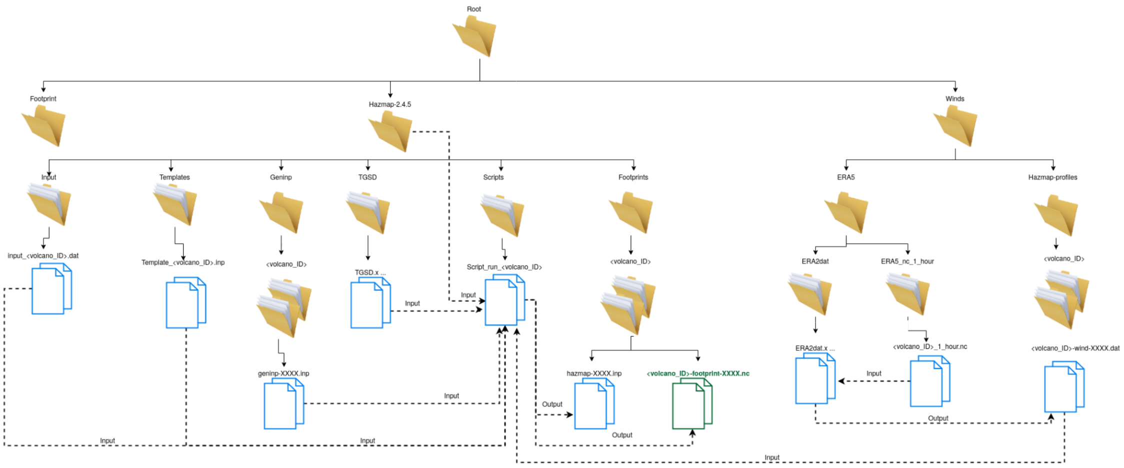
1923	11	27	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1925	**	**	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1933	1	23	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1942	**	**	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1947	1	**	1	5,00E+08	5,00E+05	0,50	2,00	1,25
1994	12	21	2	5,00E+09	5,00E+06	1,00	5,00	3,00
1996	3	5	3	5,00E+10	5,00E+07	3,00	15,00	9,00
2004	5	26	2	5,00E+09	5,00E+06	1,00	5,00	3,00
2005	1	9	2	5,00E+09	5,00E+06	1,00	5,00	3,00

APPENDIX 2. Wind profile example.

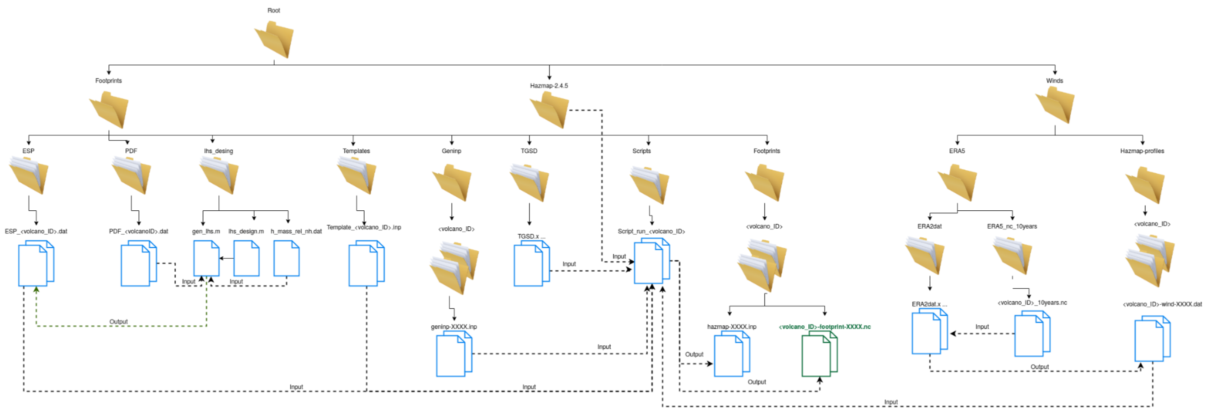
```
251 # number of pressure levels
    0. #pressure level altitude
  100.
  200.
  300.
  400.
  500.
  600.
  700.
  800.
  900.
 1000.
 1100.
 1200.
 1300.
 1400.
 1500.


...
2002 12 29   1  -0.23  -1.02 #Year Month Day Pressure level XWindSpeed YWindSpeed
2002 12 29   2  -0.23  -1.02
2002 12 29   3  -0.23  -1.02
2002 12 29   4  -0.23  -1.02
2002 12 29   5  -0.23  -1.02
2002 12 29   6  -0.43  -1.16
2002 12 29   7  -0.28  -1.04
2002 12 29   8  -0.47  -1.21
2002 12 29   9  -0.33  -1.08
2002 12 29  10  -0.51  -1.24
2002 12 29  11  -0.38  -1.13
2002 12 29  12  -0.57  -1.27
2002 12 29  13  -0.46  -1.17
2002 12 29  14  -0.35  -1.07
2002 12 29  15  -0.42  -1.22
2002 12 29  16  -0.35  -1.12
2002 12 29  17  -0.51  -1.23
2002 12 29  18  -0.41  -1.14
2002 12 29  19  -0.30  -1.05
2002 12 29  20  -0.59  -1.18
2002 12 29  21  -0.36  -1.05
2002 12 29  22  -0.88  -1.41
...
```

APPENDIX 3. Files Dependencies.

Single simulation file structure



Probabilistic approach file structure



APPENDIX 4. Hazmap.inp example.

```
0            # Mode (0=deposit, 1=hazard map, 2=barycenters)
 2           # Grid_type (0=uniform, 1=from file)
2            # Settling velocity model
1           # Settling velocity is a function of Z
0            # Write list of cells coord. (file grid.out)
0            # Flag: write particles spectra
0            # Flag: write settling velocity profile
 3           # Format of the output file
   300    300 # Number of grid cells (used if grid_type=0)
 0.20000E-01  0.20000E-01  # Dimensions of the cells (used if grid_type=0)
    -98.6      19.0        # Center of the grid (used if grid_type=0)
  1          # Number of thresholds (used if mode=1)
  100.00       # Loading threshold (used if mode=1)   1
  10000.       # Diffusion coefficient
    27        # Number of particles types
 0.32000E-01  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   1
 0.22600E-01  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   2
 0.16000E-01  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   3
 0.11300E-01  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   4
0.80000E-02  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   5
 0.56600E-02  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   6
 0.40000E-02  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   7
 0.28300E-02  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   8
 0.20000E-02  0.12000E+04  0.90000E+00 # Diam,dens,psi particle type   9
 0.14100E-02  0.12929E+04  0.90000E+00 # Diam,dens,psi particle type  10
 0.10000E-02  0.13857E+04  0.90000E+00 # Diam,dens,psi particle type  11
 0.70700E-03  0.14786E+04  0.90000E+00 # Diam,dens,psi particle type  12
 0.50000E-03  0.15714E+04  0.90000E+00 # Diam,dens,psi particle type  13
 0.35400E-03  0.16643E+04  0.90000E+00 # Diam,dens,psi particle type  14
 0.25000E-03  0.17571E+04  0.90000E+00 # Diam,dens,psi particle type  15
 0.17700E-03  0.18500E+04  0.90000E+00 # Diam,dens,psi particle type  16
 0.12500E-03  0.19429E+04  0.90000E+00 # Diam,dens,psi particle type  17
 0.88400E-04  0.20357E+04  0.90000E+00 # Diam,dens,psi particle type  18
 0.62500E-04  0.21286E+04  0.90000E+00 # Diam,dens,psi particle type  19
 0.44200E-04  0.22214E+04  0.90000E+00 # Diam,dens,psi particle type  20
 0.31300E-04  0.23143E+04  0.90000E+00 # Diam,dens,psi particle type  21
 0.22100E-04  0.24071E+04  0.90000E+00 # Diam,dens,psi particle type  22
 0.15600E-04  0.25000E+04  0.90000E+00 # Diam,dens,psi particle type  23
 0.11000E-04  0.25000E+04  0.90000E+00 # Diam,dens,psi particle type  24
 0.78100E-05  0.25000E+04  0.90000E+00 # Diam,dens,psi particle type  25
 0.55200E-05  0.25000E+04  0.90000E+00 # Diam,dens,psi particle type  26
 0.39100E-05  0.25000E+04  0.90000E+00 # Diam,dens,psi particle type  27
   100        # Number of vertical source points
    -98.6      19.0  5432.2   # Coordinates (x,y,z) of source    1
 0.18545E+09  # Total mass of source     1  (  0.333%)
 0.10821E-02  # Mass fraction particle type   1
 0.21270E-02  # Mass fraction particle type   2
```

```
0.51157E-02  # Mass fraction particle type   3
0.10679E-01  # Mass fraction particle type   4

0.30485E-01  # Mass fraction particle type   6
0.41696E-01  # Mass fraction particle type   7
0.49681E-01  # Mass fraction particle type   8
0.51689E-01  # Mass fraction particle type   9
0.47596E-01  # Mass fraction particle type  10
0.40258E-01  # Mass fraction particle type  11
0.34343E-01  # Mass fraction particle type  12
0.33915E-01  # Mass fraction particle type  13
0.41042E-01  # Mass fraction particle type  14
0.54595E-01  # Mass fraction particle type  15
0.70722E-01  # Mass fraction particle type  16
0.83972E-01  # Mass fraction particle type  17
0.89565E-01  # Mass fraction particle type  18
0.85307E-01  # Mass fraction particle type  19
0.72319E-01  # Mass fraction particle type  20
0.54706E-01  # Mass fraction particle type  21
0.36753E-01  # Mass fraction particle type  22
0.22070E-01  # Mass fraction particle type  23
0.11702E-01  # Mass fraction particle type  24

... The file contains more than 2000 lines
```