STRUCTURED EXPERT JUDGEMENT IN VOLCANOLOGY: PAST EXPERIENCES AND NEW TOOLS

A. Tadini(1), A. Bevilacqua(1), M. De' Michieli Vitturi(1), A. Neri(1), W. Aspinall(2)

(1) Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, Italia
(2) University of Bristol, School of Earth Sciences, UK

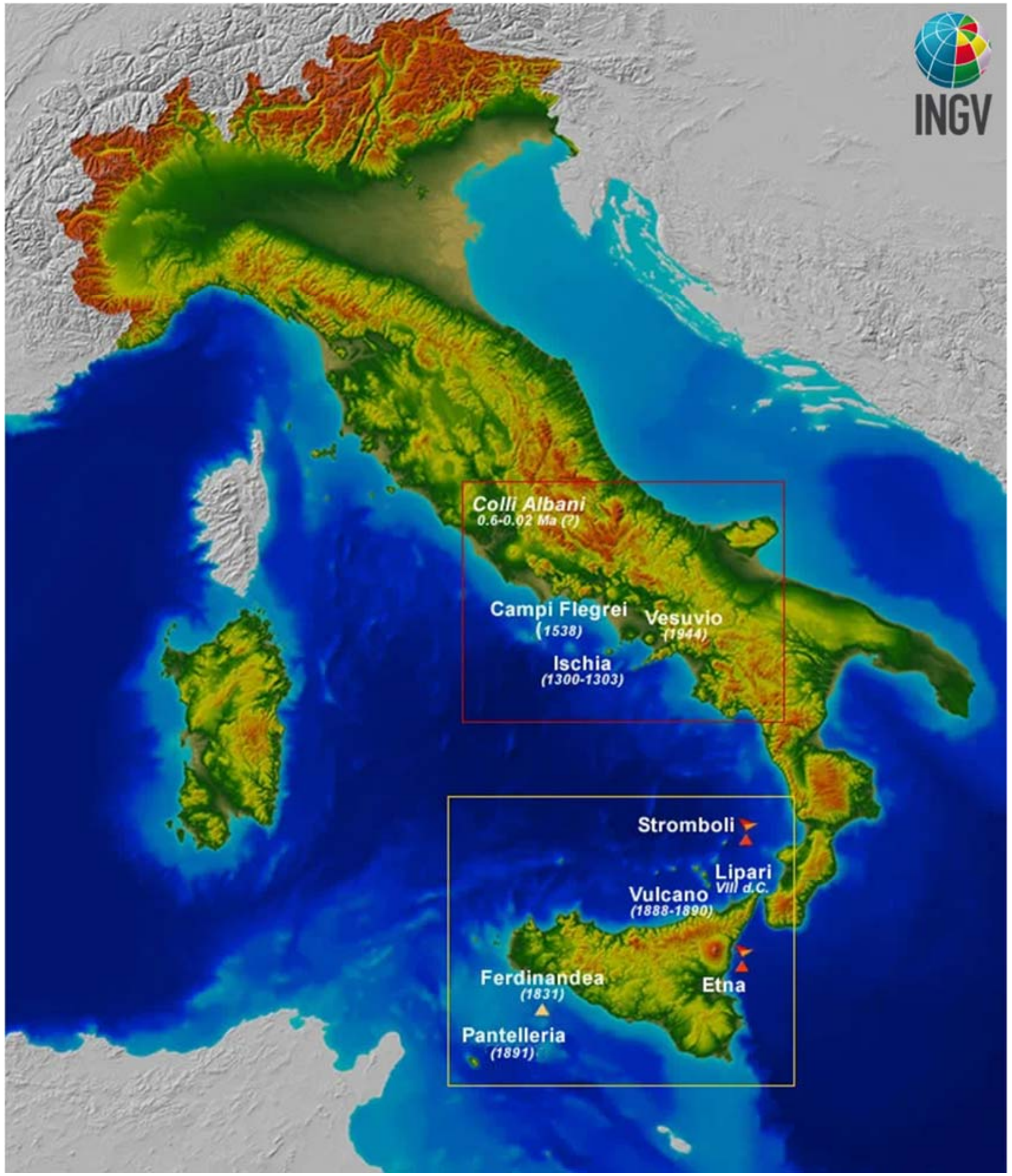
Invited SEJ Seminar, TU Delft, Netherlands, November 22nd 2022

INTRODUCTION

Active volcanoes in Italy:
Etna
Stromboli
Vulcano
Lipari

Campi Flegrei
Somma-Vesuvius
Ischia

Pantelleria
Ferdinandea
Colli Albani



Probability problems in Volcanology

Volcanic eruptions are emissions of gas and magma on the surface of the planet.
Magma is a mixture of molten rocks, crystals, and dissolved gas.

A volcano is a morphological structure created by volcanic eruptions. It can appear as a mountain, a wide depression, or a ground fissure.

Volcanic eruption forecasting is a crucial and challenging problem.

It includes the estimates of:

- Site and extension of the eruptive vent, i.e. the aperture through which the magma erupts.
- Time of the initiation (onset) and duration of the eruption
- Volume (and mass) of erupted material
- Type of the eruption and of the hazard scenarios associated.

Figure. Etna volcano (Italy), 2015.
Corsaro et al., (2017); Foto by G. Famiani.

Volcanic hazard assessment

A volcanic eruption is typically associated to various hazards:

- earthquakes (often shallow and destructive)
- flank collapses (sometimes associated to lateral blasts or tsunamis)
- ballistic projectiles (blocks and scoria ejected up to kilometric distances)
- ash fall (able to reach great distance from the volcano)
- pyroclastic density currents (hot flows/avalanches of fiery material and gas)
- lava flows (of various nature depending on the their viscosity)
- mudflows and debris flows (even centuries after the volcanic eruption)
- toxic gas and aerosols (with prolonged effects on people and animals)

The quantitative assessment of all these phenomena has been tackled by using **probabilistic approaches**.

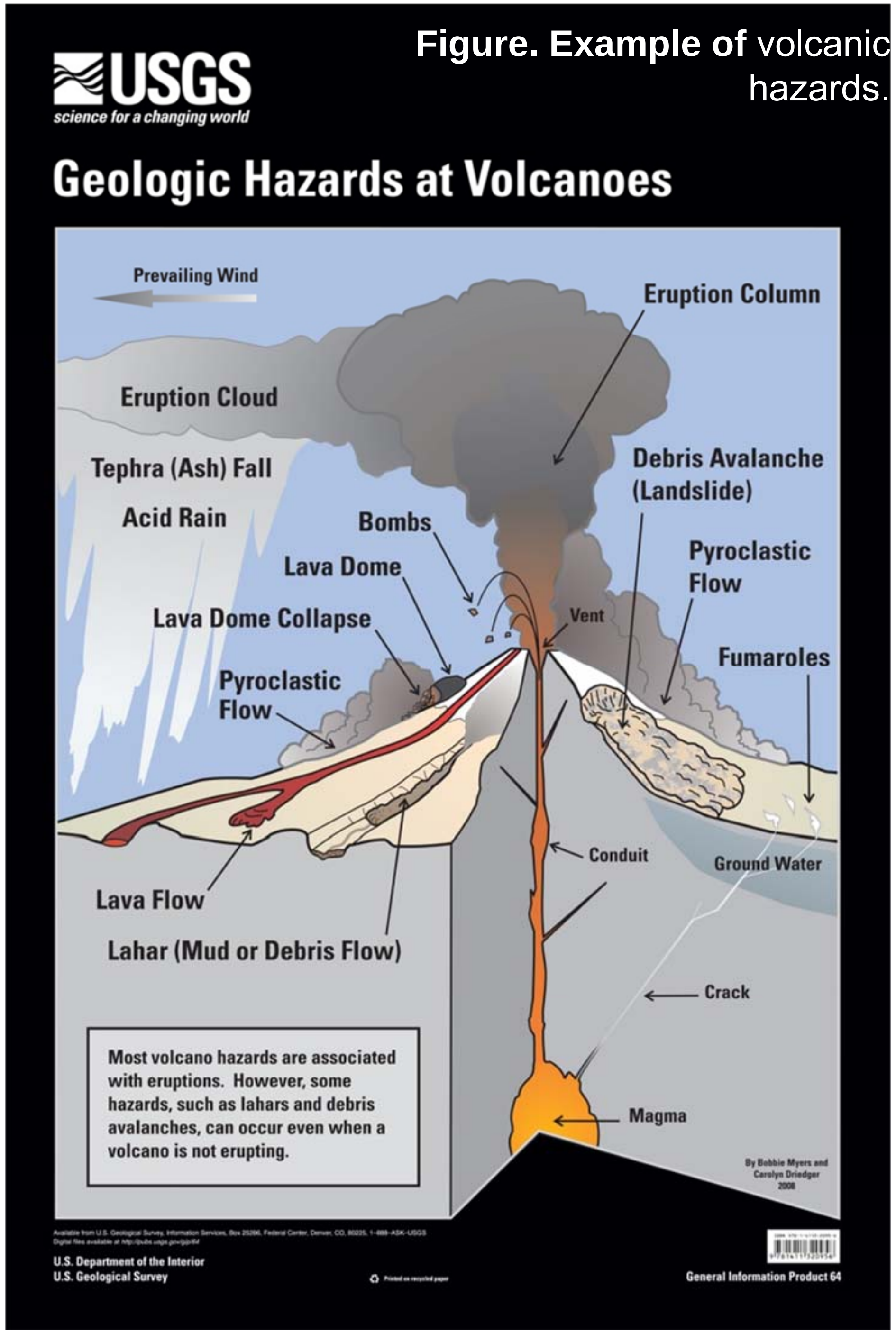


Figure. Example of volcanic hazards.

Mathematical structures for eruption forecasting

Deterministic models are a very important field of study in volcanology, but deterministic prevision is obstacled by:

- **Uncertainty** in reconstructing the deposits of past eruptions and the deep portion of the volcanic system
- **Complexity** in the geophysical processes and their nonlinear nature.

Volcanoes are therefore studied as **random systems** of which we can only get incomplete and uncertain information.

Each of the quantitative variables describing a volcanic eruption is then represented by a **random variable**:

- X the site of vent opening (a point in space, or a bounded region)
- Z the time of the eruption (the onset time or an interval of duration)
- V the volume of the eruption (total volume or a set of different volumes)

The type of the eruption is often described by a **event tree**:
it takes into account a set of possible scenarios, and provides
a model to calculate their probability of occurrence.

The use of event trees enables the decomposition of hard evaluations in a sequence of easier steps.

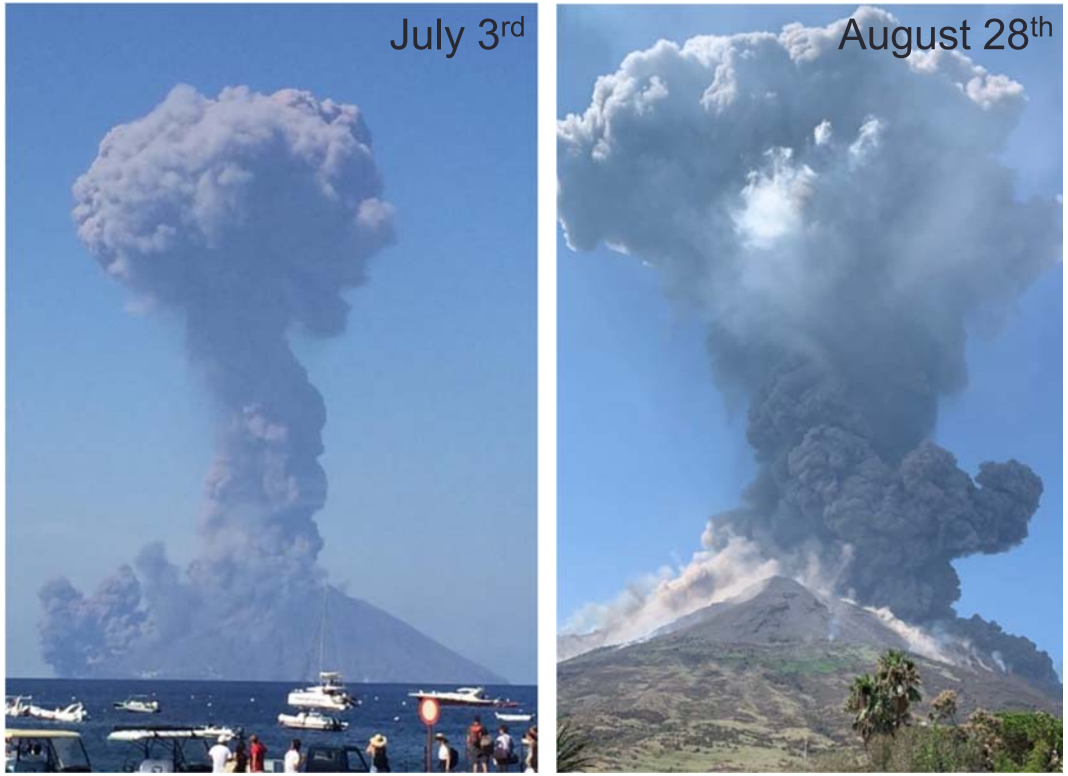


Figure. Stromboli volcano (Italy), 2019. Photo from Panarea (M. Ortenzi; T. Grillo).

Doubly stochastic approaches in volcanology

A volcano can be assumed as a **random system** that must be assessed with **uncertain information**

- Epistemic (imperfect knowledge of the system)
- Aleatoric (intrinsic randomness of the system)

The **forecast** of its behavior cannot be easily constrained by using simple probability models.

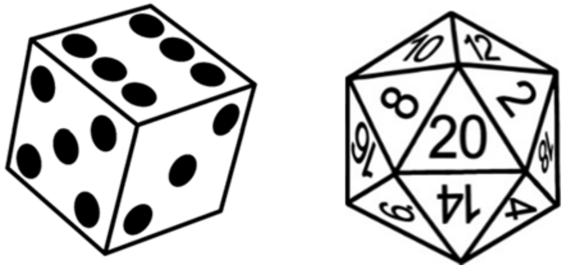
By adopting a **doubly stochastic approach,** the ill-constrained parameters of the probability models are themselves represented as additional random variables.

Therefore, all the probability estimates have their own confidence intervals.
Every probability measure is a random measure.

Example: assume to roll an unknown dice, which could have 6 or 20 faces with equal chances.

The probability P of the event of getting a number N > 3 is 50% in the first case, but it is 85% in the second.

Following a doubly stochastic approach, we might say that
P is 67.5% in mean, with an uncertainty range from 50% to 85%.



Type of dice → epistemic unc.
Number on a face → aleatoric unc.

Expert judgement methods

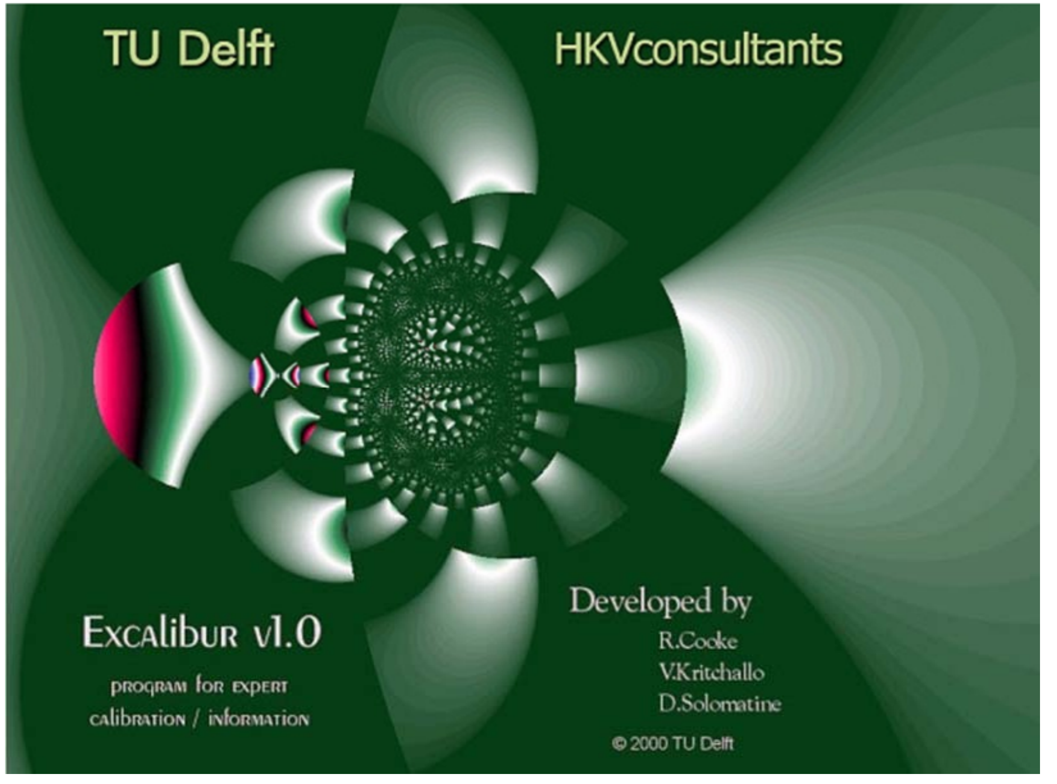
If physical models or statistical procedures are not applicable, the quantification of epistemic uncertainty can be based directly on **expert judgement**.

Expert elicitations (EE) are aimed at producing robust quantitative estimates relying on the views of a pool of experts. For example, according to our experience:

- degrees of belief on alternative conceptual models
- unknown/uncertain material quantities.

Performance-based EE include an empirical step of expert ranking, aimed at measuring their uncertainty quantification capabilities.

Figure. The EXCALIBUR software (http://www.expertsin-uncertainty.net/) was a pioneering tool to assess such performance weights following the so called 'Classical method'.



Performance-based EE

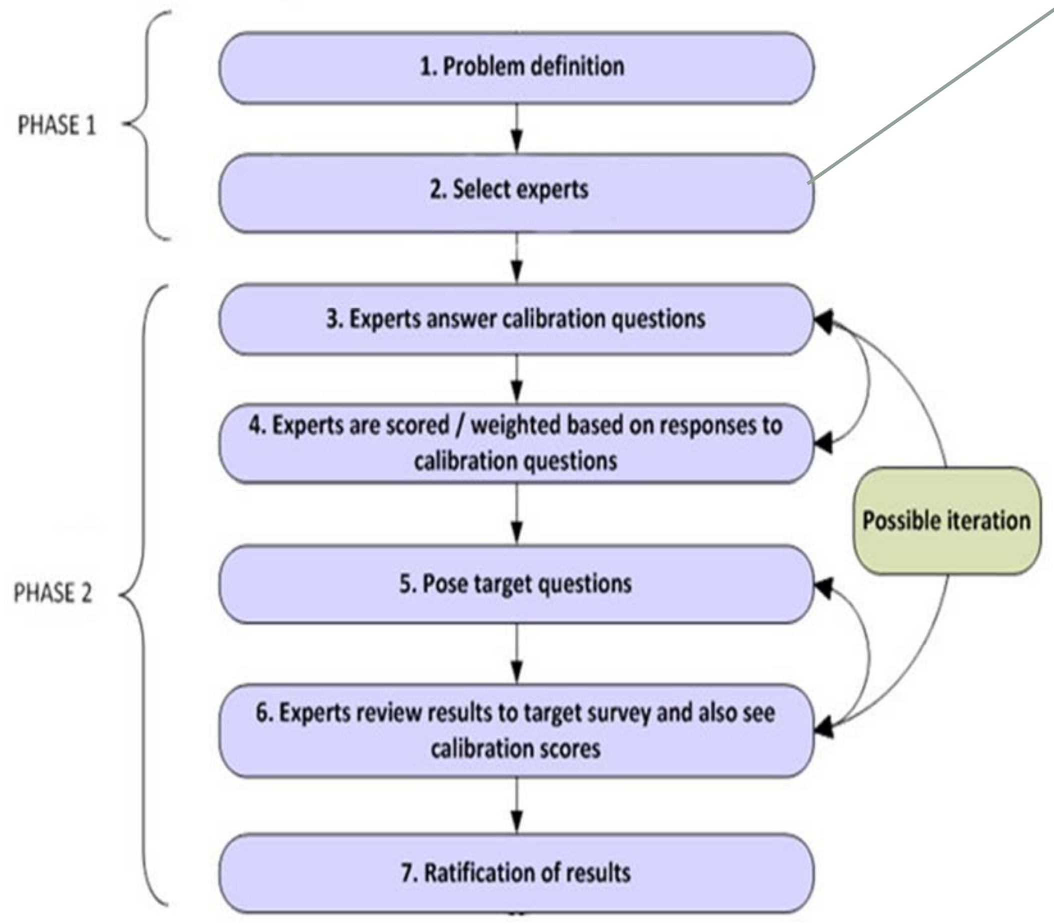


Figure. Phases of a structured elicitation session (courtesy of W. Aspinall).

At least >6 (better >10)
Basic background on at least one aspect of the problem
Worked in the study area (most)

Questionnaires:

- The **seed questions,** with known answers.
- The **target questions,** i.e. the questions of interest**.**

For each question, the experts express their views as the values of the **5th, 50th and 95th percentiles** of a probability distribution representing their uncertainty.

The **seed questions** should be similar as much as possible to the target questions.
They are adopted to score the experts' uncertainty estimation performances.

The diverse answers to the **target questions** are then pooled using the obtained scores, and their combination defines a new virtual expert:
the **global Decision Maker (DM)**.

Uncertainty distributions examples

We generally use CM as a reference and also compare it to EW and other scoring rules.

The Cooke classical method (CM) uses uniform PDF in each inter-percentile range, i.e. maximum entropy distributions. (Cooke, 1991)

The performance score in CM is the product of two values:

Calibration score: likelihood that the true results correspond to the expert distributions. It is a statistical accuracy.

Informativeness score: average relative information w.r.t. a uniform distribution. It penalizes too large uncertainty ranges.

Alternative methods implement different **scoring rules** and **uncertainty distributions**

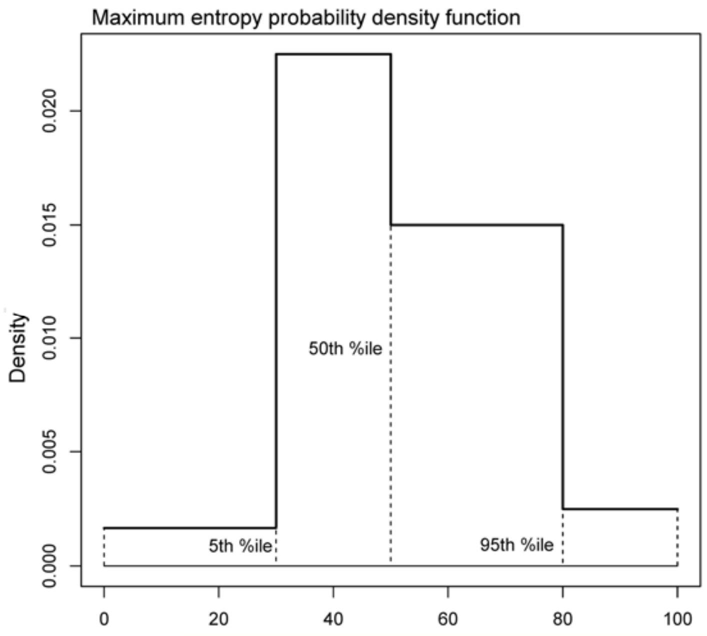


Figure A. Examples of probability distributions.

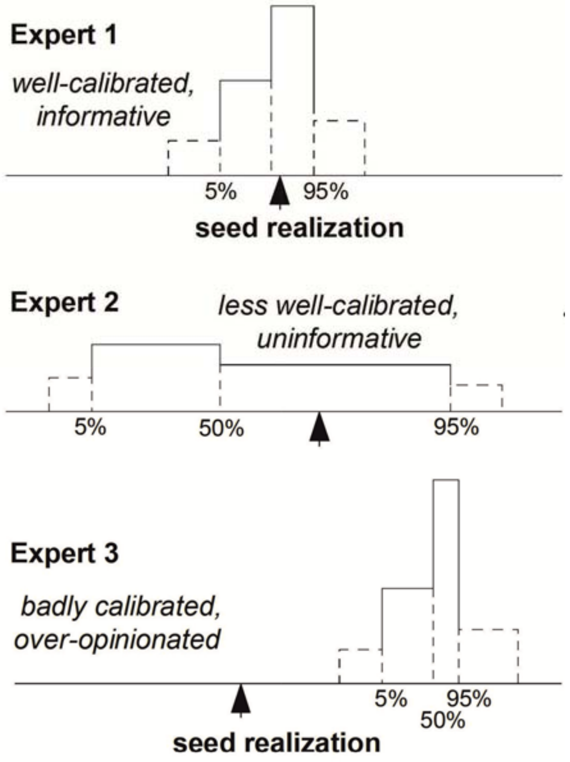


Figure B. Maximum entropy distributions associated to different performances.

Experts' pooling: the Decision Maker (DM)

Experts' answers are **pooled together** according to the weight w_i of each expert e_i

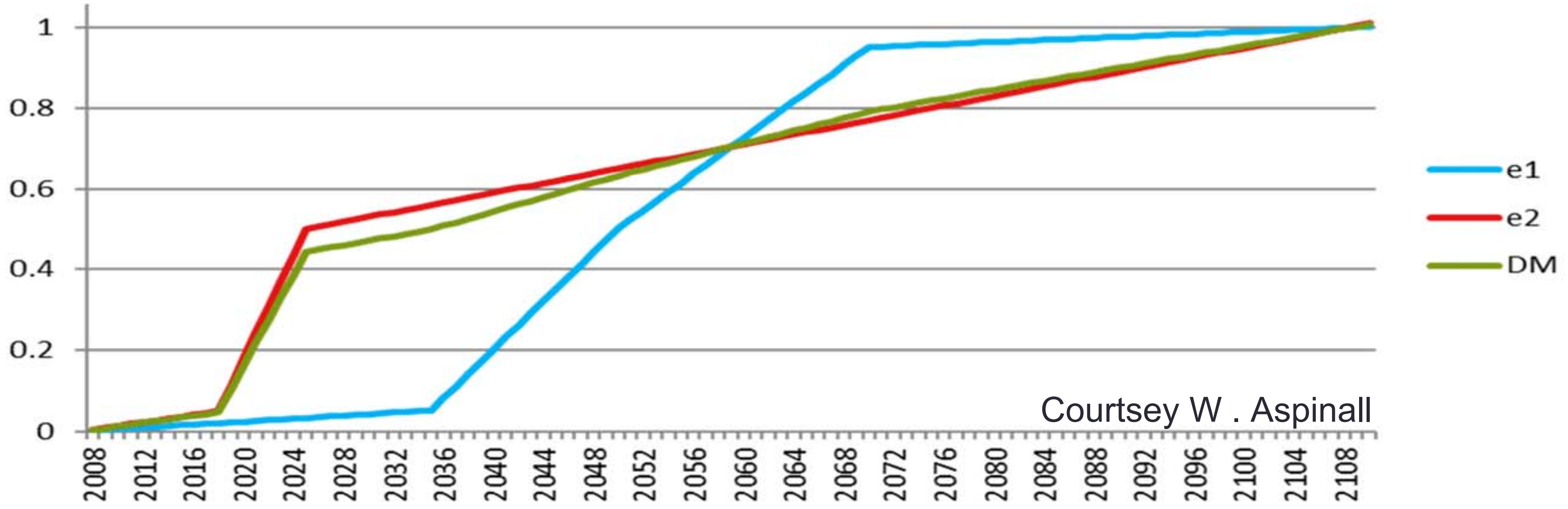
The DM is typically defined by the **weighted linear combination** of the probability distributions of all the experts, i.e. by a probability mixture.

In fact, the **statistical sampling of the DM** is performed by randomly choosing one of the experts, with a chance proportional to their weight, and then by sampling their distribution.

Figure. Example of expert pooling and DM's definition by a probability mixture.

$$w_1 = 0.12$$

$$w_2 = 0.88$$



EXPERT ELICITATIONS IN VOLCANOLOGY: CASE STUDIES

- Already done (7):
  - Somma-Vesuvius 2008
  - Campi Flegrei 2013
  - Somma-Vesuvius 2015
  - Aso 2018
  - Cotopaxi/Guagua Pichincha 2019
  - Piton de la Fournaise 2021
  - Sangay 2021 (in publication)

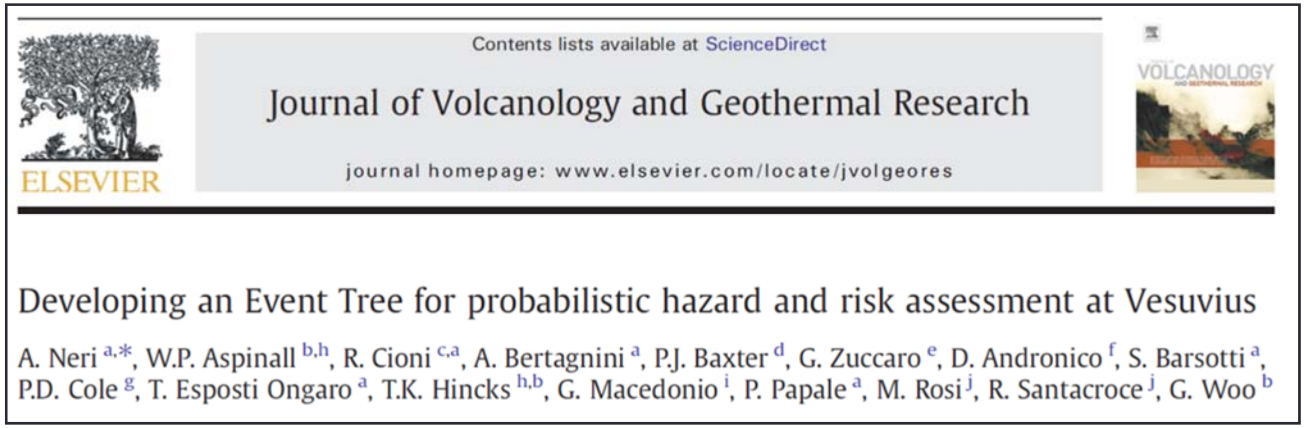
- Planned (3):
  - Stromboli
  - Kulumbo
  - World Health Organization (WHO)

Somma-Vesuvius (2008) - event tree

Performance-based EE were performed to estimate the:

- eruption scenario **recurrence rate** probabilities
- Event Tree **node probability** distributions
- **eruption time-lines** with associated probabilities, elicited for a range of different eruptive scenarios
- examples of **hazard mappings**, derived from the Event Tree.

10 **seed questions** dealing with generic aspect of volcanology and some facts specifically associated with Vesuvius; 25 experts in the ambience of EXPLORIS project.


Contents lists available at ScienceDirect

Journal of Volcanology and Geothermal Research

journal homepage: www.elsevier.com/locate/jvolgeores

ELSEVIER

Developing an Event Tree for probabilistic hazard and risk assessment at Vesuvius

A. Neri [a,*], W.P. Aspinall [b,h], R. Cioni [c,a], A. Bertagnini [a], P.J. Baxter [d], G. Zuccaro [e], D. Andronico [f], S. Barsotti [a], P.D. Cole [g], T. Esposti Ongaro [a], T.K. Hincks [h,b], G. Macedonio [i], P. Papale [a], M. Rosi [j], R. Santacroce [j], G. Woo [b]

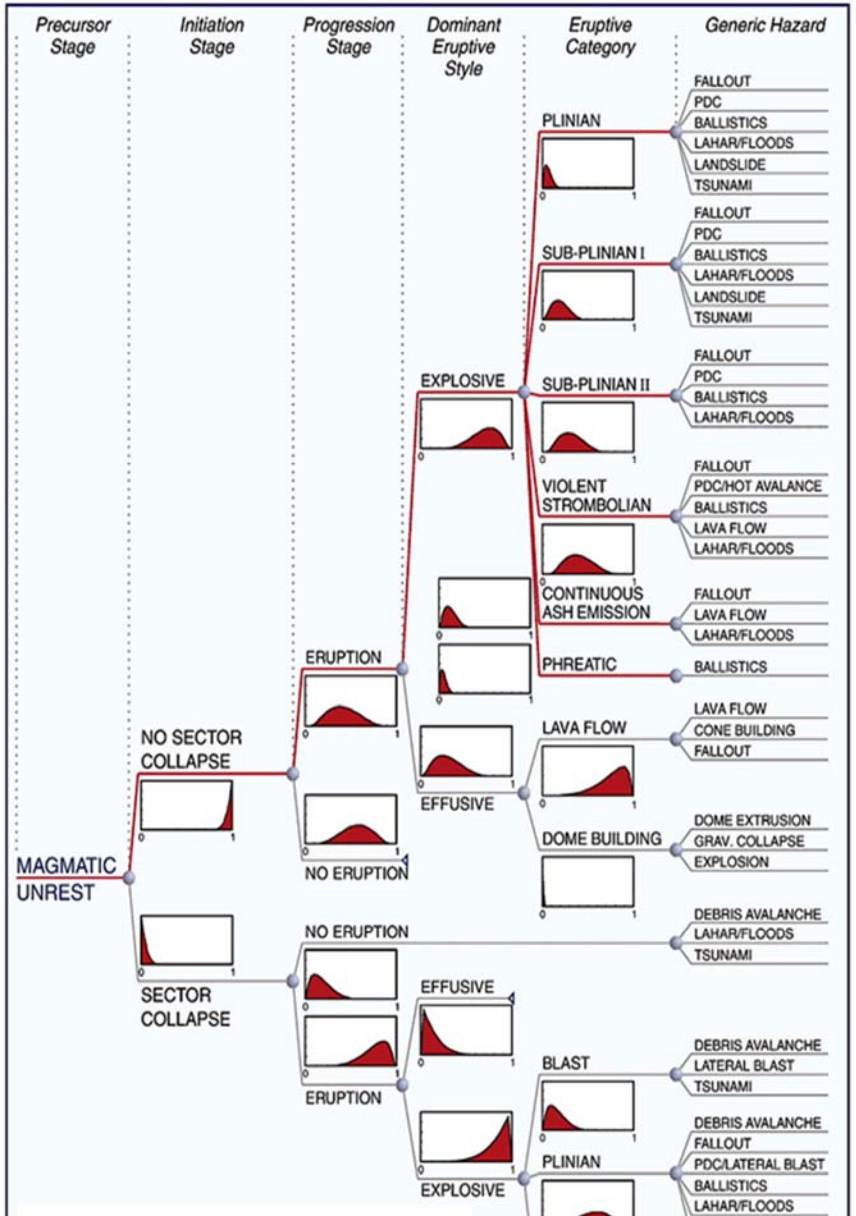


Figure. Event tree of the Somma-Vesuvius case study.

Somma-Vesuvius (2008) - 2

Figure A. Expert group elicitation outcomes for **total duration** of a future volcanic emergency at Vesuvius: cumulative probability function (grey line - referenced to right-hand axis), and PDF of any specific duration (colored distribution, referenced to left-hand axis).

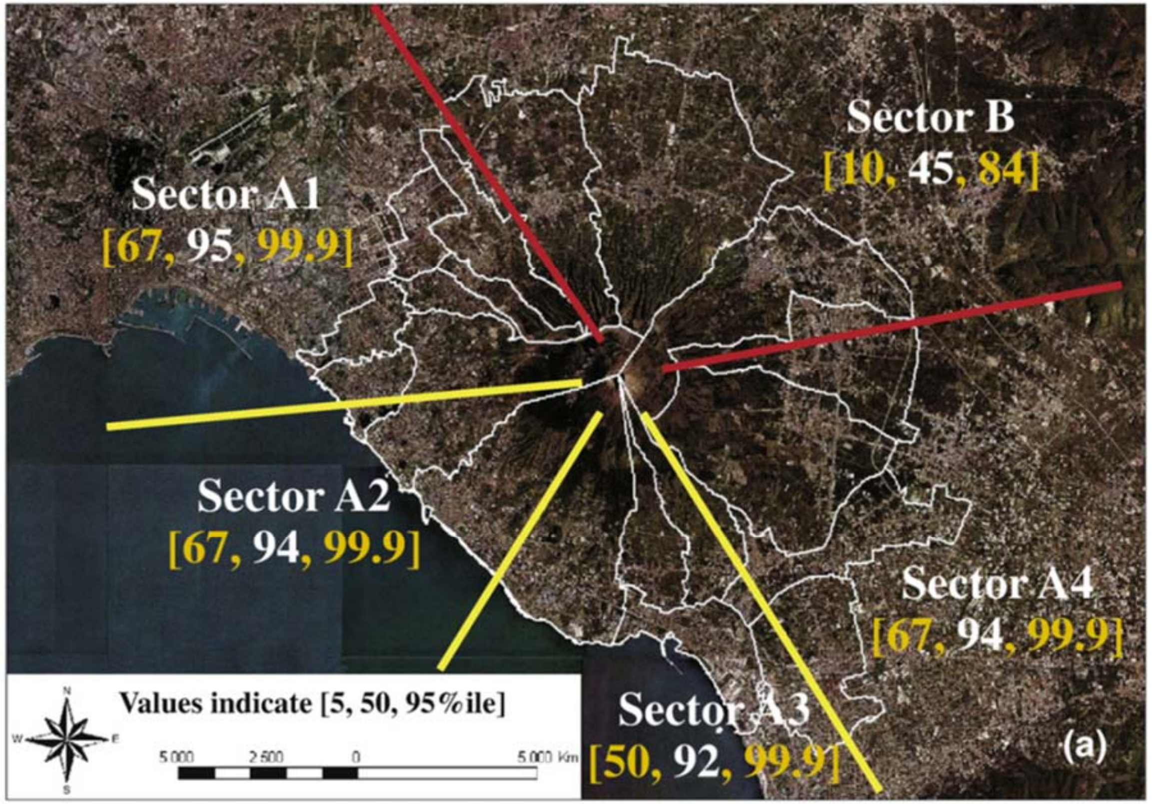
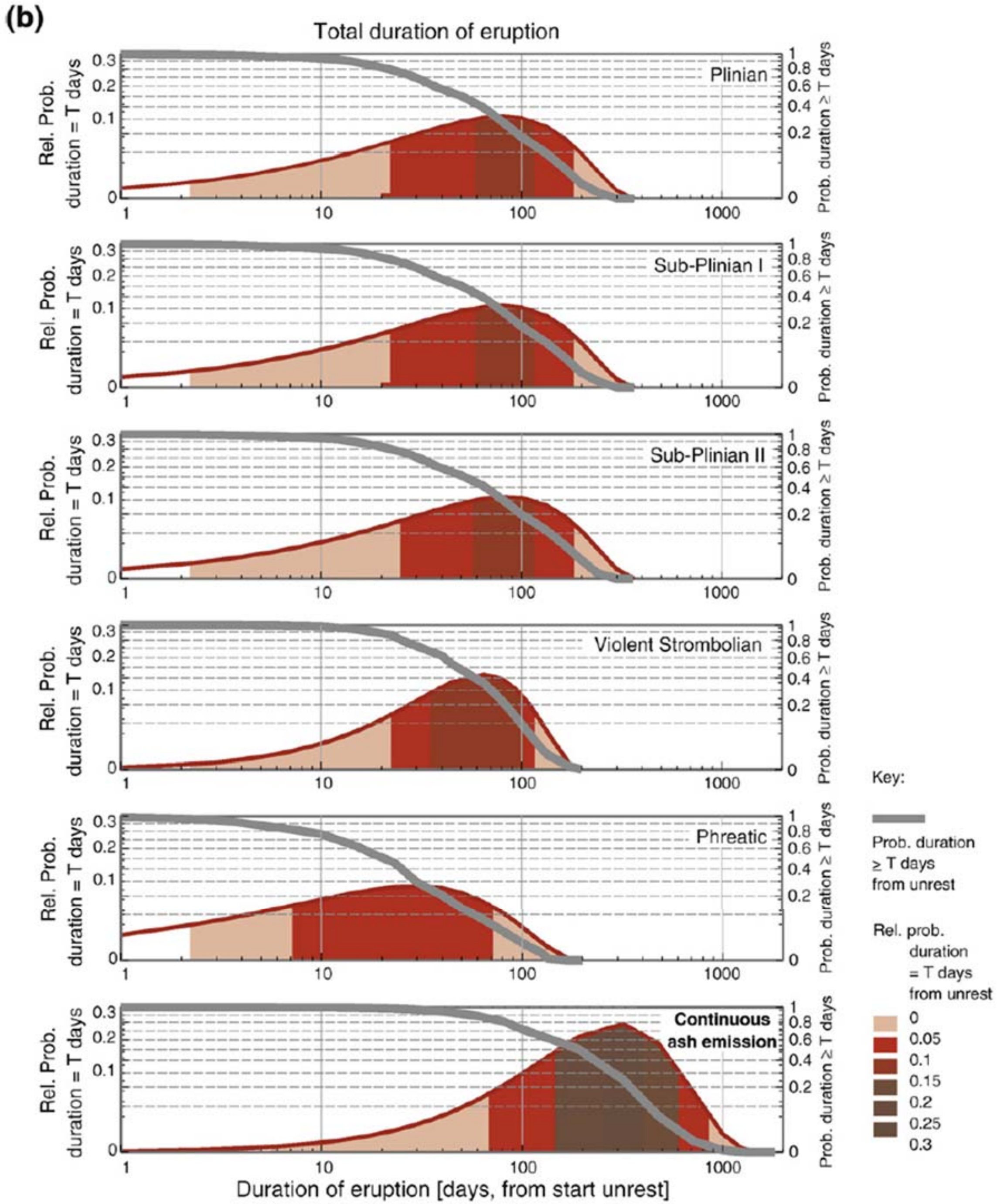


Figure B. First-order effect of Mt. Somma topography in determining areas that might be invaded by **pyroclastic density currents** of a Sub-Plinian I eruption. The values in each sector show elicited probabilities of that sector.



Campi Flegrei (2013) - Vent opening maps

Performance-based EE were performed to estimate the:

- number of past eruptions which do not correspond to presently identified locations ('**lost vents**');
- **linear weights** of the spatial distributions contributing to the vent opening map definition.

Table. The seven spatial distributions.

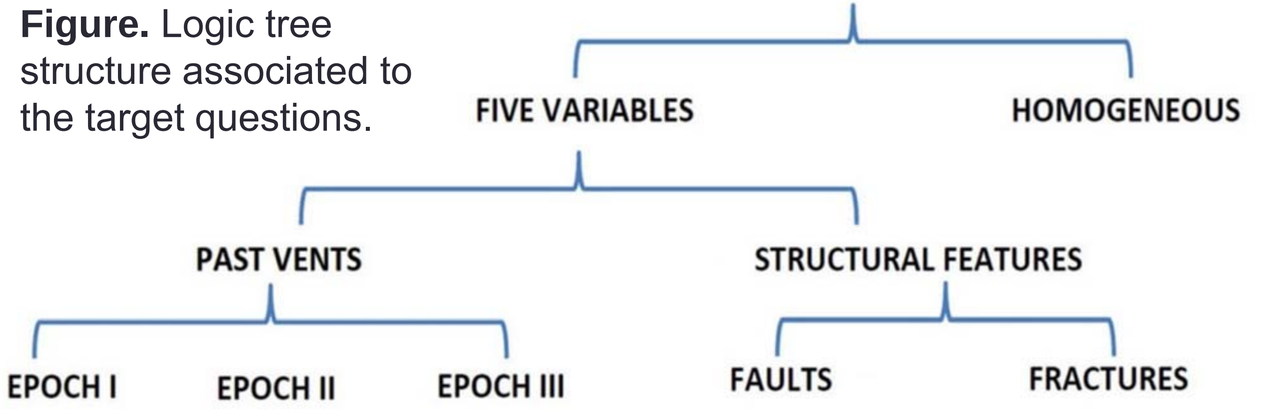
Vents Epoch I	*Vents Epoch II*	*Vents Epoch III*	*Lost vents*	*Faults*	*Fractures*	*Homog. map*

16 **seed questions** were about particular aspects of Neapolitan volcanism, other Italian volcanoes, and about explosive volcanism in general.

Several elicitation sessions, involving **8 experts** with different volcanological backgrounds were carried out through **meetings** and also email consultations.

A **hierarchical logic tree** was defined to simplify the definition of the linear weights.

Figure. Logic tree structure associated to the target questions.



Logic trees generalize event trees: the branches are not necessarily representing events, but logic choices.

Campi Flegrei (2013) - 2

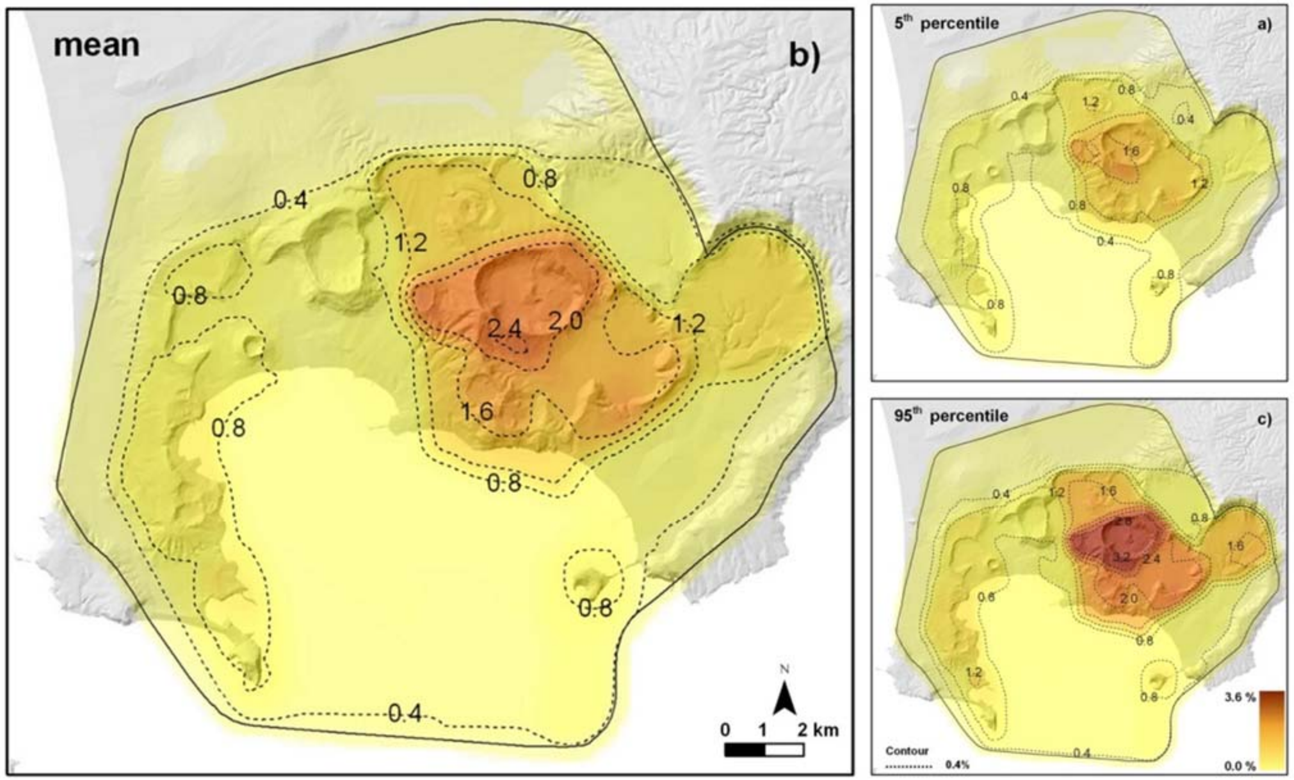


Figure B. Probability density maps of new vent opening location. Contours and colours indicate the percentage probability of vent opening per km^2 conditional on the occurrence of an eruption.

Results show a high probability region in the **central-eastern portion** of the caldera.

Quantified **uncertainty estimates** are indicative of spreads ±30% on the local mean value.

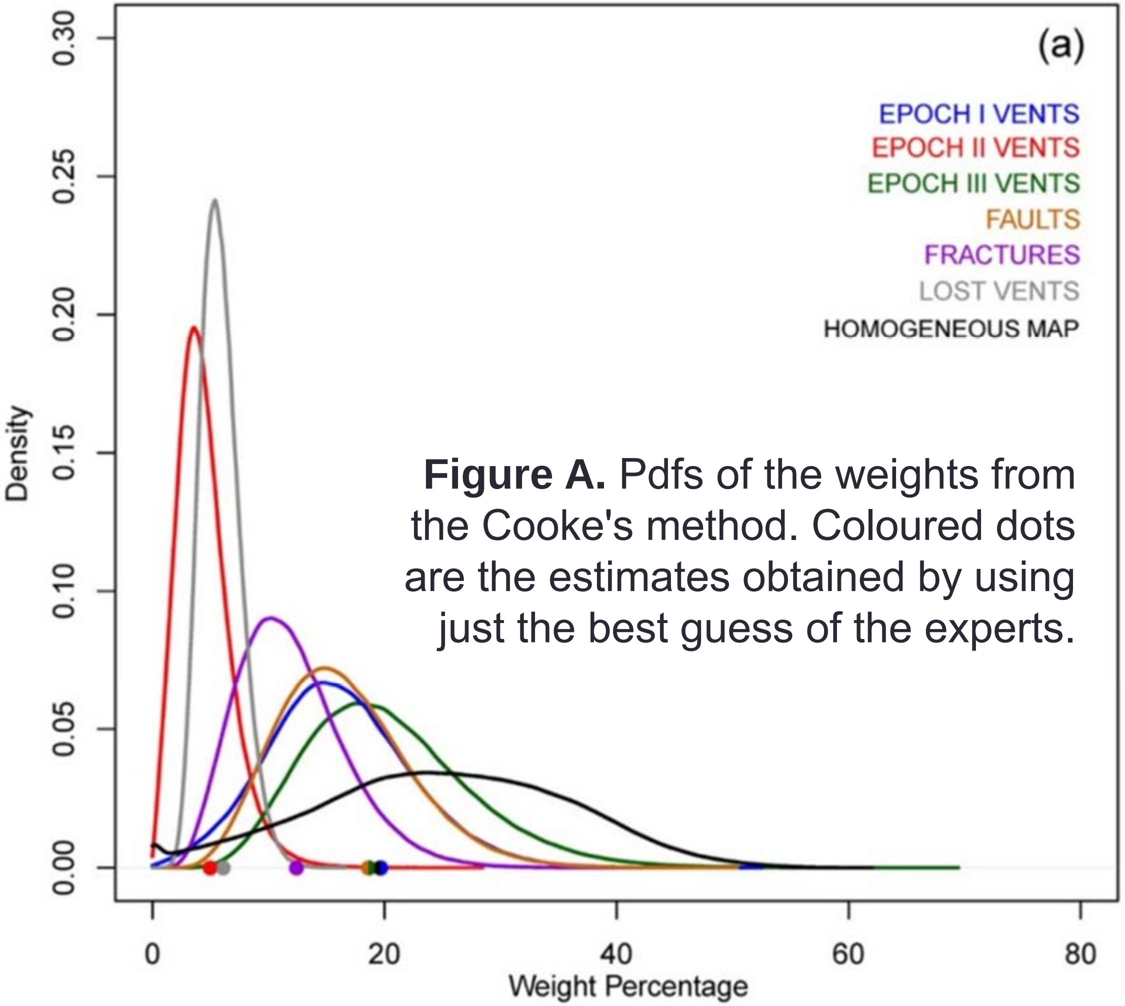


Figure A. Pdfs of the weights from the Cooke's method. Coloured dots are the estimates obtained by using just the best guess of the experts.


AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE
10.1002/2014JB011775

This article is a companion to Neri et al. [2015] doi:10.1002/2014JB011776.

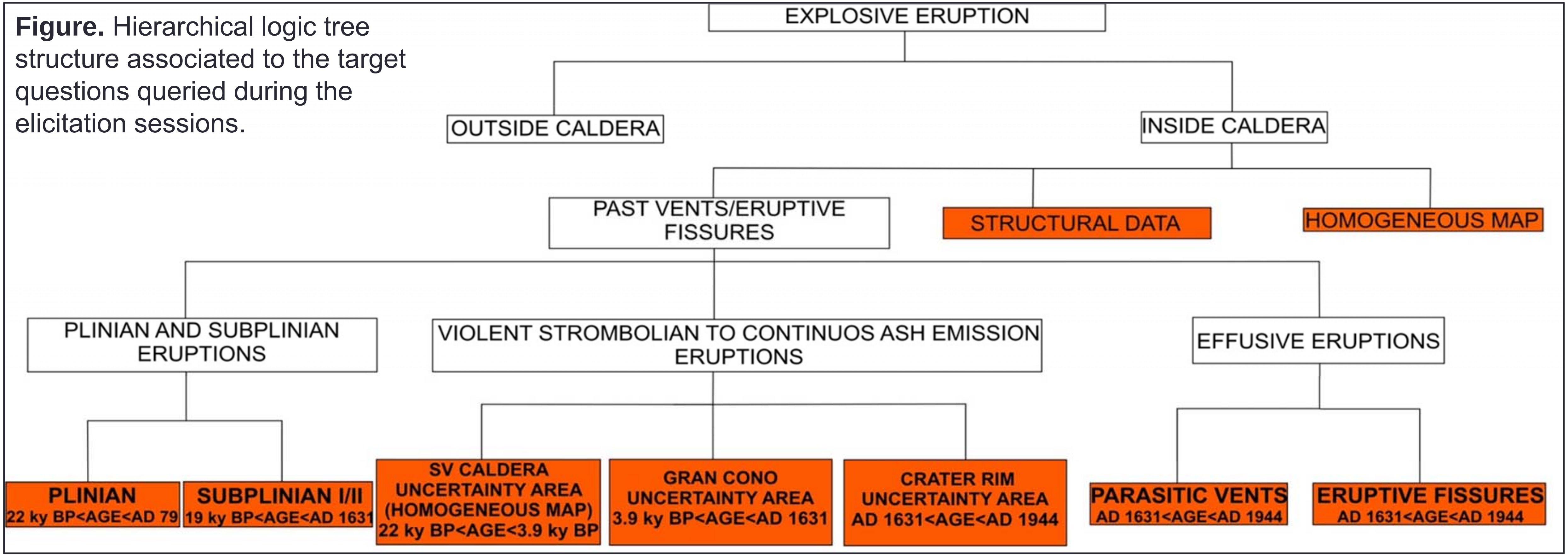
Quantifying volcanic hazard at Campi Flegrei caldera (Italy) with uncertainty assessment: 1. Vent opening maps

Andrea Bevilacqua[1,2], Roberto Isaia[3], Augusto Neri[1], Stefano Vitale[4], Willy P. Aspinall[5,6], Marina Bisson[1], Franco Flandoli[7], Peter J. Baxter[8], Antonella Bertagnini[1], Tomaso Esposti Ongaro[1], Enrico Iannuzzi[3], Marco Pistolesi[9,10], and Mauro Rosi[10,11]

Somma-Vesuvius (2015) - Vent Opening maps

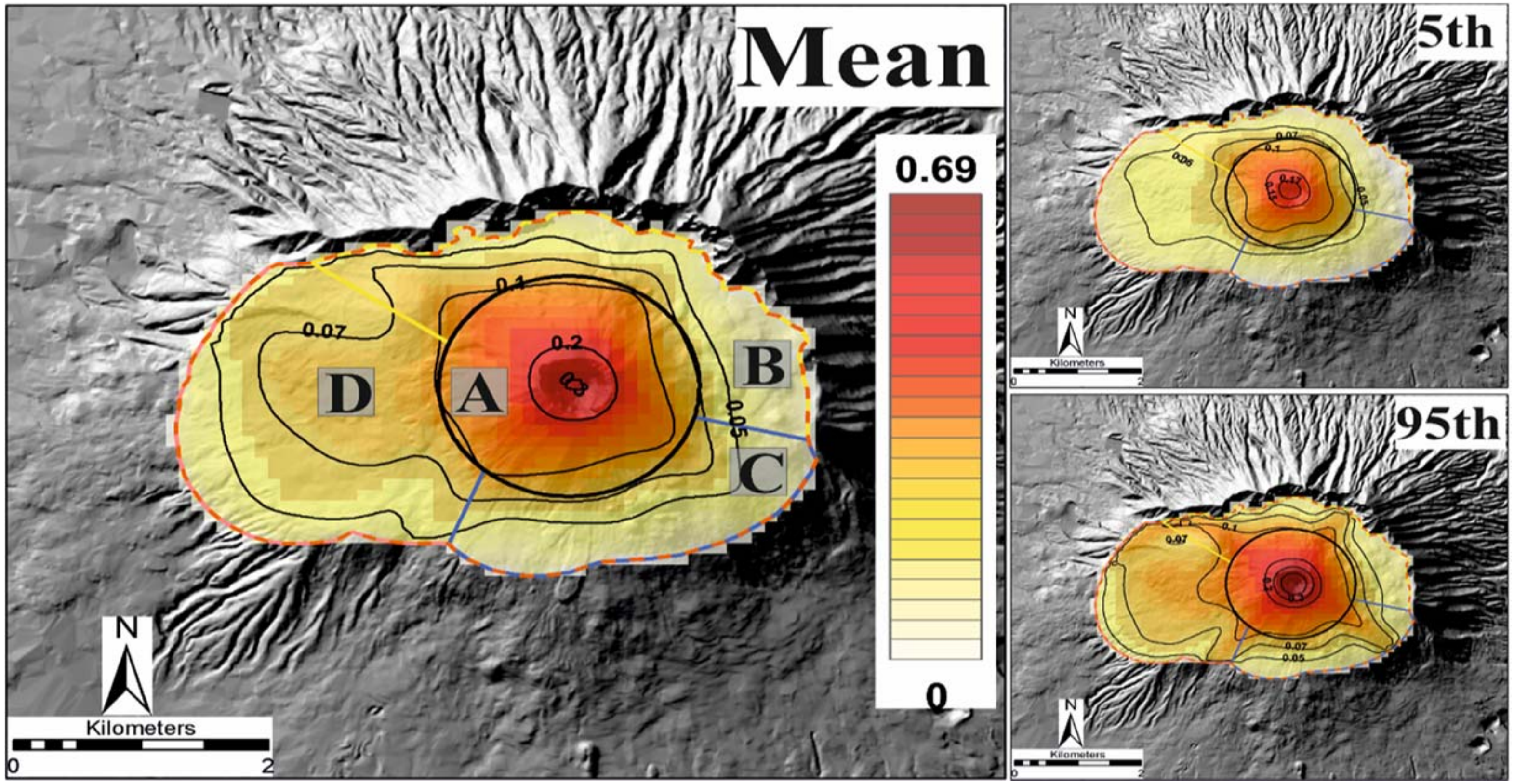
- 17 Experts were invited to provide their judgements on 16 seed items.
- 15 target items focused on the **linear weights** of the spatial distributions for the vent opening map definition.

Figure. Hierarchical logic tree structure associated to the target questions queried during the elicitation sessions.



- Sensitivity analyses were performed by considering sub-groups of experts (Juniors/Seniors, Geologists/Modelers)

Somma-Vesuvius (2015) - 2



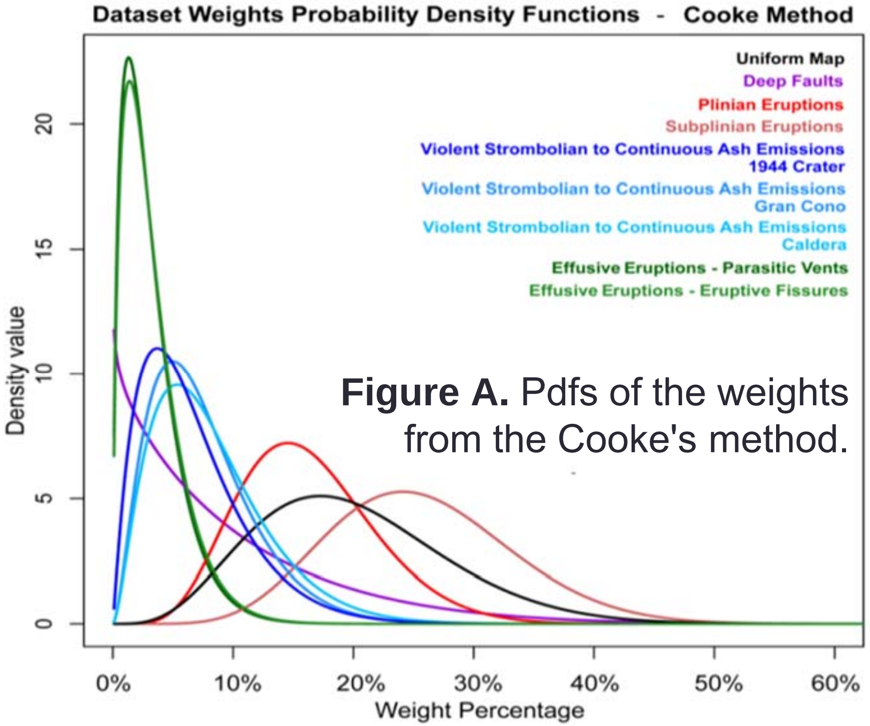


Figure A. Pdfs of the weights from the Cooke's method.

Figure B. Probability density maps of new vent opening location. Contours and colours indicate the percentage probability of vent opening per km^2 conditional on the occurrence of an eruption.

- Values of probabilities for different caldera sectors (mean)
  - Sector A (Gran Cono): 43.9%
  - Sector B (Valle del Gigante): 15.7 %
  - Sector C (Valle dell'Inferno): 10.8 %
  - Sector D (Piano delle Ginestre): 29.6 %

- Maxima located in correspondence of the present Crater values are probability percentages per cell (%/hm^2)

AGU PUBLICATIONS
JGR
Journal of Geophysical Research: Solid Earth

Assessing future vent opening locations at the Somma-Vesuvio volcanic complex: 2. Probability maps of the caldera for a future Plinian/sub-Plinian event with uncertainty quantification

A. Tadini[1,2], A. Bevilacqua[2,3,4], A. Neri[2], R. Cioni[1], W. P. Aspinall[5,6], M. Bisson[2], R. Isaia[7], F. Mazzarini[2], G. A. Valentine[8], S. Vitale[7,9], P. J. Baxter[10], A. Bertagnini[2], M. Cerminara[2], M. de Michieli Vitturi[2], A. Di Roberto[2], S. Engwell[2,11], T. Esposti Ongaro[2], F. Flandoli[12], and M. Pistolesi[1]

Cotopaxi/Guagua Pichincha (2019) - eruption type probability

Figure.
Experts at work in Clermont-Ferrand during the elicitation.

1. 20 experts (different nationalities, background, experiences)

2. 14 seed questions
   - South american volcanism
   - Plume/tephra dispersal numerical modelling

3. 55 Target questions
   Uncertainty distribution of:
   - Eruption type probability
   - Eruptive source parameters (duration, mass fallout, plume height)

Bulletin of Volcanology (2021) 83:35
https://doi.org/10.1007/s00445-021-01458-z

RESEARCH ARTICLE

Eruption type probability and eruption source parameters at Cotopaxi and Guagua Pichincha volcanoes (Ecuador) with uncertainty quantification

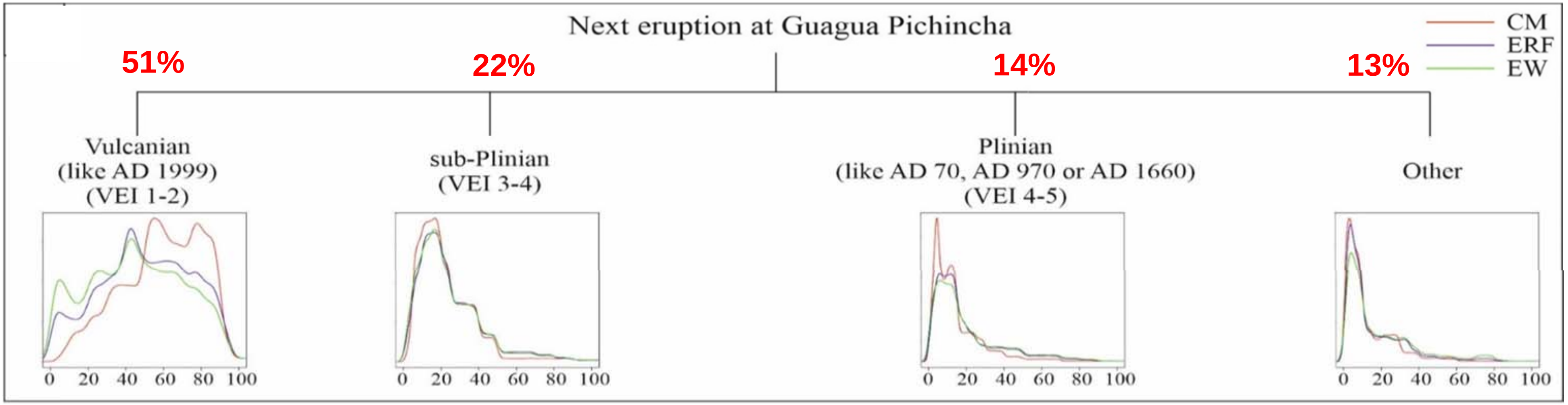
Alessandro Tadini[1] · Olivier Roche[1] · Pablo Samaniego[1,2] · Nourddine Azzaoui[3] · Andrea Bevilacqua[4] · Arnaud Guillin[3] · Mathieu Gouhier[1] · Benjamin Bernard[2] · Willy Aspinall[5] · Silvana Hidalgo[2] · Julia Eychenne[1] · Mattia de' Michieli Vitturi[4] · Augusto Neri[4] · Raffaello Cioni[6] · Marco Pistolesi[7] · Elizabeth Gaunt[2] · Silvia Vallejo[2] · Marjorie Encalada[2] · Hugo Yepes[2] · Antonio Proaño[2] · Mia Pique[2,8]

Received: 29 October 2020 / Accepted: 9 April 2021

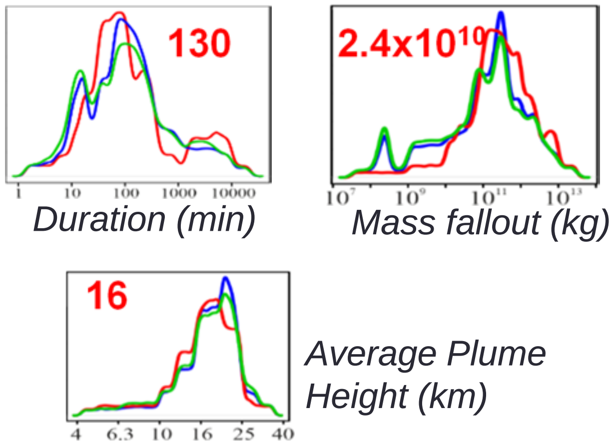
Cotopaxi/Guagua Pichincha (2019) - 2

Figure. Hierarchical logic tree structure associated to the target questions queried during the elicitation sessions

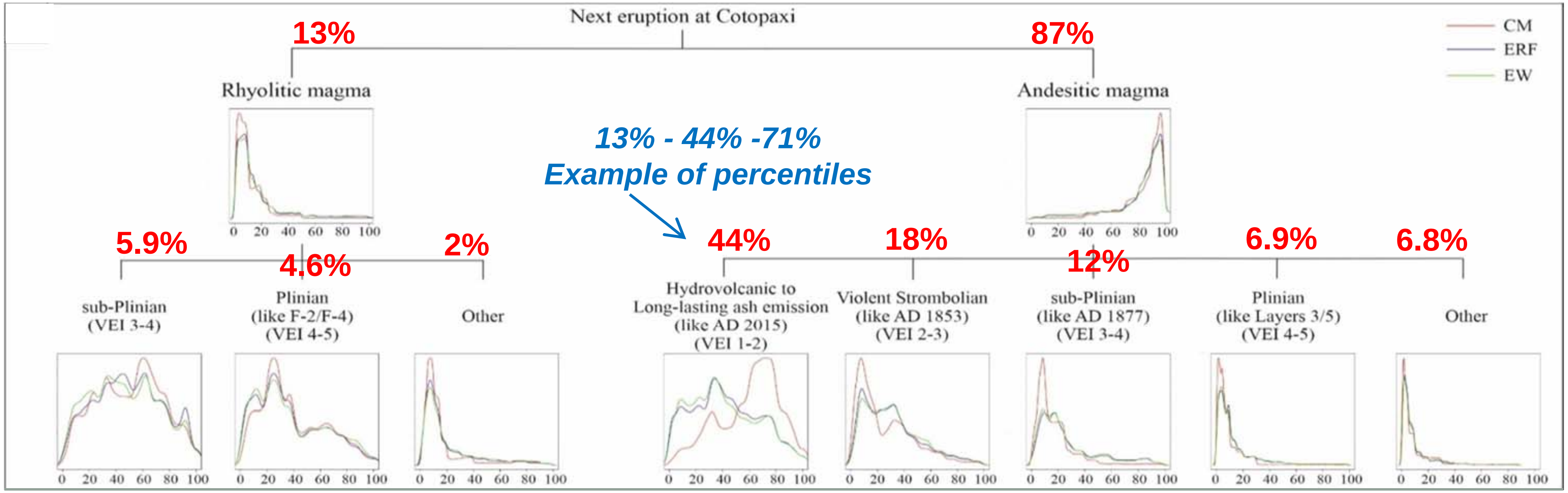
GUAGUA PICHINCHA



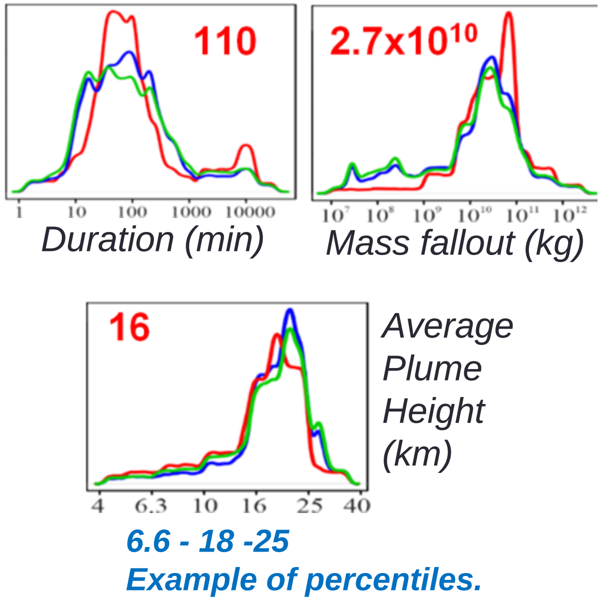
Sub-Plinian



COTOPAXI



Sub-Plinian andesitic



NEW TOOL FOR EE: ELICIPY

Questionnaire design

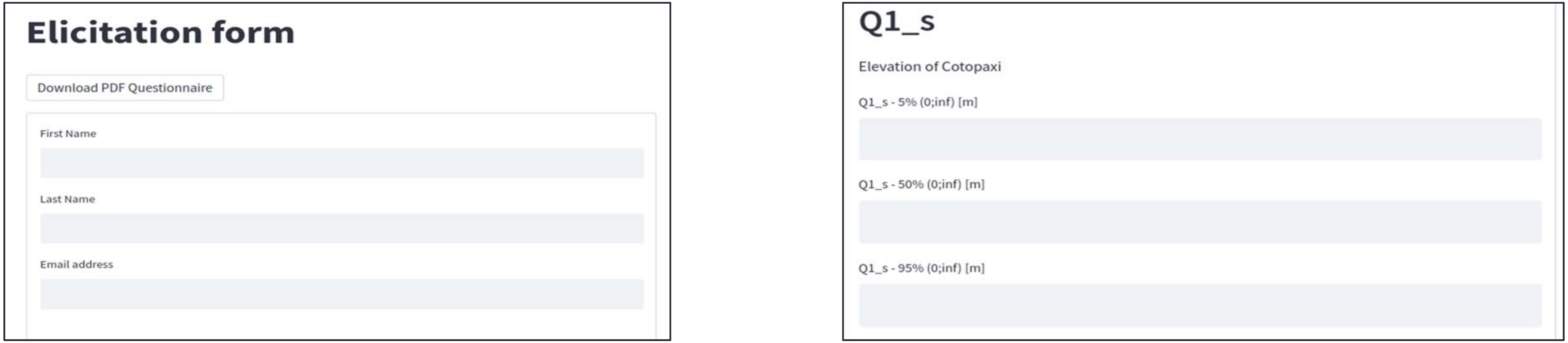


Figure. Example of the online forms

- Online questionnaires (seed/target) answered by each expert
  - Saved into local or online encripted repositories into standard csv files
- Controls on experts' answers (within bounds, sum to 100, increasing percentile values)
- Images can be added for each question
- Option of multiple languages, personalized questionnaires
- Download pdf with questions/supporting information and csv with answers

Results analysis

- Csv outputs from webforms → assembled by the analysis tool into a single csv file.

- Elicipy based on the combination of the Cooke's method scripts of the open source package Anduril (CM,EW) and the R scripts of INGV Pisa that we typically use in data processing. All translated in Python language.

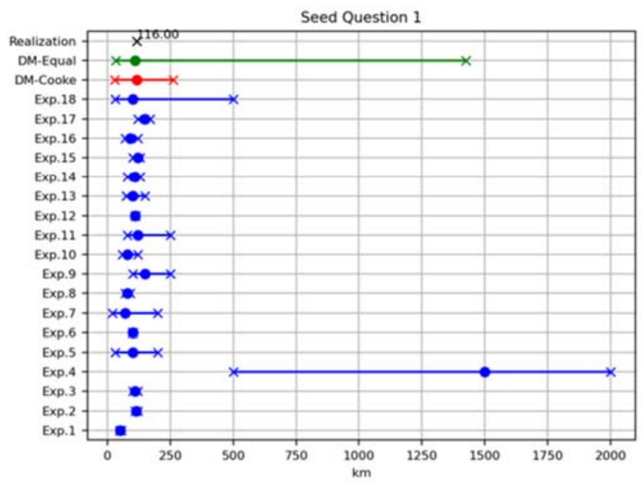
- Outputs:
  - itemwise range graphs,
  - statistical sampling of the DM responses
  - PDF and histograms for each question
  - .rls and .dtt files retro-consistent with Excalibur/Anduril

AUTOMATIC PRODUCTION → Power point presentation

Expert elicitation 25-oct.-2021

Ecuadorian arc width

(a)



Expert elicitation 25-oct.-2021

SAV (10 yrs) -

(b)

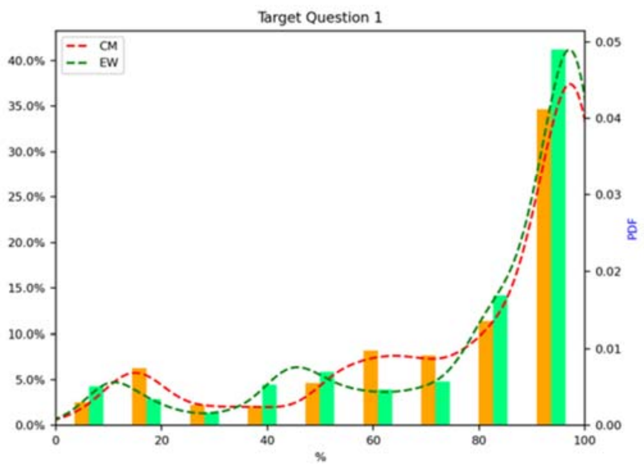


Figure. Example of (a) range graph and of (b) PDF automatically produced by ELICIPY.

Workflow

Requirements

- GitHub account
- Python 3.7
- Linux environment

→ Download tools into local directories

WEBFORM

- Link github/streamlit account, verify requirements
- Edit the «questionnaire.csv» template
- Add images (optional)
- Run python script
- Send link to participants
- Collect questionnaires (local/online repository)

ANALYSIS

- Copy «DATA» folder created by streamlit into Elicipy folder
- Edit input file for Elicipy settings (CM, EW, optionally others)
- Run python script
- Get ppt and output files

STROMBOLI ELICITATION

with: Mattia de' Michieli Vitturi, Andrea Bevilacqua, Alessandro Tadini, Tomaso Esposti Ongaro, Augusto Neri, Matteo Cerminara, Marco Pistolesi, Andrew Harris, Emmie Bonilauri, Raphaël Paris

Expert elicitation at Stromboli: target questions

Part I - annual probability of tsunamigenic landslides

- Reconstruction of the number of past tsunamigenic landslides
- Annual probability of tsunamigenic landslides (next 50 yrs)

Part II – triggering conditions of the tsunamigenic landslide

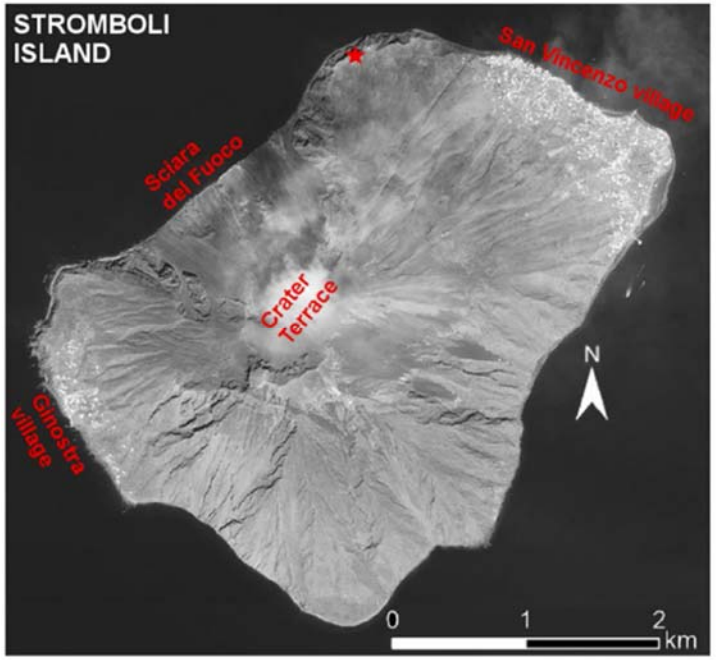


Figure A. Overview of Stromboli island.

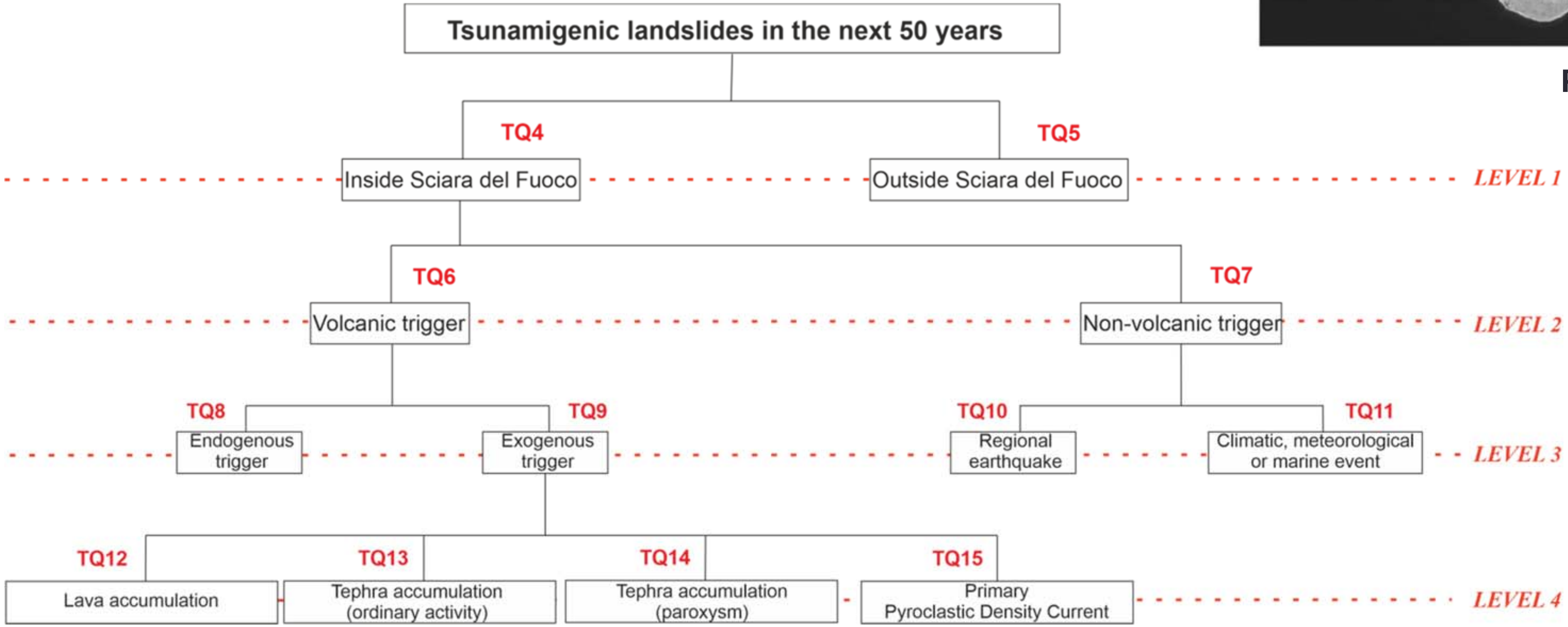


Figure B. Logic tree of Part II

Expert elicitation at Stromboli: target questions 2

Part III - spatial location and volume of the tsunamigenic landslide

- Simulations performed at INGV Pisa
- Only along the Sciara del Fuoco (SdF)

Four volume classes	Four spatial classes	
V1 = {1 < Volume ≤ 5}	Deep submarine	700-300 m BSL
V2 = {5 < Volume ≤ 14}	Shallow submarine	300-0 m BSL
V3 = {14 < Volume ≤ 30}	Lower subaerial	0-300 m ASL
V4 = { Volume ≥ 30}	Upper subaerial	300-700 m ASL
x 10^6 m^3		

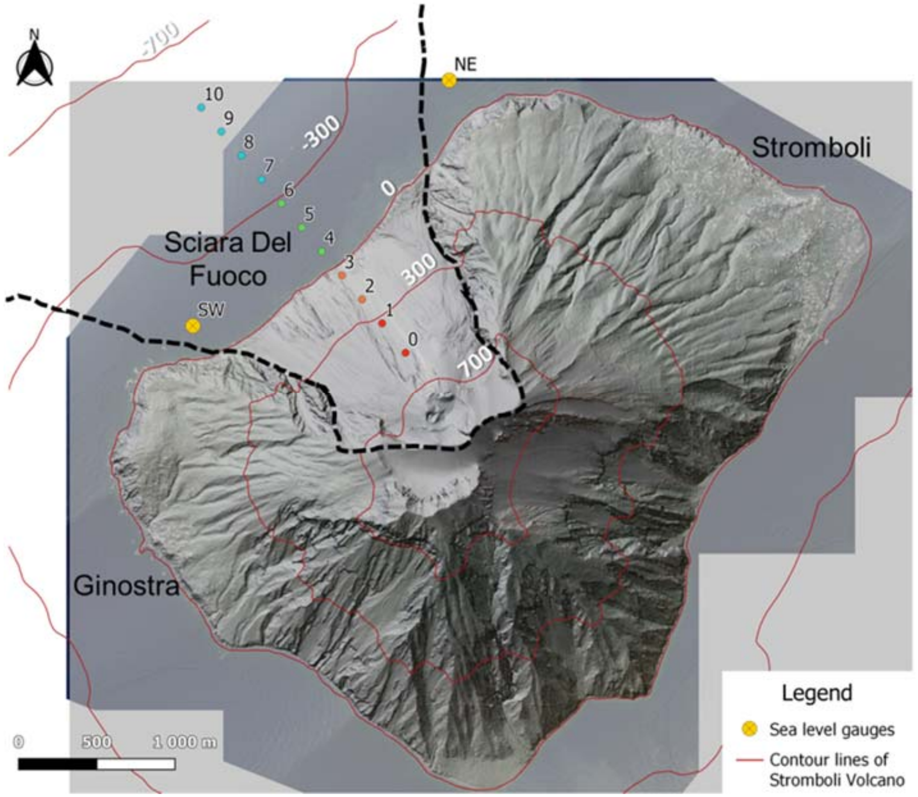


Figure A. Sciara del Fuoco (black dashed line) and positions of the center of masses for the simulations performed at INGV.

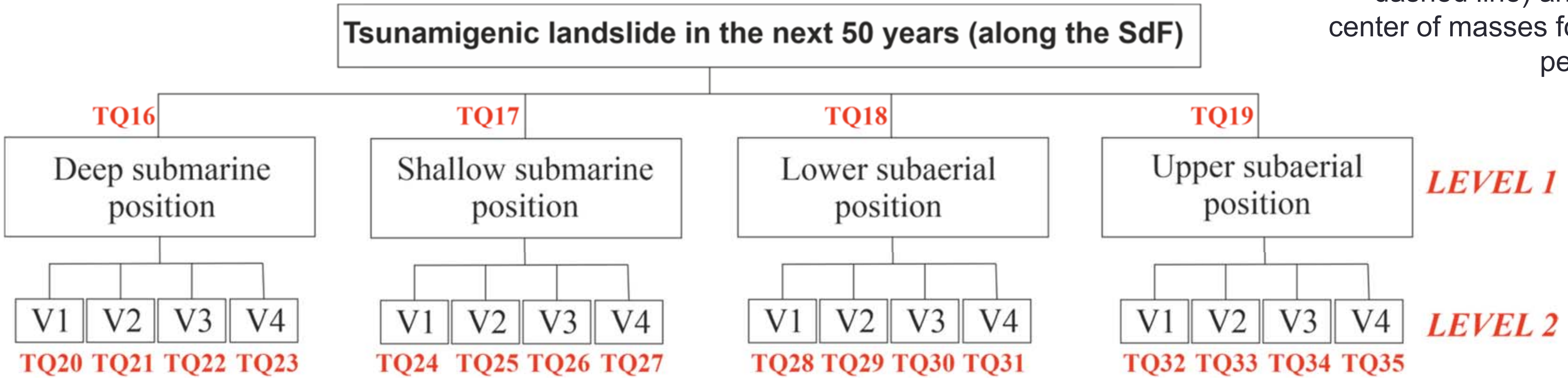


Figure B. Logic tree of Part III

THANKS FOR YOUR ATTENTION!

Questions?