

Structured elicitation of expert judgement in real-time eruption scenarios: an exercise for Piton de la Fournaise volcano, La Réunion island

Alessandro Tadini^{*α}, Andrew Harris^α, Julie Morin^β, Andrea Bevilacqua^γ, Aline Peltier^{δ,ε}, Willy Aspinall^ζ, Stefano Ciolli^η, Patrick Bachèlery^α, Benjamin Bernard^θ, Jonas Biren^{α,ι}, António Brum da Silveira^κ, Valéry Cayol^α, Oryaëlle Chevrel^α, Diego Coppola^λ, Hannah Dietterich^μ, Amy Donovan^β, Olaya Dorado^{ν,ξ}, Stéphane Drenne^ο, Olivier Dupéré^π, Lucia Gurioli^α, Stephan Kolzenburg^ο, Jean-Christophe Komorowski^δ, Philippe Labazuy^α, Domenico Mangione^η, Stefano Mannini^ζ, François Martel-Asselin^τ, Etienne Médard^α, Sophie Pailot-Bonnétat^α, Victoria Rafflin^α, Michael Ramsey^ν, Nicole Richter^φ, Silvia Vallejo^θ, Nicolas Villeneuve^{δ,κ}, Silvia Zafrilla^ν

^α *Université Clermont Auvergne, Laboratoire Magmas et Volcans, CNRS, IRD, OPGC,
6 Avenue Blaise Pascal, 63178 Aubière, France.*

**Full list of affiliations given in [Affiliations](#) section.*

ABSTRACT

Formalised elicitation of expert judgements has been used in recent years to help tackle several problematic societal issues, including volcanic crises and pandemic threats. We present an expert elicitation exercise for Piton de la Fournaise volcano, La Réunion island, held remotely in April 2021. This involved twenty-eight experts from nine countries who considered a hypothetical effusive eruption crisis involving a new vent opening in a high-risk area. The tele-elicitation presented several challenges, but is a promising and workable option for application to future volcanic crises. Our exercise considered an “uncommon” eruptive scenario with a vent outside the present caldera and within inhabited areas, and provided uncertainty ranges for several hazard-related questions for such a scenario (e.g. probability of eruption within a defined timeframe; elapsed time until lava flow reaches a critical location, and other hazard management issues). Our exercise indicated that such a scenario would probably present very different characteristics than the eruptions observed in recent decades, and that it is fundamental to include well prepared expert elicitation in updated civil protection evacuation plans to improve disaster response procedures.

Keywords: Expert elicitation; Piton de la Fournaise; Effusive volcanism;

1 INTRODUCTION

Effective management of volcanic crises is necessary to reduce, as much as possible, the number of casualties and the impacts on human infrastructure and the environment. However, large uncertainties affect the characterization of an evolving volcanic crisis, due to both the stochastic nature of eruption processes and our limited capability to conceptualize the behavior of a complex dynamic system [Aspinall and Blong 2015]. Moreover, because social and economic loss resulting from false alarms and evacuations must also be considered [Woo 2008; Hincks et al. 2014; Aspinall and Woo 2019], it is important for scientists to provide decision-makers with clear collective information about the possible evolution of a volcanic crisis, including the related uncertainties. On a broad scale, event trees [Newhall and Hoblitt 2002; Newhall and Pallister 2015] are useful for describing the main eruption types and related

hazards for any given volcano. The branches of these event trees are then populated with their relative probability of occurrence according to performance-based expert weighting techniques [e.g. Aspinall 2006; Neri et al. 2008] or Bayesian network modelling [e.g. Christophersen et al. 2018]. In some circumstances it is also necessary to provide insight on more specific questions, such as “What is the probability of an eruption within the next 6 hours?” or “When will village X be impacted by a certain phenomenon?” In such cases, expert elicitation techniques (including performance-based expert weighting) can be employed to address such issues. Although some elicitation produce sets of judgements that are more coherent than others, depending on the problem being tackled [e.g. Tyshenko et al. 2012], an elicitation-based approach can be of real prognostic value for volcanic hazards [Wadge and Aspinall 2014]. The classic example of a successful application of structured elicitation of expert judgement during a volcanic crisis is that of Soufrière Hills volcano, Montserrat, in

*Corresponding author: Alessandro.TADINI@uca.fr

1998 [Aspinall and Cooke 1998].

Whereas explosive eruptions present potentially more destructive phenomena, effusive eruptions can still be hazardous and damaging [Harris 2015a]. This is documented for instance at Nyiragongo in 1977 [Tazieff 1977], 2002 [Komorowski et al. 2002], and 2021 [GVP 2021; OCHA 2021]; Kilauea in 1960 [Macdonald 1962] and 2018 [Neal et al. 2019]; Mauna Loa in 1950 [Macdonald and Finch 1950]; and Etna in 1669 [Branca et al. 2013; 2015], 1928 [Chester et al. 1999; Branca et al. 2017], 1991–1993 [Barberi et al. 1993; Calvari et al. 1994], and 2001 [Barberi et al. 2003]. High intensity effusive crises are commonly destructive events that can evolve with rapidly extending lava flows, but for which there is commonly little experience or knowledge. Thus, we use a structured elicitation of expert judgement (abbreviated to expert elicitation hereafter) to assess the hazard associated with just such an event at Piton de la Fournaise volcano (La Réunion island, French overseas department). The goal of this exercise is to aid the responsible volcano observatory and civil protection in better preparing and planning for such an event. In doing so, we refine a methodology, used in other similar applications [Aspinall and Cooke 1998; Aspinall et al. 2020], that could be activated with a large, globally distributed and remotely connected group of experts in near-real time during an eruptive crisis or phase of unrest.

The volcano observatory on La Réunion, *Observatoire Volcanologique du Piton de la Fournaise* of the *Institut de Physique du Globe de Paris* (OVPF-IPGP), was established in 1979. The creation of OVPF-IPGP was the direct result of the 1977 eruption of Piton de la Fournaise during which lava flows entered the town of Piton Sainte Rose (see Figure 1A). Today, OVPF-IPGP manages monitoring and civil protection reporting duties for volcanic and seismic hazard on La Réunion [Peltier et al. 2022]. Most of the historical activity of Piton de la Fournaise has been confined to the unpopulated Enclos Fouqué caldera [Figure 1A; Harris et al. 2017]. However, eruptions like that of 1977 can also occur outside the caldera. In total, there have been twelve documented eruptions outside the caldera between 1708 and 2021; as such, volcanic hazards due to events outside the caldera are non-negligible [Chevrel et al. 2021]. Since the creation of OVPF-IPGP there have been (as of April 2021) 81 eruptions inside the unpopulated caldera, but only two outside of the caldera. Therefore, there is little or no experience of a high-risk eruption outside the caldera in recent memory. Indeed a gap analysis for effusive crisis response completed by OVPF-IPGP revealed that, whereas knowledge and published experience of hazard, risk and losses during “normal” effusive crises within the caldera is excellent, as well as monitoring, mitigation, and recovery efforts in those circumstances, there is a gap for “Hors Enclos” events (i.e. those occurring beyond the caldera) [Peltier et al. 2022].

Lava flow modelling efforts to improve near-real-time hazard assessment for communication to civil protection have been developing since 2014 [Harris et al. 2017; 2019; Peltier et al. 2021]. These efforts have been based on well-constrained source terms for events occurring within the caldera, characterized in terms of eruption frequency, style, and duration, as well as associated effusion rates, flow lengths, and time scales of emplacement [Chevrel et al. 2021]. However, there is no experience of, or similar data for, effusive events originating beyond the caldera. Such data are essential if we are to adequately assess and model potential events, as well as to set plausible source terms for lava flow hazard beyond the caldera. It is in this context that we conceived and implemented an expert elicitation specifically for an effusive crisis outside the caldera.

Our expert elicitation exercise was designed to assess the potential timing and outcome, in terms of hazard, of a Hors Enclos eruption in a high-risk zone of Piton de la Fournaise: the populated area of La Plaine des Palmistes (Figure 1B). In doing so, we took advantage of a global group of experts with extensive experience in managing, monitoring, modeling, and responding to effusive crises. The group involved actors from both the scientific and civil protection communities from Ecuador, France, Germany, Italy, Portugal, Spain, Switzerland, United Kingdom, and the United States.

Originally planned to be carried out with everyone present on the island of La Réunion in April 2020 within the framework of the workshop organized by the project Lava Advance in Vulnerable Areas (LAVA-ANR-16 CE39-0009*), the global SARS-COV-2 pandemic caused this event to be postponed; it was then rescheduled as a teleconference meeting for April 2021. The elicitation took place on April 13th and the results were presented on April 15th. During this period, there was an eruption underway at Piton de la Fournaise volcano (beginning on April 9th†). The eruption caused the opening of an eruptive fissure within the Enclos Fouqué to the South of Dolomieu cone and the development of a lava flow moving to the ESE‡. The eruption ended on May 24th§. The pandemic complication allowed us to develop and test a virtual form of expert elicitation that can be applied to a large, globally distributed group in near-real-time during an evolving crisis, without the need for the presence of the expert group on-site at the volcano itself.

In this study, after describing the scenario for Piton de la Fournaise that was presented to the experts, we detail the applied methodology and present the results both in terms of application for hazard assessment and

*<https://anr.fr/Project-ANR-16-CE39-0009>

†<https://www.ipgp.fr/fr/ovpf/communique-09042021-15h20-heure-locale>

‡<https://www.ipgp.fr/fr/ovpf/communique-10042021-10h00-heure-locale>

§<https://www.ipgp.fr/fr/ovpf/communique-24052021-09h00-heure-locale>

differences in thinking between scientific and civil protection groups.

2 BACKGROUND

Piton de la Fournaise is a highly active, basaltic hot-spot volcano, located on the French island of La Réunion in the Indian Ocean (Figure 1A). Its historical and recent (since the establishment of OVPF-IPGP) eruptive activity has consisted of numerous effusive eruptions with Hawaiian to Strombolian style explosive activity around the vent zones, with a mean of two events per year since 1935 [e.g. Peltier et al. 2009; Roult et al. 2012; Chevrel et al. 2021]. Ninety-five percent of the eruptions since 1708 have occurred inside the Enclos Fouqué caldera (Figure 1A), with vents opening mainly at the summit or along one of the three rift zones (N120, NS, and EW) [Chevrel et al. 2021].

The Enclos Fouqué caldera is uninhabited, but contains the island belt road, hiking trails and, on any given day, up to a few thousand visitors [CREGUR 2003a; b; Villeneuve 2020]. Populated zones are beyond the Enclos Fouqué caldera on the outer flanks of the shield and, in the past, eruptions of larger magnitude have occurred outside of the caldera, locally more than 15 km from the volcano summit, and have built large eccentric cones and extensive lava flow fields [Villeneuve and Bachèlery 2006; Chevrel et al. 2021]. Several towns and villages are now established across these flank flow fields, especially in the Le Tampon and La Plaine des Palmistes sectors (Figure 1). As a partial comparison, according to the Smithsonian Institution database*, more than 50,000 people live within 10 km of the center of Piton de la Fournaise [Harris et al. 2017]. Outer flank eruptions (also termed “Hors Enclos” eruptions) are thought to be fed by magma that may have by-passed the shallow plumbing system of the central area of the volcano. Such flank eruptions instead take a lateral and direct pathway from a deep magma storage zone located below the western outer flank of the volcano [Villeneuve and Bachèlery 2006; Boudoire et al. 2017]. A detailed list of eruptions at Piton de La Fournaise has been compiled by Staudacher et al. [2008] for the period 1998–2007, and it provides a perspective on the eruptive activity at Piton de La Fournaise for a 10-year period (Table 1).

As of 2021, OVPF-IPGP monitored the volcano with a permanent monitoring network of 107 sensors (seismometers, GNSS, tiltmeters, extensometers, gas stations, webcams, weather stations). This network allows OVPF-IPGP to provide early warning for eruptive activity, to track on-going activity [e.g. Peltier et al. 2021] and to provide the authorities with notice of any change in activity [Peltier et al. 2022]. Indeed, through the government-mandated emergency plan (*Organisation de la Réponse de Sécurité Civile* (ORSEC) - *Volcan*

du Piton de la Fournaise) OVPF-IPGP must inform the civil protection department of the *Préfecture* (i.e. the decentralized administrative service of the French government) of any changes in volcanic activity so that the authorities can change alert levels accordingly. These alert levels (currently under review) are [Peltier et al. 2022]:

1. Vigilance: Eruption “watch”: volcano is in a state of unrest;
2. Alert 1: Eruption imminent (from few dozens of minutes to few hours);
3. Alert 2.1: Eruption underway inside the summit Dolomieu crater;
4. Alert 2.2: Eruption underway in the Enclos Fouqué caldera;
5. Alert 2.3: Eruption underway outside of the Enclos Fouqué caldera.

In the case of Alerts 2.1 and 2.2, visitor access to the Enclos Fouqué caldera is closed, and in the event of Alert 2.3 all population centers under threat are evacuated. Over the last 100 years, eruptive fissures have opened only three times outside of the Enclos Fouqué caldera: in 1977, 1986, and 1998. Lava flows reached populated areas only in 1977 and 1986; no casualties were reported, but a few dozen houses and infrastructure were destroyed [Peltier et al. 2022]. Thus, apart from the 1986 eruption, OVPF-IPGP and the *Préfecture* have dealt only with eruptions that have not reached populated areas (since it was established in 1979).

In 2002 evacuations did take place because of the threat of fissures opening outside the caldera, and in 2007 an evacuation (both spontaneous and enforced) occurred following misinformation announcing a fissure opening outside the caldera [Morin 2012]. The paroxysmal eruption of March–May 2007 also gave experience of a high intensity, lower flank eruption. This was the most voluminous eruption in the last century, when discharge rates were sustained at more than $100 \text{ m}^3 \text{ s}^{-1}$ over 30 days [Staudacher et al. 2009]. However, current monitoring and mitigation experience is very much founded on relatively low intensity eruptions within the Enclos Fouqué caldera and for which protocols are well developed [Peltier et al. 2021]. It is within this context that we set up an expert elicitation for a plausible scenario but for which there is no historical experience or memory. The scenario entails a high intensity flank eruption opening in the zone of La Plaine des Palmistes, i.e. at relatively high elevation on the volcano flank and inside a populated area (Figure 1).

*http://volcano.si.edu/search_volcano.cfm

Table 1: Summary table of eruptive activity at Piton de la Fournaise for the period 1998–2007 [modified from Staudacher et al. 2008]. Seismic crisis is the period between the first seismic signal and the beginning of the eruptive activity. For eruption location, acronyms are as follows: EF = Enclos Fouqué, HE = Hors Enclos, D = Dolomieu, PO = Plaine des Osmondes (see Figure 1A), GB = Grand Brûlé (see Figure 1A). Surface area covered is related to lava flows. MDR = mass discharge rate, obtained by dividing the total volume and the eruption duration.

No.	Eruption start date	Start time (hh:mm) (local)	Seismic crisis duration (minutes)	Eruption location	Vent altitude (m asl)	Eruption duration (days)	Surface area covered (10^6 m^2)	Erupted lava volume (10^6 m^3)	Mean MDR ($\text{m}^3 \text{ s}^{-1}$)
1	09/03/1998	15:05	2100	EF (N)	2450–2050	196	8.5	60	3.5
2	12/03/1998	02:56	30	EF (W)	2200	21	0.4	1	0.6
3	09/08/1998	02:00	–	HE	1700	–	–	–	–
4	19/07/1999	18:56	39	EF (D+E)	2500–2100	12	0.65	1.8	1.7
5	28/09/1999	11:58	81	EF (D+S)	2500–1850	25	0.94	1.5	0.7
6	14/02/2000	00:18	64	EF (N)	2450–2250	19	1.4	4.1	2.5
7	23/06/2000	18:00	72	EF (ESE)	2100–1820	37	2.1	6	1.9
8	12/10/2000	05:05	57	EF (SE)	2260–2000	32	3.4	9	3.3
9	27/03/2001	13:20	26	EF (SSE)	2450–1940	8	1.5	4.8	6.9
10	11/06/2001	13:50	32	EF (SE)	2450–1800	26	3.3	9.5	4.2
11	05/01/2002	23:00	380	EF (NE)	1940	12	2.9	13	12.5
12	16/11/2002	04:33	296	EF (E)	1850–1500	18	3	8	5.1
13	23/12/2002	10:01	–	EF (D)	2480	–	–	–	–
14	30/05/2003	11:55	17	EF (D)	2490	1	0.07	0.14	1.6
15	04/06/2003	12:20	–	EF (D)	2490	5	0.18	0.7	1.6
16	13/06/2003	03:08	–	EF (D)	2490	3	0.23	0.41	1.6
17	22/06/2003	23:30	–	EF (D)	2490	15	0.19	0.36	0.3
18	22/08/2003	21:20	152	EF (N)	2590–2140	6	1.6	6.2	12.0
19	30/09/2003	23:30	65	EF (SSW)	2330–2195	0.5	0.33	1	23.1
20	07/12/2003	15:25	56	EF (ESE)	2475	18.2	0.24	1.2	0.8
21	09/01/2004	02:45	2475	EF (PO)	1500	1.4	0.6	1.9	15.7
22	02/05/2004	19:35	24	EF (S)	2525–2000	16	3.9	12.5	9.0
23	13/08/2004	02:40	25	EF (D+E)	2475–2210	64	4.9	18.5	3.3
24	17/02/2005	16:35	435	EF (PO)	1600–450	10	4.6	19	22.0
25	04/10/2005	10:26	58	EF (WSW)	2490	14	0.36	1.5	1.2
26	29/11/2005	02:25	26	EF (D+N)	2500–2300	0.6	0.4	1.5	28.9
27	26/12/2005	13:15	141	EF (NE+PO)	2100–1700	24	3.4	17.5	8.4
28	20/07/2006	03:45	324	EF (SW)	2380–2150	24	1	3	1.4
29	30/08/2006	11:35	15	EF (D+E)	1500–800	124	1.3	19	1.8
30	18/02/2007	16:38	27	EF (D)	2500	0.4	0.4	1	28.9
31	30/03/2007	22:50	145	EF (ESE)	1900	0.4	0.38	0.6	17.4
32	02/04/2007	10:00	2800	EF (GB+S)	590	30	3.8	130	50.2
Min	–	–	–	–	590	0.4	0.1	0.1	0.3
Max	–	–	–	–	2490	196	9	130	50
Mean	–	–	–	–	2124	25	2	12	9

3 METHODS

3.1 Structured expert judgement (or elicitation)

3.1.1 Background

A structured elicitation and aggregation of individual experts' judgements on uncertain topics has long been used in many different fields with different approaches [e.g. Bedford et al. 2006; Wiser et al. 2016; Barons and Aspinall 2020; Aspinall et al. 2021b]. In volcanology, examples include the production of event trees, Bayesian Belief Networks, and logic trees for eruption scenarios [Neri et al. 2008; Selva et al. 2012; Christophersen et al. 2018; Tadini et al. 2021]; vent opening probability maps [Bevilacqua et al. 2015; Tadini et al.

2017; Bebbington et al. 2018]; past PDC deposits areal extent [Neri et al. 2015; Cioni et al. 2020]; and eruptive source parameters [Aspinall and Woo 2019; Tadini et al. 2021]. Several different strategies to perform expert elicitation exist [see, e.g. Aspinall 2006; Dias et al. 2018; Hanea et al. 2021a; b; Williams et al. 2021, for a review]. Those techniques that include a performance-based procedure of a group of experts rely on the external validation of expert probability assessments. More specifically, such validation is performed for a set of unknown quantitative variables, called “target questions” or “target items.” For these variables, a performance-based algorithm produces group-synthesized uncertainty distributions, commonly called Decision Maker (DM) solutions [Cooke 1991; Aspinall 2006; 2010], through a weighting scheme (or ‘pooling method’) to

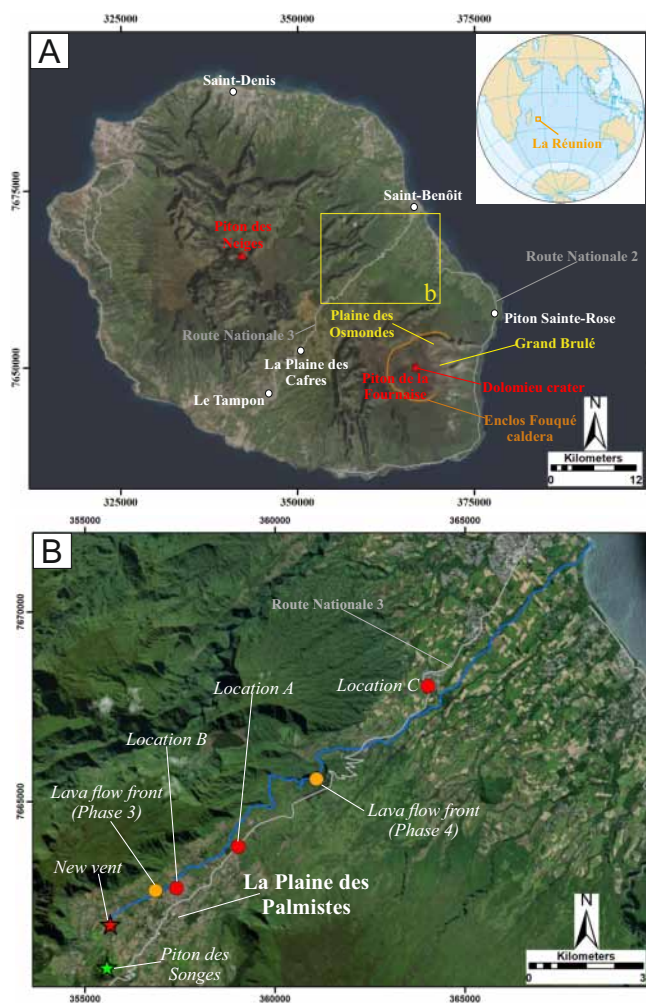


Figure 1: [A] La Réunion island; [B] enlargement of [A] with the indication of key elements for the target questions described in Section 3.1.2. Grey lines are main roads, the blue line is the line of steepest descent, the orange dashed line is the limit of the Enclos Fouqué caldera. Coordinates are expressed in the UTM-WGS84 40S system. Service Layer Credits, source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN and the GIS User Community.

combine the experts' answers, i.e. individual uncertainty distributions representing the experts' opinions [Bevilacqua 2016].

The simplest choice is to assign each expert the same weight and combine the answers linearly, an approach that is generally referred as the Equal Weight (EW) rule. In this study we compare the EW output with that of a highly selective performance-based method, i.e. the Classical Model (CM [Cooke 1991]). Performance-based methods weight the experts according to their responses to an appropriate set of “seed questions,” or “seed items,” which measure each expert's individual performance in uncertainty quantification. The seed items typically comprise factual questions with exact

answers (usually referred to as “realizations”) known to the analysts, but not to the experts. These questions are designed to be as similar as possible to the target questions. Participating experts are expected to be able to provide judgement-based credible intervals that “capture” a majority of seed item values, each expert responding according to their own knowledge, expertise, and critical reasoning. How well each expert performs over the set of seed items is the numerical basis for a personal score, which determines the weight they are given when pooling everyone's judgements.

For our exercise, we followed the Classical Model; individual experts provided estimates of 5th percentile, median, and 95th percentile for each seed and target question. This allowed us to define “maximum entropy distributions” by assuming uniform probability within each couple of quantiles. A 10 % overshoot was assumed at both ends of this percentile range and allowed us to define the group-wise minimum and maximum values of the uncertainty distribution. The 10 % overshoot is the result of the chosen percentiles, which is the range between the 5th and the 95th percentiles accounting for 90 % of the distribution. Uniform probability is also assumed across the variable 10 % intrinsic range extensions at each end [for similar applications see Bevilacqua et al. 2015; Tadini et al. 2021].

The seed items enabled us to define two scores, the statistical accuracy (also called “calibration”) and the informativeness [Cooke 1991; Aspinall and Cooke 2013]. The calibration represents an inverse distance between the empirical distribution of the real answers to the seed questions, and the probability distributions implied by the 5th, median and 95th percentiles assessed by the experts per item [Cooke 1991; Bevilacqua 2016]. Thus, a “well calibrated” expert provides answers such that the real values are symmetrically balanced with respect to their 50th percentile markers, and the majority fall between their 5th and 95th percentiles). By contrast, the expert's informativeness score is the degree to which their uncertainty distributions are concentrated; that is, the smaller the distance between the 5th and the 95th percentiles, the more an expert is informative. We note that the informativeness score is unrelated to the accuracy of the estimates with respect to the true values. The resulting weights (proportional to the product of calibration/statistical accuracy and informativeness) were then applied to linearly pool experts' answers to the target questions. The graphical outputs of the following sections show the probability density functions of the DM resulting from the application of a Gaussian kernel density estimator [Silverman 1986; Connor and Connor 2010; Tadini et al. 2017] to the weighted combination of the experts' probability distribution judgments. In doing this, we extracted a sufficient number of samples (10⁵) of expert answers to assure a robust convergence of the kernel density estimator. This kernel-based approach, already adopted in Tadini et al. [2021], has several advantages against the

classical representation of the DM through three percentiles. In fact, the DM is not a maximum entropy distribution, but a probability mixture of many experts' answers, which can possess a complex structure and multiple modes. Three percentiles would be unable to describe that information.

3.1.2 Exercise design

The aim of this exercise was to consider a hypothetical eruption, which could potentially result in hazards and risks linked to mainly effusive, but also near-vent Strombolian, activity in the event of a flank eruption at Piton de la Fournaise in or close to an inhabited area. Many pre-historic flank eruptions have formed large cones, now visible outside the Enclos Fouqué caldera and distributed all around the flanks in populated areas. One such eruption led to the formation of the Piton des Songes scoria cone [Villeneuve and Bachèlery 2006], located NW of the Enclos Fouqué caldera in the densely populated La Plaine des Palmistes area (Figure 1B). As of 2015, the 83 km² municipality of La Plaine des Palmistes had a population of 6,157 [INSEE 2018], for a population density of 75 inhabitants per km². A vent location 1 km to the NE of Piton des Songes was thus chosen as the location for a potential opening site for a flank eruption because:

1. This site is relevant for a future flank eruption;
2. It is just upslope of La Plaine des Palmistes, so lava flows may enter the town in a short period of time;
3. Slopes are steep, so flows could propagate quickly;
4. The line of steepest descent has the flow entering the ocean (at a linear distance from the vent of 16.1 km) around the town of Saint Benoît (population 37,940 [INSEE 2018]). It passes through the sectors of Pont Payet, Chemin Ceinture, and La Confiance over a linear distance of 8, 10.7, and 12.1 km, respectively;
5. Such an eruption would also cut the main road that crosses the island from southwest–northeast (Route Nationale 3, RN3), and, should it reach the coastal area, would also cut the sole belt road of the island (Route Nationale 2, RN2) (Figure 1B).

Thus, the selected scenario represents a realistic and extremely high-stakes eruption in terms of implementation of mitigation measures as quite large populations could need to be evacuated in a relatively short period of time, and infrastructure damage and loss may be extensive and severe.

To assess the potential flow path, we plotted the line of steepest descent from the chosen vent to the coast and used this as a basis for the event scenario (Figure 1B). Then, to link the hypothetical scenario with a real case, the event and exercise was subdivided into

four phases that mimicked the evolution of an actual crisis on Piton de la Fournaise. These were:

1. Phase 1: Seismic Crisis – Alert Level 1;
2. Phase 2: Eruption Underway – Alert Level 2.3;
3. Phase 3: Eruption Underway (first field observations available) – Alert Level 2.3;
4. Phase 4: Eruption Continuing (eruption has been ongoing for 1 day) – Alert Level 2.3.

For each phase, a target questionnaire with three questions (answering to the general forms of “when,” “where,” and “how/what”) was linked to an eruption bulletin with the same format as those released by OVPF-IPGP during actual crises. These documents can be found in “Data S1” in the [Supplementary Material](#). Both questionnaires and bulletins were provided to all participants in French and English, the two languages used by OVPF-IPGP for report publication. Finally, additional material such as location maps for sites mentioned and the monitoring network were given out. We remark that, for all the target questionnaires, each question asked to provide each expert's judgments through three percentiles (5th, 50th, 95th), which are then used to derive probability density functions (see following sections).

During Phase 1, we simulated a seismic swarm and ground deformation below the area of La Plaine des Palmistes (Figure 1B), preceded by a week of intense deep seismic activity and an increase in soil CO₂ degassing (see “Phase 1 – Bulletin & Questions” in Data S1 in the [Supplementary Material](#)). These simulations were set as being consistent with a flank event from a deep source whose dyke bypasses the centrally located shallow system. The three target questions in this phase aimed at exploring the likelihood of a flank eruption and asked for the probabilities that:

1. the crisis would not end in an eruption;
2. that an eruption will occur within the following six hours; and
3. that a vent will open within 2 km of the Dolomieu crater (Figure 1A).

The Phase 2 bulletin was “released” seven hours after the first bulletin, with the release in the exercise being 20 minutes after the first (as in Phases 3 and 4 also). Phase 2 envisaged a situation whereby a seismic swarm was located NW of the caldera (and off the standard tile used by OVPF-IPGP for reporting), with eruptive fissures opening in the La Plaine des Palmistes area as confirmed by resident phone calls. In addition, during actual events, the presence of thermal anomalies is reported using the MIROVA* and HOTVOLC[†]

*<https://www.mirovaweb.it/>

†<https://www.obs.univ-bpclermont.fr/SO/televolc/hotvolc/>

systems. Both also convert spectral radiance to time-averaged discharge rates [Harris et al. 2017; Peltier et al. 2021; 2022] and this information was included in the bulletins (see “Phase 2 – Bulletin & Questions”; Data S1 in the [Supplementary Material](#)). We reported a relatively high value for the discharge rate ($120 \text{ m}^3 \text{ s}^{-1}$), consistent with the paroxysmal eruption of March–May 2007 [Staudacher et al. 2009]. In this case, target questions were aimed at assessing the likely evolution of the eruption based on the starting conditions, and asked questions regarding:

1. the final length of the eruptive fissure;
2. the likely time-averaged magma discharge rate over the next hour;
3. the duration of the eruption.

Following OVPF-IPGP procedure, the Phase 3 bulletin was “released” six hours after the Phase 2 bulletin and included reports from initial field reconnaissance by OVPF-IPGP staff and further satellite data (i.e. 20 minutes after Phase 2, in our exercise), confirming vent locations and flow front locations. These were based on the typical advance rate of the lava flows in the first 11.5 hours of the March–May 2007 eruption, which was 260 m h^{-1} [from Staudacher et al. 2009]. Time-averaged discharge rates were set as increasing from $80\text{--}120 \text{ m}^3 \text{ s}^{-1}$ to $250\text{--}300 \text{ m}^3 \text{ s}^{-1}$ over the first hours of the eruption. This is consistent with the waxing trend in effusion rates for an eruption from a pressurized source [Wadge 1984] and with effusion rate time series derived from satellite data for such high-intensity effusive eruptions at Krafla and the Galápagos [Harris et al. 2000; Rowland et al. 2003]. The experts were also provided with a slope map and a map giving the vent location and the line of steepest descent (blue line in [Figure 1B](#)), which followed the line of the Ravine Sèche and on which the simulated lava was centred (see “Phase 3 – Bulletin & Questions” in Data S1 in the [Supplementary Material](#)). Questions in this case now focused on the hazard and asked for:

1. the arrival time of lava at location A on [Figure 1B](#), this being the eastern limit of La Plaine des Palmistes, meaning this is the time needed for the lava flow to cut the town from west to east;
2. the likely lava flow width at location B on [Figure 1B](#), which is approximately the level of the town centre of La Plaine des Palmistes where there are two schools and the Gendarmerie. With it being a Tuesday afternoon, the schools would have been full, and the Gendarmerie is a key centre for managing/enforcing law and order, as well as intervention during disturbances, accidents, or damaging inflicting events (natural or anthropogenic). Thus, this is the time needed to reach these two key facilities;

3. the maximum ballistic impact distance around the vent location.

In the absence of any change in activity, OVPF-IPGP typically releases situation updates every 24 hours; accordingly, the Phase 4 bulletin was “released” 24 hours after the eruption onset (i.e. 20 minutes after Phase 3, in the exercise). The event evolution followed the basis of Phase 3, i.e. continuation of an effusive event involving channel-fed lava flow and the bulletin included further observations from helicopter overflights, field surveys, and satellite observations. The final three target questions aimed at assessing the longer-term aspects of the eruption, and asked:

1. how long the eruption will continue;
2. the probability that the lava would reach the ocean at a distance of 9 km from the Phase 4 flow front ([Figure 1B](#));
3. the arrival time of lava at location C in [Figure 1B](#), which is the sector of Chemin Ceinture.

Again, slope and location maps were provided (see “Phase 4 – Bulletin & Questions” in Data S1 in the [Supplementary Material](#)).

3.1.3 Group training and calibration

Training prior to the exercise was completed the day before the elicitation and involved one day of presentations that reviewed the response protocols for Piton de la Fournaise, led by the OVPF-IPGP Scientist-in-Charge and Civil Protection. This training also included presentations focusing on monitoring efforts, satellite monitoring of Piton de la Fournaise, derivation of discharge rates, and lava flow modeling. During such presentations, the eruptions listed in [Table 1](#) were introduced in detail. At the beginning of the exercise itself, a 30-minute presentation was given reviewing the historical activity of Piton de la Fournaise and giving four eruption scenarios from the OVPF-IPGP archive (see “Workshop_Program” in the [Supplementary Material](#)).

The first 45 minutes of the exercise were devoted to the expert calibration, which was achieved through a seed questionnaire (provided in English and French). This questionnaire comprised 16 questions, with the first 14 relating to the topography and population of La Réunion island, these factors being key in influencing lava flow hazard and risk, as well as historical effusive activity at Piton de la Fournaise itself. We remark that the true answers to the seed questions had not been provided within the workshop presentations, at least not directly. This implies that the experts had to perform additional reasoning and account for additional uncertainties, which is the core of our elicitation exercise. The last two questions were based on more general topics on effusive volcanism and remote sensing,

so that the experts had to provide their judgements on questions for which the pre-exercise training was not sufficient (see “Seed questions”; Data S1 in the [Supplementary Material](#)).

3.1.4 Exercise conditions

Our exercise was held via Zoom, a video conferencing platform. Following standard procedure in expert elicitation [e.g. [Neri et al. 2008](#); [Bevilacqua et al. 2015](#); [Tadini et al. 2021](#)], each of the four phases was introduced by one of the analysts (AT) who read through the bulletin and questions. At the same time the documents relevant to the phase in hand (the bulletin, question sheet, map supplements) were distributed by upload to an online shared folder on Google Docs (a web-based collaborative word processing application) and/or the discussion channel on Zoom as an attachment. Questionnaires were returned 20 minutes after their “release” by email. The next phase was then initiated immediately with no pause or break. The experts were given a short period of time to read the bulletins, ingest the information and respond to the three questions so as to simulate a high pressure, stressful, rapidly evolving (by the hour) crisis. Experts were asked to respect closed book conditions and not use web-based search engines: this indication was furthermore enhanced by highlighting that this exercise was not an assessment of each expert’s knowledge, and that respecting such conditions is fundamental to give credibility to the results. This format allowed the five analysts (AT, AH, JM, AB, and AP) to work efficiently with a large group (28 experts in nine different countries) over a period of two hours. The analysts, who were in Clermont-Ferrand (AT and AH), Cambridge (JM), Pisa (AB), and La Réunion (AP), were also available to manage queries, but did not give information that would have biased responses. Exercise output was then processed using Andury1 [[Pieter ’t Hart et al. 2019](#)] over the ensuing 18 hours, discussed among the analysts and then presented to the expert group for open floor discussion and feedback*.

3.2 Experts involved

For this exercise, the elicited experts also participated in the “EFFUSIVE CRISIS RESPONSE VIRTUAL WORKSHOP” which was held from April 12 and 15, 2021, and therefore the minimum level of expertise required for participating in the exercise was already met considering the experts’ backgrounds. As standardized approach, experts were not involved in the design of the elicitation exercise, which was the task of the analyst(s) and problem owner(s) [[Quigley et al. 2018](#)].

Along with standard demographic data (name/surname, contact, gender, age, original na-

*see “Workshop_Program” in the [Supplementary Material](#); <https://www.youtube.com/channel/UC3E3EDtkytZsFTSnPHQ5sQg>

tionality, professional position, country of current position), experts had to:

- provide information on their years of relevant experience, number of volcanic crises they had been involved with, and previous experience of expert elicitation;
- rate, on a scale from 1 (non-existent) to 10 (excellent) both their perceived level of expertise in dealing with a volcanic crisis, and their level of knowledge on Piton de la Fournaise;
- answer three ‘test questions’ designed to assess the degree of certainty or uncertainty they would perceive as acceptable if they had to provide binary answers during a volcanic crisis (i.e. “yes” or “no”).

The information described in the following subsections was used to assess the uniformity of the group with respect to gender/experience/provenance, to highlight possible sub-groups that could be analysed separately, and to introduce the experts how dealing with probabilities and their translation from qualitative assessments.

3.2.1 Group composition

Among the thirty-one experts initially involved, twenty-eight experts completed the elicitation exercise ([Figure 2](#)) and three had to leave due to different issues (connection problems, ongoing civil protection emergency, and volcanic crises). The gender balance was 39 % female ($n = 11$) for 61 % male ($n = 17$), and a broad spectrum of age groups was represented ([Figure 2](#)).

Academic staff largely dominated the group composition, with 50 % being permanent university staff (researchers, lecturers, and professors: $n = 14$). PhD students comprised ~18 % ($n = 5$) and ~3 % were Post-Docs ($n = 1$). Volcano Observatory staff represented ~14 % of the experts ($n = 4$, two of whom also indicated university staff as second affiliation), and Civil Protection Officers represented ~11 % ($n = 3$). In addition, one journalist from a local newspaper—now retired but with more than 30 years of experience in covering eruptions at Piton de la Fournaise—also participated (counted as “others”). French institutes were overrepresented ([Figure 2](#)), making up 50 % of the participants ($n = 14$). The other experts were from American ($n = 3$), Italian ($n = 3$), Ecuadorian ($n = 2$), Spanish ($n = 2$), British ($n = 1$), German ($n = 1$), Portuguese ($n = 1$), and Swiss ($n = 1$) institutes.

3.2.2 Experience and perceived levels of expertise

The group was comprised of 18 % PhD students with no experience in crisis management ($n = 5$), and 22 % early career scientists, with up to 10 years of experience ($n = 6$). Advanced career experts, with 10 to 20 years of

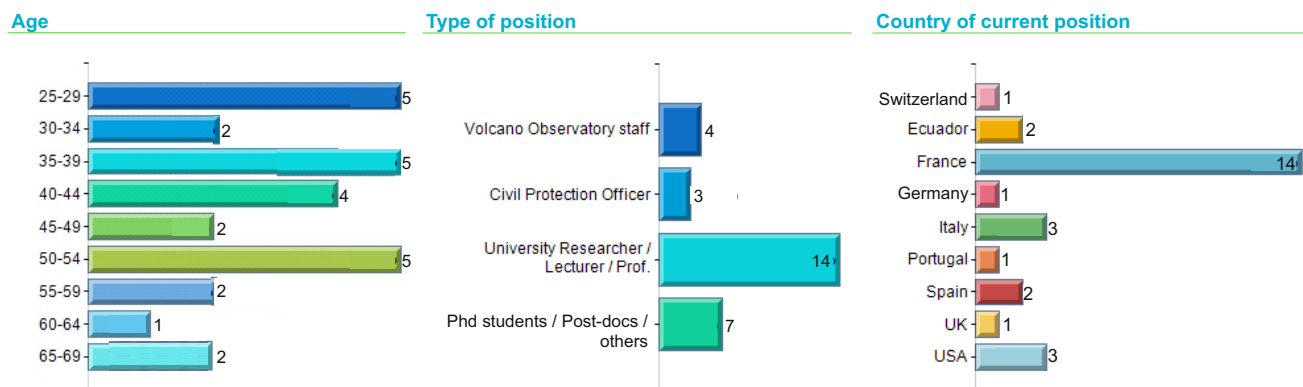


Figure 2: Group composition sorted by age, type of position and country of current position.

relevant experience ($n = 7$), made up 25 % of the group, and 36 % were senior experts with up to 40 years of relevant experience ($n = 10$). The experts had experience with between 0 to 88 volcanic crises, with a median of 10. It is worth mentioning that the three Civil Protection officers all had scientific backgrounds and multiple experiences of volcanic crisis management in collaboration with volcano observatory staff. A total of nine experts had already participated in an expert elicitation, eight of them only once, and one ten times. The self-assessed levels of expertise in dealing with a volcanic crisis, and of knowledge on Piton de la Fournaise, were asked of the participants on a scale from 1 (non-existent) to 10 (excellent). Results are broadly distributed with median values of, respectively, 6 (expertise) and 5.5 (knowledge).

The answers provided by some experts regarding their involvement in volcanic crises (“more than x years” instead of precise numbers) did not allow correlations between the various “experience” markers to be established. We have therefore created five classes associated to numerical values: (1) no experience, (2) 1 to 2 crises, (3) 3 to 5 crises, (4) 6 to 10 crises, and (5) more than 10 crises. The correlation matrix for the variables relating to the participants expertise is available in the [Supplementary Material](#) (“DataS1.zip”). There is only one strong correlation between the number of volcanic crises experienced and the self-assessed level of expertise in dealing with volcanic crises ($r = 0.83$). No other clear link between variables is evident.

3.2.3 Level of confidence required to provide binary answers during a volcanic crisis

Three “test questions” were included in the participants’ demographic sheets with two aims: a) introducing the experts to the idea of probability in their answers and b) understanding the level of confidence the experts need to be able to provide a binary—“yes” or “no”—response during a volcanic crisis. Binary data might be the type of information required by non-scientific audiences dealing with adverse events; however, volcanic crises are by essence highly uncertain

[Harris 2015b; Newhall and Pallister 2015; Donovan 2019]. It is thus interesting to see whether the experts would be reluctant or willing to provide binary or “unsure” responses given a certain degree of uncertainty, and if there is a link with the experts’ informativeness during the expert elicitation. The experts had to provide a lowest value and a highest value for each test question. The second test question (B), referring to the degree of uncertainty on the binary response “yes” to the question: “will there be an eruption by tomorrow?” was ambiguously formulated and resulted in inconsistent interpretations from about half the participants; it is therefore not reported here. The other two questions were:

- Test question A: What level of certainty would you need to provide the binary response “no” to the question (i.e. there is no chance of the event occurring): will there be an eruption by tomorrow (question posed at mid-day)?
- Test question C: What level of uncertainty you would need to provide an “unsure” response (i.e. the event may or may not happen) to the question: will there be an eruption by tomorrow (question posed at mid-day)?

Most of the experts must be at least 80% certain of their answer to be able to say that *no* eruption will happen by tomorrow (Question Test A, [Figure 3](#)). Quantifying the uncertainty associated with the word “unsure,” however, is more difficult given much larger ranges of uncertainty; i.e. 0 to 100 % (Question test C, [Figure 3](#)).

4 RESULTS

All graphs summarizing the experts’ responses to the seed questions are provided in Data S2 in the [Supplementary Material](#) along with the results for each target question. The resulting weights assigned to each expert according to the CM and a graph with the calibration/informativeness scores for both the experts and

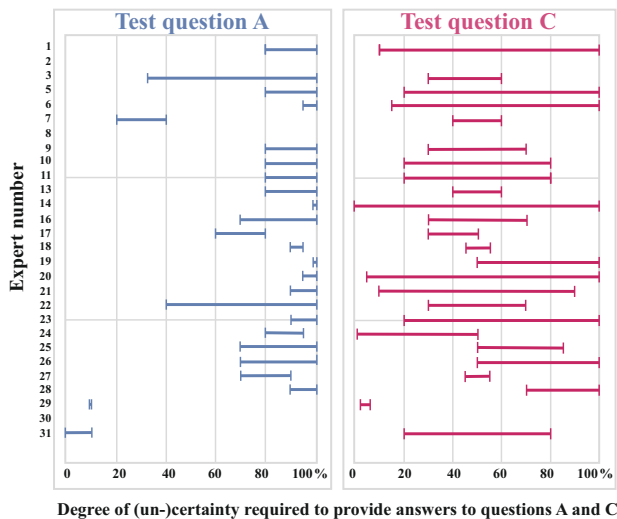


Figure 3: Degree of (un-)certainty required to provide “no/unsure” answers to the question “Will there be an eruption by tomorrow?”

the Decision Makers (CM and EW) are provided in [Table A1](#) and [Figure A1](#), respectively, of [Appendix A](#). In obtaining these weights, we decided to exclude one seed question from our analysis (Q15 of “Seed questions” in Data S1 in the [Supplementary Material](#)). To give this decision its context, in [Figure 4](#) we report an example of one seed question (Q9) for which the group generally performed well collectively, as is apparent from a relatively tight clustering of the results around the realizations. This contrasts with Q15 ([Figure 4](#)) for which the responses are highly scattered and related uncertainties extend well beyond the ranges of all other seed questions. This could be related to the fact that Q15 possibly represents an extreme case too far from the study case of Piton de la Fournaise. We thus excluded the responses for this question, improving significantly the group’s overall performance, while the remaining set of seed questions preserved a good statistical basis for performance weighting. Excluding a few problematic seed items—if, for instance, they evince gross ambiguity or fail to provide any scoring differentiation between experts—is normal practice where the analyst assesses the validity of an elicitation [[Bevilacqua et al. 2015](#); [Tadini et al. 2021](#)].

This elicitation addressed a volcanological problem attended by minimal data, big uncertainties, and divergent judgements. Full Classical Model optimization (cut-off threshold to 43 %) found only two experts would achieve real weights for contributing to the CM Decision Maker. This reflects a quite stringent P-value significance level for accepting the judgements of individual experts and identifying only two scoring experts at this level represents a small minority of the twenty-eight participants.

Instead, we chose to set the cut-off threshold to 1 %, a more accommodating statistical accuracy cut-off that

is within reason for its purpose and in line with Classical Model Decision Maker precedents for other difficult scientific problems [[Bevilacqua et al. 2015](#); [Bamber et al. 2019](#); [Cooke et al. 2021](#); [Tadini et al. 2021](#)].

Adopting the 1 % p-value criterion in the present elicitation allows real weights to be ascribed to eight of the twenty-eight participants (note that the two experts noted above naturally retain stronger, substantial weights in their own right). In this way, the judgement burden is spread over more participating experts without greatly sacrificing the statistical accuracy and information gains that a performance-based decision-maker provides over simple equal weighting. To tease out uncertainties, in a very challenging problem like this, we consider this is a matter of ‘good elicitation practice’ and not an absolute matter of pure numerical optimization.

Note that the exclusion of the experts with weights <1 % is not an indictment of their own expertise but reflects the fact that certain other experts had provided judgements that are more statistically accurate and also more informative.

The retained experts represented a good balance between observatory staff/scientists on one hand, and civil protection officers on the other.

4.1 Global

[Figure 5](#) gives the probability density functions for the DM response for all target questions, as derived both from the CM and EW scoring methods. A comparison of these two methods provides a robust assessment that allows us to a) evaluate if there are different but discrete groups of answers (i.e. different “schools of thought”) to any question, and b) highlight possible discrepancies between the best performing experts (CM) and the whole group (EW). We remark that in the CM we are not including DM optimization nor item weights. We also report the percentile values for the CM distributions in [Table 2](#) as well as those of the EW distributions in [Figure A2](#) from [Appendix A](#). From [Figure 5](#) we see that, for most of the questions, there is general consensus between the CM and EW probability density functions. That is, the location of the peaks of the distributions are similar, as are their shapes, although there are differences related to the amplitude of the peaks.

Probability density functions of the CM tend to have higher peaks, being more ‘focused’ around the median value (see for example Q3 from Phase 1 and Q1 from Phase 2; [Figure 5](#)). There are at least four cases (Q1 and Q2 from Phase 1, Q1 from Phase 3 and, partially, Q3 from Phase 1; [Figure 5](#)) in which there are differences among the CM and the EW. We highlight that for Q3 from Phase 1 and Q1 from Phase 3, the differences are limited to a small shift in the location of the main peak of the distribution, which has only minor secondary peaks (both for the CM and EW). Q1 and Q2 from Phase 1, instead, present (especially for the CM) two

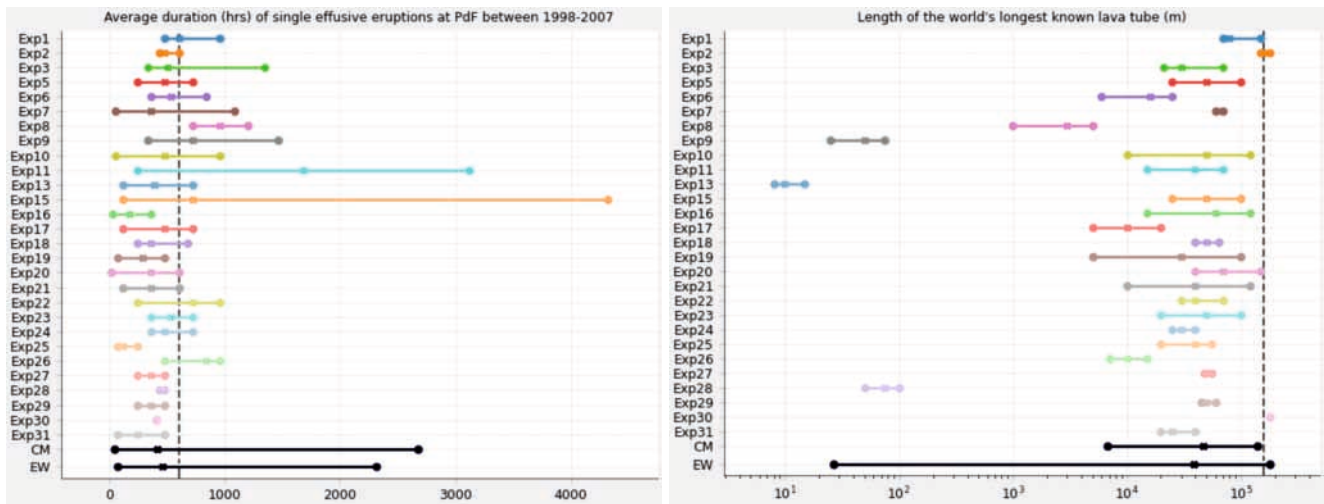


Figure 4: Itemwise graphs for Q9 (left) and Q15 (right) seed questions. “CM” and “EW” are the Decision Maker’s distribution referred to, respectively, as the Classical Model (CM) and Equal Weight (EW) methods. The vertical dashed lines represent the realizations of the seed questions.

well defined peaks (indicating two different schools of thought) and an uncertainty range that is more uniformly distributed (as expressed by the distance between the 5th and 95th percentiles and the relative position of the median; Figure 5). We note that these two questions (along with Q2 in Phase 4) involved giving answers in terms of percentages rather than actual values, as asked in the remaining questions.

4.2 Sub-groups

Analysis of the sub-groups was carried out starting from the demographic survey described in Section 3.2 to identify any possible differences in answers resulting from the differing backgrounds and expectations. Two main sub-groups have been analysed considering these latter features, i.e. that of the scientists (observatory staff, university professors, researchers, Ph.D. students) and that of the civil protection officers. Despite the second group comprising only four experts, their data are still significant because the DM of the whole group used in the CM model is significantly influenced by the civil protection sub-group. We report, in Figures 6 and 7, the probability density functions for the two sub-groups analysed with, respectively, the CM and EW methods. It is interesting to note that each one of the two different peaks evident for the CM of the whole group for Q2 from Phase 1 (Figure 6) is linked mainly to just one of the sub-groups: the upper peak to the university/observatory sub-group, and the lower peak to the civil protection sub-group. This question relates to the probability that the eruption will begin within the next six hours, with the scientists being much more risk-cautious than civil protection, with probabilities having their main peaks at ~90 % and ~20 %, respectively, i.e. at completely different ends of

the probability scale. By “risk-cautious” we mean that experts provided greater probability estimates for the hazardous phenomena.

For all other questions, the differences are less dramatic. For the EW case (Figure 7), major differences among the two sub-groups are evident in Q1 and Q2 of Phase 1, Q1 of Phase 3 and Q2 of Phase 4. Interestingly, Q1 of Phase 1, and Q2 of Phase 4, are also probability questions with the scientists being, again, more risk-cautious than civil protection (Figure 7). Although Q1 of Phase 3 is not probabilistic, it involves travel times for the lava to arrive at a given point and, still, the scientists are more risk-cautious and provide a shorter time.

5 DISCUSSION

This exercise had to be adapted to, and implemented over, an entirely remote format due to the travel restrictions resulting from the SARS-COV-2 pandemic, a situation which posed several challenges to the organization of an elicitation session. Normally, an expert elicitation is run with the experts present [e.g. Neri et al. 2008; Wadge and Aspinall 2014; Bevilacqua et al. 2015; Tadini et al. 2017], although occasionally some have been run (partially or entirely) via remote interrogation of experts [Aspinall and Cooke 1998; Baker et al. 2019; Aspinall et al. 2020; Neal and Anderson 2020; Wiser et al. 2021]. Nevertheless, these conditions provided the opportunity to refine a way of performing expert elicitation that may be useful for cases run in “normal” times. The remote format means that the elicitation involves a large group, distributed across the entire globe. For a crisis where it is important to execute the elicitation as quickly as possible, e.g. at the onset of an eruption or rapidly evolving unrest, this

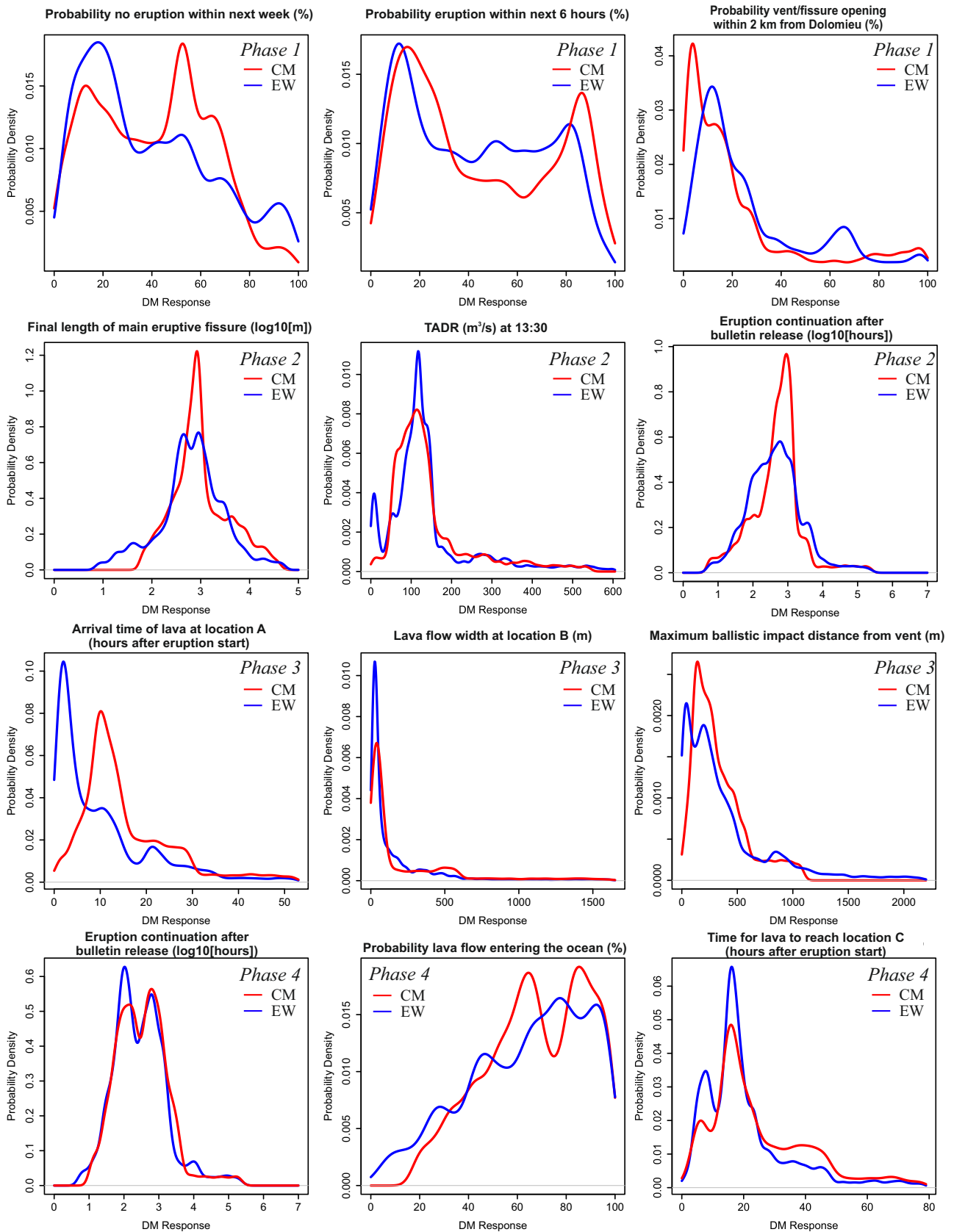


Figure 5: Probability density functions for the decision maker (DM) obtained with the Classical Model (CM—red lines) and Equal Weight (EW—blue lines) pooling methods. For Q1 & Q3 in Phase 2 and Q1 in Phase 4, the x axis is a logarithmic scale.

Table 2: Percentile values for the probability density functions of the CM of Figure 5. In contrast to Figure 5, values of questions 1 and 3 from Phase 2 and question 1 from Phase 4 have been reported here on a linear scale.

Phase	Question	5th	Median	95th
1	Probability of no eruption within next week (%)	4.8	43	78
	Probability of eruption within next 6 hours (%)	6.4	41	90
	Probability of vent/fissure opening within 2 km of Dolomieu (%)	1.3	14	89
2	Final length of main eruptive fissure (m)	213	914	17442
	Time-averaged mass discharge rate at 13:30 ($\text{m}^3 \text{s}^{-1}$)	9.3	118	387
	Eruption continuation after bulletin release (hours)	113	694	3848
3	Arrival time of lava at location A (hours after eruption start)	0.7	6.2	32
	Lava flow width at location B (m)	16	62	1087
	Maximum ballistic impact distance from vent (m)	64	265	876
4	Eruption continuation after bulletin release (hours)	61	478	3731
	Probability lava flow entering into the ocean (%)	30	68	97
	Arrival time of lava at location C (hours after eruption start)	5.1	19	60

means that the group can be virtually assembled immediately and at virtually no cost. In cases where events develop over hours to days, experts might not be able to gather in the same place within a reasonable time frame and/or travel costs for a large group gathering may be prohibitively high. Moreover, recording of the presentations (of both the eruption scenario and the target questions) could be useful in case one expert is forced to answer on a short time delay with respect to the rest of the group (as happened in our exercise).

To set up a performance-based elicitation assessment to be used in a real-time case study, it is first necessary to identify the experts involved in a specific scenario. As a rule-of-thumb, there should be at least 10 experts to guarantee statistically meaningful results [Aspinall 2006]. All experts should have basic background on the volcano and on the volcanic process(es) involved, and this list might be updated year by year. It might be good practice to include one or two experts, who are recognized experts in the field (e.g. effusive volcanism), but that lack of deeper knowledge on the volcano under scrutiny, to provide an “external” point of view. Ideally the calibration phase involving the seed questionnaire should be performed once or twice per year before the volcanic crisis, during periods of quiescence, and the pre-generated experts’ scores should be used during the volcanic crisis itself. This implies that the seed questionnaires should be repeated periodically, in order to account for possible changes in the performance of each expert, or to include new experts.

Then, key target questions that answer the key crisis-response questions “where,” “when,” and “how/what,” should be designed and prepared in advance. This needs to be done with the input of all stakeholders, civil protection officers, and decision-makers and designed according to their needs and/or gaps in knowledge, in order to minimize ambiguity. Nevertheless, we note that this phase for our exercise was the most

time-consuming and required much iteration and refinement over a preparation period of 12 months. It is worth mentioning that the global pandemic, and the resulting difficulties in knowing the exact number, backgrounds, and expectations of the participants played a role in slowing this phase.

Moreover, we remark that training the experts on the response format is fundamental and should be done carefully. Since only nine experts had already participated in other elicitation (see Section 3.2.3), “hands-on” exercises like the one here presented are useful, if repeated periodically, to provide a robust training for the experts. In the case of the prolonged Montserrat eruption [Wadge and Aspinall 2014], the scientists involved in the elicitation gained experience from participating in many repeated sessions, and the associated scientific discussions. In the present case, this first elicitation has highlighted some challenges, and thus provides a basis for refining and improving similar exercises in future. This is a strong argument for initiating and repeating elicitation well before a volcano goes critical.

5.1 Group composition and test questions

The group of experts participating in an expert elicitation should be composed of a large and representative group of individuals that have collective expertise, background and knowledge of the problem under investigation [Aspinall 2006; Tadini et al. 2021]. This assures a solid basis for final results and was fully met in our exercise (see Section 3.2.1). It is important to point out here the necessity of including a sufficient number of experts in the elicitation, so that statistical significance of the results is assured in all situations. In operational cases, where a group of experts should be assembled, it is in fact possible that some of the experts (maybe the best-performing ones) are not avail-

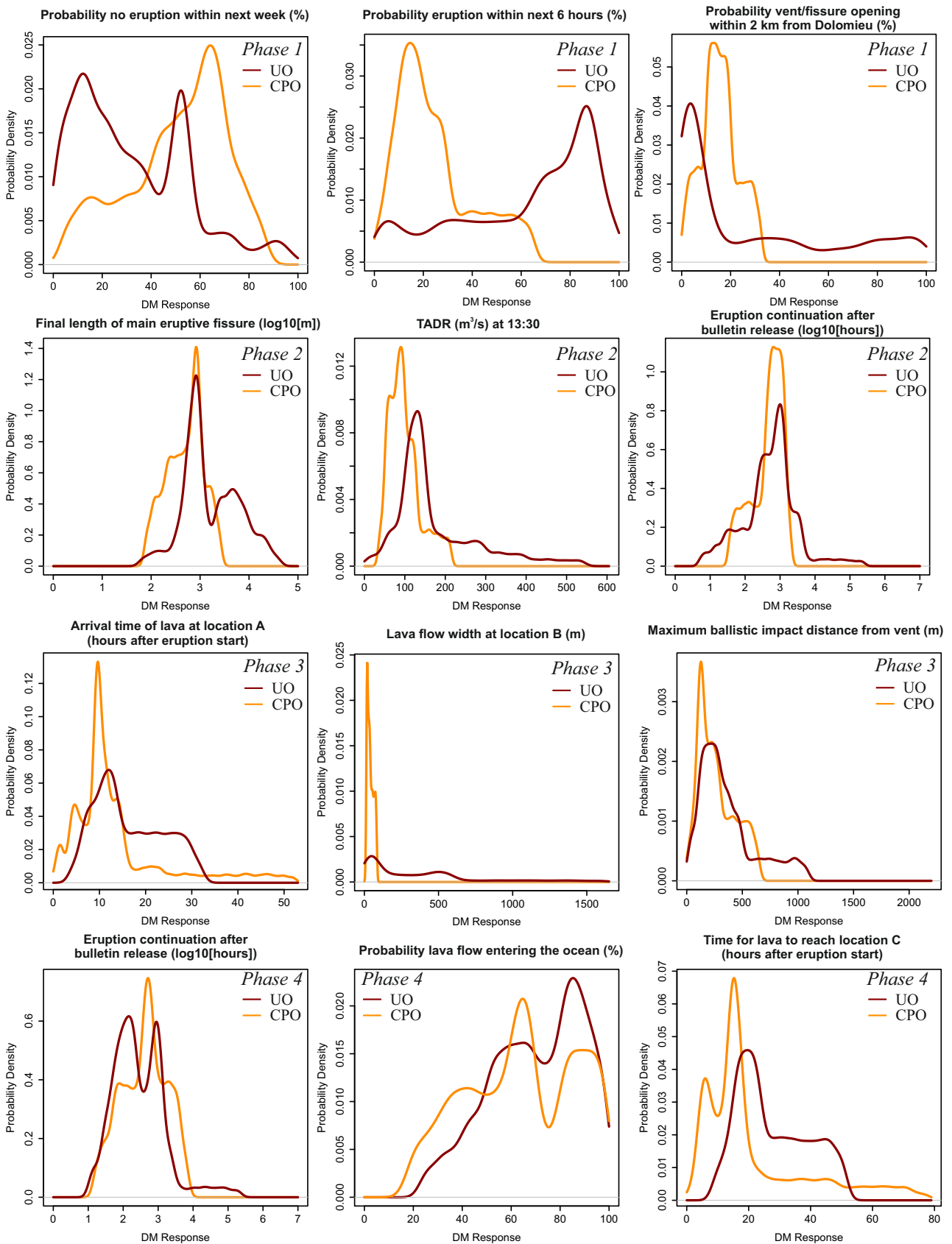


Figure 6: Probability density functions for the decision maker (DM) obtained with the Classical Model pooling method for the “University/Observatory staff” (UO—dark red lines) and “Civil Protection officers/Others” (CPO—orange lines). For Q1 & Q3 in Phase 2 and Q1 in Phase 4, the x axis is a logarithmic scale.

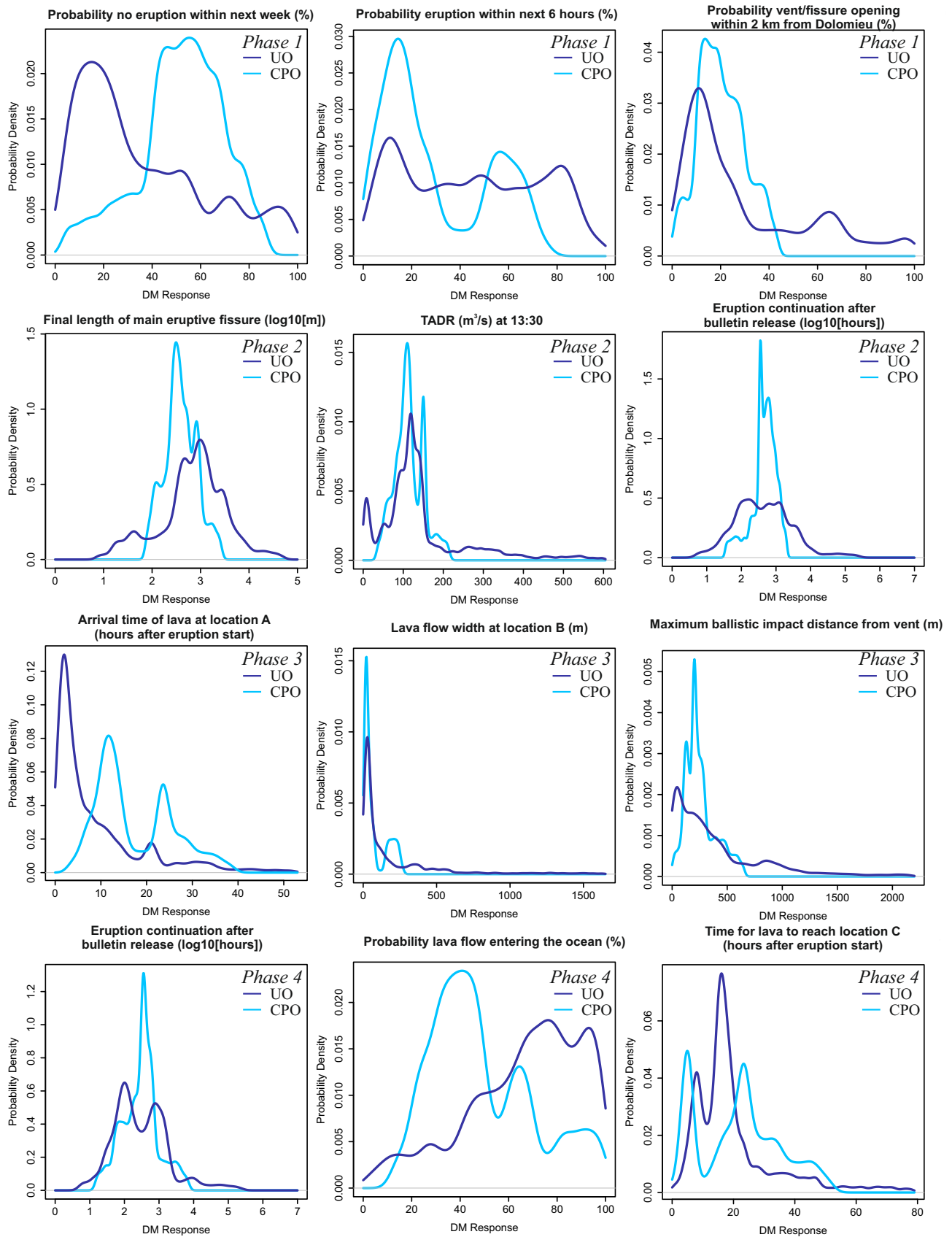


Figure 7: Probability density functions for the decision maker (DM) obtained with the Equal Weight pooling method for the “University/Observatory staff” (UO—dark blue lines) and “Civil Protection officers/Others” (CPO—light blue lines). For Q1 & Q3 in Phase 2 and Q1 in Phase 4, the x axis is a logarithmic scale.

able, and the results of the elicitation could be very different. For instance, if in our case the two best-performing experts (Exp10 and Exp31, see [Table A1](#)) were not present, then the resulting distributions for each question would have differed in the location of peaks and their uni/bi/polymodality (see Data S3 from the [Supplementary Material](#)).

In our approach, we also examined the level of confidence the experts need to be able to provide a specific type of response during a volcanic crisis, by providing three test questions with the demographic sheet (see [Section 3.2.3](#)). We found that this test was partially biased by the difficulties in interpreting the questions, particularly question B. For the other two questions, the group showed a general tendency to require a high level of certainty (i.e. >80 %) to provide a binary answer (Test question A: [Figure 3](#)), which is complementary to the large range of uncertainty associated with “unsure” (Test question C: [Figure 3](#)). These results point at the necessity of providing decision makers with a full description of the uncertainty related to a judgement, rather than a binary “yes/no,” as the required level of confidence for this type of answer is very high and rarely met with volcanic phenomena [cf. [Harris 2015a](#)]. Such assessments of quantitative assignments of qualitative uncertainty statements should be a part of any assessment of probability-based communications between, and within, groups with different backgrounds and expectations [e.g. [Sink 1995](#); [Gigerenzer et al. 2005](#); [Gill 2008](#); [Doyle et al. 2011](#)].

5.2 The “Hors Enclos” scenario

In Phase 1 (i.e. during the seismic crisis) we found consistency in the first two questions in that there is a 44 % median probability (with a 90 % credible interval ranging from 5.5 % to 85 %; [Table 2](#)) that there will be no eruption in the next week, but a 62 % median probability (with a 90 % credible interval ranging from 5.3 % to 90 %; [Table 2](#)) that there will be an eruption in the next six hours. However, there is only an 11 % probability that the eruption will have a vent opening within 2 km of the Dolomieu, revealing a mind-set that eruptions may happen at some distance from the central crater. This likely results from the information reported in the bulletin (i.e. location and distribution of ground deformation and seismicity) but also from the fact that there has been no eruption inside the Dolomieu crater since 2010. However, vent opening beyond the caldera itself may not have been in the groups thoughts as, at the issuance of the next bulletin, one of the experts asked if the tremor map was in error, as the centre of the source was located beyond the map, falling off of the NW corner (see “Phase 2 – Bulletin & Questions” in Data S1 in the [Supplementary Material](#)). As OVPF Scientist-in-Charge, AP answered that it was not in error, but did not add that the location was due to an Hors Enclos source being plotted on the standard tile used by

OVPF-IPGP, which focuses on the Enclos within which all events during the monitoring period to date have been located. In that regard, all eruptions that have happened since the tremor map was implemented have occurred inside the Enclos or very close to the Enclos (1998). However, the sense of the individual’s question revealed their latent expectation that any ensuing event would be inside the map (i.e. inside the Enclos Fouqué).

In Phase 2 (eruption has just begun), the collective view was that the eruptive fissure was most likely be around 1 km long, that effusion rates would remain roughly the same over the ensuing hour, and that the eruption would most likely continue for around three weeks ([Table 2](#)), which was approximately the case for the April 2007 eruption [[Staudacher et al. 2009](#)]. The length estimated for the eruptive fissure (960 m) is that typically associated with eruptive fissures on Piton de la Fournaise [[Soldati et al. 2018](#); [Harris et al. 2019](#)] as well as on Etna, for example during the 2002–03 eruption [[Andronico et al. 2005](#); [Fornaciai et al. 2010](#)]. Interestingly, the 95th percentile estimation of 17 km ([Table 1](#)) is more consistent with the length of dyking events during effusive events at rift-dominated systems such as Krafla and Kīlauea [e.g. [Björnsson et al. 1979](#); [Tryggvason 1984](#); [Dvorak and Dzurisin 1993](#)]; Icelandic and Hawaiian experts gave presentations on these systems during the days preceding the exercise (see “Workshop_Program” from the [Supplementary Material](#)).

In Phase 3, the arrival time for lava at location A was deemed to be between 11 and 35 hours, with the width at location B being on average 48 m, but allowing the possibility it might be up to 754 m wide. Whereas a flow moving at 260 m h⁻¹ [[Staudacher et al. 2009](#)] will reach point A, at a distance of 4.5 km from the vent in 17 hours, 50–750 m is a fairly common value range for channel-fed flow units at Piton de la Fournaise [[Rhéty et al. 2017](#); [Soldati et al. 2018](#); [Harris et al. 2019](#)]. Distances impacted by ballistics were deemed to range between 54 and 1070 m ([Table 2](#)), which is, for instance, consistent with the range of distances attained by bombs on Stromboli during major and paroxysmal explosive eruptions [[Rosi et al. 2006](#); [Gurioli et al. 2013](#); [Rosi et al. 2013](#)].

At Phase 4 (24 hours into the eruption), the anticipated continuation time of the eruption was 22.5 days, consistent with the eruption duration assessed in Phase 2; and there was deemed to be a 6 % median probability (with a 90 % credible interval ranging from 28 % to 90 %; [Table 2](#)) that lava would enter the ocean, thereby cutting through all towns between the vent and the point at which the line of steepest descent arrives at the coast ([Figure 1B](#)). However, arrival time at point C (approximately 11 km from the vent: [Figure 1B](#)) was estimated at 16 hours ([Table 2](#)), implying that—in the experts’ thinking—lava propagation velocity must have increased to 690 m h⁻¹.

The 95th percentile estimation of the duration of the

eruption of four years, as also was the case for the duration asked for in Phase 2 (Table 2), may have been the result of experience, among several of the experts, of recent long-lasting effusive eruptions such as the Pu'u Ō'ō–Kūpaianaha eruption of Kīlauea [Heliker and Wright 1991]. We note that this elicited scenario is based on and influenced by the knowledge and experience of some of the experts.

5.3 Implications for monitoring and reporting

The scenario defined by the experts can lead decision-makers to better think about low-probability scenarios. An immediate consequence of this exercise was that the local civil protection team discussed the potential need for a full-scale evacuation exercise for a scenario with a vent opening outside of the Enclos Fouqué caldera. The bulletin information and supporting material prepared can also help those communicating the information improve or refine the reporting content and format, and the way in which it is presented if the message is not being well received. Likewise, the type of information presented, and the style of presentation, can be modified if the reporting style is not effective in communicating the desired message or if the target audience does not have the background to make the correct interpretations. In the scenario followed here, it was stressed (especially by Observatory staff) that would have been more useful, in terms of interpreting the information, for sub-groups to work and discuss together as is the case in real crises. For example, different outcomes may result from seismologists, geodesists, geochemists, hazard modelers, or risk specialists interpreting each other's product or output without consultation with the information provider or specialist. However, generally the quality of the answers revealed that the format and content of the bulletins, which were based on actual OVPF product, were effective in delivering the desired information to allow scientific and civil protection actors to understand, track and think about the hazard in a correct manner. It also showed that, the content of the bulletins and the knowledge of the users were appropriate.

Our analysis of the sub-groups proves to be extremely valuable especially because the group had a very heterogeneous background. Although scientists and CP officers will not likely always be involved in the same elicitation session and/or asked to provide their judgements on the same type of questions, different sub-groups may still be identified within, for example, a group composed only of scientists [see for example Tadini et al. 2021]. In this case, a sub-group analysis allows highlighting possible differences linked to different reasoning or “schools of thoughts.” Presenting answers derived from more ‘selective’ pooling methods (e.g. the CM) has been used in several real cases [e.g. Wadge and Aspinall 2014]. This is due to the advantage that experts who are statistically better

in estimating uncertainties of known variables are less likely to perform badly for the uncertainty estimations for unknown variables (insights can be found in Cooke et al. [2021]). However, it might be important in some cases to integrate the results of the CM model with those from the EW, or even from different subgroups, at least for some target questions. If there are conflicting schools of thought for certain questions or issues linked to clear misinterpretation or misunderstanding or unclear framing of the questions, then such questions should be re-asked after clarifications. In effect, this approach constructively identifies contentious issues/questions that merit further thought, discussion and knowledge exchange among the experts. That is, to be best implemented in real response mode the elicitation should be executed with an element of interaction, discussion and, above all, aid in interpretation of data provided by specialist elements of the group. However, if, after a discussion among the experts, differences in the answers are not linked to misinterpretations, then this should be highlighted to the decision-makers—it is, in and of itself, an important form of information for reasoned decision support.

5.4 Type of question: percentages versus hard numbers

Working through an eruption crisis simulation can also help experts involved in research and/or monitoring to communicate answers with a clear and appropriate quantification of the uncertainty, in a language that is correctly understood by stakeholders and decision-makers. For our case, we found that the questions that provided answers with the largest uncertainty range were those that involved percentages (i.e. questions from Phase 1 and Q2 from Phase 4). Difficulties can be linked to the translation of qualitative uncertainty evaluations (i.e. ‘likely,’ ‘unlikely’), which are affected by uncertainty, into actual numbers [Sink 1995; Doyle et al. 2011; 2014; Cooke 2015; Harris 2015a]. While it could seem pointless to communicate “44 % median probability of no eruption with an uncertainty range from 5.5 % to 85 %” (as for example Q1 of Phase 1, see Table 2), this information could still be important for decision-makers, because it could help in conveying information in situations for which it is not possible to provide deterministic answers. In other words, while providing answers with uncertainty distributions could be more difficult to be explained to authorities and decision-makers, we think that such an approach is a fair description of the present level of confidence that a group of experts could provide. A follow up study would, then, focus on how best to communicate percent chance in a more effective way in order to deliver the correct uncertainty range [Wallsten et al. 1986; Patt and Schrag 2003; Gill 2008]. In terms of probabilities, the scientists were consistently more risk-cautious than the civil protection experts (i.e. the scientists generally

provided greater probabilities for the hazardous phenomena). This may indicate a differing mind-set when assessing probabilities but may also hint at a communication issue when giving uncertainty in qualitative terms. For example, an event for which the outcome is deemed “likely” or which has a “good chance” of occurring may be deemed to have a probability of 40 % for the risk-cautious assessor, or 85 % for the non-risk-cautious [Wallsten et al. 1986; Patt and Schrag 2003; Gill 2008; Cooke 2015].

5.5 Application for hazard assessment

Expert elicitation exercises can provide information relevant to emergency management, real-time hazard assessments, and response planning [Coppola 2010]. Elicited eruption durations can be used to help civil protection authorities identify and plan for the potential duration of a volcanic crisis, whereas the arrival times or lava flow widths at sensitive sites (e.g. exposed locations A, B, and C of Figure 1), or the area of ballistic impact, provide additional information for decisions about loss, damage, and evacuation. Although experts’ judgments are inevitably affected by human limits (e.g. stress, cognitive biases, availability), such an approach represents the best way to produce answers by consensus, which could gain further robustness if compared with separate estimations. We stress in fact that some of the outputs of this elicitation (e.g. arrival times) could be compared also with the same outputs derived from numerical models, and that the two outputs should not be used as mutually exclusive. While in fact it could be possible that elicitation/model outputs on the same problem could be different [Randle et al. 2019], the comparison between two separate results is key to properly capturing the uncertainty around a question. In this view, comparable results between elicitation/model outputs can provide robustness to both, while disagreements are useful to identify possible flaws or incorrect assumptions in both cases, leading either to an appraisal of counterpart numerical models or to an improvement of the level of expertise of the experts. In the context of Piton de la Fournaise volcano, numerical models are routinely used to perform real-time assessment of lava flow directions, arrival times, and velocities [Harris et al. 2019; Chevrel et al. 2021]; however, experts did not have access to such tools due to the time constraints of the exercise. In short, information regarding the key disaster management decision needs—what, where, and when—can be derived from such elicitation. In the present case, the assessment was set using a group of experts with experience from a variety of effusive centres (e.g. Piton de la Fournaise, Etna, Stromboli, Ecuador, Hawaii) where lava flow related hazards are common and familiar, allowing for an effective assessment that covered a range of real-life experiences.

5.6 Lessons learnt for virtual expert elicitations

Several constructive remarks were made by the experts and by the analysts during the debriefing period. This was a 45-minute open floor discussion the day after the exercise and followed a short presentation of the initial results, as well as completion of questionnaires designed as part of the EUROVOLC program*. These questionnaires were completed online, designed to assess the effectiveness and relevance of such exercises. This was supported by comments in follow-up emails and oral communications to the organizers. Key comments include:

- In a real case it will be necessary to provide a real-time exchange, between elicited experts and observatory staff, with staff members providing clarifications as to their own interpretation of the signals and limitations of the data, as was done for the second bulletin distributed here. For practical reasons, this exchange was not encouraged during this online elicitation exercise as we wanted to avoid potential bias in this particular exercise. In real scenarios, such an exchange should be facilitated and encouraged so that experts participating in real crisis-response elicitation can provide judgments based on an exchange of information. Similar experiences in other expert elicitation [e.g. Hemming et al. 2018] could provide useful insights to improve clarity.

- To provide useful information that could help civil and political authorities to deal with an effusive crisis in near-real time, obtaining and processing the answers from the experts is a task that should be performed as quickly as possible. During this exercise, the forms on which experts provided their answers and the software used to process the data were based on a system originally designed to be used during an in-person meeting. Their use during remote meetings is possible but could be greatly improved by providing experts with an online form that could be filled and then directly loaded into the software used by the analysts. In this way the analysis speed could be greatly increased so that output delivery delays could be greatly reduced, possibly to a few tens of minutes, as opposed to 48 hours as was the case here.

- Observatory staff members, in particular, felt “isolated” and “alone” during this elicitation; when in a normal situation discussion would have been carried out with, for example, the physical volcanology group seeking the opinion of the seismic group over the relevance of each other’s data. In the case here, this was hard because of the Zoom-based platform, with each member being physically alone and isolated. However, group work should be encouraged for a real scenario where expert groups could be set up and arranged (in

*<https://eurovolc.eu/>

separate virtual rooms) with each group involving a mixture of specialists to allow exchange of knowledge.

- Understanding and correctly communicating uncertainties (see, e.g. the problems in the framing of the test questions, [Section 3.2.3](#)) could be challenging even for experienced users [e.g. [Donovan and Oppenheimer 2015](#)]. Exercises like this one are therefore useful also to increase the familiarity of the participants with probability estimation, and dedicated discussions and presentations on such issues could also be useful for future exercises. For new exercises, it would be also beneficial to have a preliminary run through a training set of questions, to get the experts attuned to the elicitation concepts and to the three quantiles formulation in particular.

For this exercise, it was also suggested that 1) it could be useful to have all the data actually available to an observatory during a crisis virtually visible and accessible (at least as snapshots), and 2) some further explanation of the activity of Piton de la Fournaise volcano would have been helpful. This would involve distributing baseline datasets and monitoring data sets (e.g. seismic roll drum read outs, deformation maps, seismic location charts) for typical events, as well as statistics for historical events and available hazard maps. Particularly, it is evident that providing a summary table like [Table 1](#) is important for the experts, since this could help them in developing a statistically-based conceptual model that could allow them to translate a qualitative judgement about a development of the eruptive crisis (e.g. eruption start, lava flow reaching the ocean) into a probability value. For this latter purpose, it is also important to remark that designing appropriate seed questions is not always an easy task, since it is important that they are able to assess the accuracy of the mental reasoning that allows participants them to give probabilities to one-off events. While in our exercise we have tried to address that by asking some test questions (see [Section 3.2.3](#) and [Section 5.1](#)) and by considering some seed questions that required mental reasoning to translate some quantities into probabilities (i.e. Q4, Q6, and Q7), we acknowledge that new seed questions could be envisaged to better capture the above-mentioned criteria. Moreover, when asking experts to judge a “next event” probability at a given volcano the facilitator can suggest to participants that they think about a parallel population of very similar volcanoes, say 100 or even 1000 in number, and then ask themselves how many of the 100 or 1000 they would expect to fulfil the question condition, i.e. to give their judgement as a form of relative frequency (expressing uncertainty on this with the usual three quantiles). In practice, this is equivalent to specifying the “reference class” event and thus allows the (pooled) probability distribution to be operationalizable. For all the above reasons (and also to better analyse the bimodalities of the distributions described in [Section 4](#)), a new elicita-

tion on the same topics would be advisable. Such elicitation are in fact seldom “one-off” definitive outcomes for challenging, data-poor scientific problems. Iterations are generally needed to address and hopefully resolve the most tricky or contentious aspects, especially for safety-critical hazard/risk assessments.

Our exercise also highlighted that an Hors Enclos eruption at Piton de la Fournaise, fed by magma ascending directly from depth and bypassing the summit system, would probably present very different characteristics than the eruptions observed in recent decades. We would not expect, for example, to encounter the same effusion rates as those witnessed during other eruptions. Thus, data for historical eruptions might not be fully comparable to the hypothesized eruption of this exercise. We stress here that the above-mentioned results are thus very much the elicited expectations, based on the present state of volcanological knowledge of this and similar volcanoes, and the experience of such eruptions, accumulated over just a few recent decades. Specialists for Piton de la Fournaise were involved so that the results were tuned to that volcano and its surroundings, as well as effusive events more generally. However, the thinking of the local group may have been somewhat conditioned by collective memories of 81 “typical” eruptions on La Réunion, which they have routinely responded to over the last 40 years. In other words, the judgements of this particular expert panel are likely to have been strongly influenced by, indeed possibly following, routine experience. Importantly, this was pointed out by one local respondent, whereas another local and one non-local respondent suggested that this realization may not be a bad thing: the scenario presented in the exercise forced local participants to recognize that all participants may have to also think outside the “normal” in order to fully assess hazard and risk scenarios at Piton de la Fournaise and advise authorities accordingly.

Finally, we remark that this exercise was an introductory assay of expert elicitation for a group for the majority of whom the procedure was somewhat novel and therefore akin to a demonstration/learning process. Whereas professionally-commissioned elicitation usually have the resources to fully record datasets, discussion, model interpretations, etc., in this case available support was extremely limited and precluded scoping any effort beyond preparing and conducting the exercise.

6 CONCLUSIONS

We set up and tested a structured expert elicitation for assessing volcanic hazards that can be executed via a virtual platform allowing participation of a large, globally distributed group of experts. The system is efficient and, with slight adjustment, could support near-real-time application during a rapidly evolving volcanic crisis. Further improvements in the procedure should include data entry of responses via online forms (to speed

up the production of results), should provide the experts all the baseline information to develop conceptual or statistical models to provide answers, and should consider carefully the design of seed/target questions (to avoid ambiguities). A reliable structured expert judgement (e.g. as done at Monsterrat volcano [Aspinall and Cooke 1998]) is especially useful for “low-probability” events for which there is little or no local experience, memory, or knowledge. Combining the inputs of scientists (including observatory staff) and civil protection actors involved in the crisis allows a range of mind-sets and perspectives to be incorporated into the elicitation, and the findings can be used to assess the differing outcome expectations of the two groups. From our exercise case, we found the civil protection actors to be much less risk-cautious than the scientists, in the sense that for some questions the median values provided by the civil protection actors depicted a more “optimistic” evolution of the crisis (e.g. lower median probability that a seismic crisis could evolve into an eruption). In other cases, where civil protection officers are simply the recipients of the results of a scientific elicitation, analysis can be undertaken to check for any systematic differences in judgement among experts with different scientific backgrounds.

In parallel, utilising standard observatory reports, bulletins, and content to provide the experts with information allows the efficacy and value of such observatory documents to be assessed as a means of delivering information to the core end-users and stake holders in a volcanic crisis. Finally, the same expert elicitation approach can be applied to assess the likelihood of rare, unfamiliar, or extreme volcanic scenarios to raise awareness for, and encourage more thinking about, high-risk events for which memory or knowledge is poor [e.g. Aspinall et al. 2021a]. An important consequence of this exercise is the need for a full-scale evacuation exercise for a scenario at Piton de la Fournaise with a vent opening outside of the Enclos Fouqué caldera. When it comes to eruptions and their hazards, volcanologists are almost always blamed if they fail to advise politicians and decision makers of every plausible scenario, however unlikely; should such an event happen without the decision makers being put “on notice,” the ramifications for the scientists concerned could be potentially dangerous (for a current, on-going case see Cronin [2021]). Structured expert elicitations, like the one we trialled for the particular circumstances of Piton de la Fournaise volcano, offer a formalised basis for volcanologists to estimate the likelihoods and risks of all conceivable eruption scenarios in a fully rational and quantitatively auditable manner. This is important, especially where resources are inevitably very limited (common at most volcanoes) and, in a crisis, almost any quick elicitation is better than none. Future development of this exercise might include another similar exercise performed during a period of intense activity.

ACKNOWLEDGEMENTS

This work was funded by the Agence National de la Recherche (ANR) through project Lava Advance into Vulnerable Areas (LAVA; ANR program: DS0902 2016; project: ANR-16CE39-0009). This is ANR-LAVA contribution n°21.

This contribution is part of the European Commission grant EVE (DG ECHO Ref: 826292).

J.M. received funding from the IMAGINE ERC Grant No 804162 to support the development of this paper.

A.T. was funded by the ClerVolc project - Programme 1 “Detection and characterization of volcanic plumes and ash clouds” funded by the French government ‘Laboratory of Excellence’ initiative. This is ClerVolc contribution n°532. A.T. was also partially funded by the by the French government IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25).

We thank Roger Cooke for fruitful discussion during the revision process. Two anonymous reviewers and Heather Wright are acknowledged for insightful comments that improved the quality of the manuscript. We also thank the editorial handling of Jamie Farquharson.

The manuscript does not necessarily represent official views and policies of both the French and Italian departments of civil protection. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

AUTHOR CONTRIBUTIONS

Conceptualization: A.T., A.H., J.M., A.P., O.C.; **Data curation:** A.T., A.H., J.M., A.B.; **Formal analysis:** A.T., J.M., A.B.; **Funding acquisition:** A.H., O.C.; **Investigation:** P.B., B.B., J.B., A.B.d.S., V.C., O.C., S.C., D.C., H.D., A.D., O.DOR., S.D., O.DUP., L.G., S.K., J.C.K., P.L., D.M., S.M., F.M.A., E.M., S.P., V.R., M.R., N.R., S.V., N.V., S.Z.; **Methodology:** A.B., W.A.; **Project administration:** A.H., O.C.; **Resources:** A.H., A.P., O.C.; **Software:** A.T., J.M., A.B., W.A.; **Supervision:** A.H.; **Validation:** A.T., A.B.; **Visualization:** A.T., A.H., J.M., A.B., A.P.; **Writing – original draft:** A.T., A.H., J.M., A.B., A.P.; **Writing – review & editing:** A.T., A.H., J.M., A.B., A.P., W.A., P.B., B.B., J.B., A.B.d.S., V.C., O.C., S.C., D.C., H.D., A.D., O.DOR., S.D., O.DUP., L.G., S.K., J.C.K., P.L., D.M., S.M., F.M.A., E.M., S.P., V.R., M.R., N.R., S.V., N.V., S.Z.

AUTHOR AFFILIATIONS

α Université Clermont Auvergne, Laboratoire Magmas et Volcans, CNRS, IRD, OPGC, 6 Avenue Blaise Pascal, 63178 Aubière, France.

β University of Cambridge, Department of Geography, Downing Place, CB2 3EN Cambridge, United Kingdom.

γ Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, via Cesare Battisti 53, 56125 Pisa, Italy.

δ Université de Paris, Institut de Physique du Globe de Paris, CNRS, 1 rue Jussieu, 75238, Paris, France.

ε Observatoire Volcanologique du Piton de la Fournaise, Institut de Physique du Globe de Paris, 14 RN3, 97418, La Plaine des Cafres, France.

ζ University of Bristol, School of Earth Sciences, Wills Memorial Building, Queens Road, Bristol, BS8 1RJ, United Kingdom.

η Presidenza del Consiglio dei Ministri, Dipartimento della Protezione Civile, Via Vitorchiano 2, 00189 Roma, Italy.

θ Escuela Politécnica Nacional, Instituto Geofísico, Ladrón de Guevara E11-253 y Andalucía, Quito, Ecuador.

ι Université d'Orléans, ISTO, CNRS, BRGM, UMR 7327, F-45071, Orléans, France.

κ Universidade de Lisboa, Faculdade de Ciências, Departamento de Geologia, Campo Grande 1749-016 Lisboa, Portugal.

λ Università di Torino, Dipartimento di Scienze Mineralogiche e Petrologiche, Via Valperga Caluso, 35, 10125, Torino, Italy.

μ USGS Alaska Volcano Observatory, 4230 University Drive Suite 100, Anchorage, AK, 99508, USA.

ν Geosciences Barcelona (GEO3BCN-CSIC), C/Lluís Solé i Sabarís s/n, 08028 Barcelona, Spain.

ξ Universitat de Barcelona, Facultat de Ciències de la Terra, Departament de Mineralogia, Petrologia y Geologia Aplicada, c/Martí Franquès s/n, 08028 Barcelona, Spain.

ο Etat Major de Zone et de Protection Civil de l'Océan Indien, Protection civile, 50 boulevard Lancastel, 97400 Saint Denis, France.

π Université de la Réunion, Centre de Recherche Juridique (E.A. 14), 15 Avenue René Cassin, 97744 Saint Denis, France.

ρ University at Buffalo, Department of Geology, 126 Cooke Hall, North Campus, Buffalo, NY, 14260, USA.

ς Université de Genève, Département des sciences de la Terre, 13 Rue des Maraîchers, 1205 Genève, Switzerland.

τ Retired from Journal de l'île, 62 boulevard Chaudron, 97490 Saint Denis, France.

υ University of Pittsburgh, Department of Geology and Environmental Science, 4107 O'Hara Street, Pittsburgh, PA 15260-3332, USA.

φ Helmholtz-Zentrum Potsdam, Deutsches GeoForschungsZentrum GFZ, Wissenschaftspark "Albert Einstein", Telegrafenberg, 14473 Potsdam, Germany.

χ Université de La Réunion, Laboratoire GéoSciences Réunion, 97744 Saint Denis, France.

DATA AVAILABILITY

Data supporting the analyses and conclusions presented in this study could be found as [Supplementary Material](#).

COPYRIGHT NOTICE

© The Author(s) 2022. This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

- Andronico, D., S. Branca, S. Calvari, M. Burton, T. Caltabiano, R. A. Corsaro, P. Del Carlo, G. Garfi, L. Lodato, L. Miraglia, F. Murè, M. Neri, E. Pecora, M. Pompilio, G. Salerno, and L. Spampinato (2005). "A multi-disciplinary study of the 2002?03 Etna eruption: insights into a complex plumbing system". *Bulletin of Volcanology* 67(4), pp. 314–330. ISSN: 1432-0819. DOI: [10.1007/s00445-004-0372-8](#).
- Aspinall, W. P. (2006). "Structured elicitation of expert judgement for probabilistic hazard and risk assessment in volcanic eruptions". Ed. by H. M. Mader, S. G. Coles, C. B. Connor, and L. J. Connor. The Geological Society of London on behalf of The International Association of Volcanology and Chemistry of the Earth's Interior. DOI: [10.1144/iavcei001.2](#).
- (2010). "A route to more tractable expert advice". *Nature* 463(7279), pp. 294–295. DOI: [10.1038/463294a](#).
- Aspinall, W. P. and R. Blong (2015). "Volcanic Risk Assessment". *The Encyclopedia of Volcanoes*, pp. 1215–1231. DOI: [10.1016/b978-0-12-385938-9.00070-5](#).
- Aspinall, W. P. and R. M. Cooke (1998). "Expert Judgement and the Montserrat Volcano Eruption". *Proceedings of the 4th international conference on Probabilistic Safety Assessment and Management PSAM4*. Springer New York, pp. 13–18.
- (2013). "Quantifying scientific uncertainty from expert judgement elicitation". *Risk and Uncertainty Assessment for Natural Hazards*. Ed. by J. Rougier, S. Sparks, and L. Hill. Cambridge University Press, pp. 64–99. DOI: [10.1017/cbo9781139047562.005](#).
- Aspinall, W. P., R. S. J. Sparks, C. B. Connor, B. E. Hill, A. Costa, J. C. Rougier, H. Inakura, and S. H. Mahony (2021a). "Aso volcano: estimating the probabilistic likelihood of a future Aso4-scale eruption from stochastic uncertainty analysis of volcanological evidence using importance sampling". *5th International Workshop on Rock Mechanics and Engineering Geology in Volcanic Fields (RMEGV2021)*. Ed. by T. Ito, T. Ohta, and M. Osada. Fukuoka, Japan, p. 6.
- Aspinall, W. P., R. S. J. Sparks, R. M. Cooke, and J. Scarrow (2020). "Quantifying threat from COVID-19 infection hazard in Primary Schools in England". *medRxiv*. DOI: [10.1101/2020.08.07.20170035](#). [Preprint].

- Aspinall, W. P., R. S. J. Sparks, M. J. Woodhouse, R. M. Cooke, and J. H. Scarrow (2021b). “Pupils returning to primary schools in England during 2020: rapid estimations of punctual COVID-19 infection rates”. *Royal Society Open Science* 8(9). ISSN: 2054-5703. DOI: [10.1098/rsos.202218](https://doi.org/10.1098/rsos.202218).
- Aspinall, W. P. and G. Woo (2019). “Counterfactual Analysis of Runaway Volcanic Explosions”. *Frontiers in Earth Science* 7. ISSN: 2296-6463. DOI: [10.3389/feart.2019.00222](https://doi.org/10.3389/feart.2019.00222).
- Baker, E., C. Cruickshank, K. Jenni, and S. Davis (2019). “Comparing in-person and online modes of expert elicitation”. *Mechanical and Industrial Engineering Faculty Publication Series* 620.
- Bamber, J. L., M. Oppenheimer, R. E. Kopp, W. P. Aspinall, and R. M. Cooke (2019). “Ice sheet contributions to future sea-level rise from structured expert judgment”. *Proceedings of the National Academy of Sciences* 116(23), pp. 11195–11200. ISSN: 1091-6490. DOI: [10.1073/pnas.1817205116](https://doi.org/10.1073/pnas.1817205116).
- Barberi, F., F. Brondi, M. L. Carapezza, L. Cavarra, and C. Murgia (2003). “Earthen barriers to control lava flows in the 2001 eruption of Mt. Etna”. *Journal of Volcanology and Geothermal Research* 123(1-2), pp. 231–243. DOI: [10.1016/S0377-0273\(03\)00038-6](https://doi.org/10.1016/S0377-0273(03)00038-6).
- Barberi, F., M. L. Carapezza, M. Valenza, and L. Villari (1993). “The control of lava flow during the 1991–1992 eruption of Mt. Etna”. *Journal of Volcanology and Geothermal Research* 56(1-2), pp. 1–34. DOI: [10.1016/0377-0273\(93\)90048-v](https://doi.org/10.1016/0377-0273(93)90048-v).
- Barons, M. J. and W. P. Aspinall (2020). “Anticipated impacts of Brexit scenarios on UK food prices and implications for policies on poverty and health: a structured expert judgement approach”. *BMJ Open* 10(3), e032376. ISSN: 2044-6055. DOI: [10.1136/bmjopen-2019-032376](https://doi.org/10.1136/bmjopen-2019-032376).
- Bebbington, M. S., M. W. Stirling, S. Cronin, T. Wang, and G. Jolly (2018). “National-level long-term eruption forecasts by expert elicitation”. *Bulletin of Volcanology* 80(6). ISSN: 1432-0819. DOI: [10.1007/s00445-018-1230-4](https://doi.org/10.1007/s00445-018-1230-4).
- Bedford, T., J. Quigley, and L. Walls (2006). “Expert Elicitation for Reliable System Design”. *Statistical Science* 21(4), pp. 428–450. ISSN: 08834237.
- Bevilacqua, A. (2016). “Doubly Stochastic Models for Volcanic Hazard Assessment at Campi Flegrei Caldera”. DOI: [10.1007/978-88-7642-577-6](https://doi.org/10.1007/978-88-7642-577-6).
- Bevilacqua, A., R. Isaia, A. Neri, S. Vitale, W. P. Aspinall, M. Bisson, F. Flandoli, P. J. Baxter, A. Bertagnini, T. Esposti Ongaro, E. Iannuzzi, M. Pistolesi, and M. Rosi (2015). “Quantifying volcanic hazard at Campi Flegrei caldera (Italy) with uncertainty assessment: 1. Vent opening maps”. *Journal of Geophysical Research: Solid Earth* 120(4), pp. 2309–2329. ISSN: 2169-9313. DOI: [10.1002/2014jb011775](https://doi.org/10.1002/2014jb011775).
- Björnsson, A., G. Johnsen, S. Sigurdsson, G. Thorbergsson, and E. Tryggvason (1979). “Rifting of the plate boundary in north Iceland 1975–1978”. *Journal of Geophysical Research* 84(B6), p. 3029. ISSN: 0148-0227. DOI: [10.1029/jb084ib06p03029](https://doi.org/10.1029/jb084ib06p03029).
- Bouidoire, G., M. Liuzzo, A. Di Muro, V. Ferrazzini, L. Michon, F. Grassa, A. Derrien, N. Villeneuve, A. Bourdeu, C. Brunet, G. Giudice, and S. Gurrieri (2017). “Investigating the deepest part of a volcano plumbing system: Evidence for an active magma path below the western flank of Piton de la Fournaise (La Réunion Island)”. *Journal of Volcanology and Geothermal Research* 341, pp. 193–207. ISSN: 0377-0273. DOI: [10.1016/j.jvolgeores.2017.05.026](https://doi.org/10.1016/j.jvolgeores.2017.05.026).
- Branca, S., R. Azzaro, E. De Beni, D. Chester, and A. Duncan (2015). “Impacts of the 1669 eruption and the 1693 earthquakes on the Etna Region (Eastern Sicily, Italy): An example of recovery and response of a small area to extreme events”. *Journal of Volcanology and Geothermal Research* 303, pp. 25–40. ISSN: 0377-0273. DOI: [10.1016/j.jvolgeores.2015.07.020](https://doi.org/10.1016/j.jvolgeores.2015.07.020).
- Branca, S., E. De Beni, D. Chester, A. Duncan, and A. Lotteri (2017). “The 1928 eruption of Mount Etna (Italy): Reconstructing lava flow evolution and the destruction and recovery of the town of Mascali”. *Journal of Volcanology and Geothermal Research* 335, pp. 54–70. ISSN: 0377-0273. DOI: [10.1016/j.jvolgeores.2017.02.002](https://doi.org/10.1016/j.jvolgeores.2017.02.002).
- Branca, S., E. De Beni, and C. Proietti (2013). “The large and destructive 1669 AD eruption at Etna volcano: reconstruction of the lava flow field evolution and effusion rate trend”. *Bulletin of Volcanology* 75(2). ISSN: 1432-0819. DOI: [10.1007/s00445-013-0694-5](https://doi.org/10.1007/s00445-013-0694-5).
- Calvari, S., M. Coltelli, M. Neri, M. Pompilio, and V. Scribano (1994). “The 1991–1993 Etna eruption: Chronology and lava”. *Acta Vulcanologica* 4, pp. 1–14.
- Centre de Recherches et d’Etudes en Géographie de l’Université de La Réunion (CREGUR) (2003a). *Atlas Géographique Informatisé Régional - La Réunion dans le bassin India-Océanique*. <http://www.taglioni.net/atlas/index.htm>.
- (2003b). *Schéma d’Aménagement Régional de La Réunion. Région Réunion*.
- Chester, D. K., A. M. Duncan, C. Dikken, J. E. Guest, and P. H. Lister (1999). *Natural Hazards* 19(1), pp. 29–46. ISSN: 0921-030X. DOI: [10.1023/a:1008001003888](https://doi.org/10.1023/a:1008001003888).
- Chevrel, M. O., M. Favalli, N. Villeneuve, A. J. L. Harris, A. Fornaciai, N. Richter, A. Derrien, P. Boissier, A. D. Muro, and A. Peltier (2021). “Lava flow hazard map of Piton de la Fournaise volcano”. *Natural Hazards and Earth System Sciences* 21(8), pp. 2355–2377. DOI: [10.5194/nhess-21-2355-2021](https://doi.org/10.5194/nhess-21-2355-2021).
- Christophersen, A., N. I. Deligne, A. M. Hanea, L. Chardot, N. Fournier, and W. P. Aspinall (2018). “Bayesian Network Modeling and Expert Elicitation for Probabilistic Eruption Forecasting: Pilot Study for Whakaari/White Island, New Zealand”. *Frontiers in Earth Science* 6. ISSN: 2296-6463. DOI: [10.3389/feart.2018.00211](https://doi.org/10.3389/feart.2018.00211).

- Cioni, R., A. Tadini, L. Gurioli, A. Bertagnini, M. Mulas, A. Bevilacqua, and A. Neri (2020). “Estimating eruptive parameters and related uncertainties for pyroclastic density currents deposits: worked examples from Somma-Vesuvius (Italy)”. *Bulletin of Volcanology* 82(9). ISSN: 1432-0819. DOI: 10.1007/s00445-020-01402-7.
- Connor, C. B. and L. J. Connor (2010). “Estimating spatial density with kernel methods”. *Volcanic and Tectonic Hazard Assessment for Nuclear Facilities*. Ed. by C. B. Connor, N. A. Chapman, and L. J. Connor. Cambridge University Press, pp. 346–368. DOI: 10.1017/cbo9780511635380.015.
- Cooke, R. M. (1991). “Experts in uncertainty: Opinion and subjective probability in science”.
- (2015). “Messaging climate change uncertainty”. *Nature Climate Change* 5(1), pp. 8–10. ISSN: 1758-6798. DOI: 10.1038/nclimate2466.
- Cooke, R. M., D. Marti, and T. Mazzuchi (2021). “Expert forecasting with and without uncertainty quantification and weighting: What do the data say?” *International Journal of Forecasting* 37(1), pp. 378–387. ISSN: 0169-2070. DOI: 10.1016/j.ijforecast.2020.06.007.
- Coppola, D. P. (2010). *Introduction to International Disaster Management*. Butterworth-Heinemann. ISBN: 9780128103623. DOI: 10.1016/c2014-0-00128-1. [2nd edition].
- Cronin, S. (2021). *Whakaari tragedy: court case highlights just how complex it is to forecast a volcanic eruption*. URL: <https://theconversation.com/whakaari-tragedy-court-case-highlights-just-how-complex-it-is-to-forecast-a-volcanic-eruption-161995> (visited on 04/09/2022).
- Dias, L. C., A. Morton, and J. Quigley (2018). “Elicitation: State of the Art and Science”. *Elicitation*. Ed. by L. C. Dias, A. Morton, and J. Quigley. Springer International Publishing, pp. 1–14. ISBN: 9783319650524. DOI: 10.1007/978-3-319-65052-4_1.
- Donovan, A. R. (2019). “Critical volcanology? Thinking holistically about risk and uncertainty”. *Bulletin of Volcanology* 81(4). ISSN: 1432-0819. DOI: 10.1007/s00445-019-1279-8.
- Donovan, A. R. and C. Oppenheimer (2015). “Modelling risk and risking models: The diffusive boundary between science and policy in volcanic risk management”. *Geoforum* 58, pp. 153–165. ISSN: 0016-7185. DOI: 10.1016/j.geoforum.2014.11.005.
- Doyle, E. E. H., D. M. Johnston, J. McClure, and D. Paton (2011). “The communication of uncertain scientific advice during natural hazard events”. *New Zealand Journal of Psychology* 40(4), pp. 39–50.
- Doyle, E. E. H., J. McClure, D. M. Johnston, and D. Paton (2014). “Communicating likelihoods and probabilities in forecasts of volcanic eruptions”. *Journal of Volcanology and Geothermal Research* 272, pp. 1–15. ISSN: 0377-0273. DOI: 10.1016/j.jvolgeores.2013.12.006.
- Dvorak, J. J. and D. Dzurisin (1993). “Variations in magma supply rate at Kilauea Volcano, Hawaii”. *Journal of Geophysical Research: Solid Earth* 98(B12), pp. 22255–22268. ISSN: 0148-0227. DOI: 10.1029/93jb02765.
- Fornaciai, A., B. Behncke, M. Favalli, M. Neri, S. Tarquini, and E. Boschi (2010). “Detecting short-term evolution of Etna scoria cones: a LIDAR-based approach”. *Bulletin of Volcanology* 72(10), pp. 1209–1222. DOI: 10.1007/s00445-010-0394-3.
- Gigerenzer, G., R. Hertwig, E. van den Broek, B. Fasolo, and K. V. Katsikopoulos (2005). ““A 30% Chance of Rain Tomorrow”: How Does the Public Understand Probabilistic Weather Forecasts?” *Risk Analysis* 25(3), pp. 623–629. DOI: 10.1111/j.1539-6924.2005.00608.x.
- Gill, J. (2008). “Communicating forecast uncertainty for service providers”. *World Meteorological Organization Bulletin* 57(4), pp. 237–243.
- Global Volcanism Program (GVP) (2021). “Report on Nyiragongo (DR Congo)”. *Bulletin of the Global Volcanism Network* 46(6). Ed. by E. Venzke. DOI: 10.5479/si.gvp.bgvn202106-223030.
- Gurioli, L., A. J. L. Harris, L. Colò, J. Bernard, M. Favalli, M. Ripepe, and D. Andronico (2013). “Classification, landing distribution, and associated flight parameters for a bomb field emplaced during a single major explosion at Stromboli, Italy”. *Geology* 41(5), pp. 559–562. ISSN: 0091-7613. DOI: 10.1130/g33967.1.
- Hanea, A. M., V. Hemming, and G. F. Nane (2021a). “Uncertainty Quantification with Experts: Present Status and Research Needs”. *Risk Analysis* 42(2), pp. 254–263. DOI: 10.1111/risa.13718.
- Hanea, A. M., G. F. Nane, T. Bedford, and S. French, eds. (2021b). *Expert Judgement in Risk and Decision Analysis*. Springer International Publishing. DOI: 10.1007/978-3-030-46474-5.
- Harris, A. J. L. (2015a). “Forecast communication through the newspaper Part 1: Framing the forecaster”. *Bulletin of Volcanology* 77(4). ISSN: 1432-0819. DOI: 10.1007/s00445-015-0899-x.
- (2015b). “Forecast communication through the newspaper Part 2: perceptions of uncertainty”. *Bulletin of Volcanology* 77(4). ISSN: 1432-0819. DOI: 10.1007/s00445-015-0902-6.
- Harris, A. J. L., M. O. Chevrel, D. Coppola, M. Ramsey, A. Hrysiewicz, S. Thivet, N. Villeneuve, M. Favalli, A. Peltier, P. Kowalski, A. Di Muro, J. Froger, and L. Gurioli (2019). “Validation of an integrated satellite-data-driven response to an effusive crisis: the April–May 2018 eruption of Piton de la Fournaise”. *Annals of Geophysics* 61(Vol 61 (2018)). ISSN: 1593-5213. DOI: 10.4401/ag-7972.
- Harris, A. J. L., J. B. Murray, S. E. Aries, M. A. Davies, L. P. Flynn, M. J. Wooster, R. Wright, and D. A. Rothery (2000). “Effusion rate trends at Etna and Krafla and their implications for eruptive mecha-

- nisms". *Journal of Volcanology and Geothermal Research* 102(3-4), pp. 237–269. doi: [10.1016/s0377-0273\(00\)00190-6](https://doi.org/10.1016/s0377-0273(00)00190-6).
- Harris, A. J. L., N. Villeneuve, A. Di Muro, V. Ferrazzini, A. Peltier, D. Coppola, M. Favalli, P. Bachèlery, J.-L. Froger, L. Gurioli, S. Moune, I. Vlastélic, B. Galle, and S. Arellano (2017). "Effusive crises at Piton de la Fournaise 2014–2015: a review of a multi-national response model". *Journal of Applied Volcanology* 6(1). ISSN: 2191-5040. doi: [10.1186/s13617-017-0062-9](https://doi.org/10.1186/s13617-017-0062-9).
- Heliker, C. and T. L. Wright (1991). "The Pu'U 'O'O-Kupaianaha eruption of Kilauea". *Eos, Transactions American Geophysical Union* 72(47), pp. 521–521. ISSN: 0096-3941. doi: [10.1029/90eo00372](https://doi.org/10.1029/90eo00372).
- Hemming, V., M. A. Burgman, A. M. Hanea, M. F. McBride, and B. C. Wintle (2018). "A practical guide to structured expert elicitation using the IDEA protocol". *Methods in Ecology and Evolution* 9(1). Ed. by B. Anderson, pp. 169–180. ISSN: 2041-210X. doi: [10.1111/2041-210x.12857](https://doi.org/10.1111/2041-210x.12857).
- Hincks, T. K., J.-C. Komorowski, S. R. Sparks, and W. P. Aspinall (2014). "Retrospective analysis of uncertain eruption precursors at La Soufrière volcano, Guadeloupe, 1975–77: volcanic hazard assessment using a Bayesian Belief Network approach". *Journal of Applied Volcanology* 3(1). ISSN: 2191-5040. doi: [10.1186/2191-5040-3-3](https://doi.org/10.1186/2191-5040-3-3).
- Institut national de la statistique et des études économiques (INSEE) (2018). *Le recensement de la population 2018*.
- Komorowski, J.-C., D. Tedesco, M. Kasereka, P. Allard, P. Papale, O. Vaselli, J. Durieux, P. Baxter, M. Halbwachs, B. Baluku, P. Briole, J. Dupin, D. Garcin, H. Hamaguchi, N. Houlie, K. Kavotha, M. Kibuye, A. Lemarchand, J. Lockwood, Y. Mapendano, C. Mateso, A. Mbilizi, M. de Michele, F. Munyololo, C. Newhall, L. N'Yombo Watuk, E. Osodundu, J. Ruch, M. Serugendo, M. Tuluka, and M. Wafula (2002). "The January 2002 eruption - The January 2002 flank eruption of Nyiragongo Volcano (Democratic Republic of Congo): chronology, evidence for a tectonic rift trigger, and impact of lava flows on the city of Goma". *Acta Vulcanologica* (1/2). ISSN: 11219114. doi: [10.1400/19077](https://doi.org/10.1400/19077).
- Macdonald, G. A. and R. H. Finch (1950). "The June 1950 eruption of Mauna Loa, Part II. Special Features". *The Volcano Letter* 509, pp. 1–6.
- Macdonald, G. A. (1962). "The 1959 and 1960 eruptions of Kilauea volcano, Hawaii, and the construction of walls to restrict the spread of the lava flows". *Bulletin Volcanologique* 24(1), pp. 249–294. ISSN: 1432-0819. doi: [10.1007/bf02599351](https://doi.org/10.1007/bf02599351).
- Morin, J. (2012). "Gestion institutionnelle et réponses des populations face aux crises volcaniques: études de cas à La Réunion et en Grande Comore". PhD thesis. Université de la Réunion.
- Neal, C. A., S. R. Brantley, L. Antolik, J. L. Babb, M. Burgess, K. Calles, M. Cappos, J. C. Chang, S. Conway, L. Desmither, P. Dotray, T. Elias, P. Fukunaga, S. Fuke, I. A. Johanson, K. Kamibayashi, J. Kauahikaua, R. L. Lee, S. Pekalib, A. Miklius, W. Million, C. J. Moniz, P. A. Nadeau, P. Okubo, C. Parcheta, M. R. Patrick, B. Shiro, D. A. Swanson, W. Tollett, F. Trusdell, E. F. Younger, M. H. Zoeller, E. K. Montgomery-Brown, K. R. Anderson, M. P. Poland, J. L. Ball, J. Bard, M. Coombs, H. R. Dietterich, C. Kern, W. A. Thelen, P. F. Cervelli, T. Orr, B. F. Houghton, C. Gansecki, R. Hazlett, P. Lundgren, A. K. Diefenbach, A. H. Lerner, G. Waite, P. Kelly, L. Clor, C. Werner, K. Mulliken, G. Fisher, and D. Damby (2019). "The 2018 rift eruption and summit collapse of Kilauea Volcano". *Science* 363(6425), pp. 367–374. ISSN: 1095-9203. doi: [10.1126/science.aav7046](https://doi.org/10.1126/science.aav7046).
- Neal, C. A. and K. R. Anderson (2020). *Preliminary analyses of volcanic hazards at Kilauea Volcano, Hawai'i, 2017–2018*. doi: [10.3133/ofr20201002](https://doi.org/10.3133/ofr20201002). [Open-File Report 2020-1002].
- Neri, A., W. P. Aspinall, R. Cioni, A. Bertagnini, P. J. Baxter, G. Zuccaro, D. Andronico, S. Barsotti, P. D. Cole, T. Esposti Ongaro, T. K. Hincks, G. Macedonio, P. Papale, M. Rosi, R. Santacroce, and G. Woo (2008). "Developing an Event Tree for probabilistic hazard and risk assessment at Vesuvius". *Journal of Volcanology and Geothermal Research* 178(3), pp. 397–415. ISSN: 0377-0273. doi: [10.1016/j.jvolgeores.2008.05.014](https://doi.org/10.1016/j.jvolgeores.2008.05.014).
- Neri, A., A. Bevilacqua, T. Esposti Ongaro, R. Isaia, W. P. Aspinall, M. Bisson, F. Flandoli, P. J. Baxter, A. Bertagnini, E. Iannuzzi, S. Orsucci, M. Pistolesi, M. Rosi, and S. Vitale (2015). "Quantifying volcanic hazard at Campi Flegrei caldera (Italy) with uncertainty assessment: 2. Pyroclastic density current invasion maps". *Journal of Geophysical Research: Solid Earth* 120(4), pp. 2330–2349. ISSN: 2169-9313. doi: [10.1002/2014jb011776](https://doi.org/10.1002/2014jb011776).
- Newhall, C. G. and R. Hoblitt (2002). "Constructing event trees for volcanic crises". *Bulletin of Volcanology* 64(1), pp. 3–20. ISSN: 1432-0819. doi: [10.1007/s004450100173](https://doi.org/10.1007/s004450100173).
- Newhall, C. G. and J. S. Pallister (2015). "Using Multiple Data Sets to Populate Probabilistic Volcanic Event Trees". *Volcanic Hazards, Risks and Disasters*, pp. 203–232. doi: [10.1016/b978-0-12-396453-3.00008-3](https://doi.org/10.1016/b978-0-12-396453-3.00008-3).
- Patt, A. G. and D. P. Schrag (2003). *Climatic Change* 61(1/2), pp. 17–30. ISSN: 0165-0009. doi: [10.1023/a:1026314523443](https://doi.org/10.1023/a:1026314523443).
- Peltier, A., P. Bachèlery, and T. Staudacher (2009). "Magma transport and storage at Piton de La Fournaise (La Réunion) between 1972 and 2007: A review of geophysical and geochemical data". *Journal of Volcanology and Geothermal Research* 184(1–2), pp. 93–108. ISSN: 0377-0273. doi: [10.1016/j.jvolgeores.2008.12.008](https://doi.org/10.1016/j.jvolgeores.2008.12.008).
- Peltier, A., M. O. Chevrel, A. J. L. Harris, and N. Villeneuve (2022). "Reappraisal of gap analysis for effusive crises at Piton de la Fournaise". *Journal of Applied*

- Volcanology* 11(1). ISSN: 2191-5040. DOI: [10.1186/s13617-021-00111-w](https://doi.org/10.1186/s13617-021-00111-w).
- Peltier, A., V. Ferrazzini, A. D. Muro, P. Kowalski, N. Villeneuve, N. Richter, M. O. Chevrel, J. L. Froger, A. Hrysiewicz, M. Gouhier, D. Coppola, L. Retailleau, F. Beauducel, L. Gurioli, P. Boissier, C. Brunet, P. Catherine, F. Fontaine, F. Lauret, L. Garavaglia, J. Lebreton, K. Canjamale, N. Desfete, C. Griot, A. Harris, S. Arellano, M. Liuzzo, S. Gurrieri, and M. Ramsey (2021). “Volcano Crisis Management at Piton de la Fournaise (La Réunion) during the COVID-19 Lockdown”. *Seismological Research Letters* 92(1), pp. 38–52. DOI: [10.1785/0220200212](https://doi.org/10.1785/0220200212).
- Pieter 't Hart, C. M., G. Leontaris, and O. Morales-Nápoles (2019). “Update (1.1) to ANDURIL — A MATLAB toolbox for ANalysis and Decisions with UnceRtaInty: Learning from expert judgments”. *SoftwareX* 10, p. 100295. ISSN: 2352-7110. DOI: [10.1016/j.softx.2019.100295](https://doi.org/10.1016/j.softx.2019.100295).
- Quigley, J., A. Colson, W. P. Aspinall, and R. M. Cooke (2018). “Elicitation in the Classical Model”. *Elicitation*, pp. 15–36. ISSN: 2214-7934. DOI: [10.1007/978-3-319-65052-4_2](https://doi.org/10.1007/978-3-319-65052-4_2).
- Randle, C. H., C. E. Bond, R. M. Lark, and A. A. Monaghan (2019). “Uncertainty in geological interpretations: Effectiveness of expert elicitations”. *Geosphere* 15(1), pp. 108–118. ISSN: 1553-040X. DOI: [10.1130/ges01586.1](https://doi.org/10.1130/ges01586.1).
- Rhéty, M., A. Harris, N. Villeneuve, L. Gurioli, E. Médard, M. O. Chevrel, and P. Bachélery (2017). “A comparison of cooling-limited and volume-limited flow systems: Examples from channels in the Piton de la Fournaise April 2007 lava-flow field”. *Geochemistry, Geophysics, Geosystems* 18(9), pp. 3270–3291. ISSN: 1525-2027. DOI: [10.1002/2017gc006839](https://doi.org/10.1002/2017gc006839).
- Rosi, M., A. Bertagnini, A. Harris, L. Pioli, M. Pistolesi, and M. Ripepe (2006). “A case history of paroxysmal explosion at Stromboli: Timing and dynamics of the April 5, 2003 event”. *Earth and Planetary Science Letters* 243(3-4), pp. 594–606. DOI: [10.1016/j.epsl.2006.01.035](https://doi.org/10.1016/j.epsl.2006.01.035).
- Rosi, M., M. Pistolesi, A. Bertagnini, P. Landi, M. Pompilio, and A. Di Roberto (2013). “Stromboli volcano, Aeolian Islands (Italy): present eruptive activity and hazards”. *The Aeolian Islands Volcanoes*. Ed. by F. Lucchi, A. Peccerillo, J. Keller, C. A. Tranne; and P. L. Rossi. Vol. 37. 1. Geological Society of London, pp. 473–490. DOI: [10.1144/m37.14](https://doi.org/10.1144/m37.14).
- Roult, G., A. Peltier, B. Taisne, T. Staudacher, V. Ferrazzini, and A. Di Muro (2012). “A new comprehensive classification of the Piton de la Fournaise activity spanning the 1985–2010 period. Search and analysis of short-term precursors from a broad-band seismological station”. *Journal of Volcanology and Geothermal Research* 241–242, pp. 78–104. ISSN: 0377-0273. DOI: [10.1016/j.jvolgeores.2012.06.012](https://doi.org/10.1016/j.jvolgeores.2012.06.012).
- Rowland, S. K., A. J. L. Harris, M. J. Wooster, F. Amelung, H. Garbeil, L. Wilson, and P. J. Mouginitis (2003). “Volumetric characteristics of lava flows from interferometric radar and multispectral satellite data: the 1995 Fernandina and 1998 Cerro Azul eruptions in the western Galápagos”. *Bulletin of Volcanology* 65(5), pp. 311–330. ISSN: 1432-0819. DOI: [10.1007/s00445-002-0262-x](https://doi.org/10.1007/s00445-002-0262-x).
- Selva, J., W. Marzocchi, P. Papale, and L. Sandri (2012). “Operational eruption forecasting at high-risk volcanoes: the case of Campi Flegrei, Naples”. *Journal of Applied Volcanology* 1(1). DOI: [10.1186/2191-5040-1-5](https://doi.org/10.1186/2191-5040-1-5).
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Routledge, CRC Press. DOI: [10.1201/9781315140919](https://doi.org/10.1201/9781315140919).
- Sink, S. A. (1995). “Determining the public’s understanding of precipitation forecasts: Results of a survey”. *National Weather Digest* 19(3), pp. 9–15.
- Soldati, A., A. J. L. Harris, L. Gurioli, N. Villeneuve, M. Rhéty, F. Gomez, and A. Whittington (2018). “Textural, thermal, and topographic constraints on lava flow system structure: the December 2010 eruption of Piton de la Fournaise”. *Bulletin of Volcanology* 80(10). ISSN: 1432-0819. DOI: [10.1007/s00445-018-1246-9](https://doi.org/10.1007/s00445-018-1246-9).
- Staudacher, T., L. Ruzié, and A. Peltier (2008). “Historique des éruptions du Piton de la Fournaise de 1998 à 2007”. *Graphica (St Denis, La Réunion)* 97.
- Staudacher, T., V. Ferrazzini, A. Peltier, P. Kowalski, P. Boissier, P. Catherine, F. Lauret, and F. Massin (2009). “The April 2007 eruption and the Dolomieu crater collapse, two major events at Piton de la Fournaise (La Réunion Island, Indian Ocean)”. *Journal of Volcanology and Geothermal Research* 184(1–2), pp. 126–137. ISSN: 0377-0273. DOI: [10.1016/j.jvolgeores.2008.11.005](https://doi.org/10.1016/j.jvolgeores.2008.11.005).
- Tadini, A., A. Bevilacqua, A. Neri, R. Cioni, W. P. Aspinall, M. Bisson, R. Isaia, F. Mazzarini, G. A. Valentine, S. Vitale, P. J. Baxter, A. Bertagnini, M. Cerninara, M. de Michieli Vitturi, A. Di Roberto, S. Engwell, T. Esposti Ongaro, F. Flandoli, and M. Pistolesi (2017). “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”. *Journal of Geophysical Research: Solid Earth* 122(6), pp. 4357–4376. ISSN: 2169-9313. DOI: [10.1002/2016jb013860](https://doi.org/10.1002/2016jb013860).
- Tadini, A., O. Roche, P. Samaniego, N. Azzaoui, A. Bevilacqua, A. Guillin, M. Gouhier, B. Bernard, W. P. Aspinall, S. Hidalgo, J. Eychenne, M. de Michieli Vitturi, A. Neri, R. Cioni, M. Pistolesi, E. Gaunt, S. Vallejo, M. Encalada, H. Yepes, A. Proaño, and M. Pique (2021). “Eruption type probability and eruption source parameters at Cotopaxi and Guagua Pichincha volcanoes (Ecuador) with uncertainty quantification”. *Bulletin of Volcanology* 83(5). ISSN: 1432-0819. DOI: [10.1007/s00445-021-01458-z](https://doi.org/10.1007/s00445-021-01458-z).
- Tazieff, H. (1977). “An exceptional eruption: Mt. Niragongo, Jan. 10th, 1977”. *Bulletin Volcanologique*

40(3), pp. 189–200. issn: 1432-0819. doi: [10.1007/bf02596999](https://doi.org/10.1007/bf02596999).

Tryggvason, E. (1984). “Widening of the Krafla fissure swarm during the 1975–1981 volcano-tectonic episode”. *Bulletin Volcanologique* 47(1), pp. 47–69. issn: 1432-0819. doi: [10.1007/bf01960540](https://doi.org/10.1007/bf01960540).

Tyshenko, M. G., S. ElSaadany, T. Oraby, M. Laderoute, J. Wu, W. P. Aspinall, D. Krewski, and P. R. (2012). “Xenotropic Murine Leukemia Virus-Related Virus as a Case Study: Using a Precautionary Risk Management Approach for Emerging Blood-Borne Pathogens in Canada”. *The Continuum of Health Risk Assessments*. doi: [10.5772/38572](https://doi.org/10.5772/38572).

UN Office for the Coordination of Humanitarian Affairs (OCHA) (2021). *DR Congo: Volcanic eruption in Goma*. URL: <https://reliefweb.int/report/democratic-republic-congo/dr-congo-volcanic-eruption-goma-situation-report-no-1-25-may-2021> (visited on 04/09/2022). [Situation Report No. 1].

Villeneuve, N. (2020). *Une Géographie du massif Piton de la Fournaise – Vision pluridisciplinaire des aspects physiques de l’aléa aux applications sociétales et à la perception du risque*. [Mémoire présenté pour obtenir une Habilitation à Diriger des Recherches].

Villeneuve, N. and P. Bachèlery (2006). “Revue de la typologie des éruptions au Piton de La Fournaise, processus et risques volcaniques associés”. *Cybergeo*. issn: 1278-3366. doi: [10.4000/cybergeo.2536](https://doi.org/10.4000/cybergeo.2536).

Wadge, G. (1984). “Comparison of volcanic production rates and subduction rates in the Lesser Antilles and Central America”. *Geology* 12(9), p. 555. doi: [10.1130/0091-7613\(1984\)12<555:covpra>2.0.co;2](https://doi.org/10.1130/0091-7613(1984)12<555:covpra>2.0.co;2).

Wadge, G. and W. P. Aspinall (2014). “Chapter 24 A review of volcanic hazard and risk-assessment praxis at the Soufrière Hills Volcano, Montserrat from 1997 to 2011”. *Geological Society, London, Memoirs* 39(1), pp. 439–456. issn: 2041-4722. doi: [10.1144/m39.24](https://doi.org/10.1144/m39.24).

Wallsten, T. S., D. V. Budescu, A. Rapoport, R. Zwick, and B. Forsyth (1986). “Measuring the vague meanings of probability terms.” *Journal of Experimental Psychology: General* 115(4), p. 348.

Williams, C. J., K. J. Wilson, and N. Wilson (2021). “A comparison of prior elicitation aggregation using the classical method and SHELF”. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 184(3), pp. 920–940. issn: 1467-985X. doi: [10.1111/rssa.12691](https://doi.org/10.1111/rssa.12691).

Wiser, R., K. Jenni, J. Seel, E. Baker, M. Hand, E. Lantz, and A. Smith (2016). “Expert elicitation survey on future wind energy costs”. *Nature Energy* 1(10). issn: 2058-7546. doi: [10.1038/nenergy.2016.135](https://doi.org/10.1038/nenergy.2016.135).

Wiser, R., J. Rand, J. Seel, P. Beiter, E. Baker, E. Lantz, and P. Gilman (2021). “Expert elicitation survey predicts 37% to 49% declines in wind energy costs by 2050”. *Nature Energy* 6(5), pp. 555–565. issn: 2058-7546. doi: [10.1038/s41560-021-00810-z](https://doi.org/10.1038/s41560-021-00810-z).

Woo, G. (2008). “Probabilistic criteria for volcano evacuation decisions”. *Natural Hazards* 45(1), pp. 87–97. issn: 1573-0840. doi: [10.1007/s11069-007-9171-9](https://doi.org/10.1007/s11069-007-9171-9).

APPENDIX 1: EXPERTS’ WEIGHTS AND EQUAL WEIGHT REALIZATIONS

Table A1: Experts’ weights for both the Classical Model (CM) and the Equal Weight (EW) for the whole group and for the “Scientist/Observatory staff” (SO) and “Civil Protection/Other” (CPO) sub-groups.

Exp	All		Sub-groups		
	CM (%)	EW (%)	Exp	CM (%)	EW (%)
1	0.00	3.57	SO1	0.00	4.17
2	0.00	3.57	SO2	0.00	4.17
3	0.00	3.57	SO3	0.00	4.17
5	9.62	3.57	SO4	17.67	4.17
6	0.00	3.57	SO5	0.00	4.17
7	0.00	3.57	SO6	0.00	4.17
8	0.00	3.57	SO7	0.00	4.17
9	0.00	3.57	SO8	0.00	4.17
10	30.32	3.57	SO9	0.00	4.17
11	14.91	3.57	SO10	0.00	4.17
13	0.00	3.57	SO11	4.08	4.17
15	0.00	3.57	SO12	0.00	4.17
16	0.00	3.57	SO13	16.82	4.17
17	2.21	3.57	SO14	25.38	4.17
18	0.00	3.57	SO15	0.00	4.17
19	9.34	3.57	SO16	0.00	4.17
20	14.10	3.57	SO17	0.00	4.17
21	0.00	3.57	SO18	3.74	4.17
22	0.00	3.57	SO19	0.00	4.17
23	0.00	3.57	SO20	0.00	4.17
24	2.03	3.57	SO21	0.00	4.17
25	0.00	3.57	SO22	0.00	4.17
26	0.00	3.57	SO23	0.00	4.17
27	0.00	3.57	SO24	32.30	4.17
28	0.00	3.57	CPO1	0.00	25
29	0.00	3.57	CPO2	63.69	25
30	0.00	3.57	CPO3	36.31	25
31	17.47	3.57	CPO4	0.00	25

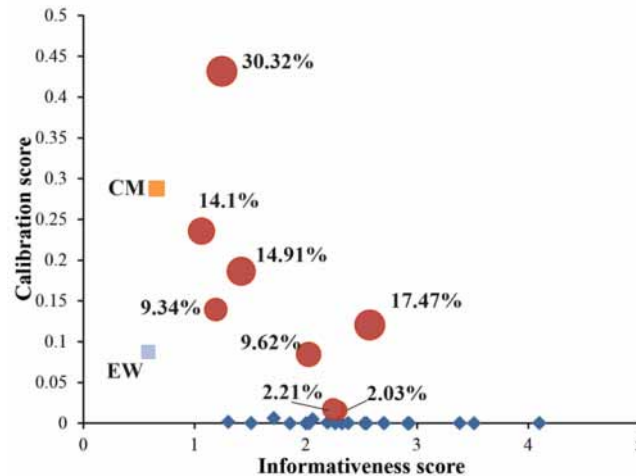


Figure A1: Graph summarizing the calibration/informativeness scores of the experts (red circles and blue diamonds) and the CM and EW Decision Makers (squares). Regarding the CM, blue diamonds are unweighted experts while the size of the red circles is proportional to the expert's weight.

Table A2: Percentile values for the probability density functions of the EW of Figure 7. In contrast to Figure 7, values of questions 1 and 3 from Phase 2 and question 1 from Phase 4 have been reported here on a linear scale.

Phase	Question	5th	Median	95th
1	Probability of no eruption within next week (%)	5.8	35	90
	Probability of eruption within next 6 hours (%)	5.2	42	87
	Probability of vent/fissure opening within 2 km of Dolomieu (%)	3.6	19	80
2	Final length of main eruptive fissure (m)	49	810	9546
	Time-averaged mass discharge rate at 13:30 ($\text{m}^3 \text{s}^{-1}$)	54	126	377
	Eruption continuation after bulletin release (hours)	94	718	33822
3	Arrival time of lava at location A (hours after eruption start)	0.7	5.9	32
	Lava flow width at location B (m)	17	177	1116
	Maximum ballistic impact distance from vent (m)	16	247	1235
4	Eruption continuation after bulletin release (hours)	37	341	30073
	Probability of lava flow entering into the ocean (%)	17	68	97
	Arrival time of lava at location C (hours after eruption start)	5	17	47